2007 IEEE International Conference on Signal Processing and Communications (ICSPC 2007), 24-27 November 2007, Dubai, United Arab Emirates

A New Strategy for Providing QoS in IEEE 802.16 Standard

H.R.Khanzadi, S.M Alavi, B.Vosoughi.Vahdat, Smaeil Sanaei
Electrical Engineering Department
Sharif University of Technology
Tehran, Iran
akhanzadi@mehr.sharif.edu, smalavi@mehr.sharif.edu, vahdat@ sharif.edu, sanaei@sharif.edu

ABSTRACT

This article presents a new QoS architecture for IEEE 802.16 Broadband Wireless MAN. The newly proposed strategy of providing QoS in packet-based networks via IEEE 802.16 has been analyzed. A mapping rule for providing layer 2.5(MPLS protocol) and MAC layer has been proposed. A fast signaling mechanism has been designed to provide QoS in Point to Multipoint (PMP) mode. Although IEEE 802.16 standards define different mechanisms to provide Quality of Service requirements, it is the responsibility of the developers to obtain efficient designs. In this paper, we introduce a novel architecture for providing QoS requirements such as bandwidth utilization and delay. Simulation result shows the high performance of our architecture for traffic classes defined by the standard

Keywords—IEEE 802.16, MPLS, CR-LDP, QoS

1. INTRODUCTION

Over the last few years, the Internet has evolved into a ubiquitous network. The new applications have driven the demand for increased and guaranteed bandwidth requirements. In addition to traditional data services, new voice and multimedia services are being developed. The Internet has emerged as the network of choice for providing these services. Transformation of the network toward a packet-based infrastructure has introduced uncertainty into what has traditionally been a fairly deterministic network. Multi-Protocol Label Switching (MPLS) is a versatile solution to address the problems faced by present-day networks-speed, QoS management, and traffic engineering. MPLS has emerged as a elegant solution to meet the bandwidth management and service requirements for Internet based network [3]. In the other hand IEEE 802.16 Wireless MAN has emerged as a last mile solution for providing services to end users [1]. It assumes a point-to multipoint topology with a Base and several Subscriber Stations. Base Station (BS) controls and manages the entire system and each Subscriber Station (SS) performs as an interface between end users and the Base Station. The IEEE 802.16 standards define a connection-oriented MAC protocol. The downlink channel (from BS to SSs) uses TDM scheme and the uplink channel in opposite direction applies TDMA scheme. IEEE 802.16 defines four types of service flows, each with different QoS requirement [1]:

Unsolicited Grant Service (UGS):

The UGS is designed to support real-time service flows that generate fixed size data packets on a periodic basis, such as T1, E1 and Voice over IP without silence suppression. The service offers fixed size grants on a real-time periodic basis.

Real Time Polling Service (rtPS):

The rtPS is designed to support real-time service flows that generate variable size data packets on a periodic basis, such as MPEG video.

Non-Real Time Polling Service (nrtPS):

This service is for non-real-time flows which require better than best effort service, e.g. bandwidth intensive file transfer like FTP applications.

Best Effort Service (BE):

This service is for best effort traffic such as HTTP. There is no QoS guarantee.

IEEE 802.16 standards use specific request and grant mechanism in which each SS indicates the amount of uplink bandwidth it needs to the BS. The BS is allowed to allocate bandwidth in two modes; Grant Per Connection (GPC), in which bandwidth is assigned to each connection, and Grant Per Subscriber Station (GPSS), in which an SS requests for transmission opportunities for all of its connections.

IEEE 802.16 Wireless MAN can support multiple types of data traffic such as data, voice and video with different QoS requirements. QoS signaling mechanisms and functions is defined in MAC layer to control data transmission

