Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Computer Communications 27 (2004) 1684–1692
computer communications
www.elsevier.com/locate/comcom

Review of quality of service performance in wireless LANs and 3G multimedia application services

Nilufar Baghaei, Ray Hunt*

Department of Computer Science, University of Canterbury, Private Bag, Christchurch 4800, New Zealand

Received 25 May 2004; accepted 25 May 2004
Available online 31 July 2004

Abstract

Although, many architectures (Integrated Services, Differentiated Services, MPLS, Traffic Engineering, etc.) have been proposed to provide service differentiation in fixed networks, research has shown that what works well in a wired network cannot be directly applied in the wireless environment where bandwidth is scarce and channel conditions are time varying. Quality of Service (QoS) is a key challenge for today's wireless IP networks and implementation of QoS, particularly for supporting voice, video, data and multimedia services in general incurs a number of difficulties that have to be analysed and resolved. A considerable amount of work has been carried out by the various standards groups in an effort to quantify and specify protocols to support QoS in wireless environments. This paper reports on these efforts outlining existing limitations, requirements and solutions proposed by organisations such as the IEEE 802.11 Task Group E for wireless LANs and the UMTS effort for 3G/wireless WANs.

Keywords: Quality of Service; IEEE 802.11a/b/g WLANs; 3G/UMTS/cdma2000 Networks; Multimedia over wireless networks

1. Introduction

Most current network architectures treat all packets in the same way—a single level of service. Applications, however, have diverse requirements and may be sensitive to latency and packet losses. Examples include interactive and real-time applications such as IP telephony; streaming services such as audio, video and bulk data streaming; and interactive services such as voice, web and transaction service processing. When the latency or the loss rate exceeds certain levels, these applications become unusable. In contrast, both reliable (tcp) and unreliable (udp) services can tolerate significant delay and loss without much degradation of perceived performance.

The capability to provide resource assurance in a network is often referred to as Quality of Service (QoS) which is critical requirement in order that new IP-based applications can operate within well-defined parameters. Resource assurance can currently be provided in a multiservice network by the use of IP service differentation (DiffServ), although resource assurance can only truly be guaranteed by the use of the Integrated Services (IntServ) model. Implementing these QoS capabilities has become one of the most difficult challenges, particularly as this often requires changes to its basic network architecture.

The requirements for each type of traffic flow can be characterised by four primary parameters: reliability, delay, jitter, and bandwidth. Several common applications are listed in Table 1 along with the stringency of their requirements. Table 2 (adopted from Ref. [1]) shows a similar set of requirements for applications operating in a wireless WAN (3G) environment.

Most IP-based networks rely on the TCP protocol to detect congestion in the network and to reduce the transmission rates accordingly. TCP-based resource allocation requires all applications to use the same congestion control scheme. Although such co-operation is achievable within a small group, in a network as large as the Internet, it can be easily abused. Furthermore, many UDP-based applications do not support TCP-like congestion control, and real-time applications typically cannot cope with large fluctuations in the transmission rate.

The service currently provided by default is often referred to as best effort. When a link is congested, packets are simply discarded as the queue overflows.

* Corresponding author. Tel.: +64-3-3642347; fax: +64-3-3642569.
E-mail address: ray@cosc.canterbury.ac.nz (R. Hunt).

0140-3664/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.comcom.2004.05.019

Table 1
Common wired-network performance characteristics

Application	Reliability	Delay	Jitter	Bandwidth
E-mail	High	Low	Low	Low
File transfer	High	Low	Low	Medium
Web access	High	Medium	Low	Medium
Remote login	High	Medium	Medium	Low
Audio on demand	Low	Low	High	Medium
Video on demand	Low	Low	High	High
Telephony	Low	High	High	Low
Videoconferencing	Low	High	High	High

Since, the network treats all packets equally, any flows could be hit by congestion and this particularly impinges on wireless and mobile connections, commonly as a result of limited bandwidth. Although best-effort service is adequate for some applications that can tolerate large delay variation and packet losses, it clearly does not satisfy the needs of many new time-sensitive multimedia-based applications.

Resource assurance is critical for many new wireless applications. Although, the Integrated Services (IntServ) and Differentiated Services (DiffServ) paradigms figure predominantly as QoS solutions, they focus on the IP layer and it is necessary for the underlying layers to be able to respond to and configure such IP-based service requirements. The following sections address the specification and provisioning of these underlying QoS-based requirements for both wireless LAN and wireless WAN (3G) architectures.

