JYVÄSKYLÄ STUDIES IN COMPUTING
70

Kari Luostarinen

Resource Management Methods for QoS Supported Networks

JYVÄSKYLÄN YLIOPISTO

JYVÄSKYLÄ STUDIES IN COMPUTING 70

Kari Luostarinen

Resource Management Methods for QoS Supported Networks

Esitetään Jyväskylän yliopiston informaatioteknologian tiedekunnan suostumuksella
julkisesti tarkastettavaksi yliopiston Villa Ranan Blomstedtin salissa
joulukuun 15. päivänä 2006 kello 12.

Academic dissertation to be publicly discussed, by permission of
the Faculty of Information Technology of the University of Jyväskylä,
in the Building Villa Rana, Blomstedt Hall, on December 15, 2006 at 12 o'clock noon.

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2006

Resource Management Methods for QoS Supported Networks

JYVÄSKYLÄ STUDIES IN COMPUTING 70

Kari Luostarinen

Resource Management Methods for QoS Supported Networks

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2006

Editors
Tommi Kärkkäinen
Department of Mathematical Information Technology, University of Jyväskylä
Irene Ylönen, Marja-Leena Tynkkynen
Publishing Unit, University Library of Jyväskylä

URN:ISBN:9513927164
ISBN 951-39-2716-4 (PDF)

ISBN 951-39-2711-3 (nid.)
ISSN 1456-5390

Jyväskylä University Printing House, Jyväskylä 2006

ABSTRACT

Luostarinen, Kari
"Resource Management Methods for QoS Supported Networks"
Jyväskylä: University of Jyväskylä, 2006, 60 p.
Jyväskylä Studies in Computing
ISSN 1456-5390; 70)
ISBN 951-39-2716-4 (PDF), 951-39-2711-3 (nid.)
Finnish summary
Diss.

Quality of Service is important in modern networks for several reasons. Critical applications, such as real-time audio and video, should be given priority over less critical ones, such as file transferring and web surfing. Furthermore, many multimedia applications require delay or delay variation guarantees for acceptable performance. While using the network services, each user is concerned with his own quality of service that he derives from the network. In the absence of novel Connection Admission Control and charging methods, the users try to maximize their usage of the network resources.

Modern wireless and wired networks have to provide support to a number of different types of traffic, each with it's own particular characteristics and Quality of Service parameters including e.g. guaranteed bandwidth, jitter and latency. While the network service provider is optimizing the revenue, rightful service should be given to the customers based on different service classes because the customers of different classes pay different prices to the service provider.

This work studies network QoS, control and management issues concerning wireless and wired network environments. Most of the work is based upon analysis of the different simulations and the corresponding results.

Keywords: QoS, Quality of Service, next generation networks, scheduling, revenue

Author's address	Kari Luostarinen Metso Paper Oy P.O.Box 587, FIN-40101 Jyväskylä, Finland Email: kari.luostarinen@metso.com
Supervisors	Professor Timo Hämäläinen Department of Mathematical Information Technology University of Jyväskylä Finland Professor Jyrki Joutsensalo Department of Mathematical Information Technology University of Jyväskylä Finland
Reviewers	Prof. Alexander Kravchuk Moscow State University, GSP-2, Leninskie Gory, Moscow, 119992, Russian Federation Professor Frantisek Jakab Technical University of Košice Dept. of Computers and Informatics Letná 9, 042 00 Košice, Slovak Republic
Opponent	Dr. Niels Aakvaag ABB Corporate Research Bergerveien 12 N-1375 Billingstad Norway

ACKNOWLEDGEMENTS

This thesis is based on work completed at University of Jyväskylä during 1999-2006 while at the same time working for Metso Paper Inc.

I wish to express my gratitude to professors Timo Hämäläinen and Jyrki Joutsensalo for their co-operation, advice, discussions and help during this work. I also want to thank Jarmo Siltanen of his time and effort during this dissertation project.

This thesis consists of publications that have been published in co-operation with Timo Hämäläinen, Jyrki Joutsensalo, Jarmo Siltanen, Eero Wallenius, Oleg Gomzikov, Lari Kannisto, and Timo Nihtilä.

I would like to give my special thanks to Mr Ted StJohn who has given me help with the English linguistics of my thesis.

Finally, many thanks to my wife Erja and my son Mikael for their patience and support on my dissertation project.

Jyväskylä 1st December 2006

Kari Luostarinen

CONTENTS

ABBREVIATIONS

3G	3^{rd} Generation cellular networks
3GPP	3G Partnership Project
4G	4^{th} Generation cellular networks
ACK	Acknowledgement
AIFS	Arbitrary Interframe Space
AP	Access Point
BER	Binary Error Rate
BSS	Base Station System
CAC	Connection Admission Control
	Call Admission Control
CBR	Constant Bit Rate
CCI	Controlled Contention Interval
CCO	Controlled Contention Opportunity
CFP	Contention Free Period
CN	Core Network
CP	Contention Period
CS	Circuit Switched
CSCW	Computer Supported Cooperative Work
CW	Contention Window
DCF	Distributed Coordination Function
DiffServ	Differentiated Services
DIFS	DCF Interframe Space
DSCP	DiffServ Code Point
EDCF	Enhanced Distributed Coordination Function
EAI	Enterprise Application Integration
ERP	Enterprise Resource Planning
ETSI	European Telecommunications Standards Institute
GCRA	Generic Cell Rate Algorithm
GPRS	General Packet Radio Service
HC	Hybrid Coordinator
HCF	Hybrid Coordination Function
HSRP	Hot Standby Router Protocol
ICT	Information and Communication Technology
IEEE	Institute of Electrical and Electronics Engineers
IMS	IP Multimedia System
IntServ	Integrated Services
IP	Internet Protocol
ITU-T	International Telecommunication Union,
	Telecommunication Standardization Sector
LB	Leaky Bucket
MAC	Medium Access Control
MMPP	Markov Modulated Poisson Process
MPLS	Multi Protocol Label Switching

MSDU	MAC Service Data Unit
NAV	Network Allocation Vector
NE	Network Element
OFDM	Orthogonal Frequency Division Multiplexing
PBC	Policy Based Control
PBM	Policy Based Management
PCF	Policy Control Function
PDP	Policy Decision Point
PF	Persistence Factor
PHB	Per-Hop Behavior
PIFS	PCF Inter Frame Space
PS	Packet Switched
QBSS	QoS-supporting Basic Service Set
QoS	Quality of Service
RED	Random Early Detection
RLC	Radio Link Control
RSVP	Resource Reservation Protocol
RTS	Request To Send
SCFQ	Self Clocked Fair Queuing
SDU	Service Data Unit
SIFS	Short Inter Frame Space
SOA	Service Oriented Architecture
TC	Traffic Category (802.11e)
THP	Traffic Handling Priority
TMN	Telecommunications Management Network
TXOP	Transmission Opportunity
UMTS	Universal Mobile Telecommunication System
UTRAN	UMTS Terrestrial Radio Network
VBR	Variable Bit Rate
VOD	Video On Demand
VRRPP	Virtual Router Redundancy Protocol
VS	Virtual Scheduling
WFQ	Weighted Fair Queuing
Wi-Fi	Wireless Fidelity
WiMAX	Worldwide Interoperability for Microwave Access
WLAN	Wireless Local Area Network

LIST OF PUBLICATIONS

This thesis includes nine research papers, which structurally follow the selected research approach:

[PI] E. Wallenius, T. Hämäläinen, T. Nihtilä, K. Luostarinen, J. Joutsensalo, "3G Interworking with WLAN QoS 802.11e", Proceedings of the IEEE Vehicular Technology Conference (VTC2003), volume 3, 6th – 9th October 2003, Orlando, USA, pp.1803-1806.

[PII] T. Hämäläinen, E. Wallenius, K. Luostarinen, and T.Nihtilä, "End-to-End QoS Issues at the Integrated WLAN and 3G Environments", proceedings of the 9th IEEE Asia-Pacific Conference on Communications (APCC2003), volume 2, 21th -24th September 2003, Penang, Malaysia, pp. 639 – 643.

[PIII] T. Hämäläinen, E. Wallenius, T. Nihtilä, K. Luostarinen, "Providing QoS at the Integrated WLAN and 3G Environments", proceedings of the 14th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'2003), volume 3, 7th-10th September 2003, Beijing, China, pp. 2014-2018.

[PIV]J. Joutsensalo, O. Gomzikov, T. Hämäläinen and K. Luostarinen, "Enhancing Revenue Maximization with Adaptive WRR". Proceedings of Eighth IEEE Symposium on Computers and Communication, (ISCC 2003), Turkey, 2003, pp. 175-180.

[PV] J. Zhang, K. Luostarinen, T. Hämäläinen and J. Joutsensalo, "Enabling QoS Support in Global Replicated Web Systems". In proceedings of the 9th IEEE Asia-Pasific Conference on Communications (APCC2003), volume 2, 21th – 24th September 2003, Penang, Malaysia pp. 735–739.

[PVI] Jyrki Joutsensalo, Timo Hämäläinen, Kari Luostarinen and Jarmo Siltanen, "Adaptive Scheduling Method for Maximizing Revenue in Flat Pricing Scenario", AEU - International Journal of Electronics and Communications, volume 60, issue 2, February 2006, pp. 159-167.

[PVII] J. Joutsensalo, T. Hämäläinen, J. Siltanen, and K. Luostarinen, "Delay Guarantee and Bandwidth Allocation for Network Services", proceedings of the 1st Conference on Next Generation Internet Networks, Italy, April 2005, pp. 129-134.

[PVIII] K. Luostarinen, J. Joutsensalo, L. Kannisto, T. Hämäläinen, "Delay Minimization and Pricing Method for the Network Services", to be

published in ACTA Electrotechnica et Informatica Journal in Košice, on Slovak Republic , 2006.

[PIX] K. Luostarinen, O. Gomzikov, J. Joutsensalo, and T. Hämäläinen, "Delay Guarantee, Bandwidth Allocation, and Pricing in Multinode Network", Proceedings of The IASTED International Conference on Networks and Communication Systems (NCS2006), March 29 th - 31 th, 2006, Chiang Mai, Thailand.

CONTRIBUTIONS OF THE AUTHOR IN COMMON PUBLICATIONS

In publications [PI], [PII] and [PIII] the author has designed test cases. In publication [PI] the author also designed the simulation scenario for the end-to-end quality of service tests. The author also attended to the result analysis. In publication [PIV] the author made the simulation tool and ran the tests. In the above publications, all authors have also attended to the written work.

In [PV], the author designed the used simulation scenario and analyzed the obtained results. In publications [PVI] and [PVII], the first and second author contributed their idea to the algorithm design and all the co-authors participated to the result analyses and written work. In publications [PVIII] and [PIX], the author was main responsible for designing and implementing simulation arrangements. Also, in [PIX] the author contributed to the idea of adaptive scheduling and its extension to the multinode case.

1 INTRODUCTION

Currently a number of industrial companies are moving towards service oriented business operations. In other words, they try to identify new business opportunities by offering different kinds of support services to the customers. These include, among others, maintenance and repairing services, analysis of problems or potential problems, and fine-tuning and modifications for improved performance of the machine. For this need, different kinds of information and communication technology (ICT) solutions, e.g. GroupWare and Computer Supported Cooperative Work (CSCW) applications, knowledge management and expert systems, databases, information hotels etc. have already been developed. [8].

Web technologies enable today information and communication technology supported business processes across enterprise boundaries [9], [57]. However, there are numerous questions to be solved before the information systems of collaborating partners can be set up and made accessible for the exchange of information, independent of the geographic location of the sites of the partners. A core issue in building e-services partnerships is mutual trust [25]. Quality of Service (QoS) is becoming more and more important or even a critical issue as the communication infrastructure is changing to a resource for performing e-business or e-commerce between (industrial) companies.

To build trust in cases where the provisioning of the service requires exchange of data between the business information systems of the partners, one of the first thresholds to overcome is the protection of the information systems and data from unauthorized access. However, remote services feature a promising business opportunity. Monitoring and diagnosis of machinery run with automated controlling systems, the possibility for remotely accessing of the data residing in the information systems of a client organization means a win-win for both parties. Instead of sending staff with maintenance expertise to visit for controls and fault diagnosis on-site, the scarce expertise can be offered without delays and at cheaper total cost by remotely accessing the paper mill site information systems.

1.1 Networked Remote Services at Metso Paper Inc.

Metso Paper is one of the world's leading suppliers of papermaking, tissue, boardmaking and pulping lines. Metso Paper's range of products and services covers the whole process life cycle, from production lines and their rebuilds to various service and maintenance operations. Of Metso Paper's business operations, about one third consists of new lines, one third of rebuilds, and one third of process improvements and other aftermarket services.

Metso Paper has its own operations and production in 28 countries. In addition, its products and services are sold by 24 sales units, more than 40 service centers in different parts of the world and the logistics centers in Finland, the USA and China. There are altogether 12 technology centers in Finland, Sweden, Italy and the USA. [68]

The Service business line consists of networks for mill maintenance units, service centers, logistics centers and remote monitoring centers. The business line is responsible for the continuous development and production of expert services related to the maintenance of customer production lines, processes and equipment. In the initial stage, the remote monitoring centers offers problem solving and instant analysis assistance to contract customers.

Service is a growing business with the bottlenecks of scarce expertise, long distances to the client sites on all continents, and trust: the clients trusting the services provider with access to their ERP (Enterprise Resource Planning) or process automation systems. Remote access of business critical systems by partners means a risk. The challenge is to provide secure ways and means to manage the user data and access control firstly within a geographically distributed enterprise and secondly, ensure for the collaboration safe access to the systems of the client organizations at any location globally. In case of a malfunction of a paper mill, remote access to the control systems enables immediate diagnosis by service experts. This may mean saving even days of downtime for a paper producer, when their loss in the value of the production can be even over 15 000 € per hour when the machine is down. For the services experts this would mean saving days of possibly unnecessary traveling across time zones, and for their employer, improved efficiency of the services function [77].