1-4244-1236-6/07/$25.00 © 2007 IEEE

between BS and SSs. Since IEEE 602.16 MAC protocol is connection oriented, the application must establish the connection with the BS as well as the associated service flow. These service flows characterized by QoS parameters such as maximum sustained traffic rate, minimum reserved traffic rate, throughput, jitter, latency. These service flows can be created, changed or deleted through the issue of DSx-messages such as Dynamic Service Addition (DSA), Dynamic Service Change (DSC), and Dynamic service Deletion (DSD). In [1], some rules to classify IP packets based on Type of Service (ToS) field in IP header has been proposed. Therefore, in general, the QoS architecture of IEEE 802.16 under PMP mode can support QoS in network (IP) layer. Each SS in IEEE 802.16 standard has mandatory Basic and Primary management connection (uses for exchanging MAC management messages), Secondary management connection (for exchanging Protocol-based messages such as DHCP,SNMP,TFTP,etc) and one or more transport connections for transmitting traffic data. MPLS signaling messages classified as protocol-specific messages and will be transmitted in the secondary management connection to provide layer 3 QoS. On the other hand DSA/DSC/DSD messages will be transmitted in the primary management connection to provide layer 2 QoS. Since the secondary management connection is defined for delay traffic and there are many other IP protocol related messages (DHCP, SNMP, TFTP, etc) sharing the same queue the whole QoS provision will be rather slow. In this paper we propose a mapping strategy between MPLS messages and DSx-messages that efficiently improve bandwidth utilization.

2. MPLS AND COSTRAINT-BASED LDP

Label distribution protocol (LDP) is defined for distribution of labels inside one MPLS domain. One of the most important services that may be offered using MPLS is support for Constraint-based Routing of traffic [2]. CR-LDP offers the opportunity to extend the information used to setup paths beyond what is available for the routing protocol. For instances an LSP can be set up based on explicit route constraint, QoS constraint, etc. Explicit Routing is a subset of Constraint-based Routing. Like any other LSP a CR-LSP is a path through an MPLS network. The difference is that while other paths are setup solely based on information in routing tables, Constraint-based Route is setup based on some criteria for better support of the traffic sent over the LSP. The traffic characteristics of a path are described in the traffic parameters TLV of messages distributed through network in terms of a peak rate, committed rate, and service granularity. The peak and traffic rates describe the bandwidth constraints of a path while the service granularity can be used to specify a constraint on a delay variation. There is a field in Label request and label mapping messages named traffic parameters TLV that is used to signal the parameters values. These parameters are Peak Data Rate (PDR), Peak Burst Size (PBS), Committed Data Rate (CDR), Committed Burst Size (CBS) and Excess Burst Size (EBS).

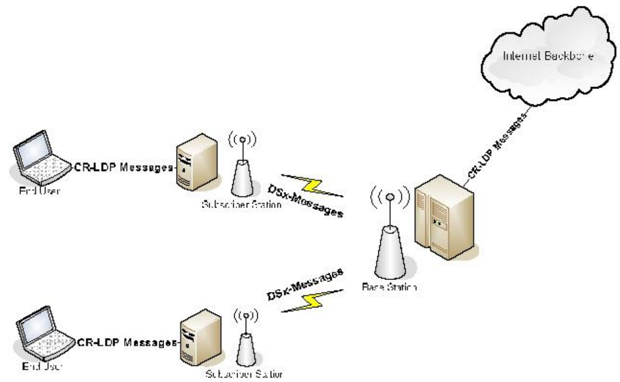

Figure 1. Traffic classification and Mapping

3. MAPPING STRATEGIES

To specify the QoS characteristics of a data service there are a number of mandatory QoS parameters that have to be included in the service flow definition when the scheduling service is enabled for it. Each scheduling service has a minimum number of associated parameters which are listed in Table I [4].

Since there are so many similarities between providing QoS requirement in MPLS using CR-LDP and MAC layer QoS using DSA/DSC/DSD, we can map the traffic parameters TLV in CR-LDP messages into traffic parameters in DSx-messages (Figure 1). This mechanism is superior to the traditional way in providing high efficiency and fastness.