2. QoS in IEEE 802.11 Wireless LANs

In its current form, the IEEE 802.11 wireless LAN standard [2] cannot provide QoS support for the increasing number of applications which demand QoS parameters—typical of many multimedia applications. A number of IEEE 802.11 QoS enhancement schemes have been proposed, each focusing on a particular mode of operation. This section first analyses the QoS limitations of the IEEE 802.11 MAC layer and then summarises the QoS enhancement schemes that have been proposed and experimented with to date. Finally, it discusses the new IEEE 802.11e QoS enhancements.

Table 2
Common wireless WAN (3G) network performance characteristics

Application	Reliability	Delay	Jitter	Bandwidth
E-mail	Low	High	–	Low
File transfer	Low–medium	High	–	High
Web access	Low–medium	Medium	–	Medium
Remote login	Low	Low	–	Low
Control	Null	Low	–	–
Real time	Low–medium	Low	Low	Medium–high

2.1. An Overview of IEEE 802.11 MAC Operation

In general, the IEEE 802.11 WLAN standard covers the MAC sub-layer and the physical layer of the OSI network reference model. The Logical Link Control (LLC) sub-layer is specified in the IEEE 802.2 standard. This architecture provides a transparent interface to higher layer users: stations may move, roam through an IEEE 802.11 WLAN and still appear as stationary to the IEEE 802.2 LLC sub-layer and above. This allows existing network protocols (such as TCP/IP) to transparently operate over IEEE 802.11 WLANs without any special considerations.

At the PHY layer, the IEEE 802.11 standard provides three alternatives for operation in the 2.4 GHz band:

- Infrared (IR) base-band PHY
- Frequency Hopping Spread Spectrum (FHSS) radio
- Direct Sequence Spread Spectrum (DSSS) radio.

All three PHY layers support both 1 and 2 Mbps operation. In 1999, the IEEE defined an 11 Mbps 802.11b standard designed to operate in the 2.4 GHz free ISM (Industrial, Science, and Medical) band and a 54 Mbps 802.11a OFDM (Orthogonal Frequency Division Multiplexing) standard for operation in the 5 GHz frequency band.

The IEEE 802.11 MAC sub-layer defines two related medium access coordination functions, the Distributed Coordination Function (DCF) and the optional Point Coordination Function (PCF) (Fig. 1, adopted from Ref. [3]).

The IEEE 802.11 MAC protocol supports two types of transmission: Asynchronous and Synchronous [2]. Asynchronous transmission is provided by the DCF, which implements the basic access method for the IEEE 802.11 MAC protocol. DCF is based on the Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) protocol, and is the default implementation. The Synchronous service (also called contention free service) is provided by PCF and implements a polling-based access method. The PCF uses a centralised polling approach that requires an Access Point (AP) to act as a Point Coordinator (PC). The AP cyclically polls stations to give them the opportunity to transmit packets. Unlike the DCF, the implementation of the PCF is not mandatory. Furthermore, the PCF itself relies on the underlying asynchronous service provided by the DCF.

Although providing different service functions, neither DCF nor PCF + DCF have the ability to offer true QoS to Wireless LAN applications.

2.2. QoS limitations of IEEE 802.11 MAC

In addition to providing channel access (via DCF or PCF + DCF), the wireless LAN MAC layer needs to

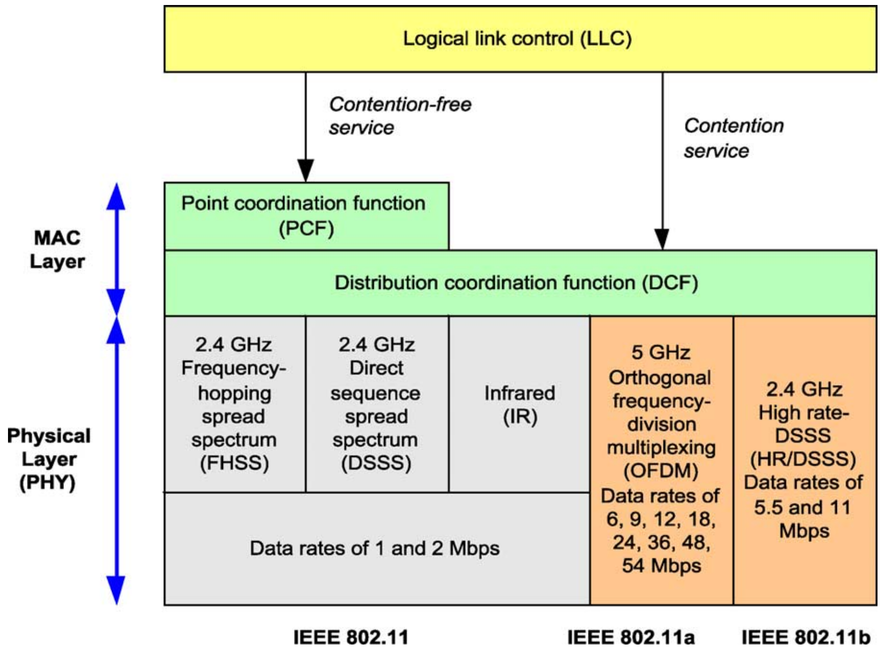

Fig. 1. PCF and DCF in IEEE 802.11 MAC layer.

provide facilities for:

- maintaining QoS
- providing security.