In conjunction with service agreements, the center is a channel of systematic data collection and analysis. The services of the remote monitoring center may include condition monitoring and maintenance systems as well as production management and process information systems. In addition to customer mill systems, the center can also be connected to customers or Metso Paper's measuring equipment. Video connection and mobile technology allow the transfer of video or still images. Advances in wireless data transfer will in the future open up opportunities for the development of new applications as part of the center's services. Most of the new industrial services are based on intelligent devices, embedded software and automation systems with broadband connections so that remote connections to the paper mills are enabled. Around this core enabling technology, companies have tried to develop solutions for different purposes: e.g. data collection systems, monitoring and maintenance systems, diagnostic and process analysis tools, reporting systems, among tens of others. This kind of fragmented development has resulted in an ICT architecture that forms very complicated combinations. Architectural ICT design (not to mention QoS) has simply been forgotten.

Demand for total service is increasing through continuing specialization. Globalization in the pulp and paper industry, combined with increasing challenges in profitability and return on investments lead to concentration on core competencies. New technologies support this change. An increasing amount of automation and distributed intelligence demand expert service. Communication technology will be integrated with single products, making remote diagnostics more focused and improving reliability dramatically and simultaneously helping to optimize maintenance actions. Communication products make inventory management automatic, simplifying internal processes and further freeing up expertise for the more essential core business.

The demand is changing towards more complete service. Single components and services no longer fulfil the needs of tomorrow's papermaker. Smart investments are based on improved tests and a real-time information flow of needs and available solution between the papermaker and machine supplier. Metso´s Life Cycle Management concept includes all the above mentioned elements and is our way to fulfil our customers future needs and requirements. Applying information and communication technology to information and service logistics is one of the strategic fields of know-how to become a knowledge based technology company. As mentioned above networked services will be increasingly interesting and essential in the future [27].

The nature of the wireless communication has changed quickly. So far assurance of Quality of Service has not been important in wireless networks. Data transfer rate has been small compared to normal wired networks. For example mobile telephone networks have been used mostly to transfer speech and the traffic of Wireless Local Area Networks (WLAN) has mostly been non-time critical like www browsing or reading emails. The sales of wireless local area network devices have grown rapidly and it is predicted that the boom is going to go on [27]. At the same time the use of time critical applications has grown. New generation networks such as GPRS (General Packet Radio Service) and UMTS (Universal Mobile Telecommunication System) are able to transmit moving picture (live video) and music among other things, which demands considerable better QoS in the network [76], [84]. This development leads to higher communication rates and smaller delays and error rates are needed in wireless communication. Quality of Service assurance will be more and more important. Interesting subject has also been the competition for the technology of future local area networks. Studies concerning especially the IEEE 802.11a standard and the ETSI BRAN (Broadband Radio Access Networks) HiperLan/2 standard have been published [26], [65], [70]. The studies concern functioning of these competing technologies in the same network, their efficiency and how good QoS it is possible to offer with these technologies.

ICT environment of the business network has the architecture based on central and site business hubs for the remote maintenance of intelligent paper machines with embedded control and condition monitoring systems (see Figure 1). Paper machines of Metso Paper contain control systems delivered by Metso Automation but in the general case could be operated using control systems of other vendors. The control system of Metso Automation follows enterprise application integration (EAI) technologies using service-oriented architecture (SOA) [78]. Web services and a business process management system are especially in use for condition monitoring purposes of Metso experts. Together with maintenance functions, business hubs provide user management (at the central hub) and message security. With the current solutions Metso assures that access to a customer site is limited to only Metso users, the traffic to a customer site can be filtered at the Metso end, the traffic is logged and user traffic can be traced and reported to a customer.

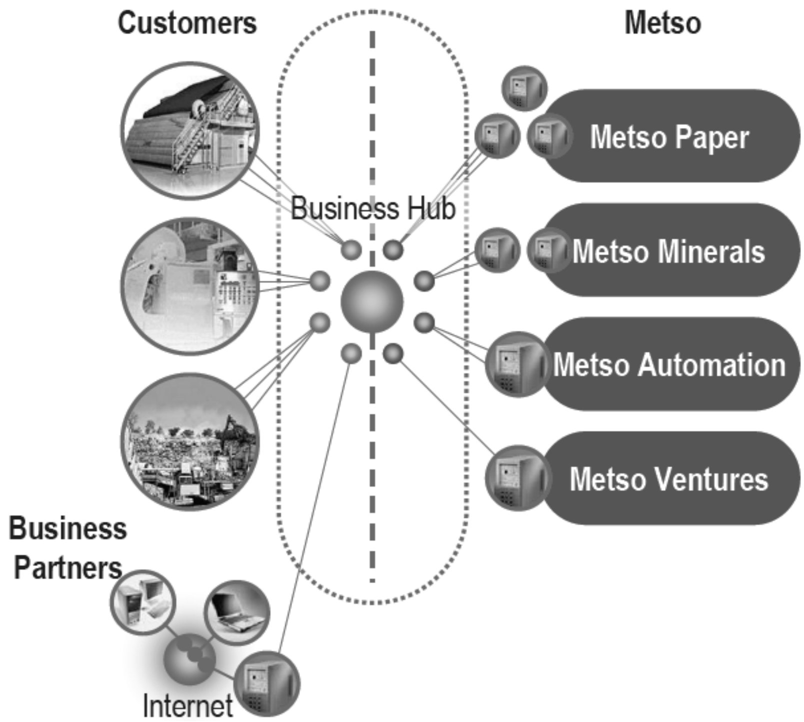

FIGURE 1 Metso's ICT environment [78]

Constant development of communication networks is leading up to several parallel wireless networks in future and the performance is good enough for normal users. The user has freedom to choose case by case which system to use. Anyway, it is not assumed that the end user has an extra terminal for each system and that is why it must be possible to use the systems with one terminal.

The multiple protocol use of the terminals will be important in future. That is why it is important to check up co-operation of different protocols (e.g. WLAN, UMTS, GPRS) and also how the different traffic classes behave in wireless local area networks and how the QoS of the traffic classes is fulfilled.

1.2 Problem Statement

This research work concerns the next generation networks and services and their quality and resource management. Different research problem issues have been discussed in this thesis to find out the best possible understanding of the nature of the problems in next generation networks and service areas. This work has been divided into 3 sections:

- Quality of Service (QoS)
- 3G and WLAN interworking
- revenue-aware resource allocation schemes

One of the main issues that has been raised is the Quality of Service e.g. how to guarantee similar bandwidth, delay, error quality to Packet Switched (PS) connections as Circuit Switched (CS) connections have had.

Quality of network services is the other important matter within next generation networks. Services will be the next big drivers for the network applicability and usage. Most of the services are still on the way but are yet to come, so the service penetration is still quite low and development is slow. The main reason for this is the limited capacity of the connection, leading to unacceptable quality of. The end user equipment properties are also still at a very low level, with too small screens and slow connection rates for proper streaming and video services. The level and restrictions still present with the current technology show clearly we are still along way from such a wireless service environment, whereby subscribers could themselves control the data and information received or sent by their wireless equipment. There are many problem mechanisms that should be solved to ensure QoS requirements in networks. Furthermore, solutions for these problems must work in co-ordination with each other, which means that they have to be configured appropriately. It is a challenging task to configure one certain mechanism because it may have several parameters, the values of which may depend on the configuration of other entities. Failing to provide the correct values may lead to a decreased QoS.

This work aims to simplify the process of finding the optimal configuration for the QoS in wireless and wired network environments. Adaptation of the wired fair queuing algorithms to the wireless domain is non-trivial because of the unique properties in wireless channels such as location dependent and bursty errors, channel contention, and joint scheduling of uplink and downlink flows in a wireless cell. In the past few years, several wireless fair queuing algorithms have been developed [11], [59], [60], [72], [79]. These algorithms provide varying degrees of short-term and long-term fairness, short-term and long-term throughput bounds, average case and worst case delay bounds, and graceful degradation for flows in the presence of channel error. The problem statement of this thesis can be formulated as follows: try to build analytical models that will be capable of adapting working parameters of the scheduling disciplines by taking into account the QoS requirements and a provider's resources.

1.3 Structure of the Thesis

Chapter 2 presents QoS issues at the wireless networks. We will especially focus on IEEE's (Institute of Electrical and Electronics Engineers) 802.11e standard, because it will give mechanisms, which can be used to prioritize traffic flows. In chapter 3 we present our research results. We concentrate on the effects of the air interfaces with errors at the WLAN side to the third-generation (3G) traffic classes We have built an extensive simulation environment where different kind of traffic scenarios can be simulated and analyzed. Related research work has been focused mainly to QoS issues in 3G or WLAN and beyond networks [5], [47], [48], [56], [66]. The simulation results have confirmed the functioning of the models. They are capable of ensuring the QoS guarantees of several service classes with various requirements and pricing schemes. At the same time, free resources are allocated in such a way that the total revenue is maximized. As considered theoretically in chapter 3 and presented in the simulation results, the models can be applied easily to the Differentiated Services (DiffServ) and Integrated Services (IntServ) QoS frameworks.

In chapter 4 we present scheduling method that optimizes the weighted mean delay of the network in a general case. Chapter 5 is the summary of the research results and discusses ideas for future work. Different delays are allocated for different pricing classes, say gold, silver, and bronze classes. Our purpose is to minimize weighted mean delay for connections, where weights are the pricing factors of different classes. An adaptive and optimal solution is derived for weights of the scheduler. Simulations show that in addition to the mean delay minimization, the revenue of the service provider is also maximized in the linear pricing scenario. In addition, adaptive updating rule is simple to implement, since it converges fast. Our scenario is independent on the statistical assumptions, and therefore it is robust against possible erroneous estimates of the customers' behaviour.

2 QUALITY OF SERVICE

Extensive evolution and growth of Internet through the last years has urged the wireless network community to support the deployment of Internet Protocol (IP) multimedia services with guaranteed quality of service in 3G wireless networks. [67]

Wireless networks offer possibility to connect to wired network in a small area e.g. in buildings, stadiums, university campus areas, mass events and ships. Here wireless network means standard IEEE 802.11 or in other words WLAN that is nowadays the most popular local area network standard.

In the wireless local area network all users divide the transmission path (air interface) so that only one user at time can send. Dividing the right to use the transmission path based on the priority makes it easier to offer Quality of Service (QoS) in wireless local area network. In the traditional 802.11 standard offering QoS is not very well implemented, but the standard is under development and nowadays there are several technologies that can offer diverse QoS to different applications [31], [33], [56], [66].

During the past few years, the use of Internet and its technologies, have gone through a significant increase in such extension, that it is now considered the universal communication platform. The Internet is based on the simplicity of the Internet Protocol, which only offers best-effort services typically based on flat-rate and the users can use the network's recourses as much as their connection allows. New multimedia applications with audio and video need much more bandwidth compared to traditional Internet applications such as email or file transfer. The new applications also have a need of widely different operational requirements with the demands of accurate timing of e.g. real-time applications. The lack of any Quality of Service mechanisms has urged the Internet Engineering Task Force (IETF) community to seek QoS support through network layer mechanisms. The most apparent mechanisms are integrated services and differentiated services [34]. The next-generation Internet must exploit these mechanisms, among others, to provide QoS-enabled services. With these services it has to be possible e.g. separate traffic with accurate timing requirements from services which can withstand delay and packet loss and enable different service levels for different applications. This means that the preferable structure of the best effort network model must first be able to provide different Classes of Service (CoS) to indicate the needed service level of individual packets or traffic flows being transported. Then it has to be able to allocate network resources among these service classes supported based on their QoS requirements.

In parallel to the growth of the Internet, wireless and mobile network technologies have also been extensively advanced to the point that mobile and wireless data services are deployed very fast worldwide. In this context, third-generation wireless infrastructures have already adopted IP as the core network (CN) protocol in their data subsystems, and they also promise guaranteed quality for IP multimedia service, in both the access and core networks.

2.1 QoS parameters

The fundamental QoS parameters can be defined as a set of parameters where other sets can always be mapped to [13], [17], [53]. Following this definition, a sufficient set of QoS parameters includes delay, throughput, jitter, packet loss ratio and availability [16], [36], [63].

Delay is the QoS parameter that is critical for the real-time applications. From the user's point of view, end-to-end delay is the measure that matters. Specifically, an end-to-end delay guarantee along a network path can usually be partitioned into the local delay guarantees in each individual network element (NE) across that network path [58], [71], [85].

Throughput specifies the amount of data that an application can send during a time unit without losses. It equals to the bandwidth guarantee of the traffic. The traffic entitled to this bandwidth guarantee can be modelled simply as constant bit rate, or models like leaky buckets can be used to better capture the nature of traffic distribution. As examples of mechanisms to handle throughput, we refer to two equivalent versions for the Generic Cell Rate Algorithm (GCRA) as defined in International Telecommunication Union, Telecommunication Standardization Sector (ITU-T) Recommendation I.371 [35], namely the Virtual Scheduling (VS) and continuous-state Leaky Bucket (LB) algorithm. They are defined for cell based networking, but the algorithms can also be enhanced for packet based networking.

Jitter or delay variation between the packets, is the most critical parameter for the interactive applications such as on-line audio and video conversations. There is a number of ways to define how to calculate jitter. It is quite common to calculate it accumulatively and to have a smoothing equation to give more weight to the latest samples [81], [82].

Packet loss ratio indicates the probability for a packet to get lost in the network. This parameter is critical for those applications that perform guaranteed data delivery because every time a router drops a packet, a sending application has to retransmit it, which results in ineffective bandwidth usage. It is also important for some real-time applications since packet drops reduce the quality of transmitted video and/or audio data [15].

Availability is usually directly related to the services of the network. It has a close relation to reliability, and there are lots of mechanisms to enhance the availability of the Internet, such as router redundancy protocols Virtual Router Redundancy Protocol (VRRP) [52] and Hot Standby Router Protocol (HSRP) [55].