The Traffic Parameters TLV contains a Flags field, a Frequency, a Weight, and the five Traffic Parameters PDR, PBS, CDR, CBS, EBS. (Figure 2)

<table>
<tr><td>0</td><td>0</td><td colspan="2">Type = 0x0810</td><td colspan="2">Length =24</td></tr>
<tr><td colspan="3">Flags</td><td>Frequency</td><td>Reserved</td><td>Weight</td></tr>
<tr><td colspan="6">Peak Data Rate(PDR)</td></tr>
<tr><td colspan="6">Peak Burst Size(PBS)</td></tr>
<tr><td colspan="6">Committed Data Rate (CDR)</td></tr>
<tr><td colspan="6">Committed Burst Size(CBS)</td></tr>
<tr><td colspan="6">Excess Burst Size(EBS)</td></tr>
</table>

Figure 2. Traffic Parameters TLV format

We use the peak and committed rates for bandwidth and resource (buffer) reservation and the Frequency field for delay and jitter specification. The Frequency field is coded as an 8 bit integer with the following code points defined:

0: Unspecified

1: Frequent

2: Very Frequent

3-255: reserved

The value "Very Frequent" means that the available rate should average at least the CDR when measured over any time interval equal to or longer than the shortest packet time at the CDR. The value "Frequent" means that the available rate should average at least the CDR when measured over any time interval equal to or longer than a small number of shortest packet times at the CDR. The value "Unspecified" means that the CDR may provided at any granularity.

4. ADMISSION CONTROL AND SCHEDULING IN BS

BS based on DSA/DSC/DSD messages that receives, calculates the available bandwidth (BW_a). Since each connection is associated with a single data service, suppose there are I connections and *rate* is the Minimum reserved traffic rate so the available bandwidth is calculated as follows [6]:

$$BW_a = BW_{total} - \sum_{i=0}^{I} rate(i) \tag{1}$$

In which $rate(i)$ is the Minimum Reserved traffic rate of i^{th} connection. For UGS services the Minimum Reserved traffic rate is equal to Maximum sustained traffic rate.

For receiving DSA/DSC massages the admission control policy has the following principle:

$$BW_a \geq 0 \tag{2}$$

We use a packet scheduling with these following properties:

- For rtPS traffics it uses Rate-Controlled Static Priority (RCSP). An RCSP server consists of two components: a rate controller and a static priority scheduler. When a packet arrives at the server, an eligibility time is calculated and assigned to the packet by the regulator. The packet is held in the regulator until its eligibility time.
- For nrtPS, first in according to traffic priority field in DSx-messages, packets are assigned to two priority queues. This field take values from 0 to 7 that higher numbers indicate higher priorities. According to this field, we divide these kinds of traffics into a high priority queue (with traffic priority field 4 to 7) and the lower priority queue (with traffic priority field 0 to 3). Afterward, a WFQ is deployed on these two queues.

TABLE I. MAPPING RULES

Scheduling service	**Parameters**	CR-LDP Parameters	**Possible applications**
UGS	Maximum Sustained Traffic Rate	PDR and PBS	ATM CBR; E1/T1 over ATM; TDM Voice; T1/E1; VoIP without silence suppression
	Maximum Latency	Frequency	
	Tolerated Jitter	Frequency	
	Minimum Reserved Traffic Rate	CDR and CBS	
rtPS	Maximum Sustained Traffic Rate	PDR and PBS	MPEG video; VoIP with silence suppression
	Maximum Latency	Frequency	
	Minimum Reserved Traffic Rate	CDR and CBS	
nrtPS	Minimum Reserved Traffic Rate	CDR and CBS	ATM GFR; TFTP;HTTP; FTP
	Maximum Sustained Traffic Rate	PDR and PBS	
	Traffic Priority	Used for queue priority	
BE	Maximum Sustained Traffic Rate	PDR and PBS	E-Mail; P2P file sharing
	Traffic Priority	Used for queue priority	

- For BE traffic in first level the use of traffic priority is applied then an RR is deployed on the queues. (Figure 3)

