國立交通大學

網路工程研究所

碩 士 論 文

一個 IEEE 802.16 基地台上考慮延遲與調變之排程機制

A Latency and Modulation Aware Scheduling Solution for IEEE 802.16 Base Stations

研 究 生：張原誌

指導教授：陳耀宗 教授

中 華 民 國 九 十 七 年 八 月

一個 IEEE 802.16 基地台上考慮延遲與調變之排程機制
A Latency and Modulation Aware Scheduling Solution for IEEE 802.16 Base Stations

研 究 生：張原誌　　Student：Yuan-Chih Chang

指導教授：陳耀宗　　Advisor：Yaw-Chung Chen

國 立 交 通 大 學
網 路 工 程 研 究 所
碩 士 論 文

A Thesis
Submitted to Institute of Network Engineering
College of Computer Science
National Chiao Tung University
in partial Fulfillment of the Requirements
for the Degree of
Master
in

Computer Science

August 2008

Hsinchu, Taiwan, Republic of China

中華民國九十七年八月

A Latency and Modulation Aware Scheduling Solution for IEEE 802.16 Base Stations

Student：Yuan-Chih Chang

Advisor：Yaw-Chung Chen

A Thesis

Submitted to Institute of Network Engineering

College of Computer Science

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Computer Science

August 2008

Hsinchu, Taiwan, Republic of China

一個 IEEE 802.16 基地台上考慮延遲與調變之排程機制

學生：張原誌　　　　　　　　　指導教授：陳耀宗 博士

國立交通大學網路工程研究所

摘　　要

雖然行動WiMAX系統提供了高速的資料傳輸速率以及較廣的頻寬。然而，在一般空氣介質中仍有許多干擾因素存在，因此由於這些干擾使得資料傳輸通道品質在長距離的傳輸上較為不穩定。在WiMAX系統上，若發現傳輸通道品質發生變化，則WiMAX無線網路節點可以改變使用的調變與編碼機制。因此，我們考慮調變與編碼的機制提出了在WiMAX基地台上的排程解決方案。此排程解決方案主要包含三大機制：許可控管機制，排程演算法，頻寬分配機制。除了考慮調變與編碼機制之外，我們針對即時的應用服務也將延遲納入排程考量。

在我們所提出的許可控管機制中，根據節點所使用之根據調變與編碼的機制以及最低保留速率要求計算出進入許可之條件。我們提出精準的許可控管機制不僅可以提昇整個系統的資料吞吐量，亦可以幫助排程器得以保證所有被允許進入連線的最低保留速率。我們考慮調變與編碼機制所提出的排成演算法以及頻寬分配機制，藉由將服務品質的要求轉換為所需的時間來達到保證所有被允許進入連線之最低服務品質需求，而且更能在有限的訊框時間中有效率地分配頻寬。最後列舉出幾項整個排程解決方案的特性： 考慮調變與延遲，在基地台上易於實做，保證 UGS,rtPS,nrtPS 這些服務之最低服務品質，有效率地分配頻寬。最後，模擬結果顯示出我們所提出之排程機制不僅可以提昇整體網路資料吞吐量而且保證了所有被允許進入網路連結之最低服務品質要求。

A Latency and Modulation Aware Scheduling Solution for IEEE 802.16 Base Stations

Student：Yuan-Chih Chang　　　Advisor：Dr. Yaw-Chung Chen

Institute of Network Engineering

National Chiao Tung University

Abstract

The mobile WiMAX system provides high data rate, wide bandwidth for wireless network. However, there are too many interference in the air interface. So, the link quality is not stable owing to air interference and long transmission. When link quality change, mobile node can revise the modulation and coding scheme dynamically. Thus, we propose the modulation-aware scheduling solution for WiMAX Base station. It contains three primary schemes: admission control scheme, scheduling algorithm, and bandwidth allocation scheme. In addition to modulation and coding scheme, we also take latency for real-time service into account.

The proposed admission control scheme calculate the admit condition based on the different kinds of modulation and coding schemes and the minimum reserved rate requirements. Our precise admission control scheme can improve the total throughput of entire network and help the scheduler guarantee the minimum reserved rate to each admitted connection. The modulation aware scheduling algorithm and bandwidth allocation scheme can allocate bandwidth more efficiently in the finite frame duration and guarantee the minimum QoS requirements to each admitted connection by translating the QoS requirements into time slots. Some properties of proposed scheduling solution are as follows: modulation and latency aware, easily implemented in BS, guarantee minimum QoS for some specific services (UGS, rtPS, nrtPS), and allocate bandwidth efficiently. Finally, our simulation results show that the proposed scheduling solution can improve the throughput of the entire network and guarantee the minimum QoS requirements to each admitted connection.

Keyword: 802.16, WiMAX, QoS, scheduling, modulation aware.

Acknowledgement

First of all, I would like to express my sincerity to Prof. Yaw-Chung Chen for not only providing valuable suggestions and ideas during my graduate study but let me not always think with narrow sight. Thanking to Prof. Yaw-Chung Chen for leading me successive growing.

Besides, I appreciate plenty to the members of Multimedia Communication Laboratory, including REX, Cheng-Yuan Ho, amegoo, ydoon, Hz, and Alex, especially to REX, for their cheers so that I have had a great time in my school life.

Finally, I would like to express my appreciation to my family, without their support behind; I won’t be with current status.

Contents

Table List

Figure List

Chapter 1 Introduction

The IEEE 802.16 standard defines a wireless broadband access network technology called Worldwide Interoperability for Microwave Access (WiMAX). The main advantages of WiMAX are the higher data rate and the longer transmission range. Besides, it introduces several characteristics, such as the support for Quality-of- Service (QoS) at the MAC layer and mobility.

In 2001, IEEE 802.16-2001 standard defines the specification of WiMAX. In WiMAX system, the transmission rate can achieve 70Mbps or even 100Mbps, the transmission range is up to 20km and the operating frequency within the interval 10–66 GHz. In January 2003, IEEE 802.16a-2003 standard adds some specification, the two of which are centre frequency within the interval 2–11 GHz and the support for Non-Line-of-Sight (NLOS) environment. Until October 2004, IEEE 802.16-2004 [1] standard was published. It consolidates IEEE standards 802.16-2001, 802.16a-2003 and 802.16c-2002. These early documents based on IEEE 802.16-2004 are designed for fixed applications so we refer to these as fixed WiMAX.

In December 2005, the IEEE 802.16e-2005 [2] was proposed. It is an amendment to the IEEE 802.16-2004 standard that adds the support for mobility. The IEEE 802.16e-2005 becomes a fundamental for WiMAX solution and mobile applications and is referred to as mobile WiMAX. The basic characteristics of the various IEEE 802.16 standards are summarized in Table 1.1

The WiMAX standard defines two basic operational modes to achieve medium sharing: point to multipoint (PMP) and Mesh. In the PMP mode, a subscriber station (SS) can communicate with other stations only through the base station (BS). In the Mesh mode, a SS can communicate with other subscriber stations or BS directly. The

PMP mode is illustrated in Figure 1.1.

document	Description
Dec. 2001, 802.16	10–66 GHz; line-of-sight(LOS); transmission rate: 70Mbps
Jan. 2003, 802.16a	2-11 GHz; non-line-of-sight (NLOS)
Oct. 2004, 802.16-2004	consolidates previous 802.16 standards;
Dec.2005, 802.16e amendment	Mobility; OFDMA (SOFDMA)

Table 1.1 The basic characteristics of the various IEEE 802.16 standards [3]

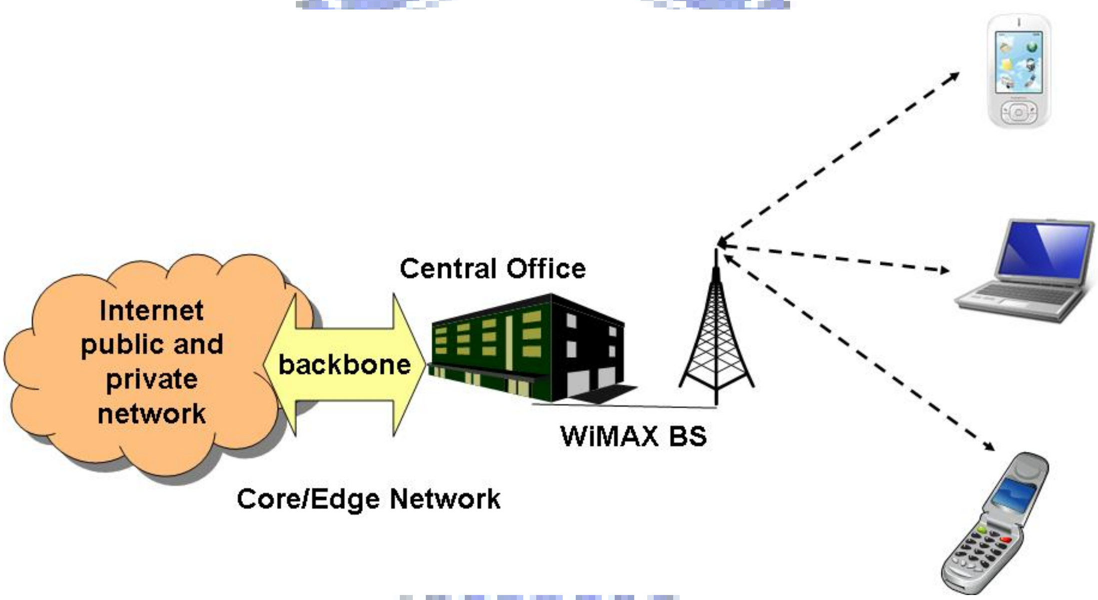

Figure 1.1 IEEE 802.16 PMP mode.