2.2 802.11e QoS support mechanisms

To support QoS, there are priority schemes currently under discussion [80]. IEEE 802.11 Task Group E currently defines enhancements to the above-described 802.11 MAC (Medium Access Control), called 802.11e, which introduces EDCF (Enhanced Distributed Coordination Function) and HCF (Hybrid Coordination Function). Stations, which operate under 802.11e, are called enhanced stations. An enhanced station, which may optionally work as the centralized controller for all other stations within the same QBSS (QoS-supporting Basic Service Set), is called the Hybrid Coordinator (HC). A QBSS is a Base Station System (BSS), which includes an 802.11e -compliant HC and stations. The hybrid coordinator will typically reside within an 802.11e access point (AP). In the following, by stations we mean 802.11e -compliant enhanced stations. With 802.11e, there may still be the two phases of operation within the superframes, i.e., a Contention Period (CP) and a Contention Free Period (CFP), which alternate over time continuously. The EDCF is used in the CP only, while the HCF is used in both phases, which makes this new coordination function hybrid [66].

2.2.1 Enhanced Distributed Coordination Function (EDCF)

The EDCF in 802.11e is the basis for the HCF. The QoS support is realized with the introduction of Traffic Categories (TC 's). Inside WLAN the categories act as some kind of virtual stations that compete for the TXOP (Transmission Opportunity) against other virtual stations and WLAN stations. For example if there is a video application on the station and in the background email is sent from the server, these two applications compete for transmission opportunity inside the station. Competition situation is shown in Figure 2.1.

Transmission opportunity is defined as an interval of time when a station has the right to initiate transmissions, defined by a starting time and a maximum duration. TXOPs are allocated via contention (EDCF-TXOP) or granted through HCF (polled-TXOP). The duration field inside the poll frame limits the duration of an EDCF-TXOP [74].

In the contention period, each traffic category within the stations contends for a TXOP and independently starts a backoff after detecting the channel being idle for an Arbitration Interframe Space (AIFS). The AIFS is at least DIFS (DCF Interframe Space), and can be enlarged individually for each traffic class. Each category has an own AIFS, CWmin (Contention Window minimum) and Persistence Factor (PF). CWmax parameter is optional. Contention Window is an integer within the range of values of the PHY characteristics CWmin and CWmax, that is CWmin ≤ CW ≤ CWmax. SlotTime equals the value of the corresponding PHY characteristics. CW parameter shall take an initial value of CWmin. The CW shall take the next value in the series after each unsuccessful transmission until the CW reaches the value of CWmax. Once it reaches CWmax, the CW shall remain at the value of CWmax until it is reset. This improves the stability of the access protocol under high-load conditions. The CW shall be reset to CWmin after each successful attempt to transmit a packet. The set of CW values shall be sequentially ascending integer powers of 2, minus 1, beginning with a PHY specific CWmin value, and continuing up to CWmax value. The purpose of use of backoff procedure is to reduce the chances of collision by letting stations to select a random number as backoff time.

The class with higher priority class has naturally smallest values of parameters, in other words the backoff window is the smallest. It means that it has the best probability to get the transmission opportunity. The values of the parameters (waiting times) are the bigger the smaller priority the traffic category has.

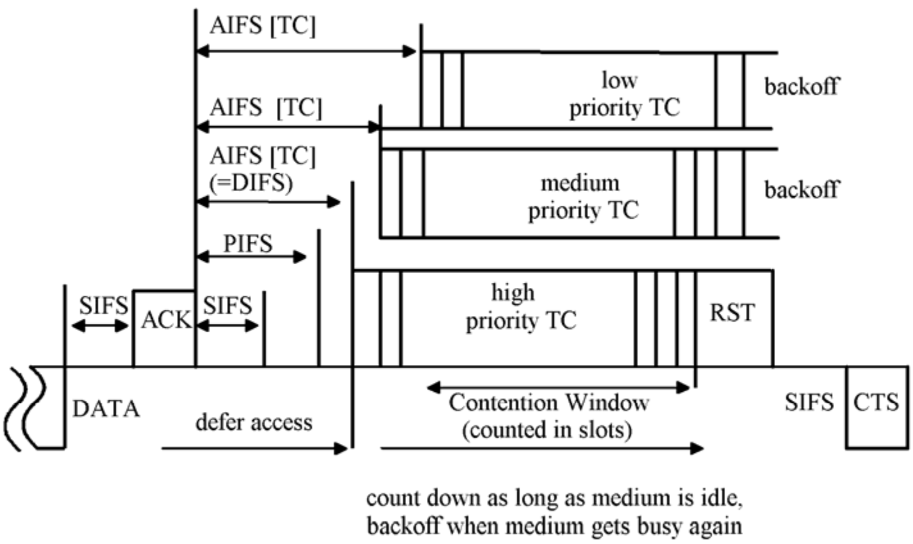

FIGURE 2.1 Transmission opportunity competition of traffic the categories.

After waiting for AIFS, each backoff sets a counter to a random number drawn from the interval [1,CW+1]. Traditional Distributed Coordination Function (DCF) stations are skipped by setting CWmin < 15 and AIFS=DIFS. See Figure 2.1. for illustration of the EDCF parameters.

If the medium, e.g. a channel, is determined busy before the backoff counter reaches zero, the backoff has to wait for the medium being idle for AIFS again, before continuing to count down the counter. A big difference from the legacy DCF is that when the medium is determined as being idle for the period of AIFS, the backoff counter is reduced by one, beginning the last slot interval of the AIFS period. Note that with the legacy DCF, the backoff counter is reduced by one beginning the first slot interval after the DIFS period.

After any unsuccessful transmission attempt a new CW is calculated with the help of the persistence factor PF and another uniformly distributed backoff counter is drawn out of this new, enlarged CW to reduce the probability of a new collision. Whereas in legacy 802.11 CW is always doubled after any unsuccessful transmission (equivalent to PF=2), 802.11e uses the PF to increase the CW different for each TC:

$$newCW \geq [(oldCW + 1) * PF] - 1 \tag{1}$$

The CW never exceeds the parameter CWmax, which is the maximum possible value for CW.

A single station may implement up to eight transmission queues realized as virtual stations inside a station, with QoS parameters that determine their priorities. If the counters of two or more parallel TC's in a single station reach zero at the same time, a scheduler inside the station avoids the *virtual collision*. The scheduler grants the TXOP to the TC with highest priority, out of the TC's that virtually collided within the station, as illustrated in Figure 2.2. There is then still a possibility that the transmitted frame collides at the wireless medium with a frame transmitted by other stations.

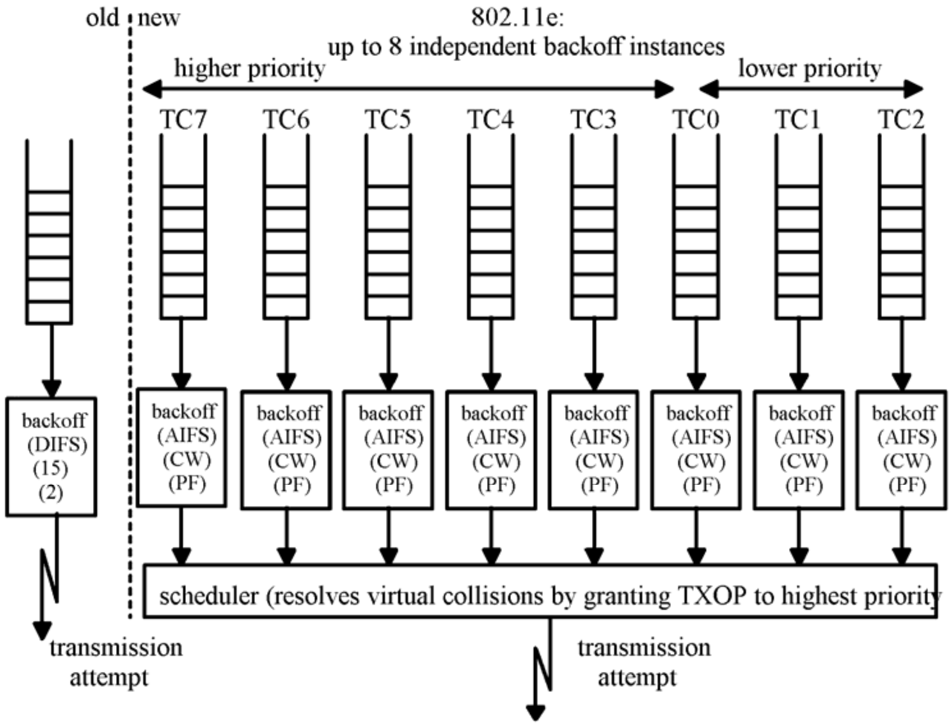

FIGURE 2.2 Virtual backoff of eight traffic categories: (1) left one: legacy DCF, close to EDCF with AIFS=34us, CWmin=15, PF=2; (2) right one: EDCF with AIFS[TC]>=34us, CWmin[TC]=0-255, PF[TC]=1-16.

2.2.2 Hybrid Coordination Function

The HCF extends the EDCF access rules. The hybrid coordinator may allocate TXOP' s to itself to initiate MAC Service Data Unit (MSDU) deliveries whenever it wants, however, only after detecting the channel as being idle for Policy Control Function (PCF) Inter Frame Space (PIFS), which is shorter than DIFS. To give the HC priority over the EDCF, AIFS must be longer than PIFS and can therefore not have a value smaller than DIFS.

During contention period, each TXOP begins either when the medium is determined to be available under the EDCF rules, i.e., after AIFS plus backoff time, or when the station receives a special poll frame, the QoS CF-Poll, from the HC. During the CFP, the starting time and maximum duration of each TXOP is specified by the HC, again using the QoS CF-Poll frames. Stations will not attempt to get medium access on its own during the CFP, so only the HC can grant TXOP' s by sending QoS CF-Poll frames. Although the poll frame is a new frame with 802.11e, also the stations using old standard can set their Network Allocation Vector (NAV) –parameter according to the time from the poll frame.

The CFP ends after the time announced in the beacon frame or by a CF-End frame from the HC. See Figure 2.3. for an example of an 802.11e superframe.

NAV –parameter is used to reduce the hidden station problem with Request-to-Send/Clear-to-Send (RTS/CTS) mechanism. Before transmitting data frames, a station has the option to transmit a short RTS frame, followed by the CTS transmission by the receiving station. The RTS and CTS frames include the information of how long it does take to transmit the next data frame, i.e., the first fragment, and the corresponding acknowledgement (ACK) response. Thus, other stations close to the transmitting station and hidden stations close to the receiving station will not start any transmissions their timer, Network Allocation Vector is set.

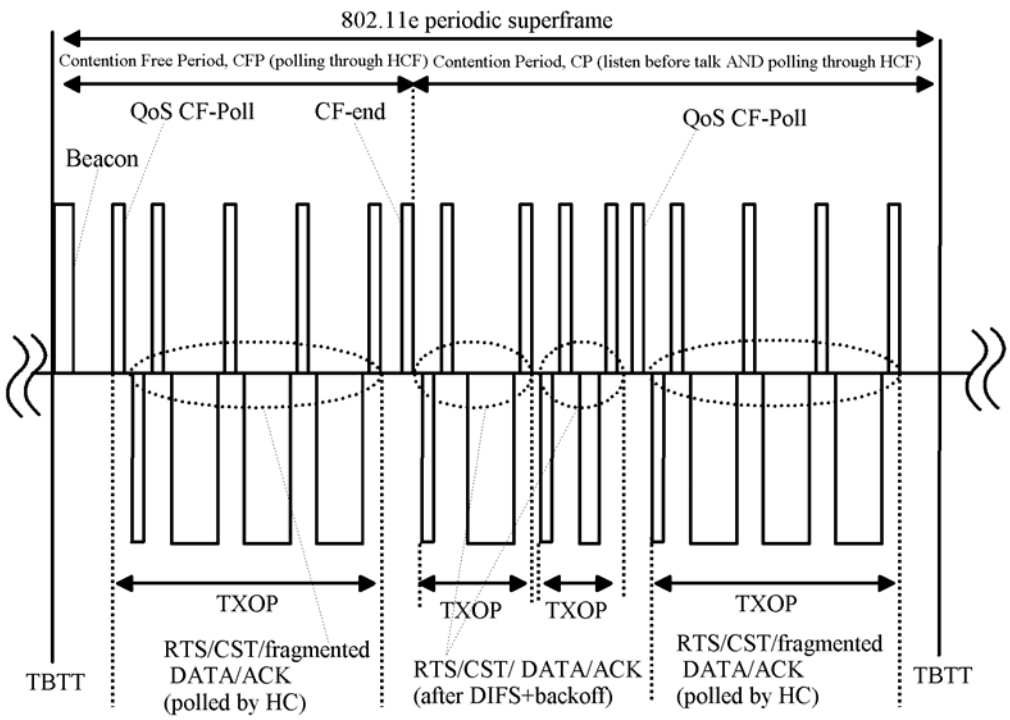

FIGURE 2.3 A typical IEEE 802.11e superframe.

2.2.3 Controlled Contention

The HC polls stations for MSDU Delivery. As part of 802.11e, an additional random access protocol that allows fast collision resolution is defined. For this, the HC requires information that has to be updated by the polled stations from time to time. Controlled Contention Opportunity (CCO) is a way for the HC to learn which station needs to be polled, at which times, and for which duration. The controlled contention mechanism allows stations to request the allocation of polled TXOP 's by sending resource requests [20].

Each instance of controlled contention occurs during the controlled contention interval, (CCI) which is started when the HC sends a specific control frame. This control frame forces legacy stations to set their NAV until the end of the controlled contention interval, thus they remain silent during the controlled contention interval. The control frame defines a number of Controlled Contention opportunities, short intervals separated by Short Inter Frame Space (SIFS), and a filtering mask containing the TC 's in which resource requests may be placed. Each station with queued traffic for a TC matching the filtering mask chooses one opportunity interval and transmits a resource request frame containing the requested TC and TXOP duration, or the queue size of the requested TC. For fast collision resolution, the HC acknowledges the reception of request by generating a control frame with a feedback field so that the requesting stations can detect collisions during controlled contention [20].