Wireless links have specific characteristics such as reliability, bandwidth, packet delay and jitter. Furthermore, the wireless link characteristics are not constant and may vary over time and place. Mobility of users may cause the end-to-end path to change when users roam, and further, users will expect to receive the same QoS as they change from one Access Point (AP) to another. This implies that the new path should also support the existing QoS by service reservation, and problems may arise when the new path cannot support such requirements.

There are two ways to characterise QoS in WLANs, viz.: parameterised or prioritised QoS [4,5]. Parameterised QoS is a strict QoS requirement, which is expressed in terms of quantitative values, such as data rate, delay bound, and jitter bound. In a Traffic Specification (TSpec)—such as is used in the IntServ model, these values are expected to be met by the MAC data service in support of the transfer of data frames between peer stations. In a prioritised QoS scheme, the values of QoS parameters such as data rate, delay bound, and jitter bound—typically resulting from a DiffServ model, may vary during the transfer of data frames. In this instance, there is no need to reserve the required resources by negotiating the TSpec between the station and the AP as the DiffServ queue architecture is relied upon to manage the QoS.

2.2.1. QoS limitations of DCF

DCF can only support best-effort services and does not provide any QoS guarantees. Typically, time-bounded services such as Voice over IP, audio and videoconferencing require specified bandwidth, delay and jitter, but can also tolerate some loss. However, in DCF mode, all the stations in one BSS or all the flows in one station compete for the resources and channel with the same priority. There is no differentiation mechanism to guarantee bandwidth, packet delay and jitter for high-priority stations or multimedia flows [6].

2.2.2. QoS limitations of PCF

Although PCF has been designed by the IEEE Working Group to support time-bounded multimedia applications, this mode has some major problems, which leads to poor QoS performance. In particular, the central polling scheme is inefficient and complex and causes deterioration of the performance of PCF high-priority traffic under load. Additionally, all communications have to pass through the AP which degrades the bandwidth performance [7].

2.3. QoS enhancement schemes for IEEE 802.11 MAC

QoS issues in wired Ethernet have been neglected due to the relative ease with which the physical layer bandwidth has improved Normally, the IP layer assumes that a LAN scarcely drops or delays packets. However, in WLANs, the challenges of the wireless channel make physical layer data rate improvements more difficult to achieve, particularly as the IEEE 802.11 WLAN was originally designed for best-effort services. The physical layer's error rate can be more than three orders of magnitude larger than that of wired LAN. Further, high collision rate and frequent retransmissions cause unpredictable delay and jitter, which further degrade the quality of real-time voice and video transmission. To address these issues, a number of proposals have been made which are detailed in the following sections.

2.3.1. Service differentiation based enhancement schemes

QoS enhancement can be supported by adding service differentiation into the MAC layer This can be achieved by modifying the parameters that define how a station or a flow should access the wireless medium. Current service differentiation-based schemes can be classified with respect to a multitude of characteristics. For example, a possible classification criterion can be based upon whether the schemes base the differentiation on per-station or per-queue (per-priority) parameters. Another classification depends on whether they are DCF (distributed control) or DCF + PCF (centralised control) enhancements. Fig. 2 shows this classification. Previous research work has mainly focused on the station-based DCF enhancement schemes [6,8,9], while other recent work has focused on queue-based hybrid coordination (combined PCF and DCF) enhancement schemes [10–12], since queue-based schemes perform more efficiently.

2.3.2. Error control based enhancement schemes

In parallel, QoS enhancement can also be obtained by error control mechanisms Since, the network may occasionally drop, corrupt, duplicate or reorder packets, the transport protocol (e.g. TCP) or the application itself (e.g. if UDP is being used) must recover from these errors on an end-to-end basis. Error recovery in the sub-network is justified only to the extent that it can enhance overall performance. However, some sub-networks such as wireless links require link layer error recovery mechanisms to enhance the performance, but these enhancements need to be lightweight [4]. For example, wireless links normally require link-layer error recovery (such as IEEE 802.2 LLC) and MAC-level error recovery in the sub-network.