We use the PMP mode in our proposed scheduling algorithm because in this case, we can control the environment easily to ensure the QoS customers' requirements. QoS is an important issue in WiMAX network especially for real-time service flows belonging to voice or video applications. The 802.16e standard MAC Layer provides QoS differentiation for the different types of applications, it defines five scheduling service types, i.e. Unsolicited Grant Service (UGS), real-time Polling Services (rtPS), non-real-time Polling Service (nrtPS), and Best Effort (BE).

To ensure meeting the QoS requirements for different service flows, we take the QoS parameters which defined in 802.16e standard into account. In order to utilize resource of network efficiently, we consider the modulation and coding scheme used in different SS, the latency requirement, the maximum sustained traffic rate, and the minimum reserved traffic rate in our scheduling solution.

Although WiMAX system provides high data rate and wide bandwidth, the channel quality is unstable due to the long-distance and electromagnetic interference. Thus, the scheduling solution we proposed is modulation-aware. The scheduling solution that we proposed contains admission control, scheduling and bandwidth allocation mechanism. Furthermore we take latency parameter into account for real-time service flow to calculate deadline of packets.

Several research works [5, 6, 7, 8] propose using complex scheduling algorithm, such as weighted fair queueing and worst-case weighted fair queueing, and even a hierarchy schedulers. These scheduling algorithms are not so simple to implement. It is a difficult task to use a hierarchy of scheduler when the frame duration is very brief in time. Thus, we propose one level easy-implemented scheduling solution.

We test the scheduling solution by choosing NIST WiMAX module in the NS-2 simulator that provides a better platform to simulate realistic network topologies, service flow characteristics, and behavior of the WiMAX system.

The rest of this article is organized as follows. Chapter 2 presents an overview of the WiMAX system architecture and the background about scheduling in WiMAX. Next, chapter 3 describes the scheduling solution we proposed in detail. The numerical analysis and simulation evaluation are presented in chapter 4. The conclusions are stated in chapter 5.

Chapter 2 Background

In this chapter, we first introduce some procedures and characteristics at 802.16 MAC layer, including the bandwidth allocation, service classes, dynamic service flow procedure, and classification and mapping for different connections. Before scheduling, we add algorithm into dynamic service flow procedure to achieve admission control. Regarding scheduling, design for MAP, and bandwidth allocation are very important issues.

Next, we briefly express the WiMAX frame architecture and MAC support of PHY layer. The 802.16 standard contains two main duplexing techniques: Time Division Duplexing (TDD) and Frequency Division Duplexing (FDD). We choose TDD duplexing in our proposed scheduling solution due to the TDD duplxing is more suitable when data rates are asymmetrical.

2.1 Review of IEEE 802.16 MAC

2.1.1 Mac layer of WiMAX

The MAC protocol layers architecture of the 802.16 standard is shown in Figure 2.1, is divided into three distinct components: the service specific convergence sub-layer (CS), the common-part sub-layer, and the security sub-layer. The CS is the top sub-layer of the MAC Layer in WiMAX/802.16, which accepts higher-layer PDUs from the higher layers and transmits them to the MAC CPS where classical type MAC procedures are applied. These higher-layer packets are known as MAC service data units (SDU). The CS performs operations that are dependent on the nature of the higher-layer protocol, such as address mapping and header compression.

The WiMAX MAC layer is connection oriented and identifies a logical connection between the BS and the SS by the unidirectional connection identifier (CID). In order to map the higher-layer address to the CID, the CS needs to keep track of the mapping between the destination address and the respective CID. It is quite likely that higher-layer SDUs belonging to a specific destination address might be carried over different connections, depending on their QoS requirements, in which case the CS determines the appropriate CID, base on not only the destination address but also various other factors, such as service flow ID (SFID), and source address.

The common-part sub-layer performs all the packet operations that are independent of the higher layers, such as transmission of MAC protocol data units (PDU), concatenation of SDUs into MAC PDUs, ARQ, and QoS control. The security sub-layer is responsible for encryption, authorization and proper exchange of encryption keys between the BS and the SS.

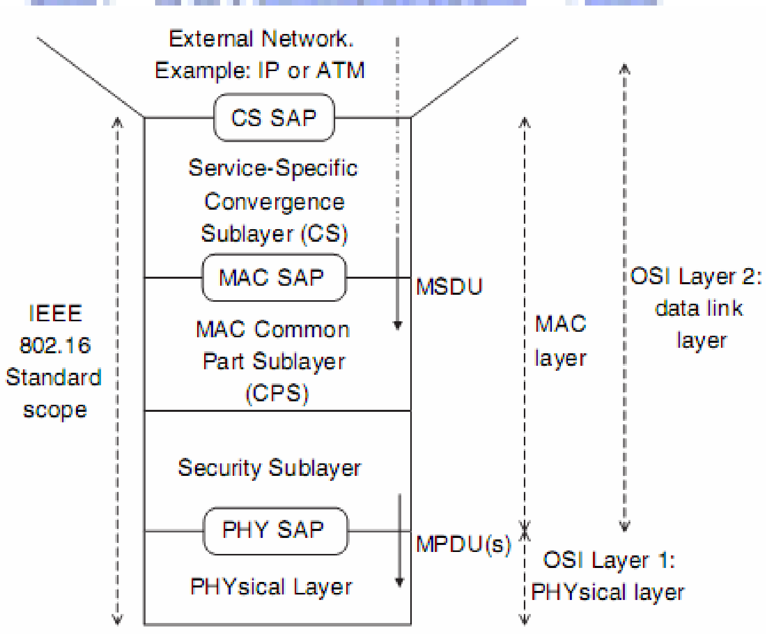

Figure 2.1 MAC layers of the 802.16 standard [3]

The common-part sub-layer performs all the packet operations that are independent of the higher layers, such as transmission of MAC protocol data units (PDU), concatenation of SDUs into MAC PDUs, ARQ, and QoS control. The security sub-layer is responsible for encryption, authorization and proper exchange of encryption keys between the BS and the SS.

2.1.1 Connections and Service Flow

The CS provides transformation or mapping of external network data received through the CS sub-layer into MAC Service Data Units (SUDs) received by the common-part sub-layer. This includes classifying external network Service Data Units (SDUs) and associating them with the proper MAC Service Flow Identifier (SFID) and Connection Identifier (CID). There are two 802.16 MAC layer fundamental concepts about classification and mapping. They are introduced as follows:

- Connection

  A connection is a MAC level connection between a BS and an SS. It is a unidirectional mapping between a BS and SS MAC peers for the purpose of transporting a service flow's traffic. A connection is only for one type of service. A connection is identified by a CID (Connection Identifier). The CID is coded on 16bits.

- Service flow

  A service flow is a MAC transport service that provides unidirectional transport of packets either to uplink packets transmitted by the SS or to downlink packets transmitted by the BS. A service flow is characterized by a set of QoS Parameters such as latency, jitter, maximum sustained rate, minimum reserved rate, and throughput assurances. A service flow is partially characterized by the following attributes:

- Service Flow ID: An SFID is assigned to each existing service flow. The SFID serves as the principal identifier for the service flow in the network.
- CID: Mapping to an SFID that exists only when the connection has an admitted or active service flow.
- ProvisionedQoSParamSet: A QoS parameter set provisioned via means outside of the scope of this standard, such as the network management system.
- AdmittedQoSParamSet: Defines a set of QoS parameters for which the BS (and possibly the SS) is reserving resources. The principal resource to be reserved is bandwidth, but this also includes any other memory or time-based resource required to subsequently activate the flow.
- ActiveQoSParamSet: Defines a set of QoS parameters defining the service actually being provided to the service flow. Only an Active service flow may forward packets.
- Authorization Module: A logical function within the BS that approves or denies every change to QoS Parameters and Classifiers associated with a service flow. As such, it defines an “envelope” that limits the possible values of the AdmittedQoSParamSet and ActiveQoSParamSet.