2.3 Scheduling

Fairness is an important issue when accessing a shared wireless channel. With fair scheduling, different flows sharing a wireless channel can be allocated bandwidth in proportion of their "weights". However, the fair scheduling algorithms have been made for centralized protocols, a designated host (base station or access point) coordinates access to the wireless medium.

Distributed Coordination Function (DCF) in IEEE 802.11 is an example of the distributed approach. DFS algorithm in distributed fair scheduling enables fair scheduling in such WLAN.

Fair scheduling algorithms

Consider the system shown in Figure 2.4, where a node maintains several queues (or flows) which store packets to be transmitted on an output link. A fair queuing algorithm is used to determine which flow to serve next, so as to satisfy a certain fairness criterion.

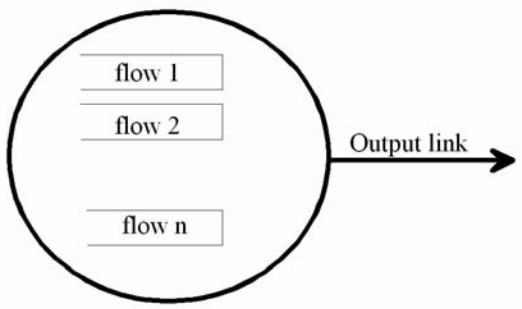

FIGURE 2.4 A node with several flows sharing a link.

Fair queuing algorithms in literature typically attempt to approximate the Generalized Processor Sharing (GPS) discipline [76]. When using the GPS discipline, a server serves, say, n flows each characterized by a positive weight; let Φ_i denote the weight associated with flow i (i = 1,…, n). Let $W_i\ (t_1, t_2)$ be the amount of flow i traffic served in the interval $[t_1, t_2]$. Then, for a GPS server, if flow i is backlogged (flow is not empty) throughout $[t_1, t_2]$, the following condition holds:

$$\frac{W_i(t_1,t_2)}{W_j(t_1,t_2)} \geq \frac{\Phi_i}{\Phi_j}, \forall j \tag{2}$$

Equality holds above if flow j is also backlogged in interval $[t_1, t_2]$. Note that the above condition is valid regardless of how small the interval $[t_1, t_2]$ is. The GPS discipline cannot be accurately implemented in practice, since data transmitted on real networks is packetized. This observation led to development of several packet fair queuing algorithms which approximate GPS under the constraint that each packet must be transmitted as a whole [6], [14], [22], [49], [76], [83].

Many centralized fair queuing algorithms behave as follows:

- The coordinator maintains a "virtual clock" – different algorithms differ in how the virtual clock is updated.
- *Start* and *finish* tags are assigned to each packet arriving on each flow. Packets are scheduled for transmission in the order of either finish tags [6], [14], [21] or start tags [22].

Two significantly different approaches are used for updating the virtual clock:

- In one approach, the *rate* of increase of virtual time is a function of the set of backlogged flows [6], [14].
- In the alternative approaches, the virtual clock is updated to be equal to either start tag [22] or finish tag [21] of the most recent packet in service.

The first approach above (for updating virtual clocks) potentially allows the fair queuing algorithms to match GPS more closely [6]. However, in the second approach, the virtual clock can be updated without knowing which flows are backlogged [21], [22]. Due to this property, the second approach is more suitable for a distributed implementation like WLAN.

Function of Distributed Fair Scheduling

Distributed Fair Scheduling (DFS) protocol is based on the SCFQ (Self-Clocked Fair Queueing) algorithm [49] and DCF protocol. SCFQ reduces the complexity of GPS for updating its virtual time on a packet arrival. The SCFQ algorithm can achieve easier implementation as well as maintaining the fairness property by introducing a new virtual time function. The DFS protocol borrows on SCFQ's idea of transmitting the packet whose finish tag is smallest, as well as SCFQ's mechanism for updating the virtual time. A distributed approach for determining the smallest finish tag is employed, using the DCF backoff interval mechanism. The essential idea is to choose a backoff interval that is proportional to the finish tag of packet to be transmitted. Several implementations of this idea are possible.

Let's assume that all packets at a node belong to a single flow. Each node i maintains a local virtual clock, $v_i(t)$, where $v_i(0) = 0$. Now, P_i^k represents the k-th packet arriving at the flow at node i on the LAN.

- Each transmitted packet is tagged with its finish tag.
- When at time *t* node *i* hears or transmits a packet with finish tag Z, node *i* sets its virtual clock v_i equal to *max* $(v_i(t); Z)$.

Start and finish tags for a packet are not calculated when the packet arrives. Instead, the tags for a packet are calculated when the packet reaches the front of its flow. When packet P_i^k reaches the front of its flow at node *i*, the packet is stamped with start tag $S_i^k = v(f_i^k)$, where f_i^k denotes the real time when packet P_i^k reaches the front of the flow.

Finish tag F_i^k is calculated as follows, where appropriate choice of the *Scaling Factor* allows us to choose a suitable scale for the virtual time.

$$F_i^k = S_i^k + \textit{Scaling Factor} * (L_i^k / \Phi_i) \tag{3}$$

$$= v(f_i^k) + \textit{Scaling Factor} * (L_i^k / \Phi_i) \tag{4}$$

The objective of the next step is to choose a backoff interval such that a packet with smaller finish tag will ideally be assigned a smaller backoff interval. This step is performed at time f_i^k. Specifically, node *i* picks a backoff interval B_i for packet P_i^k, as a function of F_i^k and the current virtual time $v_i(f_i^k)$, as follows:

$$B_i = F_i^k - v(f_i^k) \tag{5}$$

$$= \textit{Scaling Factor} * (L_i^k / \Phi_i) \textit{ time slots.} \tag{6}$$

To reduce the possibility of collisions, we randomize the B_i value chosen above as follows:

$$B_i = \rho * B_i \tag{7}$$

where ρ is a random variable with mean 1. When this step is performed, a variable named *Collision Counter* is reset to 0.

If a collision occurs (because backoff intervals of two or more nodes count down to 0 simultaneously), then the following procedure is used. Let node *i* be one of the nodes whose transmission has collided with some other node(s). Node *i* chooses a new backoff interval as follows:

- Increment CollisionCounter by 1.
- Choose new B*i* uniformly distributed in $[1, 2^{CollisionCounter-1} * CollisionWindow]$, where *CollisionWindow* is a constant parameter.

If CollisionWindow is chosen to be small, the above procedure tends to choose a relatively small B_i (in the range [1,*CollisionWindow*]) after the first collision for a packet. The motivation for choosing small B_i after the first collision is as follows: The fact that node *i* was "a potential winner" of the contention for channel access indicates that it is node *i*'s turn to transmit in the near future.

Therefore, B_i is chosen to be small to increase the probability that node *i* wins again soon. However, to protect against the situation when too many nodes collide, the range for B_i grows exponentially with the number of consecutive collisions.

Problems with DFS

The DFS protocol can exhibit short-term unfairness for some nodes when their packets collide. For instance, assume that, at the beginning of time, nodes 1, 2 and 3 pick backoff intervals of 25, 25, and 26 slots, respectively. Nodes 1 and 2 would collide when their backoff intervals count down to 0 (the backoff interval of node 3 would count down to 1 slot by this time). After collision, nodes 1 and 2 pick new backoff intervals of, say, 2 and 3 slots, respectively. In this case, node 3 would end up transmitting a packet before nodes 1 and 2, even though these two nodes should have transmitted earlier (since their original backoff intervals were smaller).

To eliminate such unfairness, a collision resolution protocol that guarantees colliding stations access prior to access by any other node (or, a protocol, which ensures this with a high probability) must be used. Guaranteeing "near-perfect" collision resolution, however, may add the "handling data" of the packets and decrease the efficiency of the algorithm [90].

3 3G AND WLAN INTERWORKING

3.1 Introduction

Developing next generation networks and their environments, the most important challenges for success are Services and QoS. At the moment the development of new next generation network service core needs the most development effort. QoS is very tightly tied together with services by guaranteeing their usability for the subscribers and this is why the control and management of this functionality must be developed as fast and with as high quality as possible for the present 3G environment. ITU-T M.3010 recommendation defines the management requirements of administrations to plan, provision, install, maintain, operate and administer telecommunications networks and services. It provides management functions for telecommunication networks and services, offering communications between itself and the telecommunication networks, services, and other Telecommunications Management Network (TMN) [10].

In publication's [P1] – [P3] we present our research work related to WLAN and 3G interworking. [29], [30], [91]. WLAN connections [1], [2] where they are available, can offer faster connections than 3G (2Mb/s) with cheaper prices. For example the new 802.11a version can achieve data rates of up to 54 Mbit/s at the wireless medium using the Orthogonal Frequency Division Multiplexing (OFDM) modulation technique in the unlicensed 5 GHz band [64].

QoS supported WLAN accesses will be soon available and it makes possible to use these access networks with 3G mobile devices. The problem with WLAN networks is the high error rate probability. It can rise up to 40% causing trouble especially to streaming type of applications. 802.11e standard is trying to correct the situation by enabling the use maximum of eight separate priority queues for prioritizing higher priority traffic compared to other traffic [31]. QoS supported WLAN uses the Enhanced Distributed Co-ordination Function (EDCF). It is the basis for the Hybrid Co-ordination Function (HCF)

[31]. The QoS support is done with the Traffic Categories (TC). MAC service data units will be delivered through multiple back off instances within one station, each back off instance parameterized with TC-specific parameters. A single station can have up to eight transmission queues realized as virtual stations inside a station, with QoS parameters that determine their priorities. If the counters of two or more parallel TC's in a single station reach zero at the same time, a scheduler inside the station avoids the virtual collision. The scheduler grants the Transmission Opportunity (TXOP) to the TC with highest priority, out of the TCs that virtually collided within the station. There is then still a possibility that the transmitted frame collides at the wireless medium with a frame transmitted by other station [31].

3.1.1 QoS issues in 3G and 802.11e

The 3G Partnership Project (3GPP) has defined four traffic (QoS) classes that have their own QoS attributes. All traffic in the 3/4G network will be handled according to the operator's requirements at the each of the traffic classes. The main QoS method to be used is proposed to be DiffServ [2]. Also Resource Reservation Protocol (RSVP) [3] could be used optionally.

3.1.2 3GPP defined traffic classes

The 3G traffic classes are Conversational class for voice and RT multimedia messaging. Streaming class for streaming type of applications, Video On Demand (VOD) etc. Interactive class for interactive type of applications, eCommerce, WEB browsing, etc. Background class for background type applications, email, FTP, etc. Table 3.1. presents the current QoS values for the radio bearer [3].

TABLE 3.1 3GPP QoS attributes for radio access bearer services.

Traffic class	Conversational class	Streaming class	Interactive class	Background class
Maximum bitrate (kbps)	<= 16 000 (2) (7)	<= 16 000 (2) (7)	<= 16 000 - overhead (2) (3) (7)	<= 16 000 - overhead (2) (3) (7)
Delivery order	Yes/No	Yes/No	Yes/No	Yes/No
Maximum SDU size (octets)	<=1 500 or 1 502 (4)	<=1 500 or 1 502 (4)	<=1 500 or 1 502 (4)	<=1 500 or 1 502 (4)
SDU format information	(5)	(5)		
Delivery of erroneous SDUs	Yes/No/-	Yes/No/-	Yes/No/-	Yes/No/-
Residual BER (Binary Error Rate)	$5*10^{-2}$, 10^{-2}, $5*10^{-3}$, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}	$5*10^{-2}$, 10^{-2}, $5*10^{-3}$, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}	$4*10^{-3}$, 10^{-5}, $6*10^{-8}$ (6)	$4*10^{-3}$, 10^{-5}, $6*10^{-8}$ (6)
SDU error ratio	10^{-2}, $7*10^{-3}$, 10^{-3}, 10^{-4}, 10^{-5}	10^{-1}, 10^{-2}, $7*10^{-3}$, 10^{-3}, 10^{-4}, 10^{-5}	10^{-3}, 10^{-4}, 10^{-6}	10^{-3}, 10^{-4}, 10^{-6}
Transfer delay (ms)	80 – maximum value	250 – maximum value		
Guaranteed bit rate (kbps)	<= 16 000 (2) (7)	<= 16 000 (2) (7)		
Traffic handling priority			1,2,3	
Allocation/Retention priority	1,2,3	1,2,3	1,2,3	1,2,3
Source statistic descriptor	Speech/unknown	Speech/unknown		
Signalling Indication			Yes/No	

1) Void.

2) The granularity of the bit rate attributes shall be studied. Although the UMTS network has capability to support a large number of different bitrate values, the number of possible values shall be limited not to unnecessarily increase the complexity of for example terminals, charging and interworking functions. Exact list of supported values shall be defined together with S1, N1, N3 and R2.

3) Impact from layer 2 protocols on maximum bitrate in non-transparent RLC protocol mode shall be estimated.

4) In case of PDP type = PPP, maximum SDU size is 1502 octets. In other cases, maximum SDU size is 1 500 octets.

5) Definition of possible values of exact SDU sizes for which UTRAN can support transparent RLC protocol mode, is the task of RAN WG3.

6) Values are derived from CRC lengths of 8, 16 and 24 bits on layer 1.

7) In case of GERAN the highest bitrate value is 473.6 kbps.

3.1.3 802.11e QoS issues and mapping to 3G

The Enhanced Distributed Co-ordination Function is the basis for the Hybrid Co-ordination Function [31]. The QoS support is done with the Traffic Categories (TC). MAC service data units will be delivered through multiple back off instances within one station, each back off instance parameterized with TC-specific parameters. A single station can have up to eight transmission queues realized as virtual stations inside a station, with QoS parameters that determine their priorities. If the counters of two or more parallel TC' s in a single station reach zero at the same time, a scheduler inside the station avoids the virtual collision. The scheduler grants the Transmission Opportunity (TXOP) to the TC with highest priority, out of the TC' s that virtually collided within the station. There is then still a possibility that the transmitted frame collides at the wireless medium with a frame transmitted by other stations [31]. When we are using strict priority queuing the mapping can be done as follows. We can use four strict priority queues of which each corresponds one of the

3GPP traffic classes. We can also use the three THP' s (Traffic Handling Priority) used for Interactive class to further sub classify Interactive class traffic by inserting it to three separate queues. All control plane traffic can be handled similarly mapping it into one or several 802.11e queues. Mapping process can be policy based controlled (PBC) and the mapping can be indicated at the IP level by the DSCP (DiffServ Code Point) inserted to the TOS field by DS classifier/marker mechanism or by the actual application that generates the control plane traffic. Table 3.2. shows the Per-Hop Behavior (PHB) actions with DSCP mapping [30].