2.3.3. IEEE 802.11e QoS enhancement standards

The focus of the IEEE 802.11 Task Group e (802.11e) is to enhance the IEEE 802.11 MAC (DCF, PCF) to support QoS, providing classes of service, enhanced security and authentication mechanisms. It aims to enhance the ability of all the physical layers (IEEE 802.11b, 802.11a, 802.11g) to

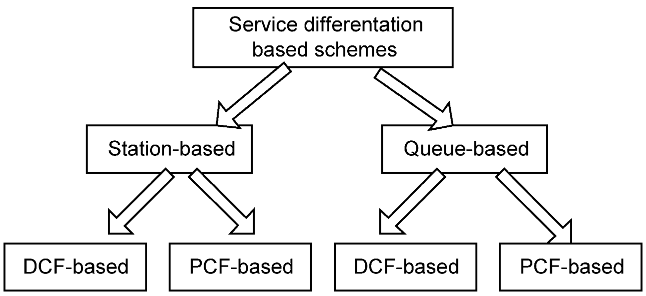

Fig. 2. Classification of service differentiation based schemes.

deliver time-critical multimedia data, in addition to a best effort data service. There are many new features in the IEEE 802.11e draft 3.0 [13] which enhance the existing DCF and PCF + DCF functionality in order to support new QoS applications [4] and the most important of these are briefly described.

- EDCF (Enhanced Distribution Coordination Function)
- HCF (Hybrid Coordination Function)
- Direct communication in infrastructure mode
- AP mobility
- MAC-level FEC (Forward Error Correction).

2.3.3.1. Enhanced distribution coordination function (EDCF) and hybrid coordination function (HCF). Since, DCF and PCF do not differentiate between traffic types or sources, the IEEE 802.11 Task Group e is proposing QoS enhancements via two alternative modes of operation These changes would let critical service requirements be fulfilled while maintaining backward-compatibility with current IEEE 802.11 standards.

The proposed enhancement to DCF—Enhanced Distribution Coordination Function (EDCF)—introduces the concept of traffic categories. Each station has eight traffic categories, or priority levels. Using EDCF, stations try to send data after detecting that the medium is idle and after waiting a period defined by the corresponding traffic category called the Arbitration Interframe Space (AIFS). A higher-priority traffic category will have a shorter AIFS than a lower-priority traffic category.

To avoid collisions within a traffic category, the station counts down an additional random number of time slots, known as a contention window, before attempting to transmit data. If another station transmits before the countdown has ended, the station waits for the next idle period, after which, it continues the countdown where it left off. No guarantees of service are provided, but EDCF establishes a probabilistic priority mechanism to allocate bandwidth based on traffic categories.

Another way IEEE 802.11e aims to extend the polling mechanism of PCF is with the Hybrid Coordination Function (HCF). A hybrid controller polls stations during a contention-free period. The polling grants a station a specific start time and a maximum transmit duration. EDCF appears to be gaining more acceptance than HCF [13,14]. In addition to the IEEE 802.11e standard, a group of vendors have proposed Wireless Multimedia Enhancements (WME), to provide an interim QoS solution for IEEE 802.11 networks.[1]

2.3.3.2. Additional IEEE 802.11e daft 3.0 features. IEEE 802.11e adds a capability for stations to send traffic directly

[1] For more information, refer to http://grouper.ieee.org/groups/802/11/Documents/D2T551-600.html.

to each other in the infrastructure mode, which significantly improves the bandwidth in certain networks.[2]

In IEEE 802.11e, AP mobility [13] is introduced by a concept of QAP-Capable and such a station can operate either as a real QoS AP (QAP) or a wireless station (WSTA). Therefore, in the IEEE 802.11e draft 3.0 [13], AP mobility implies the transfer of the AP functionality between different QAP-Capable stations, i.e. a station can become an AP and subsequently change back to a station again.

3. QoS in 3G Wireless Networks

2G networks such as Global System for Mobile communications (GSM)/code-division multiple access (CDMA), have essentially only one QoS option—i.e. speech at full-rate coding in GSM. Subsequently, a half-rate service was introduced, thus offering a new QoS. In reality, however, this was done to save network capacity and therefore serving more users in congested hot spots, rather than offering a new grade of service to users. The user was not offered the choice of full/half rate, but, more often, those with half-rate capable mobile phones were put onto half rate, without the subscriber knowing that the speech quality has been deliberately lowered by the network being used.

In 2.5G networks such as general packet radio service (GPRS), there has been a deliberate attempt to introduce mechanisms whereby the subscriber can request a different QoS (average/peak data throughput, packet delay, etc.). In principle, this QoS requirement can be established at the beginning of the data transfer session (at Packet Data Protocol (PDP) context set-up). For example, a user intending to use an interactive service (such as web surfing) may want to use a service with a faster reaction time/lower round-trip delay. They can then ask for a smaller packet delay at PDP context set-up time, and the network can confirm whether this request is accepted or rejected.