As defined in the standard, classification is the process by which a MAC SDU is mapped on to a particular connection for transmission between MAC peers. This process allows 802.16 BWA to deliver MAC SDUs with the appropriate QoS constraints. Classification and mapping mechanisms exist in the uplink and downlink. In the case of a downlink transmission, the classifier will be present in the BS and in the case of an uplink transmission it is present in the SS. The MAC CS Layer classifies each application. If a packet matches the specified packet matching criteria, it is then delivered to the SAP for delivery on the connection defined by the CID. The service flow characteristics of the connection provide the QoS for that packet. The classification mechanism is shown in Figure 2.2.

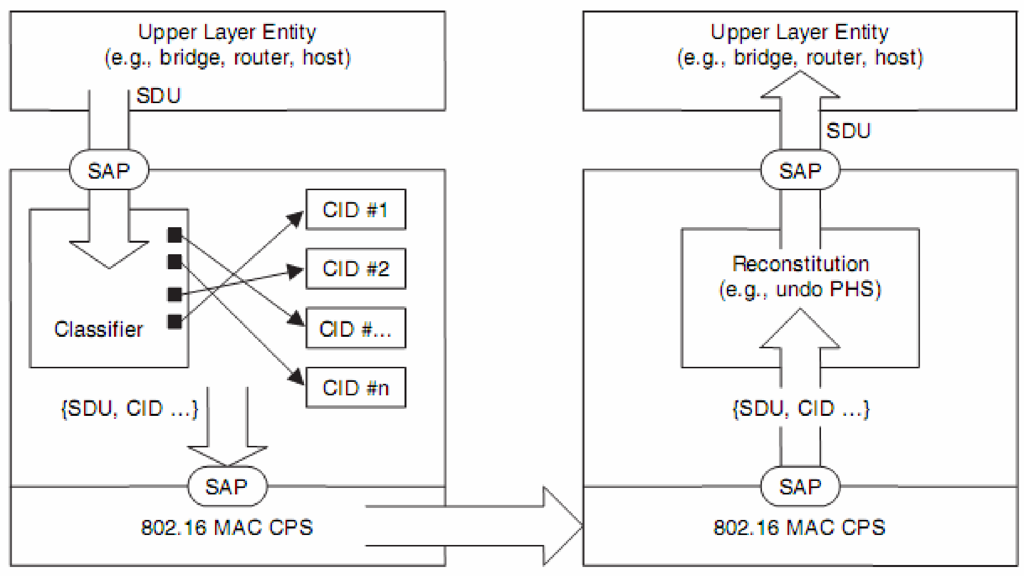

Figure 2.2 The classification and CID mapping mechanism [3]

Dynamic Service Flow Creation

Creation of service flows may be initiated by either BS or by SS. The SS-initiated protocol is illustrated in Figure 2.3. A Dynamic Service Addition Request message (DSA-REQ) from an SS contains a service flow reference and QoS Parameter set (marked either for admission-only or for admission and activation). In our scheduling solution, the QoS Parameter set of DSA-REQ message includes maximum latency, maximum sustained rate, and minimum reserved rate. We make use of the minimum reserved rate to calculate the condition of admission. BS responds with DSA-RSP indicating acceptance or rejection. In the case when rejection was caused by presence of non-supported parameter of non-supported value, specific parameter may be included into Dynamic Service Addition Response message (DSA-RSP).

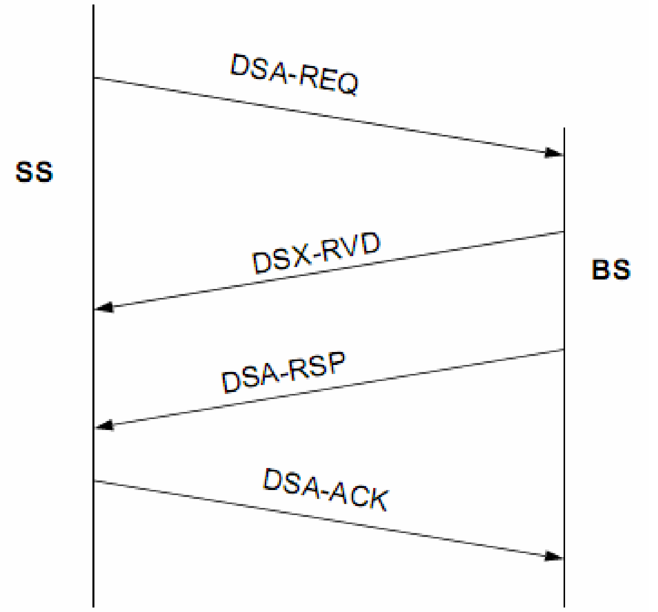

Figure 2.3 DSA message flow—SS-initiated

The BS-initiated protocol is illustrated in Figure 2.4. A DSA-REQ from a BS contains an SFID for either uplink or downlink Service flow, possibly its associated CID, and a set of active or admitted QoS Parameters. SS responds with DSA-RSP indicating acceptance or rejection. In the case when rejection was caused by presence of non-supported parameter of non-supported value, specific parameter may be included into DSA-RSP.

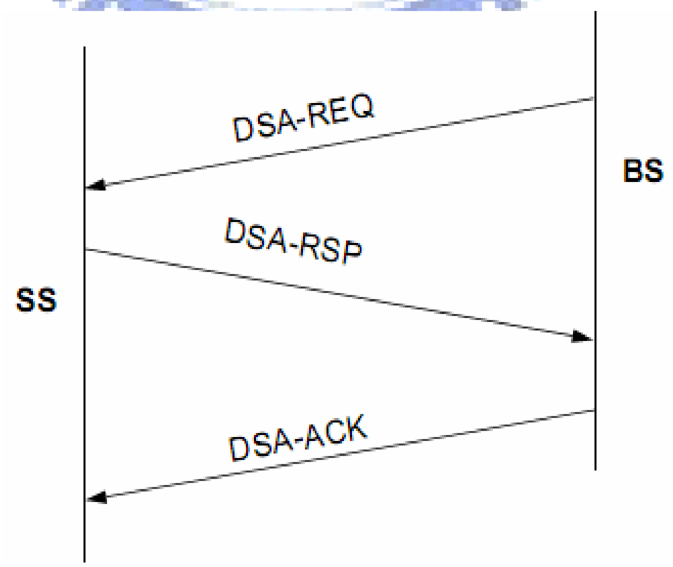

Figure 2.4 DSA message flow—BS-initiated

2.1.2 Scheduling services

The 802.16 standard MAC Layer provides QoS differentiation for the different types of applications that might operate over 802.16 networks, through defined scheduling service types. Every user has a quality of scheduling service class, also known as QoS class. According to this parameter, the BS scheduler allocates the necessary amount of bandwidth required for each application.

Scheduling services represent the data handling mechanisms supported by the MAC scheduler for data transport on a given connection. Each connection is associated with a single scheduling service. These scheduling services are determined by a set of QoS parameters about their behaviors. These parameters are managed by using the Dynamic service addition (DSA) and Dynamic service change (DSC) message dialogs. Five services are supported: Unsolicited Grant Service (UGS), Real-time Polling Service (rtPS), Non-real-time Polling Service (nrtPS), Best Effort (BE), and extended real-time Polling service (ertPS).

Uplink request / grant scheduling is performed by the BS with the intent of providing each SS with bandwidth for uplink transmissions or opportunities to request bandwidth. By specifying a scheduling service and its associated QoS parameters, the BS scheduler can anticipate the throughput and provide polls and/or grants at the appropriate times. Table 2.1 summarizes the scheduling services and the poll / grant options. The following sections define flow scheduling services for uplink operations.

Unsolicited Grant Service (UGS)

The UGS scheduling service type is designed to support real-time data streams consisting of fixed-size data packets issued at periodic intervals. This would be the case, for example, for TI/EI classical PCM (Pulse Coded Modulation) phone signal transmission and Voice over IP without silence suppression.

In a UGS service, the BS provides fixed-size data grants at periodic intervals. This eliminates the overhead and latency of SS requests. The BS provides Data Grant Burst IEs to the SS at periodic intervals based upon the maximum sustained traffic rate of the service flow. The size of these grants is sufficient to hold the fixed-length data associated with the service flow. Figure 2.5 illustrates the UGS mechanism.