TABLE 3.2 3G Traffic class mappings into DSCP

Traffic class	PHB actions (mapped into DSCP)
Conversational	EF
Streaming	AF 1
Interactive (THP 1)	AF 21
Interactive (THP 2)	AF22
Interactive (THP 3)	AF 23
Background	AF 3

3.2 Results

3.2.1 General Issues

The goal for the simulations was to study how much the throughput of the traffic decreases with various error rates while changing the traffic mix and average packet sizes of the individual traffic class.

We used network simulator NS-2 with IEEE 802.11 EDCF functionality implemented by Ni Qiang et al. of the Planete Project-INRIA [73]. To emulate the process of packet transmission errors we extended the simulator by implementing a two-state Markov model in the air interface.

MMPP (Markov Modulated Poisson Process) is a doubly stochastic process where the intensity of a Poisson process is defined by the state of a Markov chain (see Figure 3.1.). The transition matrix Q of the modulating Markov chain is defined by

$$Q = \begin{bmatrix} -\omega_1 & \omega_1 \\ \omega_2 & -\omega_2 \end{bmatrix}. \tag{8}$$

In our error scenario, the channel switches between a "good state" (λ_1) and a "bad state" (λ_2). Packets are transmitted correctly when the channel is in state λ_1, and errors occur when the channel is in state λ_2. When the channel is in state λ_1, it can either remain in this state, with probability ω_1 or make the transition to state λ_2, with probability 1-ω_1. Likewise, if the channel is in state λ_2, it remains in this state with probability ω_2 and changes state with probability 1-ω_2.

.

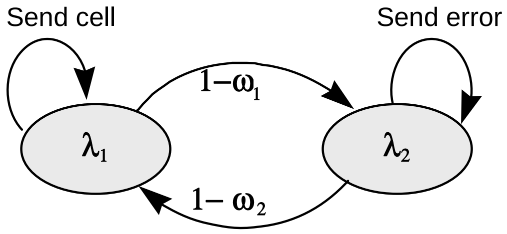

FIGURE 3.1 State machine for a 2-state MMPP

3.2.2 3G Interworking with WLAN QoS 802.11e

In publication 1 [PI] the simulation scenario is presented in Figure 3.2. It includes five WLAN stations. One of them is the access point. Station 1 is the highest and station 4 is the lowest in priority. In our simulations, each station generates the same amount of traffic.

The total traffic load is 2.5 Mbit/s per station. Station 1 and station 3 have both five flows all generating 500 kbit/s constant bit rate (CBR) traffic. Stations 2 and 4 have also five flows each generating an average of 500 kbit/s variable bit rate (VBR) traffic. VBR parameters of both of these stations are as shown in Table 3.3.

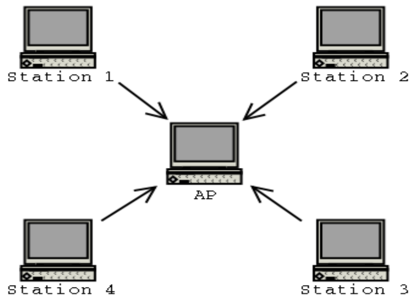

FIGURE 3.2 Simulation scenario

TABLE 3.3 WLAN station traffic properties

Rate	500 kbit/s
Rate deviation	0.25
Rate time	2.0
Burst time	1.0
No. of changes	10
Time deviation	0.5
Maxrate	648 kbit/s

Four priority levels were used in our simulations. EDCF parameters of different TC' s are shown in the Table 3.4.

TABLE 3.4 Used EDCF Parameters.

Traffic class	1	2	3	4
CWMin	7	10	15	127
CWMax	7	31	255	1023
AIFS (CWOffset)	2	4	7	15

We varied the packet sizes of the traffic classes by iterating through all the packet size combinations starting from 100 bytes to 1500 bytes with 280 byte steps. The possible packet sizes were thus 100, 380, 660, 940, 1220 and 1500 bytes. The iteration was following: at the iteration round 0 packet size of all the classes is 100 bytes. The lowest priority traffic class packet size is increased by 280 bytes at every iteration. If the packet size is 1500 bytes, it is reset to 100 bytes and the next lowest traffic class packet size is increased by 280 bytes. This way the lowest priority traffic class iterates through all the packet size possibilities within 6 iterations and the next lowest priority traffic class within 6^2=36 iterations. The second highest traffic class goes through all the packet sizes within 6^3=216 iterations and the highest within 6^4=1296 iterations. The radio channel packet error rate was varied also from 0% to 60% with steps of 20%.

Table 3.5. shows the transition probabilities we used to achieve different packet error rates in the air interface.

TABLE 3.5 Transition probabilities for 2- state MMPP.

Error rate	ω_1	ω_2
20%	0.16	0.63
40%	0.40	0.60
60%	0.90	0.60

As can be seen from the results EDCF works quite well with different channel error rates. Highest traffic class will get significantly more capacity than lower priority classes in all our simulation scenarios. This shows also the fact that EDCF is suitable for using with 3G to support QoS at the situations when 3G devices will use WLAN access to the core IP network. Interesting issue arises with the fact that highest priority traffic class will get more bandwidth when the packet size and the channel error rate are increasing. With the proposed 3G traffic classification specifications this can lead to the situation that the highest class uses all the capacity under certain network conditions. Our simulation results proved this issue clearly. Based on that, it is important to find out the limits where packet size and channel error rate can change. If we think for example about the situation where we have small size high priority packets (voice traffic) under heavily loaded network i.e. channel error rate is high, we will loose also those small size high priority packets. Another interesting fact is that the packet size that gives the best throughput to the highest class is dependent on the error rate.

3.2.3 End-to-End QoS Issues at the WLAN and 3G environments

In publication 2 [PII] the simulation scenario is presented in Figure 3.3. It includes eight WLAN stations. Stations 1, 2, 3 and 4 are in contact with access point AP1. Station 1 is the highest and station 4 is the lowest in priority. Stations 5, 6, 7 and 8 are connected to access point AP2. The link between AP2 and the core network is a 10 Mbit/s link with 5 ms delay.

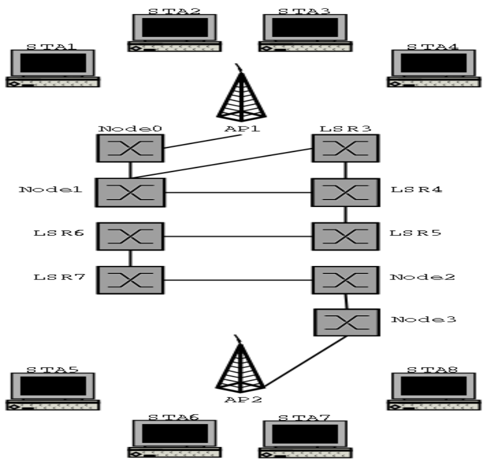

FIGURE 3.3 Used network configuration.

All other wired links in core network are 10 Mbit/s links with 0 ms delay except between LSR5 and LSR6 routers there is a 15 Mbit/s link with 13 ms delay and between LSR6 and LSR7 the delay is 14 ms.

DiffServ is used in core network. LSR4 is a DiffServ Edge router and LSR5 is a DiffServ Core router. In LSR4 there are four physical DropTail Random Early Detection (RED) queues with two precedence levels in each queue. One level for packets that are in profile and one for packets that are out of profile. The queue sizes are:

TABLE 3.6 Queue sizes.

Traffic class	Conversational	Streaming	Interactive	Background
In profile	30	60	30	30
Out of profile	0	30	30	0

A Token bucket policer is used. The parameters of this policer are as shown:

TABLE 3.7 Token bucket policer parameters.

Traffic class	Conversational	Streaming	Interactive	Background
CIR	3.0 Mbit/s	2.5 Mbit/s	2 Mbit/s	1.5 Mbit/s
MBS	3 KB	2 KB	1.5 KB	1.5 KB

In our simulations, stations 1, 2, 3 and 4 are sending nodes and stations 5, 6, 7 and 8 are receiving nodes. Station 1 sends to station 5, and so forth. Each station generates the same amount of traffic. The total traffic load is 2.5 Mbit/s per station. Station 1 and station 3 have both five flows all generating 500 kbit/s constant bit rate (CBR) traffic. Stations 2 and 4 have also five flows each generating an average of 500 kbit/s variable bit rate (VBR) traffic. VBR parameters of both of these stations are as shown:

TABLE 3.8 WLAN station traffic properties.

Rate	500 kbit/s
Rate deviation	0.25
Rate time	2.0
Burst time	1.0
No. of changes	10
Max. rate	648 kbit/s

We used four priority levels in our simulations. EDCF parameters of different TC' s are shown in the following table.

TABLE 3.9 Used EDCF Parameters.

Traffic class	Conversational	Streaming	Interactive	Background
CWMin	7	10	15	127
AIFS (CWOffset)	2	4	7	15
CWMax	7	31	255	1023

In this case we varied the packet sizes of the traffic classes iterating through all the packet size combinations starting from 500 bytes to 1500 bytes with 500 byte steps. The possible packet sizes were thus 500, 1000 and 1500 bytes. The iteration was following: at iteration 0 packet size of all the classes is 500 bytes. The lowest priority traffic class packet size is increased by 500 bytes every iteration. If the packet size is 1500 bytes, it is reset to 500 bytes and the next lowest traffic class packet size is increased by 500 bytes. This way the lowest priority traffic class iterates through all the packet size possibilities in 3 iterations and the next lowest priority traffic class in 3^2=9 iterations. The second highest traffic class goes through all the packet sizes in 3^3=27 iterations and the highest in 3^4=81 iterations. The radio channel packet error rate was varied also from 0% to 60% with 30% steps.

Table 3.10. shows the transition probabilities we used to achieve different packet error rates in the air interface.

TABLE 3.10 Transition probabilities for 2- state MMPP.

Error rate	ω_1	ω_2
30%	0.27	0.62
60%	0.90	0.60

EDCF works quite well with different channel error rates. Highest traffic class will get significantly more capacity than lower priority classes in all our simulation scenarios. This shows also the fact that EDCF is suitable for using with 3G to support QoS at the situations when 3G devices will use WLAN access to the core IP network. Interesting issue arises with the fact that the highest priority traffic class will get more bandwidth when the packet size and the channel error rate are increasing. With the proposed 3G traffic classification specifications this can lead to the situation that the highest class uses all the capacity under certain network conditions. Our simulation results proved this issue clearly. Based on that it is important to find out the limits where packet size and channel error rate can change. If we think for example the situation where we have small size high priority packets (voice traffic) under heavily loaded network i.e. channel error rate is high, we will loose also those small size high priority packets.

3.2.4 Providing QoS at the Integrated WLAN and 3G Environments

In publication 3 [PIII] the aim for the simulations is to find out optimal packet sizes for each 3G traffic classes from the viewpoint of throughput and delay. We ran two different simulation scenarios, where the packet sizes of the traffic classes were varied by iterating through all the packet size combinations starting from 100 bytes to 1500 bytes with 200 byte steps. The possible packet sizes were thus 100, 300, 500, 700, 900, 1100, 1300 and 1500 bytes. The iteration was following: at iteration 0 packet size of all the classes is 100 bytes. The lowest priority traffic class packet size is increased by 200 bytes every iteration. If the packet size is 1500 bytes, it is reset to 100 bytes and the next lowest traffic class packet size is increased by 200 bytes. This way the lowest priority traffic class iterates through all the packet size possibilities in 8 iterations and the next lowest priority traffic class in 8^2=64 iterations. The second highest traffic class goes through all the packet sizes in 8^3=512 iterations and the highest in 8^4=4096 iterations. In these simulations we used two different simulation configurations which where the same as presented in publications 1 and 2.

The best packet size combination was found from the simulation results by calculating an emphasized factor to all packet size combinations. The factor was calculated by using the following formula:

$$S_i = D_i + T_i, \tag{9}$$

where

$$D_i = \sum_{j=1}^{4} 100 * \delta_j * d_j^i \tag{10}$$

and

$$T_i = \sum_{j=1}^{4} 100 * \tau_j * b_j^i. \tag{11}$$

In these D_i is a total delay factor and T_i is a total throughput factor calculated from the results of all classes. δ_j is the weight coefficient of the delay of class *j* and d_j^i is the factor calculated to the delay of class *j* at the iteration *i*. τ_j is the weight coefficient of the throughput of class *j* and b_j^i is the factor calculated to the throughput of class *j* at the iteration *i*. The delay factor d_j^i is a factor between 0 and 1. It was calculated in the following way:

$$d_j^i = \frac{d_j^{\max} - d_i}{d_j^{\max} - d_j^{\min}}, \tag{12}$$

where d_i is the measured delay at iteration *i*, $d_j^{\max}$ is the maximum limit of delay of class *j* and $d_j^{\min}$ is the minimum limit of delay of class *j*. The factor of all delays above the maximum limit is 0 and the factor of all delays below the minimum limit is 1. The throughput factor b_j^i is also a factor between 0 and 1. It was calculated in the following way:

$$b_j^i = \frac{b_i - b_j^{\min}}{b_j^{\max} - b_j^{\min}}, \tag{13}$$

where b_i is the measured throughput at iteration *i*, $b_j^{\max}$ is the maximum limit of throughput of class *j* and $b_j^{\min}$ is the minimum limit of throughput of class *j*. The factor of all throughputs above the maximum limit is 1 and the factor of all throughputs below the minimum limit is 0. The delay and throughput limits are presented in Table 3.11.

TABLE 3.11 Delay and throughput limits for each traffic class.