3G is a wireless industry term for a collection of international standards and technologies aimed at improving the performance of mobile wireless networks 3G wireless services offer packet data enhancements to applications and these include higher speeds, increased capacity for voice and data services as well as QoS facilities. The two main 3G technologies for which QoS is being standardised are:

- UMTS (Universal Mobile Telecommunications System)
- cdma2000 (Code Division Multiple Access 2000).[3]

Table 3 (adopted from Ref. [15]) shows QoS-based application requirements in terms of bandwidth, delay, and losses for different categories such as data, real-time traffic, non-real time traffic, and network services in 3G networks.

Table 3
Typical QoS application requirements in 3G

Type of application and example			(Kbps)	Losses (%)	Delay (ms)
Data		FTP	Limitless	0	TCP timer
Real time	Audio	Voice	≤ 64	10^{-4}	≤ 300
		Voice over IP	10–64	$5-10^{-2}$	≤ 300
	Video	MPEG-4	≤ 2000	10^{-2}	≤ 40
		H.320	≤ 64	10^{-4}	≤ 40
Non-real time	Audio	CD	150	10^{-4}	Buffer length
	Video	MPEG-4	Limitless	0	Buffer length
Network services			Limitless	0	–

3.1. UMTS/3GPP-defined QoS

3GPP,[4] Third Generation Partnership Project, has standardised a common QoS framework for IP-based data services. They have defined a comprehensive framework for end-to-end QoS covering all subsystems in a UMTS network, including core network, wireless and universal terrestrial radio access networks, etc. UMTS is the first wireless data service, which offers a comprehensive QoS specification across a wireless wide area network infrastructure. In addition, the specification provides for control signaling, user plane transport and QoS management functionality.

QoS enables a network to deliver classes of service (CoS), i.e. different prioritised treatments to different services or to different groups of users. QoS allocates network capacity according to the type of traffic required for a certain type of service, while CoS provides preferred allocation of the network resources in a similar manner to which DiffServ does for IP-based services. CoS is implied in a QoS policy associated with a subscriber. It is used by the network to provide differential QoS treatments to different services subscribed by different users.

UMTS defines QoS classes [16], but cdma2000—at least in its early formulations—does not. Users of these services may communicate with both fixed networks and other mobiles, thus end-to-end performance is also influenced by the features of these networks on which other parties may be situated.

The 3GPP end-to-end QoS specification, which includes the definition of UMTS QoS architecture, bearer services, and recommendations for supporting QoS mechanisms, also

[2] This would be the same as ad hoc mode, however in ad hoc mode, there is no AP defined in the architecture.

[3] 3GPP is responsible for the UMTS standards specification, while 3GPP2 is responsible for the cdma2000 standards specification. This has resulted in two 3G standards being released, i.e. UMTS and cdma2000.

[4] www.3GPP.org.

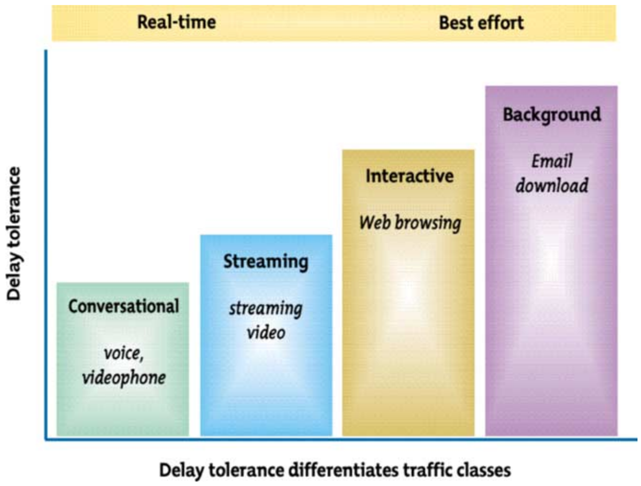

Fig. 3. UMTS traffic classes.

establishes four overriding UMTS QoS classes or traffic classes for wireless data, taking into account the restrictions and limitations of the air interface. The characteristics of these four QoS classes are described in the following sections.

3.1.1. UMTS QoS Basic Classes

The basic classes defined by UMTS/3GPP are [17]:

1. Conversational
2. Streaming
3. Interactive
4. Background.

The main distinguishing factor between these traffic classes lies with sensitivity to delay (Fig. 3 (adopted from Ref. [18])).