Scheduling type	**PiggyBack request**	**Bandwidth stealing**	**Polling**
UGS	Not allowed	Not allowed	PM bit is used to request a unicast poll for bandwidth needs of non-UGS connections.
rtPS	Allowed	Allowed	Scheduling only allows unicast polling.
nrtPS	Allowed	Allowed	Scheduling may restrict a service flow to unicast polling via the transmission / request policy; otherwise all forms of polling are allowed.
BE	Allowed	Allowed	All forms of polling allowed.
ertPS	Allowed	Allowed	All forms of polling allowed.

Table 2.1 Scheduling services and usage rules [4]

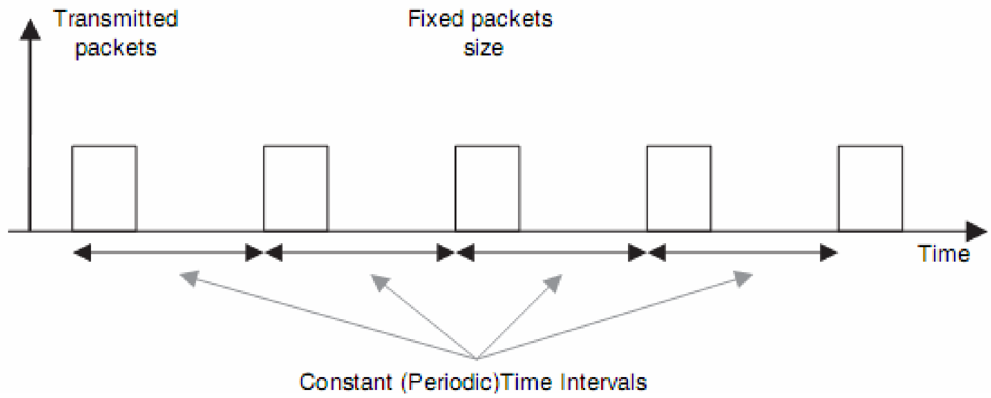

Figure 2.5 UGS service uplink grants allocation mechanism [3]

Real-time Polling Service (rtPS)

The rtPS scheduling service type is designed to support real-time data streams consisting of variable-sized data packets that are issued at periodic intervals. This would be the case, for example, for MPEG (Moving Pictures Experts Group) video transmission.

This service provides real-time, periodic, and unicast request opportunities which conform to the flow's real-time requirement and allow the SS to specify the size of the desired grant. In this service, the BS provides periodic unicast (uplink) request opportunities, which meet the flow's real-time needs and allow the SS to specify the size of the desired grant. This service requires more request overheads than UGS, but supports variable grant sizes for optimum real-time data transport efficiency. Figure 2.6 shows the rtPS mechanism.

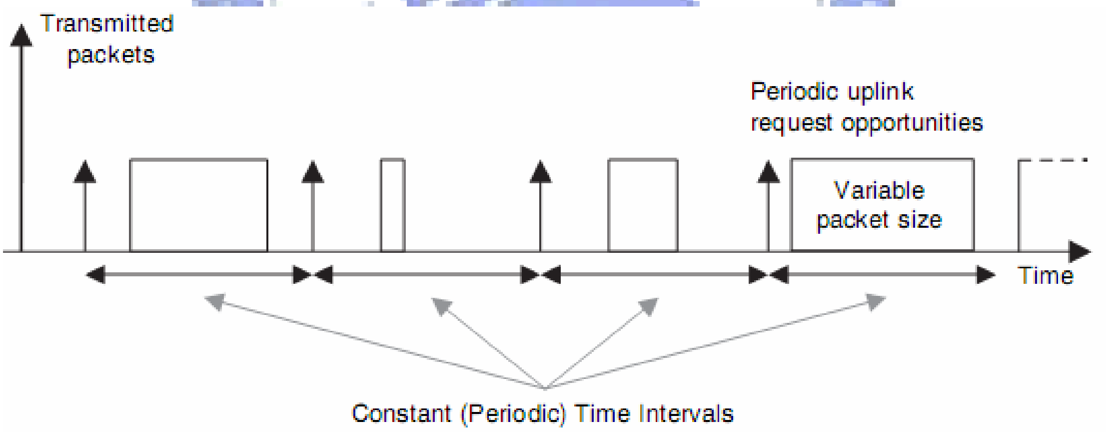

Figure 2.6 rtPS service uplink grants allocation and request mechanism [3]

Non-real-time Polling Service (nrtPS)

The nrtPS is designed to support delay-tolerant data streams consisting of variable-sized data packets for which a minimum data rate is required, for example, FTP may use nrtPS.

In the nrtPS scheduling service, the BS provides unicast uplink request polls on a ‘regular’ basis, which guarantees that the service flow receives request opportunities even during network congestion. The standard states that the BS typically polls nrtPS CIDs on an interval on the order of one second or less. Furthermore, the SS is allowed to use contention request opportunities. Figure 2.7 shows the nrtPS mechanism.

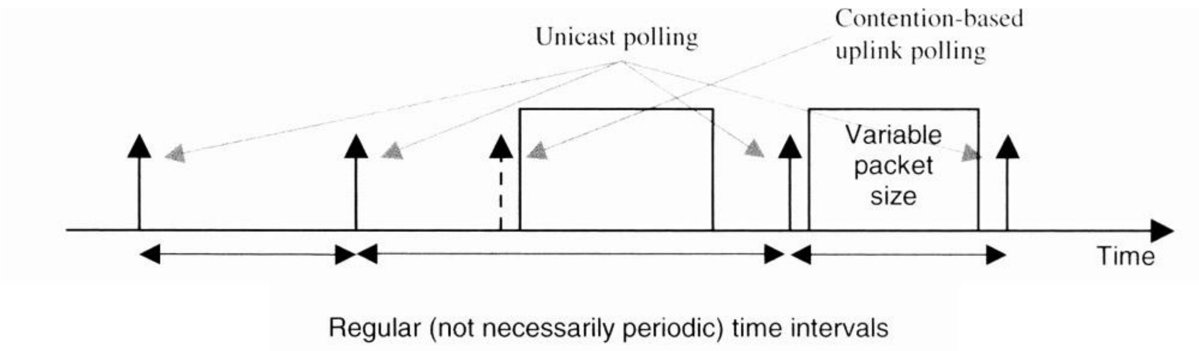

Figure 2.7 nrtPS service uplink grants allocation and request mechanism [3]

Best Effort Service (BE)

The BE service is designed to support data streams for which no minimum service guarantees are required and therefore may be handled on a best available basis. This would be the case, for example, for web browsing or E-mail. The SS may use contention request opportunities as well as unicast request opportunities when the BS sends any. The BS does not have any unicast uplink request polling obligation for BE SSs, so a long period can run without transmitting any BE packets when the network is in the congestion state. Figure 2.8 shows the BE mechanism.

Extended real-time Polling service (ertPS)

The ertPS is suitable for variable rate real-time applications that have data rate and delay requirements. An example is Voice over IP without silence suppression.

The ertPS is a scheduling mechanism that builds on the efficiency of both UGS and rtPS. The BS shall provide unicast grants in an unsolicited manner like in UGS, thus saving the latency of a bandwidth request. However, UGS allocations are fixed in size, but ertPS allocations are dynamic. The BS may provide periodic UL allocations that may be used for requesting the bandwidth as well as for data transfer. By default, size of allocations corresponds to the current value of Maximum Sustained Traffic Rate of the connection.

The SS may request changing the size of the UL allocation by either using an extended piggyback request or using a Bandwidth Request (BR). The BS shall not change the size of UL allocations until receiving another bandwidth change request from the SS. When the bandwidth request size is set to zero, the BS may only provide allocations for bandwidth request header or no allocations at all.