Traffic class	$d_{\min}$ (ms)	$d_{\max}$ (ms)	$b_{\min}$ (kbit/s)	$b_{\max}$ (kbit/s)
Conversational (class 1)	100	500	-	2400
Streaming (2)	250	1000	-	2200
Interactive (3)	-	-	-	2100
Background (4)	-	-	-	2000

If there was no maximum and/or minimum limit defined for the delay or the throughput, the lowest result was the minimum limit and the highest result was the maximum limit. Coefficients δ_j and τ_j define how much weight are given to the delay and the throughput of a class *j* when calculating the total delay and the throughput factors. The coefficients of all the classes of different error rate scenarios are shown in Table 3.12.

TABLE 3.12 Weight coefficients of traffic classes in different error scenarios.

Error rate	0%		20%		40%		60%	
	δ	τ	δ	τ	δ	τ	δ	τ
Class 1	0	0.1	0.1	0.1	0.15	0.15	0.5	0.5
Class 2	0	0.1	0.1	0.2	0.3	0.4	0	0
Class 3	0	0.4	0	0.25	0	0	0	0
Class 4	0	0.4	0	0.25	0	0	0	0

With low error rates the weights on high priority class traffic characteristics are low, because when error rate is 0% or 20%, the delays and the throughputs of high priority classes are very satisfactory with all packet sizes. Thus, with these error scenarios low priority classes traffic characteristics can be emphasized, especially throughput, because delay is not an important issue with best effort traffic classes. As the error rate increases, high priority traffic class traffic becomes more important, and their weights are raised. When the radio channel packet error rate is as high as 40%, best effort traffic class traffic characteristics are given no weight. In addition, the delays of conversational and streaming classes are so low that the emphasis is more on classes throughputs. When the error rate is 60%, the highest priority traffic class gets all the weights.

Interesting issue arises from the results with the fact that the highest priority traffic class will get more bandwidth when the packet size and the channel error rate is increasing. With the proposed 3G traffic classification specifications this can lead to the situation that the highest class uses all the capacity under certain network conditions. Our simulation results proved this issue clearly. Based on that it is important to find out the limits where packet size and channel error rate can change. If we think for example the situation where we have small size high priority packets (voice traffic) under heavily loaded network i.e. channel error rate is high, we will loose also those small size high priority packets. In 0% and 20% error scenario, the best results are achieved when the highest-class packet size is about 1500 bytes, but if the error is very high (we tested with 60%), smaller packet size gives the best results.

4 REVENUE-AWARE RESOURCE ALLOCATION SCHEMES

A significant amount of work has been done in resource scheduling for traditional network. Network resources in traditional networks mainly refer to bandwidth [7], [89]. Packet Fair Queuing (PFQ) disciplines such as Weighted Fair Queuing (WFQ) and WF2Q [7] provide perfect fairness among contending network flows. However, WFQ and WF2Q cannot readily be used for processor scheduling because they require precise knowledge of the execution times for the incoming packets at time of their arrival in the node. Another PFQ algorithm for bandwidth scheduling is Start-Time Fair Queuing (SFQ) [23], [24], which does not use packet lengths for updating virtual time, and therefore seems suitable for scheduling computational resources since it would not need prior knowledge of the execution times of packets [24]. However, the worst-case delay under SFQ increases with the number of flows and it tends to favor flows that have a higher average ratio of processing time per packet to reserved processing rate [75]. A significant amount of work has also been done on CPU scheduling [12], [37], [54], but most of them are on CPU scheduling for end systems and work on task level, not on packet level.

In this chapter a scheduling model is presented that optimizes the weighted mean delay of the network, not just in the worst case as in [63], but in a general case. The proposed algorithm ensures less delay for the users paying more for the connection (i.e. higher service class) than those paying less. This work extends our previous pricing and QoS research [46], to take into account scheduling issues by introducing fair delay guaranteeing mechanism. Described scheduling mechanism achieves this goal without requiring knowledge of the users utility functions and without requiring any explicit feedback from the network. An adaptive form algorithm for updating weights is obtained. It is fast - converging typically in one iteration for given number of connections - and robust against erroneous estimates of customer's behavior.

4.1 The packet scheduler and delays

In this section, we formulate expression for delays of the data traffic. Consider the packet scheduler in Figure 4.1. There are now two service classes. Gold class customers pay most of money while getting best service, and silver class customer's pay least of

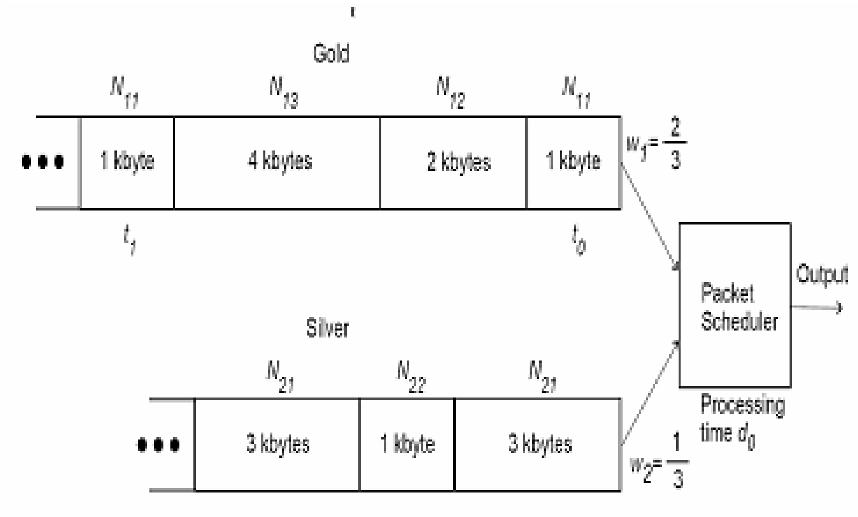

FIGURE 4.1 An example of a packet scheduler with two classes.

money. Parameter Δt_i denotes time which passes when data is transferred through the queue i to the output in the switch, when $w_i = 1$. If the queue is almost empty, delay is small, and when the buffer is full, it is large. Variable w_i is the weight allocated for class i. Constraint for weight w_i is

$$\sum_{i=1}^{m} w_i = 1, w_i > 0. \tag{14}$$

Variables w_i give weights, how long time queues i are served per total time. Therefore, delay d_i in the queue i is actually

$$d_i = \frac{\Delta t_i}{w_i} \tag{15}$$

Without loss of generality, only non-empty queues are considered, and therefore

$$w_i \neq 0, i = 1, \ldots, m, \tag{16}$$

where m is number of service classes. When one queue becomes empty, m -> m – 1. Parameter N_i denotes the number of connections in the i-th service class. Mean overall delay is formulated as follows:

$$E(d) = E\left(\frac{\Delta t_i}{w_i}\right) = \frac{1}{\sum_{i=1}^{m} N_i} N_i \frac{\Delta t_i}{w_i} + \lambda\left(1 - \sum_{i=1}^{m} w_i\right). \tag{17}$$

Weighting factors are r_i, $i = 1, \ldots, m$. In addition, $r_i > r_j$, where the class i (e.g. gold class) has higher priority than the class j (e.g. silver class). Weighted mean delay has the form

$$E(r_i d) = E\left(\frac{r_i \Delta t_i}{w_i}\right) = \frac{1}{\sum_{i=1}^{m} N_i} \sum_{i=1}^{m} N_i r_i \frac{\Delta t_i}{w_i} + \lambda\left(1 - \sum_{i=1}^{m} w_i\right). \tag{18}$$

Weighted mean delay is minimized by putting the derivative to zero:

$$\frac{\partial E(r_i d)}{\partial w_i} = \frac{N_i r_i \Delta t_i}{\sum N_l w_i^2} - \lambda = 0. \tag{19}$$

Thus

$$w_i = \sqrt{-\frac{1}{\lambda} \frac{N_i r_i \Delta t_i}{\sum_{l=1}^{m} N_l}}. \tag{20}$$

Penalty factor λ is solved out as follows:

$$\lambda = -\frac{N_i r_i \Delta t_i}{\sum_{l=1}^{m} N_l w_l^2}. \tag{21}$$

On the other hand:

$$\lambda = \lambda \sum_{i=1}^{m} w_i = -\frac{1}{\sum_{l=1}^{m} N_l} \sum \frac{N_i r_i \Delta t_i}{w_i}. \tag{22}$$

Then

$$w_i = \sqrt{\sum_{l=1}^{m} N_l \frac{1}{\sum_{k=1}^{m} \frac{N_k r_k \Delta_k}{w_k}} \frac{1}{\sum_{l=1}^{m} N_l} N_i r_i \Delta t_i}$$

$$= \sqrt{\frac{1}{\sum \frac{N_k r_k \Delta t_k}{w_k}} N_i r_i \Delta t_i}. \tag{23}$$

Uniqueness of the solution is seen by taking second order derivative. First order derivative is

$$\frac{\partial E(r_i d)}{\partial w_i} = -\frac{N_i r_i \Delta t_i}{\sum_{l=1}^{m} N_l w_i^2} + \frac{1}{\sum_{l=1}^{m} N_l} \sum \frac{N_k r_k \Delta t_k}{w_k}. \tag{24}$$

Second order derivative is

$$\frac{\partial^2 E(r_i d)}{\partial w_i^2} = \frac{2 N_i r_i \Delta t_i}{\sum_{l=1}^{m} N_l w_i^3} - \frac{N_i r_i \Delta t_i}{\sum_{l=1}^{m} N_l w_i^2} > 0. \tag{25}$$

Therefore, mean weighted delay is convex, and has global unique minimum. Fixed point type algorithm for optimizing the weights is performed as follows:
1. At time step t, update the weights

$$v_i(t) = \sqrt{\frac{1}{\sum_{k=1}^{m} \frac{N_k(t) r_k(t) t_k(t)}{w_k}} N_i(t) r_i(t) \Delta t_i(t)}. \tag{26}$$

2. Perform scaling

$$w_i(t+1) = \frac{v_{i(t)}}{\sum_{k=1}^{m} v_k(t)}. \tag{27}$$

3. If the weights are converged using some predetermined criterion - e.g. $|w_i(t+1) - w_i(t)| < \delta$, where δ is some small positive number - stop the iteration; otherwise, go to step 1.

4.2 Pricing and revenue

We concentrate on the pricing and fair resource guarantee from the point of view of the customers. On the other hand, from the point of view of the service provider, we try to maximize revenue. First, we introduce the concept of pricing functions. For delay, pricing functions are denoted by $f_i(d)$, where d is the delay, and f_i is decreasing with respect to d. In addition, $f_i(d)$ is (strictly) convex with respect to d. Revenue obtained for one user in the class i is just $f_i\left(\frac{\Delta t_i}{w_i}\right)$.. Because there are N_i connections in the class i with all having the same delay, revenue corresponding to the delays in the class i is

$$R_i(w_i) = N_i F_i\left(\frac{\Delta t_i}{w_i}\right). \tag{28}$$

Total revenue is then

$$R = \sum_{i=1}^{m} R\,(w_i). \tag{29}$$

In our study, we use *linear* pricing function for delays d. Then

$$f_i(d) = r_i d + k_i, k_i > 0, \tag{30}$$
$$f'_i(d) = -r_i, \tag{31}$$
$$f''_i(d) = 0. \tag{32}$$

Here $k_i > 0$ guarantees positive revenue with minimum delay. For the classes that have better service, factors r_i are larger compared with those classes that have service of lower priority. Notice that we use here the same pricing factors than those in the weighted mean delay calculation. In addition, for classes having better service, k_i is larger. Notice that the pricing function may be even negative, when the delay is too large. However, Call Admission Control (CAC) mechanism takes care that this situation is prevented. From equations. (4.15)-(4.17) we see that the total revenue in the linear pricing scenario is

$$R = -\sum_{i=1}^{m} \frac{N_i r_i \Delta t_i}{w_i} + \sum_{i=1}^{m} N_i k_i. \tag{33}$$

4.3 Extension to the multinode case

In this section, we extend our approach for multinode case. In the multinode case, there are n nodes with m service classes. Number of connections in the switch i and class j is denoted by N_{ij}. Weights for switch and class (*i, j*) is denoted by w_{ij}, and the constraint

$$\sum_{j=1}^{m} w_{ij} = 1 \tag{34}$$

must be satisfied. The delays are denoted by Δt_{ij}. By using these notations, we obtain the weighted mean delay as follows:

$$E(r_i d) = \frac{1}{\sum_{i=1}^{n} \sum_{j=1}^{m} N_{ij}} \sum_{i=1}^{n} \sum_{j=1}^{m} \frac{N_{ij} r_j \Delta t_{ij}}{w_{ij}} + \sum_{i=1}^{n} \lambda_i \left(1 - \sum_{j=1}^{m} w_{ij}\right). \tag{35}$$

The first order derivative is

$$\frac{\partial E(r_j d)}{\partial w_j} = -\frac{N_{ij} r_j \Delta t_{ij}}{\sum_{k=1}^{n} \sum_{l=1}^{m} N_{kl} w_{ij}^2} - \lambda_i . \tag{36}$$

Solving λ_i out, we obtain the updating rule

$$w_{ij} = \sqrt{\frac{N_{ij} r_j \Delta t_{ij}}{\sum_{h=1}^{m} \frac{N_{ih} r_h \Delta t_{ih}}{w_{ih}}}} . \tag{37}$$

4.4 Results

Publication [PIV] extends previous pricing and QoS research [38], [39], [42] to take into account queuing issues. Most popular queuing algorithms are priority queue [51] and weighted fair queuing (WFQ) [76]. Service differentiation and management are issues, which need careful QoS design in wireless networks. The service differentiation can be implemented with a stateless core architecture that uses QoS parameter carried by packet header to provide fair-queuing [88] or guaranteed delay [87]. The weights in gradient type WRR algorithm are adapted using revenue as a target function to be maximised. Adaptive WRR can also be used as an admission control mechanism, but this is the topic in the future research. The close research is [28], where adaptive WFQ algorithm was investigated, but revenue criterion was not used. Our algorithm uses the following input parameters: (1) minimum processing time of the classifier for transmitting data from one queue to the output; (2) number of customers in the queues; (3) delay parameters such as insertion delay, transmission delay etc. Due to the adaptive nature of the algorithm, it can operate in the non-stationary environments. In addition, it is non-parametric in the sense that any call densities or duration distributions are assumed. The performance of the algorithm is investigated in the three-service class case.