3.1.1.1. Conversational class. This class applies to any application that involves real-time person-to-person communication such as audio voice, videophone etc. The basic qualities required for speech are low delay, low jitter, reasonable clarity (common codecs and quality), and absence of echo. In the case of multimedia applications, such as videoconferencing, it is also necessary to maintain synchronisation of the different media streams. Failure to provide low enough transfer delay will result in unacceptable lack of quality. This class is tolerant of some errors, e.g. voice packet corruption lasting for up to 20 ms. However, the degree of error protection required varies with applications. (Table 4)

3.1.1.2. Streaming class. The streaming class consists of real-time applications that exchange information between viewer and listener, without any human response. Examples of this include video-on-demand, live MPEG4 listening, Web-radio, news streams, and multicasts. Because of the absence of interaction, there is no longer a need for low delay, but the requirements for low jitter and media synchronisation remain. Error tolerance is a function of the application. The removal of the low delay criterion makes it possible to use buffering techniques in the end-user equipment, so the acceptable level of network jitter is higher than for the conversational class. (Table 5).

3.1.1.3. Interactive class. This class covers both humans and machines that interact with another device Examples of this include some games, network management systems polling for statistics, and Web-browsing or database retrieval. Applications in this class are characterised by the request-response pattern of the end-user. Round-trip delay and tolerance to packet loss are key QoS characteristics. (Table 6).

3.1.1.4. Background class. The background class covers all applications that either receive data passively or actively request it, but without any immediate need to handle this data Examples of this include e-mails, short message service, and file transfers. The only requirement is for data integrity, although large file transfers will also require an adequate throughput.

Table 7 summarises the characteristics of each of the above four classes.

Table 4
End-user performance expectations—conversational/real-time services

Medium	Application	Degree of symmetry	Data rate (Kbps)	Key performance parameters and target values		
				End-to-end one way delay	Delay variation within call	Information loss
Audio	Conversational voice	Two-way	4–25	<150 ms pref. <400 ms limit	<1 ms	<3% Frame error rate
Video	Videophone	Two-way	32–384	<150 ms pref. <400 ms limit Lip-sync <100 ms		<1% Frame error rate
Data	Telemetry—two way control	Two-way	<28.8	<250 ms	NA	Zero
Data	Interactive games	Two-way	<1	<250 ms	NA	Zero
Data	Telnet	Two-way—asymmetric	<1	<250 ms	NA	Zero

Table 5
End-user performance expectations—streaming services

Medium	Application	Degree of Symmetry	Data rate (Kbps)	Key performance parameters and target values		
				Startup delay (s)	Transport delay variation	Packet loss at session layer
Audio	Speech/music medium/high quality	Primarily one-way	5–128	<10	<2 s	<1% Packet loss ratio
Video	Movie clips surveillance real-time video	Primarily one-way	20–384	<10	<2 s	<2% Packet loss ratio
Data	Bulk data transfer/retrieval, layout, synchronisation information	Primarily one-way	<384	<10	NA	Zero
Data	Interactive games	Primarily one-way		<10	NA	Zero

Table 6
End-user performance expectations—interactive services

Medium	Application	Degree of Symmetry	Data rate	Key performance parameters and target values		
				One-way delay	Delay variation	Information loss
Audio	Voice messaging	Primarily one-way	4–13 Kbps	<1 s for playback <2 s for record	<1 ms	<3% Frame error rate
Data	Web-browsing—HTML	Primarily one-way		<4 s/page	NA	Zero
Data	Transaction services—high priority, e.g. ATM	Two-way		<4 s	NA	Zero
Data	E-mail (server access)	Primarily one-way		<4 s	NA	Zero

3.1.2. UMTS QoS targets for mobility categories

UMTS defines specific QoS targets to provide an adequate service to mobile wireless users [17] (Table 8).

In UMTS, the maximum speed envisaged for the high mobility category is 500 km/h using terrestrial services to cover all high-speed train services and 1000 km/h using satellite links for aircraft. The data-rate bands are also related to the environment and cell size in UMTS with rural environments and satellite links being restricted to 144 Kbps aggregate for a single mobile and with 2 Mbps only available as an instantaneous rate rather than a guaranteed rate.

3.2. cdma2000 QoS

cdma2000 does not set such explicit QoS targets nor define its own classes analogous to UMTS. In practice, however, it supports the same general range of applications and provides appropriate degrees of support through the use of radio link level features and the capabilities of Mobile IP.

3.2.1. cmda2000 QoS control plane

The IP Multimedia Subsystem (IMS) framework, end-to-end QoS support requires signalling, traffic regulation and resource allocation capabilities [19]. QoS signalling is used to provision and enforce QoS parameters between endpoints and is handled in the application layer, network layer and link layer. The Session Initiation Protocol (SIP) [20] is used as the application level signalling protocol to establish sessions while the Session Description Protocol (SDP) [21] parameters are carried as part of the SIP payload, and contain session specific information.