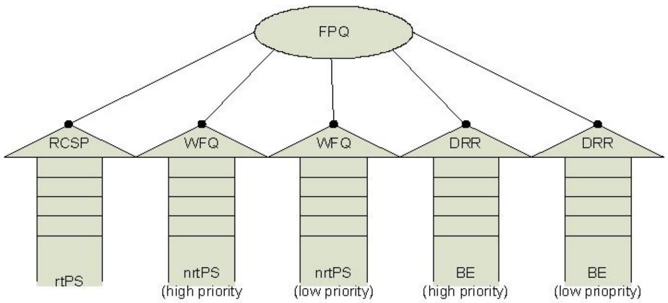

Figure 3. packet scheduling and bandwidth allocation

5. SIMULATION AND RESULTS

In order to evaluate and analyze the performance of our strategy, we have developed a simulation using NS-2.28 (Network Simulator) program. The objective of the experiment is to show that the proposed strategy can provide QoS support in terms of delay and bandwidth utilization. Our configuration consists of one BS and four SS with total bandwidth of 60 Mbps in each direction (uplink and downlink).Each SS has four services with these traffics: 2 Mbps UGS traffics, 2 Mbps rtPS traffics, 2 Mbps nrtPS traffics and 9 Mbps BE traffics. Since Pareto distribution is heavy-tailed distribution that can be used to model the bursty nature of traffic or self-similar properties of HTTP traffic, we used this distribution for modeling of BE traffic. For nrtPS services we use FTP traffic that used normal distribution for file size modeling. For rtPS services we used MPEG video that used constant distribution for frame inter-arrival time and normal distribution for packet size [5].

Results were evaluated using Average Delay and Bandwidth Utilization metrics. Average Delay is defined as average latency between the reception of a packet by the BS or SS on its network interface and the forwarding of the packet to its physical interface [1]. Bandwidth Utilization specifies the proportion of the total bandwidth granted to each SS that is utilized for data and message transmission. Figure 4 shows Bandwidth Utilization, in *SS*. We achieved 7ms Average Delay. While the bursty nature of best effort traffic makes bandwidth utilization more inefficient than other services, the result figured in 4 shows that our method always gains approximately full bandwidth utilization.

6. CONCLUSION

In this paper, we proposed a novel architecture to support Quality of Service mechanisms in IEEE 802.16 standards together with a design to implement such architecture. A fast signaling mechanism using combination of MPLS signaling and Dynamic Service signaling has been designed to provide QoS. The simulation results proved the appropriate performance of the proposed architecture.

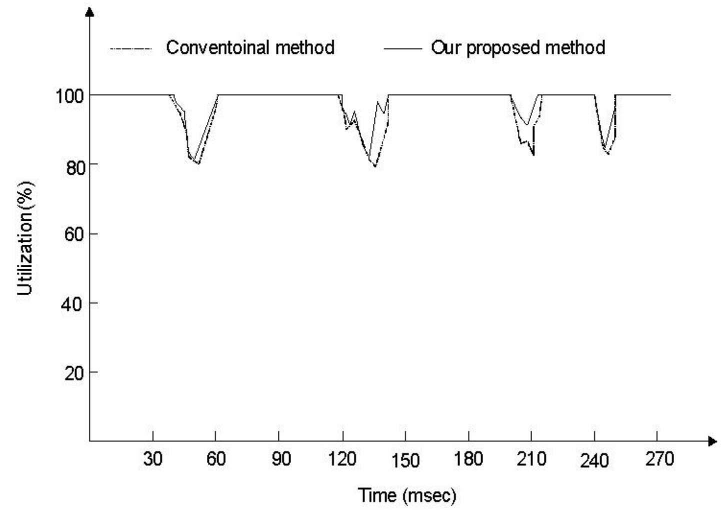

Figure 4. Banwidth Utilization

REFERENCES

[1] IEEE P802.16-REVd/D5-2004: "Air Interface for Fixed Broadband Wireless Access Systems"

[2] Constraint-Based LSP Setup using LDP, RFC 3212,January2002,http://www.faqs.org/rfcs/rfc3212.html

[3] Multiprotocol Label Switching (MPLS),The International Enginnering Consertium , http://iec.org

[4] Carlos Bilich, Wireless access network project number II: "TCP over WIMAX networks", University of TRENTO, department of information and communication technology, June 16 , 2005

[5] B. Ryu and S. Lowen, "Fractal Traffic Models for Internet Simulation" July 2000 edition of IEEE Int'l Symposium on Computer Communication

[6] Jianfeng Chen, Wenhua Jiao, Hongxi Wang "A Service Flow Management Strategy for IEEE 802.16 Broadband Wireless Access Systems in TDD Mode",2005