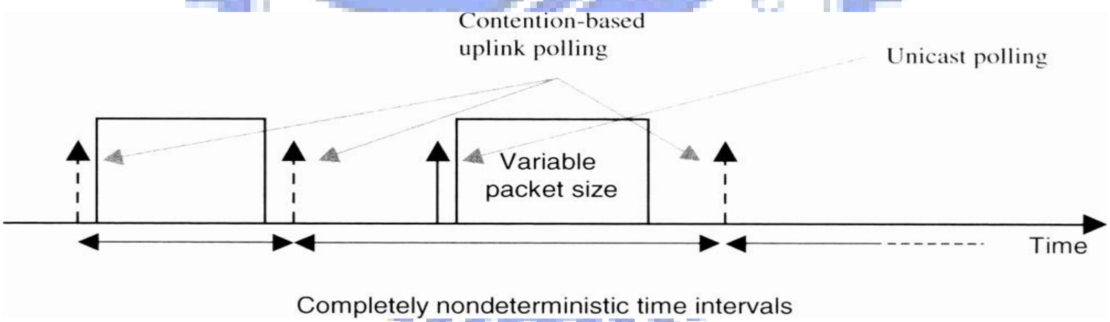

Figure 2.8 BE service uplink grants allocation and request mechanism [3]

2.1.3 Bandwidth allocation and request mechanisms

Downlink and uplink bandwidth allocations are completely different. The 802.16 standard has a scheduling centralized architecture where the BS scheduler controls all the system parameters, including the radio interface. The downlink allocation of bandwidth is a process accomplished by the BS according to different parameters that are determinant in the bandwidth allocation. Taking into consideration the QoS class for the connection and the quantity of traffic required, the BS scheduler supervises the link and determines which SS will have downlink burst(s) and the appropriate burst profile. In the uplink of WiMAX, the SS sends a message to the BS containing the immediate requirements when it needs to ask for bandwidth on a connection with its scheduling service. QoS for the connection was established at connection establishment and is looked up by the BS.

Requests

Requests refer to the mechanism used by SS(s) to inform to the BS regarding their need of uplink bandwidth allocation. A Request may come as a stand-alone bandwidth request header or it may come as a PiggyBack Request.

Because the uplink burst profile can change dynamically, all requests for bandwidth shall be made in terms of the number of bytes needed to carry the MAC header and payload, but not the PHY overhead. The Bandwidth Request message may be transmitted during any uplink allocation, except during any initial ranging interval.

Bandwidth Requests may be incremental or aggregate. When the BS receives an incremental Bandwidth Request, it shall add the quantity of bandwidth requested to its current perception of the bandwidth needs of the connection. When the BS receives an aggregate Bandwidth Request, it shall replace its perception of the bandwidth needs of the connection with the quantity of bandwidth requested. The Piggybacked Bandwidth Requests should be always incremental.

Polling

Polling is the process by which the BS allocates to the SS(s) bandwidth specifically for the purpose of making bandwidth requests. These allocations may be to individual SS or to groups of SS(s). This allocation technique is used when bandwidth resource demand is not relevant enough to have unsolicited bandwidth grants for all users; the BS can then directly assign the request amount to the SS(s) as needed. The 802.16 standard defines two main grant-request methods, unicast polling, and contention-based polling.

- **Unicast polling**

When an SS is polled individually, it is so-called unicast polling. In the case of unicast polling, there is no explicit message that can be sent to poll SS. Rather, the BS actively allocate SS slots in UL-MAP (introduced in section 2.1.4) for sending bandwidth request. For any individual uplink allocation, the SS may optionally decide to use the allocation for data, requests or requests piggybacked in data transmission. Unicast polling could only be used if the bandwidth is sufficient for polling whole individual SSs. Figure 2.9 represents the unicast polling mechanism.

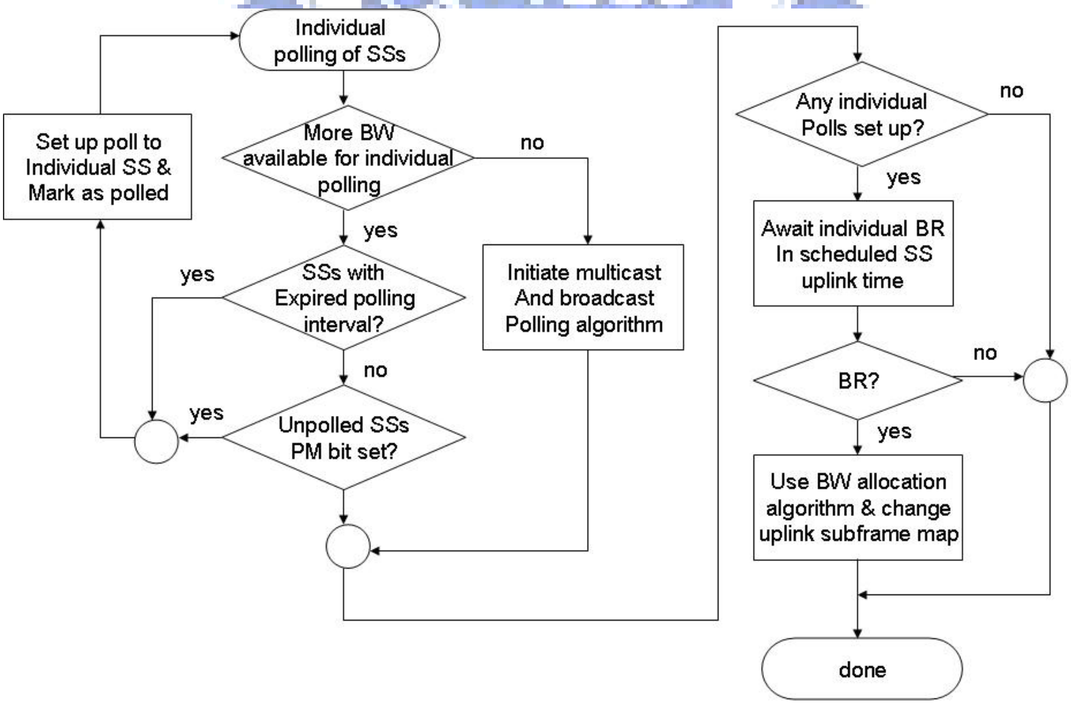

Figure 2.9 Unicast Polling mechanism.

First, we have to check whether more BW is available for individual polling or not in the WiMAX frame duration. If it is enough, we keep on making some necessary checking procedures, such as expiration for polling interval, PM bit, and so on. However, if there is not enough available bandwidth for polling, the multicast and broadcast polling algorithm should be initiated. Finally, after the necessary checking procedure, the SS would be polled.

- **Contention-based Group (Multicast or Broadcast) Polling**

The available bandwidth may not be sufficient to individually poll all inactive SS(s). Contention-based grant-request mechanisms are allocated a small part of each uplink frame (in the FDD mode) or sub-frame (in the TDD mode), known as the bandwidth requests contention slot. With this contention slot, an SS can access the network by asking the BS for an uplink slot.

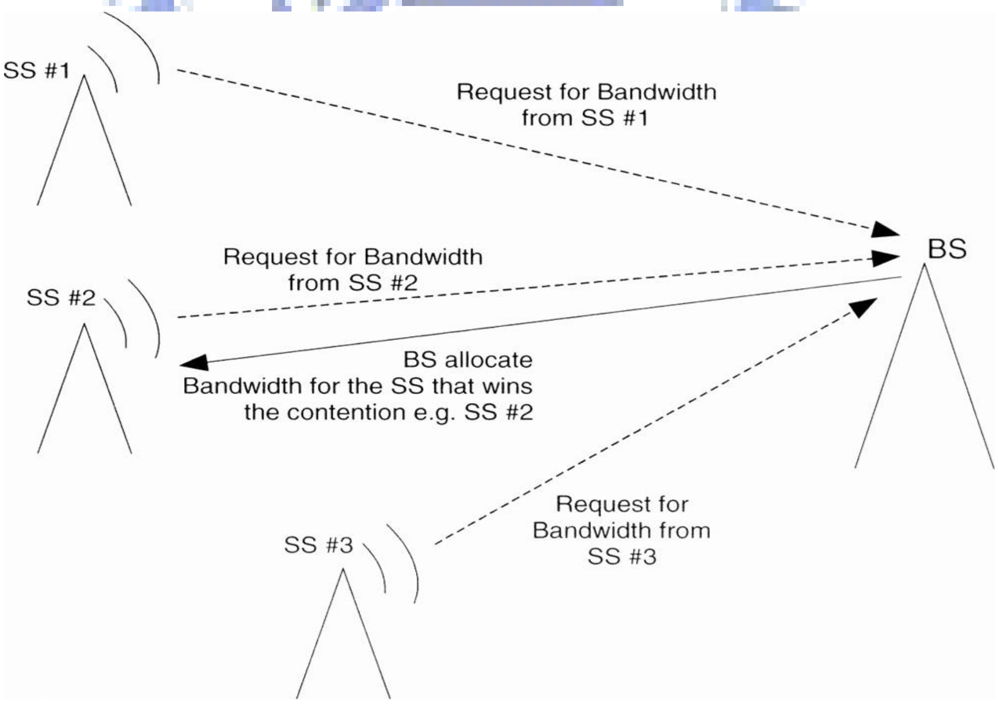

Figure 2.10 Illustration of contention-based group polling. [3]

Group (multicast or broadcast) polling works as follows. When the poll is directed at a multicast CID or the broadcast CID, an SS belonging to the polled group may request a bandwidth during any request interval allocated to that CID in a UL-MAP. In order to reduce the probability of collision with multicast and broadcast polling, only SS(s) needing a bandwidth reply. Figure 2.10 represents an illustration of the contention-based group polling mechanism.