The three sequential steps for QoS signalling and traffic regulation are as follows:

- Mobile Station (MS) or user registration
- Session initiation, including media and QoS negotiation
- Traffic regulation.

SIP is used for user registration, session management and media and QoS negotiation. Session management involves initiation and termination of a SIP session using a three-way handshake with INVITE, OK and ACK messages. Media

Table 7
UMTS QoS traffic classes

Traffic class	Conversational class	Streaming class	Interactive class	Background class
Fundamental characteristics	Preserve timing of stream conversational pattern—stringent, low-delay	Preserve timing of stream	Request response pattern preserve payload content	Destination does not care about arrival time Preserve payload content
Application Example	Voice	Streaming video	Web Browsing	Background —e.g. e-mails

Adopted from 3GPP 23.107.

Table 8
IMT-2000 mobility categories

Category	Physical speed	Data rate
Limited mobility	Up to 10 km/h	2 Mbps
Full mobility	Up to 120 km/h	384 Kbps
High mobility	More than 120 km/h	144 Kbps

and QoS negotiation is conducted using a SDP payload of SIP messages [22]. SDP negotiation contains parameters that are related to QoS, such as media stream codec and bandwidth requirements.

SIP is an application layer control and signalling protocol used for creating, modifying, and terminating multimedia sessions with one or more participants. These participants could be two mobile stations or a mobile station and an application server. SIP messages (e.g. INVITE) carry session descriptions that allow participants to agree on a set of compatible media types and QoS requirements. SIP operates over different transport protocols such as TCP or UDP.

QoS parameters are negotiated between endpoints running SIP user agents through the SIP proxy and AAA server. The Policy Decision Point (PDP) implements the SIP proxy to determine the allowed QoS parameters based on SIP negotiation and local policy of the network. Session specific QoS parameters are exchanged via SDP while SIP header fields containing the QoS parameters are enforced using the Policy Enforcement Point (PEP) as part of the Packet Data Serving Node (PDSN) in cdma2000 [19].

3.2.2. cmda2000 QoS data plane

QoS-based data in cdma2000 is handled by the DiffServ architecture at the IP layer. These may result from a SIP-based signalled request or be based on user profile 3GPP2 differentiated services class options registered with a home RADIUS server. In the case of Mobile IP, the service class has to be copied by the home agent from the IP packets to the differentiated services field of the mobile tunnel.

The QoS targets (so far) [23] for audio and video streaming are:

- *Synchronisation*. For transmission of combined audio and video streams, the inter-media skew should be kept below 20 ms.
- *Minimum bandwidth*. Minimum bandwidth for audio and video streaming applications are defined.
- *Play-out delay*. The video streaming service shall be able to provide service of reasonable end-to-end delay to accommodate data transfer from the source to the mobile terminal, and shall support buffering at the terminal to accommodate transmission path degradations to a specific level. The recommended maximum play-out delay is 30 s.
- *Delay jitter*. The system shall be able to operate under delay jitter of three times the RLP (Radio Link Protocol) retransmission time in the network with retransmission activated.
- *Error rate*. The service shall operate over channels with end-to-end BER (Bit Error Rate) in the order of 10^{-3} (for circuit-switched network services) and FER (Frame Error Rate) in the order of 10^{-2} (for packet-switched network services).

For videoconferencing, the targets are similar, except the play-out delay has to be much less so that end-to-end delay does not exceed 400 ms. The degree of jitter that must be compensated is up to 200 ms. Throughput must range from 32 Kbps upwards, including the specific rates of 384 and 128 Kbps for packet and circuit switching, respectively.

4. Conclusions

Providing QoS for modern audio and video-based multimedia applications is a key challenge for today’s wireless mobile networks. Limited bandwidth, varying channel conditions, mobility, as well as QoS interface requirements between a variety of wireless and wired network infrastructures is a very complex problem to solve. The specification of current standards focus on Layer 2 for wireless LANs (typically implying upper layer support of a differentiated-based IP architecture) while for wireless WANs the protocol specification encompasses multiple layers of the protocol stack.

This paper has addressed the fundamental concepts of QoS provisioning in wireless LANs and 3G networks. Much work has yet to be carried out in order to offer this same service across a concatenation of fixed/mobile and wired/wireless networks. Further, addition of security functionally (defined individually for WLANs via IEEE802.1x or IPSec and WWANs via IPSec, WAP2.0, etc) across these various network infrastructures will require a considerable amount of further research.