In Publication 5, [PV] we extend the scheduling methods to support QoS in local replicated web systems. We present a framework for building global replicated Web systems and select/design the scheduling algorithms used in the framework. This work extends previous QoS-aware scheduling algorithm research [97]. Through the simulations, we demonstrate that our presented framework and scheduling algorithms can succeed in implementing differentiated services in global replicated Web systems under different workload levels. We also investigate the effects of some important parameters on the performances of the framework.

Publication 6, [PVI] presents an adaptive scheduling algorithm for the traffic allocation. We use flat pricing scenario in our model, and the weights of the queues are updated using revenue as a target function. Due to the closed form nature of the algorithm, it can operate in the non-stationary environments. In addition, it is non-parametric and deterministic in the sense that any assumptions about connection density functions or duration distributions are not made. This paper extends our previous pricing and QoS research [40], [41], [44], [45] to take into account queuing scheduling issues by introducing dynamic weight tracking algorithm in the scheduler. QoS and revenue aware scheduling algorithm is investigated. It is derived from Lagrangian optimization problem, and optimal closed form solution is presented.

Publication 7, [PVII] extends our previous studies [46], [96], where only delay was considered as QoS parameter. We present adaptive algorithms for optimizing the network operator revenue and for ensuring QoS requirements. In this paper, we add bandwidth as another QoS parameter to be observed. Pricing models and revenue maximization is discussed, as well as gradient and fixed-point algorithms are introduced for fair service and revenue optimization.

Publication 8, [PVIII] proposes a scheduling model for ensuring delay as a Quality of Service requirement in the communications network. Different delays are allocated for different pricing classes, say gold, silver and bronze classes. Our purpose is to minimize weighted mean delay for connections, where weights are the pricing factors of different classes. An adaptive and optimal solution is derived for weights of the scheduler. Simulations show that in addition to the mean delay minimization, the revenue of the service provider is also maximized in the linear pricing scenario. In addition, adaptive updating rule is simple to implement, since it converges fast. Our scenario is independent on the statistical assumptions, and therefore it is robust against possible erroneous estimates of the customer's behavior. In this paper we propose a scheduling model that optimizes the weighted mean delay of the network, not just in the worst case as in [63], but in a general case. The proposed algorithm ensures less delay for the users paying more for the connection (i.e. higher service class) than those paying less. This work extends our previous pricing and QoS research, to take into account scheduling issues by introducing fair delay guaranteeing mechanism. We describe a scheduling mechanism that achieves this goal without requiring knowledge of the users' utility functions and without requiring any explicit feedback from the network. An adaptive form algorithm for updating weights is obtained. It is fast - converging typically in one iteration for given number of connections - and robust against erroneous estimates of customer's behaviour.

Publication 9, [PIX] presents a packet-scheduling scheme for ensuring delay and bandwidth as a Quality of Service requirement in multinode case. For customers, rightful service is given while optimizing revenue of the network service provider. A fixed-point type algorithm for updating the weights of a packet scheduler is derived from a revenue-based optimization problem. We compared algorithm with optimal bruteforce method. Especially fixed-point algorithm converges very fast to the optimal solution, when number of classes is three. The weight updating procedures are independent on the assumption of the connections statistical behaviour, and therefore they are robust against erroneous estimates of statistics. Also, a Call Admission Control (CAC) is implemented in context of our scenario.

5 CONCLUSIONS AND FUTURE STUDIES

5.1 General about the thesis

In this thesis quality of service issues have been studied at the integrated wireless LAN and 3G environments. The behavior of 802.11e WLAN QoS used for 3GPP type services and service environment is presented widely under different kind of traffic scenarios. As the WLAN is probably the first 3G complementary access solution, these kinds of feasibility studies are important and their purpose is to show methods how to use different 3G WLAN accesses.
Our simulation results showed that a throughput optimizing packet size could always be found within different channel error rate scenarios. Presented results depicted also that the mobile stations operating in EDCF can guarantee throughput even the channel error rate is high. We can also see from the results that EDCF can be implemented with 3G to support QoS at the situations when 3G devices will use fast WLAN access to the core network. Based on the obtained results the packet sizes can be changed according to the priority class and the channel error rate to achieve optimal throughput for the 3G traffic classes.

5.2 Future Studies

The consolidating paper industry is looking for new competitiveness by raising the cost efficiency of production and business processes as well as by developing more effective ways to manage information and knowledge. The suppliers, including Metso Paper, are looking for new growth in the service business enabling the customers' new requirements. The realization of these new growth strategies calls for strong information and communications technology and related service know-how.

One problem on this way (and generally on developing wireless services) is that the design of 3G it is based on the needs of transferring speech and the network is not optimized for data transferring. Because of that even the developed 3G is not capable to transfer e.g. live video as well as fixed wide band network. Maybe next generation networks have answers to these problems. Today's wireless standards battle is being waged most vociferously in the 802.11n committee, which has agreed, in principle, on next generation WLAN specification that promises to push throughput beyond 100 Mbps. The technical foundation for 802.11n's fast throughput is MIMO (multiple-input, multiple-output), a wireless system design that uses multiple radios and antennas, together with spatial multiplexing, to boost performance. MIMO as a technology is becoming more general because it is possible to apply it also to traditional 802.11a/b/g networks [69].

The WiMAX (Worldwide Interoperability for Microwave Access), technology, based on the IEEE 802.16-2004 Air Interface Standard is rapidly proving itself as a technology that will play a key role in fixed broadband wireless metropolitan area networks. In December 2005 the IEEE ratified the 802.16e amendment [32] to the 802.16 standard. This amendment adds the features and attributes to the standard necessary to support mobility. The WiMAX Forum is now defining system performance and certification profiles based on the IEEE 802.16e Mobile Amendment and, going beyond the air interface, the WiMAX Forum is defining the network architecture necessary for implementing an end-to-end Mobile WiMAX network [93]. WiMAX represents a potentia competitor to UMTS and perhaps even to Wi-Fi (Wireless Fidelity). Unlike competing platforms, WiMAX was designed from the start as an IP network technology, providing an optimal mobile network environment that is apperently more hospitable to emerging mobile applications.

In the future, we expect to see an array of wireless systems, including 3G, Wi-Fi and WIMAX tied together by higher-layer standards like IP Multimedia System (IMS). Integrating those technologies in a way that allows a wide range of devices to move transparency across networks won't be easy, but as standards evolve, that hope is likely to be realized. Eventually, it's likely that most wide-area voice and data services will converge around an IP transport infrastructure combined with the IMS multimedia architecture. That will likely make lower-layer protocols less important to end-users while enhancing interoperability, not only among alternative mobile networks but also between mobile and fixed networks. IMS, which is being driven by 3GPP, is widely considered to be the foundation for next generation fixed/mobile convergence [69].

Based on availability and other constrains of 3G our future work will concentrate on seamless interworking between the 3G and WLAN, and defining the end-to-end QoS functionalities. One important issue is to study whether our adaptive scheduling algorithm, which was derived from revenue target function, can be implemented into wireless base stations.

REFERENCES

[1] 3GPP DRAFT TS 23.234: "3GPP System to Wireless Local Area Network (WLAN) Interworking: System Description", v1.0.0, September 2002. (Release 6), www.3gpp.org/specs/specs.htm

[2] 3GPP TS22.934, "Feasibility study on 3GPP system to Wireless Local Area Network (WLAN) interworking", Release 6.0, www.3gpp.org/specs/specs.htm

[3] 3GPP TS23.107, "QoS Concept and Architecture", Rel. 5.10.0,www.3gpp.org/specs/specs.htm

[4] 3GPP TS23.207, "End-to-End QoS concept and Architecture", Release 5.0, www.3gpp.org/specs/specs.htm

[5] Kumar Balachandran, "Convergence of 3G and WLANs", Keynote speech at IEEE International Conference on Communications ICC'2002 (http://easy.intranet.gr/paper_19.pdf)

[6] J. C. R. Bennett and H. Zhang, "WF2Q: Worst-case Fair Weighted Fair Queueing", Proceedings of IEEE INFOCOM'96, Vol. 1. 1996, pp.: 120-128.

[7] J.C.R. Bennett and H. Zhang, "WFQ: Worst-case Fair Weighted Fair Queuing", In Proceedings of the IEEE INFOCOM, San Francisco, March 1996.

[8] Biehl, M., E. Prater, J.R. McIntyre, "Remote repair, diagnostics, and maintenance", Communications of the ACM 47(11). (2004) pp. 100-106.

[9] Britton, C., and Bye, P. "IT Architectures and Middleware", (2nd Edition). Addison-Wesley, Boston, (2004).

[10] CCITT/ITU-T Recommendation M.3010, "Principle for a telecommunications management network"

[11] H. M. Chaskar and U. Madhow, "Fair Scheduling with Tunable Latency: A Round Robin Approach", Proceedings of IEEE Globecom'99, 1999, pp. 1328-1333.

[12] Clifford W. Mercer, Stefan Savage, Hideyuki Tokuda, "Processor Capacity Reserves: Operating System Support for Multimedia Applications", In Proc. of the IEEE International Conference on Multimedia Computing and Systems, May 1994, pp. 90-99.

[13] Crawley E., Nair R., Rajagopalan B., Sandick H., "A Framework for QoS-based Routing in the Internet", RFC 2386, Aug 1998, www.ieee.org/rfc.html.

[14] A. Demers, S. Keshaw, S. Shenker: “Analysis and Simulation of a Fair Queuing Algorithm”. Journal of Internetworking: Research and Experience, Vol. 1, 1990, pp. 3-26.

[15] C. Dovrolis and P. Ramanathan, “Proportional differentiated services Part II: Loss rate differentiation and packet dropping,” Proc. of International Workshop on Quality of Service, Pittsburgh, June 2000, pp. 53-61.

[16] C. Dovrolis, D. Stiliadis, and P. Ramanathan, “Proportional differentiated services: Delay differentiation and packet scheduling,” IEEE/ACM Transactions on Networking, vol. 10, No. 1, Feb. 2002, pp. 12-26.

[17] M. A. ElGendy, A. Bose, K. G. Shin. .Evolution of the Internet QoS and support for soft real-time applications. Proceedings of the IEEE. Issue 7, Volume 91, July 2003, pp. 1086 - 1104.

[18] El-SayedM., Jaffe J., "A View of Telecommunications Network Evolution", IEEE Communications Magazine, Vol. 40, Issue: 12, Dec 2002, p. 74-81.

[19] S. Floyd and V. Jacobson, “Link-sharing and resource management models for packet networks,” IEEE/ACM Transactions on Networking, Volume: 3 No: 4, Aug. 1995, pp. 365 -386.

[20] Ganfollo P.T., "Trade-Off Analysis (802.11 versus 802.15.3 QoS mechanism)", XtremeSpectrum White Paper, June 2002, http://www.techonline.com/community/tech_group/comm/tech_paper / 21357.

[21] S. J. Golestani, ”A self-clocked fair queueing scheme for broadband applications”, Proceedings of IEEE INFOCOM’ 94, Vol. 2., 1994, pp.636-646.

[22] P. Goyal, H. M. Vin, and H. Cheng, "Start-time fair queuing: A scheduling algorithm for integrated services packet switching networks," IEEE/ACM Transactions on Networking, vol. 5, Oct. 1997, pp. 690-704.

[23] P. Goyal, H.M. Vin and H. Cheng, "Start-time fair queuing: A scheduling algorithm for integrated services packet switching networks", IEEE/ACM Transactions on Networking, vol. 5, no. 5, pp. 690-704, October 1997.

[24] P. Goyal, X. Guo, and H. M. Vin, "A hierarchical CPU scheduler for multimedia operating systems", In Proceedings of the Second Symposium on Operating Systems Design and Implementation (OSDI)." USENIX, October 1996, pp. 107-121.

[25] Heikkilä J., Heikkilä M., Lehmonen J., "Sharing for Understanding and Doing for Learning: An Emerging Learning Business Network", The ICFAI Journal of Knowledge Management, Vol. III, No. 1, (2005), pp. 28- 45.

[26] Hettich A., Schrötter M., "IEEE 802.11 or ETSI BRAN HIPERLAN/2: Who will win the race for high speed wireless LAN standard?", Proceedings of European Wireless '99, Munich, Germany, Oct. 1999, pp. 169-174.

[27] C. L. Horney, "WLAN: Wireless Data´s Silver Bullet", Forward Concepts Co., 225 pages, Dec. 2002.

[28] M-F. Horng et al., "An Adaptive Approach to Weighted Fair Queue with QoS Enhanced on IP Network," IEEE Catalogue No. 01CH37239, 2001.

[29] T. Hämäläinen, E. Wallenius, T. Nihtilä, K. Luostarinen and J.Joutsensalo "Providing QoS at The Integrated WLAN and 3G Environments", Proceedings of the 14th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, Beijing, China, 7-10 September 2003, pp. 2014-2018.

[30] T. Hämäläinen, E. Wallenius, T. Nihtilä, K. Luostarinen: "End-to-End QoS Issues at the Integrated WLAN and 3G Environments", Proceedings of the 9th IEEE Asia-Pacific Conference on Communications (APCC), Penang, Malaysia 21-24 September 2003 pp. 639-643.

[31] IEEE 802.11 WG, Draft Supplement to STANDARD FOR Telecommunications and Information Exchange Between Systems - LAN/MAN Specific Requirements - Part 11: wireless Medium Access Control (MAC) and physical layer (PHY) specifications: Medium Access Control (MAC) Enhancements for Quality of Service (QoS), IEEE 802.11e/D2.0, Nov. 2001.

[32] IEEE P802.16e/D12, "Air Interface for Fixed and Mobile Broadband Wireless Access Systems", February, 2005.

[33] IEEE, "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications", ANSI/IEEE Std 802.11, 1999 Edition.