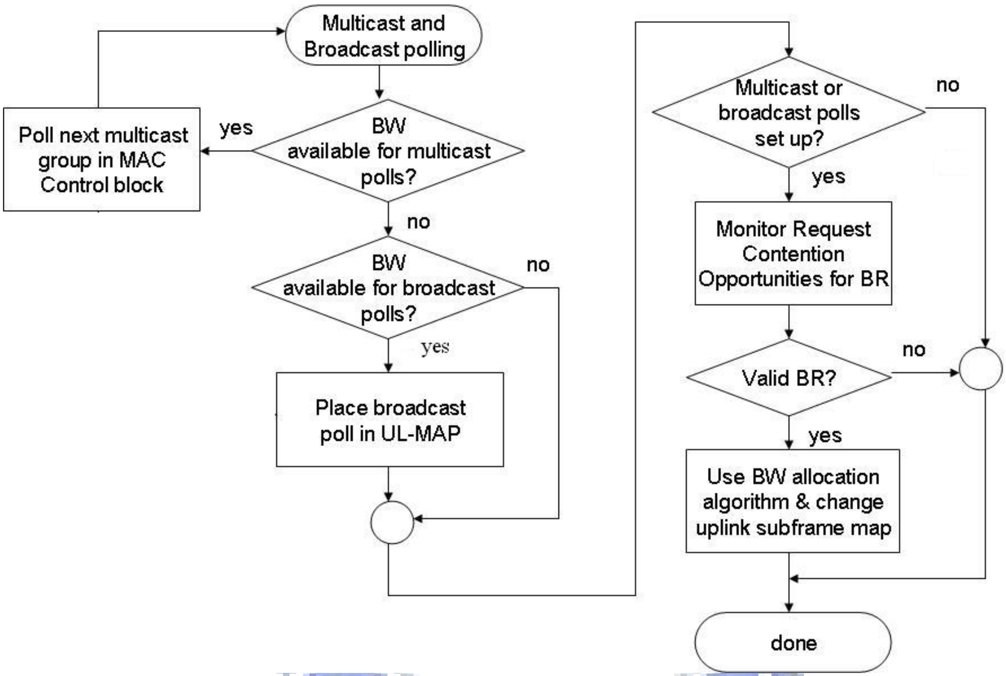

Figure 2.11 Information exchange procedure for contention-based polling

The information exchange procedure for multicast and broadcast polling is shown in Figure 2.11. In order to avoid collision, the procedure use contention resolution algorithm to decrease the collision probability. If no grant has been received during the specified timeout duration, we can assume that the transmission was failed. As illustrated in Figure 2.11, we have to check whether more BW is available for multicast and broadcast polling or not. If it is enough, we will poll the next multicast group. However, if we have no sufficient available bandwidth, we keep on making some necessary checking procedures, such as set up for contention-based polls, validity for bandwidth request (BR), and so on.

2.2 MAC support of PHY

After we introduce some necessary background for the MAC layer of WiMAX, we introduce duplexing techniques and frame structure in this section. The 802.16 standard includes the two main duplexing techniques: Time Division Duplexing (TDD) and Frequency Division Duplexing (FDD). The choice duplexing technique may affect certain PHY parameters as well as impact on the features that can be supported.

2.2.1 Time Division Duplex (TDD)

In the case of TDD, the uplink and downlink transmissions share the same frequency but they take place at different times. A TDD frame (see Figure 2.12) has a fixed duration and contains one downlink and one uplink sub-frame. The frame is separated into an integer number of Physical Slots (PSs), which help to partition the bandwidth easily. The TDD framing is adaptive in that the bandwidth allocated to the downlink versus the uplink can change. The split between the uplink and downlink is a system parameter and the 802.16 standard illustrates that it is controlled at higher layers within the system.

2.2.2 Frequency Division Duplex (FDD)

In an FDD system, the uplink and downlink channels are located on separate frequencies. For this reason, we can transmit and receive simultaneously in full-duplex mode. A fixed duration frame is used for both uplink and downlink transmissions. It also allows simultaneous use of both full-duplex SSs which can transmit and receive simultaneously and optionally half-duplex SSs which cannot. A full-duplex SS is capable of continuously listening to the downlink channel, while a half-duplex SS can listen to the downlink channel only when it is not transmitting on the uplink channel. Figure 2.13 illustrates different cases of the FDD mode of operation.

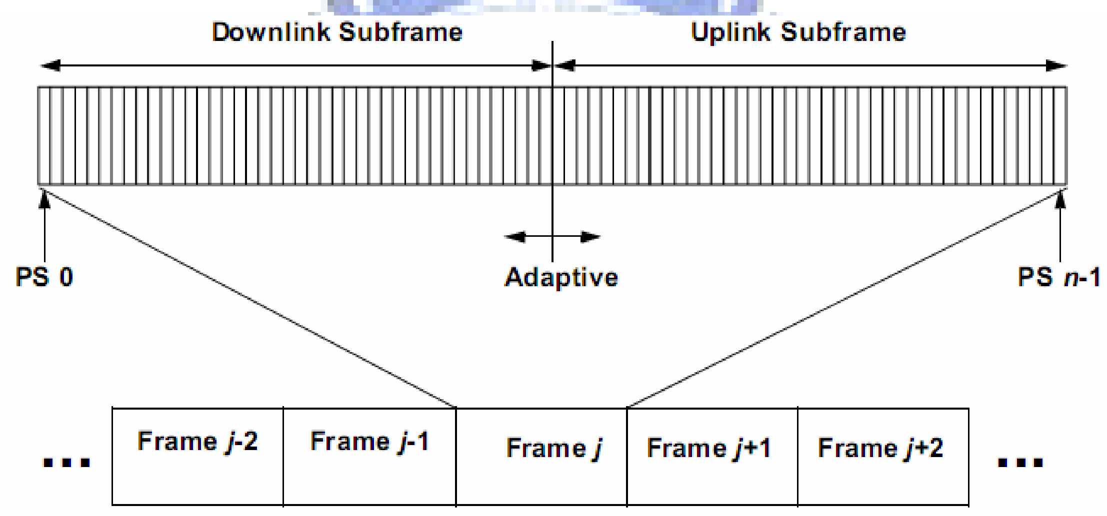

Figure 2.12 TDD frame structure

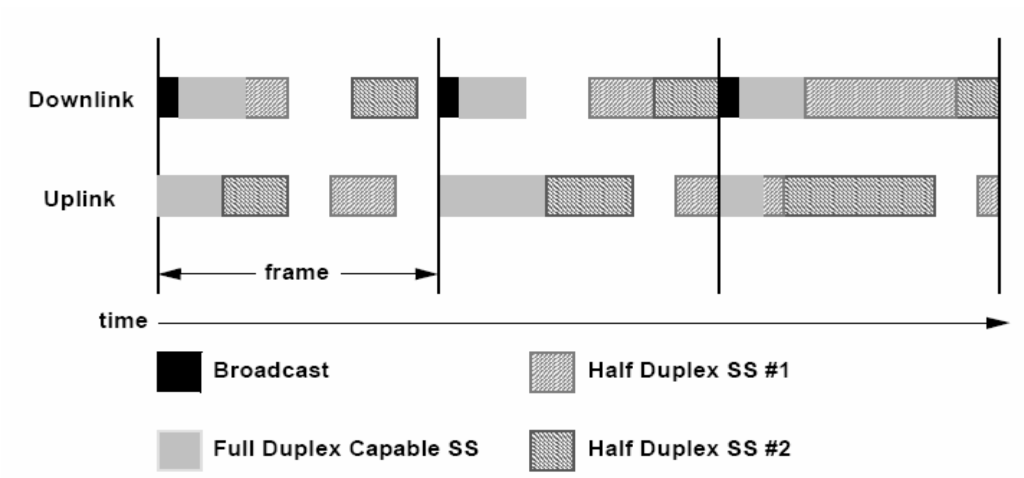

Figure 2.13 Illustration of different FDD mode operations

Comparing the two duplexing modes, a fixed duration frame is used for both uplink and downlink transmissions in FDD while the TDD distribution is adaptive. Therefore TDD duplexing is more suitable when data rates are asymmetrical between the uplink and downlink. TDD is the preferred duplexing mode for wireless network. We also choose the TDD duplexing in proposed scheduling solution.