Although advances have been made in the bandwidth capacity of 3G networks (UMTS networks can now offer up to 2 Mbps with limited mobility), the problems become much more complex as issues of diverse and multiple networks are added. Network operators rarely have end-to-end control over a data path, and the problems of guaranteeing IP-based QoS service across multiple networks remains. While mechanisms such as RSVP offer QoS guarantees, this still relies upon these mechanisms being implemented by the service providers across multiple wired/wireless networks and the expectation that the underlying lower layer infrastructure can respond to these stringent requirements.

References

[1] S.V. Vasconcellos, J.F. Rezende, QoS and mobility in 3G Core Networks, Proceedings of Workshop on QoS and Mobility, Brazil, Nov. 2002, www.gta.ufrj.br/publicacoes.

[2] IEEE 802.11 WG, Reference number ISO/IEC 8802-11:1999(E) IEEE Std 802.11, 1999 ed.

[3] G. Held, The ABCs of IEEE 802.11, IT Professional, 3(6), Institute of Electrical and Electronics Engineering, Inc., Nov. 2001, pp. 49–52.

[4] Q. Ni, L. Romdhani, T. Turletti, I. Aad, QoS Issues and Enhancements for IEEE 802.11Wireless LAN, INRIA, France, 2002.

[5] J.M. Ho, Some comments on 802.11e Draft 2.0, IEEE 802.11e working document 80211-02/005r0, January 16, 2002.

[6] I. Aad, C. Castelluccia, Differentiation mechanisms for IEEE 802.11, IEEE Infocom 2001, Anchorage, Alaska, USA, April 2001.

[7] A. Lindgren, A. Almquist, O. Schelen, Evaluation of quality of service schemes for IEEE 802.11 wireless LANs, Proceedings of the 26th Annual IEEE Conference on Local Computer Networks, Florida, USA, Nov. 15–16, 2001.

[8] J. Deng, R.S. Chang, A priority scheme for IEEE 802.11 DCF access method, IEICE Transactions in Communications 82-B (1) (1999).

[9] A. Veres, A.T. Campbell, M. Barry, L.H. Sun, Supporting service differentiation in wireless packet networks using distributed control, IEEE Journal of Selected Areas in Communications (JSAC), Special Issue on Mobility and Resource Management in Next-Generation Wireless Systems 19 (10) (2001) 2094–2104.

[10] I. Aad, C. Castelluccia, Remarks on per-flow differentiation in IEEE 802.11, in: Proceedings of European Wireless (EW2002), Florence, Italy, Feb. 2002.

[11] S. Mangold, S. Choi, P. May, O. Klein, G. Hietz, L. Stibor, IEEE 802.11e wireless LAN for quality of service, Proceedings of European Wireless (EW2002), Florence, Italy, Feb. 2002.

[12] L. Romdhani, Q. Ni, T. Turletti, AEDCF: enhanced service differentiation for IEEE 802.11 wireless ad-hoc networks, INRIA Research Report No. 4544, Sept. 2002.

[13] 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/Draft 3.0, May 2002.

[14] A. Ganz, A. Phonphoem, Z. Ganz, Robust super-poll with chaining protocol for IEEE 802.11 wireless LANs in support of multimedia applications, Wireless Networks 7 (2001) 65–73.

[15] F. Fitzek, M. Krishnam, M. Reisslein, Providing application-level QoS in 3G/4G wireless systems: a comprehensive framework based on multi-rate CDMA, IEEE Wireless Communications 9 (2) (2002) 42–47.

[16] ETSI TS 123 107, QoS Concepts and Architecture, V4.3.0, 2001, http://www.etsi.org.

[17] ETSI TS 122 105, Services and Service Capabilities, V5.2.0, 2002, http://www.etsi.org.

[18] Nortel Networks white paper, Benefits of Quality of Service (QoS) in 3G wireless Internet, www.nortelnetworks.com, 2001.

[19] M.A. Siddiqui, K. Guo, S. Rangarajan, S. Paul, End-to-End QoS Support for SIP Sessions in CDMA2000 Networks, Bell Labs Technical Journal 9 (3) (2004).

[20] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, E. Schooler, SIP: Session Initiation Protocol, RFC 3261, Jun, 2002.

[21] S. Gaast, A. Hajjaoui, E. Meeuwissen, Quality of Service for SIP Sessions in 3GPP-based Networks, Bell Labs Technical Journal 9 (2004) 3.

[22] G.Ed. Camarillo, W.Ed. Marshall, J. Rosenberg, Integration of Resource Management and Session Initiation Protocol (SIP), RFC-3312 (Oct. 2002).

[23] 3GPP2 S.R0021/R0022, Video Conferencing Services—Stage 1, 2002. http://www.3gpp2.org.