[34] IETF Web site: http://www.ietf.org, DiffServ and IntServ Working Groups.

[35] ITU-T Recommendation I.371. "Traffic control and congestion control in BISDN," ITU-T Study Group 13, April 1996, Geneva.

[36] Jacobson V., Nichols K., Poduri K., "An Expedited Forwarding PHB", RFC 2598, Jun 1999, www.ieee.org/rfc.html.

[37] N. Jason, G. James, J. Duane, A. Gerard, "SVR4 UNIX Scheduler Unacceptable for Multimedia Applications", Proceedings of the 4th International Workshop on Network and Operating System Support for Digital Audio and Video, Lancaster, UK Nov. 1993, pp 41-53. ISBN 3-540-58404-8.

[38] J. Joutsensalo and T. Hamalainen, "QoS-Aware Adaptive Pricing for Network Services," Proc. of IEEE Globecom, Vol. 4, USA, Nov. 2001, pp. 2455-2459.

[39] J. Joutsensalo and T. Hämäläinen, "Optimal Link Allocation and Revenue Maximization," Journal on Communications and Networks, Vol. 4, No. 2, June 2002, pp. 136-147.

[40] J. Joutsensalo and T. Hämäläinen, ”Optimal Link Allocation and Revenue Maximization”. Journal of Communications and Networks, Vol. 4, No. 2, June 2002, pp. 136–147.

[41] J. Joutsensalo, O. Gomzikov, T. Hamalainen, and K. Luostarinen: "Enhancing Revenue Maximization with Adaptive WRR".Proceedings of Eighth IEEE International Symposium on Computers and Communication, (ISCC 2003), Turkey, 2003, pp. 175 -180.

[42] J. Joutsensalo, T. Hämäläinen and J. Zhang, ”Revenue maximization-based adaptive WFQ,” In Proceedings of SPIE symposium of Asia-Pacific Optical and Wireless Communication (APOC2002), Volume 4907, pp. 108-117, October 13-18, 2002, China.

[43] J. Joutsensalo, T. Hämäläinen, J. Siltanen, and K. Luostarinen, "Delay Guarantee and Bandwidth Allocation for Network Services", proceedings of the 1st Conference on Next Generation Internet Networks, Italy, April 2005, pp. 129 - 134.

[44] J. Joutsensalo, T. Hämäläinen, M. Pääkkonen, and A. Sayenko: ”QoS- and Revenue Aware Adaptive Scheduling Algorithm”, Journal of Communications and Networks, Vol. 6, No.1 March 2004, pp. 68-77.

[45] J. Joutsensalo, T. Hämäläinen, Mikko Pääkkönen, and Alexander Sayenko, ”Adaptive Weighted Fair Scheduling Method for Channel

Allocation". In Proceedings of IEEE International Conference on Communications, ICC'03, 2003. Volume 1, 11-15 May 2003, Anchorage, Alaska, Page(s): 228 - 232.

[46] Jyrki Joutsensalo, Timo Hämäläinen, Kari Luostarinen, and Jarmo Siltanen, "Adaptive Scheduling Method for Maximizing Revenue in Flat Pricing Scenario", AEU - International Journal of Electronics and Communications, vol. 60, issue 2, February 2006, pp. 159-167.

[47] Kassler A., Schulthess P., "An End-to-End Quality of Service Management architecture for Wireless ATM Networks", Proceedings of the 32nd Hawaii International Conference on System Sciences, 1999.

[48] Wolfgand Keller, Mathias Wagner, Robert Hirschfeld, Josef Noll, Stein Sveat, Jaime Ferreira, Olavi Karasti, Teddy Hudginson, Rafaelle Giaffreda, Stewart Fallis, John Charles Farancis, Christian Fischer, "Systems beyond 3G– Operator's vision", Eurescom Project P1203, December 2002.

[49] S. Keshav, "On the efficient implementation of fair queuing," Journal of Internetworking: Research and Experience, Sep. 1991, vol. 2, pp. 57-73.

[50] Jong-Seon Kim; Lee, D.C., Weighted round robin packet scheduler using relative service share, Proceedings of IEEE of MILCOM'01, Vol.2, 2001, pp. 988-992.

[51] L. Kleinrock, "Queueing System Volume I: Theory", John Wiley & Son Inc., 1975.

[52] S. Knight, D. Weaver, D. Whipple, R. Hinden, D. Mitzel, P. Hunt, P. Higginson, M. Shand and A. Lindem, "Virtual Router Redundancy Protocol," RFC 2338, April 1998.

[53] H. Knoche, H. De Meer. "QoS parameters: A comparative study", University of Hamburg. Tech. Rep. 1997, Hamburg, Germany.

[54] Lakshman K., Yavaykar R. and Finkel R., "Integrated CPU and Network I/O QoS Management in an Endsystem", Proc. 5 th International Workshop on Quality of Service (IWQOS97), Columbia University, New York, USA, pp. 167-178.

[55] T. Li, B. Cole, P. Morton and D. Li, "Cisco Hot Standby Router Protocol (HSRP)," RFC 2281, March 1998.

[56] Lindgren A., Almquist A., and Schel´en O., "Quality of Service schemes for IEEE 802.11 - a simulation study". Proceedings of the Ninth International Workshop on Quality of Service (IWQoS'2001), Jun 2001.

[57] Linthicum, D. "Next Generation Application Integration: From Simple Information to Web Services", Addison-Wesley, Boston, 2004.

[58] D. H. Lorenz and A. Orda, "Optimal partition of QoS requirements on unicast paths and multicast trees," In Proc. of IEEE INFOCOM'99, pp. 246-253, March 1999.

[59] S. Lu, T. Nandagopal, and V. Bharghavan, "Fair scheduling in wireless packet networks," Proceedings of ACM/IEEE international conference on Mobile computing and networking (MOBICOM'98), October 1998, pp 10-20.

[60] S. Lu, V. Bharghavan, and R. Srikant, "Fair Scheduling in Wireless Packet Networks," IEEE/ACM Trans. Networking, vol. 7, no. 4, Aug. 1999, pp. 473-489.

[61] K. Luostarinen, J. Joutsensalo, L. Kannisto, T. Hämäläinen, "Delay Minimization and Pricing Method for the Network Services", To be published in ACTA Electrotechnica et Informatica Journal in Kosice, Hungary, 2006.

[62] K. Luostarinen, O. Gomzikov, J. Joutsensalo, and T. Hämäläinen, "Delay Guarantee, Bandwidth Allocation, and Pricing in Multinode Network", Proceedings of The IASTED International Conference on Networks and Communication Systems (NCS2006), March 29 th - 31 th, 2006, Chiang Mai, Thailand.

[63] E. Magana, D. Morato, and P. Varaiya, "Router scheduling configuration based on the maximization of benefit and carried best effort traffic", Telecommunication systems, 24 (2-4): 275292, Oct-Dec 2003.

[64] Mangold S. and Choi S. and Esseling N., "An Error Model for Radio Transmissions of Wireless LANs at 5GHz," in Proc. Aachen Symposium'2001, Aachen, Germany, pp. 209-214, Sept. 2001.

[65] Mangold S., Hmaimou M., Wijay H., "Coexistence of IEEE 802,11a and ETSI BRAN HiperLan/2: The problem of fair recourse sharing in the license exempt band at 5 GHz" Proceedings of IEEE International Conference on Third Generation Wireless Communications, Jun 2000.

[66] Mangold, S. Choi, P.May, O. Klein, G. Hiertz, L. Stibor, , “IEEE 802.11e Wireless Lan for Quality of Service”, Proceedings of the European Wireless, Italy, Feb. 2002.

[67] S. I. Maniatis, E. G. Nikolouzou, and I. S. Venieris, " QoS Issues in the Converged 3G Wireless and Wired Networks ". IEEE Communications Magazine, pp. 44-53, Aug.2002.

[68] Metso Annual Review 2005

[69] Dave Molta, "A Future Without Cables", Journal of Network Computing, Vol. 17, No. 14, July 20, 2006, pp 40-44.

[70] Myers M.B., "Predicting and measuring Quality of Service for mobile multimedia", 3G Mobile Communication Technologies, Conference Publication No. 471 IEEE 2000.

[71] R. Nagarajan, J. F. Kurose and D. Towsley, ”Allocation of Local Quality of Service Constraints to Meet End-to-End Requirements,” IFIP workshop on the Performance Analysis of ATM Systems, Martinique, Jan. 1993.

[72] T.S. Ng, I. Stoica and H. Zhang, "Packet fair queueing algorithms for wireless networks with location-dependent errors," Proceedings of IEEE INFOCOM '98, 1998, pp. 1103-1111.

[73] Ni Qiang, Lamia Romdhani, Thierry Turletti and Imad Aad, A Survey of QoS Enhancements for IEEE 802.11 Wireless LAN, Journal of Wireless Communication and Mobile Computing (JWCMC), 2003 <http://www-sop.inria.fr/planete/qni/Research.html>

[74] Nihtilä T., "Quality of Service in Wireless Communication Networks", Master Thesis, University of Jyväskylä, 2003, 107 pages.

[75] Prashanth Pappu and Tilman Wolf, "Scheduling Processing Resources in Programmable Routers", Department of computer science, Washington University, St. Louis, MO, USA, WUCS-01-32, July 25, 2001.

[76] A. K. Parekh and R. G. Gallager, "Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks: The Single Node Case", IEEE/ACM Transactions on Networking,vol. 1, no. 3, June 1993, pp. 344-357.

[77] Mirja Pulkkinen, Anton Naumenko, Kari Luostarinen, "Managing Information Security in a Business Network of Machinery Maintenance Services Business – Enterprice Architecture as a Coordination Tool", to be

published in Journal of Systems and Software, Special issue: Methodology of Security Engineering for Industrial Security Management Systems.

[78] Pyötsiä, J. "ICT opportunities and challenges for remote services."Proceedings of the 1-st International IFIP/WG12.5 Working Conference on Industrial Applications of Semantic Web, August 25-28, 2005, Jyvaskyla, Finland, Springer, IFIP, (2005) 213 – 225.

[79] P. Ramanathan and P. Agrawal, "Adapting Packet Fair Queueing Algorithms to Wireless Networks," Proceedings of ACM/IEEE international conference on Mobile computing and networking (MOBICOM'98), 1998, pp. 1-9.

[80] Schreiber, F., "Effective Control of Simulation Runs by a New Evaluation Algorithm for Correlated Random Sequences," in AEÜ International Journal of Electronics and Communications, vol. 42, no. 6, pp. 347–354, 1988.

[81] H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", Audio-Video Transport Working Group, RFC 1889. January 1996.

[82] H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson. "RTP: A Transport Protocol for Real-Time Applications," RFC 3550, July 2003.

[83] M. Shreedhar, G. Varghese, "Efficient fair queuing using deficit round-robin", IEEE/ACM Transactions on Networking, Vol. 4 No. 3 , June 1996, pp. 375 - 385.

[84] Sobrinho J.L., Krishnakumar A.S., "Quality-of-Service in Ad Hoc Carrier Sense Multiple Access Wireless Networks", IEEE Journal on selected areas in communications, Vol 17, Nr. 8, Aug 1999, pp. 1353-1368.

[85] J.L. Sobrinho, A.S. Krishnakumar, "Real-Time Traffic over the IEEE 802,11 Medium Access Control Layer", Bell Labs Technical Journal, Autumn 1996, pp. 174-187.

[86] M. Srivastava, C. Fragouli, and V. Sivaraman, "Controlled Multimedia Wireless Link Sharing via Enhanced Class-Based Queueing with Channel-State-Dependent Packet Scheduling," Proceedings of IEEE INFOCOM'98, March 1998.

[87] I. Stoica and H. Zhang, "Providing guaranteed services without per flow management," Proc.of ACM SIGCOMM 1999, Boston, Sept. 1999, pp. 81-94.

[88] I. Stoica, S. Shenker, and H. Zhang, “Core-stateless fair queuing: Achieving approximately fair bandwidth allocations in high speed networks,” IEEE/ACM Transactions on Networking, Vol. 11, No: 1 , Feb. 2003, pp. 33 – 46.

[89] Ion Stoica, Scott Shenker and Hui Zhang, "Core-stateless Fair queueing: A Scalable Architecture to Approximate Fair Bandwidth Allocations in High Speed Networks", in SIGCOMM 98.

[90] Vaidya N.H., Bahl P., Gupta S., "Distributed Fair Scheduling in a Wireless LAN", Proceedings of MOBICOM’00, 2000.

[91] E. Wallenius, T. Hämäläinen, T. Nihtilä, K. Luostarinen, "3G Interworking with WLAN QoS 802.11e", Proceedings of IEEE Vehicular Technology Conference, (VTC Fall), 6-9 October 2003, Orlando, USA, pp. 1803-1806

[92] A. Westerinen, J. Schnizlein, J. Strassner, M. Scherling, B. Quinn, S. Herzog, A. Huynh, M. Carlson, J. Perry, S.Waldbusser, "Terminology for Policy-Based Management", RFC 3198, November 2001.

[93] WiMAX Forum: "Mobile WiMAX – Part I: A Technical Overview and Performance Evaluation", August, 2006.

[94] Wroclawski J., "The Use of RSVP with IETF Integrated Services", RFC 2210, Sep 1997, www.ieee.org/rfc.html.

[95] XII Valmet Paper Technology Days, June 14 – 15, 2000, Turku, Finland, pp 171.

[96] J.Zhang, K.Luostarinen, T.Hämäläinen and J.Joutsensalo, "Enabling QoS Support in Global Replicated Web Systems". In proceedings of the 9th IEEE Asia-Pasific Conference on Communications (APCC2003), volume 2, 21th – 24th September 2003, Penang, Malaysia pp. 735 – 739.

[97] J. Zhang, T. Hämäläinen and J. Joutsensalo, ”A New Mechanism for Supporting Differentiated Services in Cluster-based Network Servers,” In Proceedings of 10th IEEE International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunications Systems (MASCOTS’02), pp. 427-432, October 12-16, 2002, Fort Worth, Texas.