Next, we introduce the OFDM PHY frame structure in TDD mode (see Figure 2.14). The OFDM PHY frame is separated into downlink sub-frame (DL subframe) and uplink sub-frame (UL sub-frame). A downlink sub-frame has only one downlink PHY PDU. An uplink sub-frame contains contention intervals for initial ranging and bandwidth request and one or multiple uplink PHY PDUs, each transmitted from a different SS. Some filed of OFDM PHY frame is illustrated as follows.

- Preamble: It is used for PHY synchronization.
- Frame Control Head (FCH): The FCH burst is one OFDM symbol long. It contains DL_Frame_Prefix to specify burst profile and length of one or several downlink bursts immediately following the FCH.
- DL-MAP and UL-MAP: They are two MAC management messages that defines burst start time and profiles. Design for DL-MAP and UL-MAP is very important with regard to scheduling algorithm.
- Contention slots allowing initial ranging: Via the Initial Ranging IE, the BS specifies an interval in which new stations may join the network. Ranging period is an interval, equivalent to the maximum round trip propagation delay plus the transmission time of the RNG-REQ message, shall be provided in some UL-MAPs to allow new stations to perform initial ranging.
- Contention slots allowing bandwidth requests: Via the Request IE, the BS

specifies an uplink interval in which requests may be made for a bandwidth for uplink data transmission.

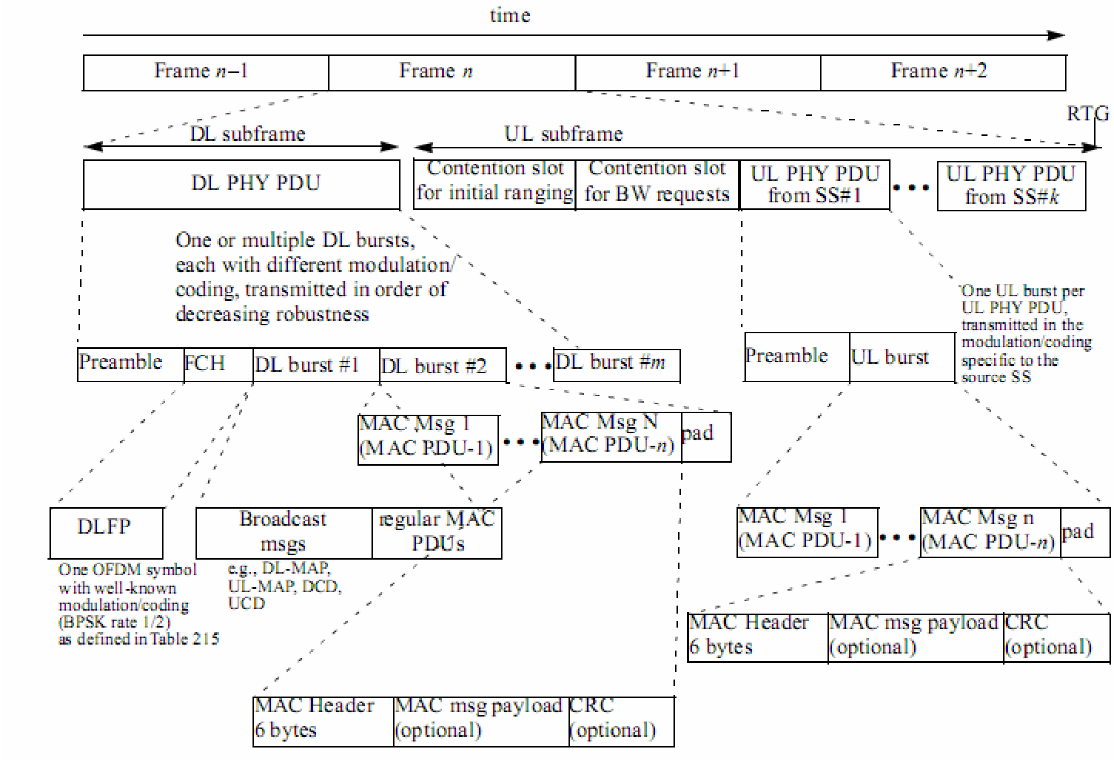

Figure 2.14 OFDM frame structure with TDD

2.2.3 DL_MAP and UL_MAP

The primary part of design for WiMAX scheduling is DL/UL MAP. The DL_MAP is a MAC management message that defines burst start times on the downlink. Equivalently, the UL_MAP is a set of information that defines the uplink access for all SSs during a scheduling interval. Then DL_MAP and UL_MAP are directories, and broadcasted by the BS.

- DL_MAP message

The DL-MAP is a MAC management message that defines burst start time and profiles on the downlink. Each burst start time is indicated by a DL-MAP_IE (DL-MAP Information Elements). The DL-AMP_IE format is PHY layer-dependent. The BSs generate OFDM PHY DL-MAP messages in the format shown in Figure 2.15, including all of the following parameters:

- MAC management message type: 2 for DL_MAP
- DCD count: The value of the Configuration Change Count (CCC field) of the DCD, which describes the downlink burst profiles concerned by this map.
- Base Station ID: The Base Station ID is a 48-bit long field identifying the BS.

Management message type of DL-MAP (= 2); (8 bits)	DCD count (8 bits)	Base Station ID (48 bits)	CID= i_1 DIUC= j_1 Start Time= t_1	CID= i_2 DIUC= j_2 Start Time= t_2	CID= i_3 DIUC= j_3 Start Time= t_3	…
			DL-MAP IE_1	DL-MAP IE_2	DL-MAP IE_3	

Figure 2.15 DL_MAP message general form

The remaining part of a DL-MAP is the encoding of the DL-MAP IEs that are PHY-specification dependent. The DL-MAP IE of the OFDM PHY Layer has the format shown in Figure 2.16 and includes all of the following parameters:

- Connection Identifier (CID): Represents the assignment of the IE to a broadcast, multicast or unicast address.
- DIUC: The 4-bit DIUC defines the burst type associated with that burst time interval. It also defines the modulation and coding scheme.
- Preamble present: If set, the indicated burst will start with the short

preamble.

- Start Time: This indicates the start time, in units of OFDM symbol duration, relative to the start of the first symbol of the PHY PDU where the DL-MAP message is transmitted.

CID	16 bits
DIUC	4 bits
Preamble present	1 bit (0=not present, 1=present)
Start Time	11 bits

Figure 2.16 DL_MAP IE fields

- UL_MAP message

The UL-MAP message allocates access to the uplink channel. The general form of the UL-MAP message is almost similar to DL-MAP and is shown in Figure 2.17. The differentiation between DL_MAP and UL_MAP is the Allocation Start Time field. The Allocation Start Time is the start time of the uplink allocation. The unit of the Allocation Start Time is the Physical Slot (PS) starting from the beginning of the downlink frame in which the UL-MAP message is placed.

Management message type of UL-MAP (= 3); (8 bits)	UCD count (8 bits)	Base Station ID (48 bits)	Allocation Start time (32 bits)	CID=i_1 UIUC=j_1 Start Time=t_1 Duration= D_1	CID=i_2 UIUC=j_2 Start Time=t_2 Duration= D_2	…
				UL-MAP IE_1	UL-MAP IE_2	

Figure 2.17 DL_MAP message general form

Chapter 3 Proposed Approaches

In this chapter, we introduce proposed WiMAX scheduling solution in detail. The WiMax system provides high data rate, wide bandwidth, and long transition range for wireless network. However, there are too many interference in the air interface. So, the link quality is not stable owing to air interference and long transmission. When link quality change, mobile node can revise the modulation and coding scheme dynamically. The principle is rather simple: when the radio link is good, use a high-level modulation; when the radio link is bad, use a low-level, but also robust, modulation. Therefore, the WiMAX scheduling solution is required to be modulation-aware.

Based on the above statement, we propose the modulation-aware scheduling solution for WiMAX Base station. It contains three primary schemes: admission control policy, scheduling algorithm, and bandwidth allocation scheme. In addition to modulation and coding scheme, we also take latency for real-time service into account. Several research works [5, 6, 7, 8] propose using complex scheduling algorithm, such as weighted fair queueing, and worst-case weighted fair queueing, and even a hierarchy schedulers. All of them are not so simple to implement. It is a challenging to use a hierarchy of scheduler because the frame duration is very brief in time. This is precisely reason why we propose one level scheduling solution. Some properties of proposed scheduling solution are as follows: modulation and latency aware, easily implemented in BS, guarantee minimum QoS for some specific services (UGS, rtPS, nrtPS), and allocate bandwidth efficiently.

3.1 QoS architecture of WiMAX

Figure 3.1 presents a basic QoS architecture of WiMAX in proposed scheduling solution. Unlike in [9], we do not separate the BS scheduler into downlink scheduler and uplink scheduler. There are two kinds of directional queue in BS, i.e., downlink queue which deal with incoming packets from each downlink connection, and uplink virtual queue which deal with bandwidth requests correspond to the uplink resource demands of SS. Having the information on bandwidth demands and the QoS requirements, BS can make scheduling decisions to allocate available slots for downlink and uplink connections.

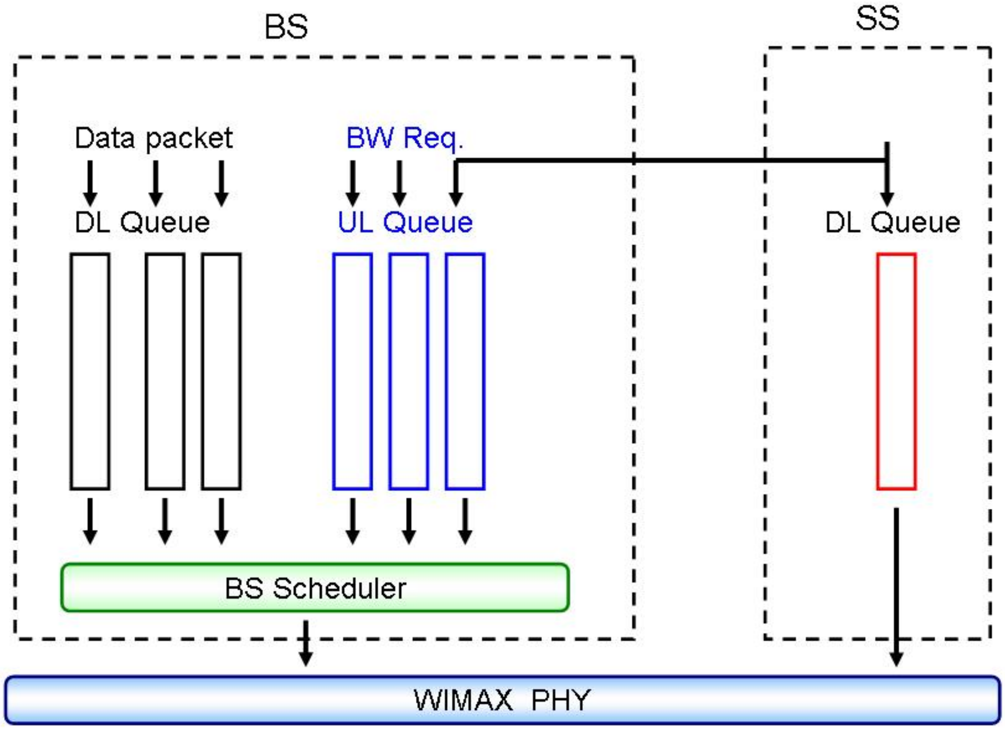

Figure 3.1 QoS architecture of WiMAX

Next, we will introduce the proposed latency and modulation aware WiMAX scheduling solution in detail. It includes admission control, scheduling algorithm, and bandwidth allocation. The scheduling solution we proposed concerns only the scheduler at the BS that use TDD duplexing in PMP mode. We expect that the BS will be connected to the wired medium with a link whose bandwidth is larger than the maximum bandwidth of the WiMAX network.

3.2 Admission control scheme

The first step to design scheduling solution is a precise admission control scheme. A precise admission control is very important to the scheduling algorithm, although it is not the primary part of design for scheduling solution. Because WiMAX system may adjusts the modulation and coding scheme dynamically when the link quality changes, so the admission control scheme is required to be modulation-aware.

In [5], it assumes that all of WiMAX mobile nodes use the same modulation and coding schemes. Therefore, it assumes that the WiMAX frame is a fixed-size data frame. However, the definition for WiMAX frame in standard [1] is that the TDD frame duration is a constant interval in time. Obviously there are some problems in the situation which some mobile nodes use different kinds of modulation and coding schemes. In [5], if it assumes that all mobile node use high-level modulation, the BS may allow too many SSs to access WiMAX network. On the other hand, the BS may allow too less SSs to access WiMAX network if it assumes that all mobile node use low-level modulation. Based on the above statement, we proposed a modulation aware admission control scheme.

The admit condition of the modulation aware admission control scheme is calculated base on the minimum reserved rate for each admitted data connection. So, we can guarantee the minimum QoS requirement for different kinds of service connections. In the TDD node, the transmission unit in WiMAX system is symbol whose capacity is depends on the modulation and coding scheme. Therefore, when a new frame starts, the required data size in downlink or the bandwidth request size in uplink must be firstly translated to some amount of symbol as expression (1).

$$number\_of\_symbols = \frac{data\_size}{bits\_per\_symbol} \qquad (1)$$

Since an OFDM symbol contains 192 data sub-carriers and 8 pilot sub-carriers in WiMAX PHY, and the modulation and coding scheme decide the number of bits which carried in a sub-carrier. We calculate the number of bits per symbol as expression (2).

$$bits\_per\_symbol = 192 * efficiency * coding\_rate - 8 \qquad (2)$$

The efficiency is illustrated in Table 3.1. Thus we can calculate the bits_per_symbol for different kinds of modulation and coding schemes and represent it in Table 3.2.

Modulation and coding scheme	Efficiency
OFDM_BPSK_1_2	1 bps/Hz
OFDM_QPSK_1_2	2 bps/Hz
OFDM_QPSK_3_4	2 bps/Hz
OFDM_16QAM_1_2	4 bps/Hz
OFDM_16QAM_3_4	4 bps/Hz
OFDM_64QAM_2_3	6 bps/Hz
OFDM_64QAM_3_4	6 bps/Hz

Table 3.1 The efficiency for different modulation and coding.

Modulation and coding scheme	Bits_per_symbol
OFDM_BPSK_1_2	88
OFDM_QPSK_1_2	184
OFDM_QPSK_3_4	280
OFDM_16QAM_1_2	376
OFDM_16QAM_3_4	578
OFDM_64QAM_2_3	760
OFDM_64QAM_3_4	856

Table 3.2 Bits per symbol for different modulation and coding.

Since the TDD frame duration is a constant interval in time, we define the FPS stands for the frames per second in expression (3). Suppose $BW\_\min(i)$ is the minimum reserved rate of the ith connection. Also suppose the $bits\_per\_symbol(i)$ stands for the bits per symbol in the ith connection. Based on the parameters introduced above, we can calculate the average required symbols per frame for the ith connection $T(i)$ based on minimum reserved rate and modulation and coding scheme by using the expression (4). We allow a connection to access the WiMAX network when the sum of $T(i)$ is smaller than the available symbols of a WiMAX frame. The admit condition of proposed admission control scheme is presented in expression (5).

$$FPS = \frac{1}{frame\_duration} \tag{3}$$

$$T(i) = \frac{BW\_\min(i) / FPS}{bits\_per\_symbol(i)} \tag{4}$$

$$if \sum_{n=1}^{i} T(n) < avalible\ symbol\ in\ a\ \ frame \tag{5}$$

The modulation-aware admission control scheme we proposed can provide a precise scheme to calculate the number of SS which can be admitted. So, we can allocate bandwidth in scheduling more efficiently.

3.3 Scheduling algorithm

The bandwidths are allocated in the form of data bursts in the WiMAX system. Each burst is composed of an integer number of symbols. All symbols in one burst use the same profile and modulation and coding scheme. Therefore, a symbol is the basic fixed allocation unit in WiMAX scheduling. The job of the BS scheduler is to ensure the connection requirements by allocating a required amount of symbols based on the resource demands and the QoS parameters. Suppose $BW\_\max$ stands for the maximum sustained traffic rate, $BW\_\min$ stands for minimum reserved traffic rate, and MAX_d stands for the maximum latency. The QoS parameters we considered for different kinds of service classes in the proposed scheduling algorithm are presented in Table 3.3.

Scheduling service	$BW\_\max$	$BW\_\min$	MAX_d
UGS	✓	✓	
rtPS	✓	✓	✓
nrtPS	✓	✓	
BE	✓		

Table 3.3 QoS parameters for different kinds of service classes

The characteristics of the proposed scheduling solution are as follows: it is latency and modulation aware, one-level scheduling, easily implemented, allocation for bandwidth efficiently, and comply with the IEEE 802.16 standard. According to the WiMAX standard, the BS scheduler should provide grants (a mount of symbols) at periodic interval to the UGS service connections to access data because UGS service do not need bandwidth requests. For nrtPS and rtPS, we should also provide periodic polls to request bandwidth in the uplink. Besides, the BS scheduler should guarantee the maximum latency and bandwidth requirements for different kinds of service classes. The BS generates DL_MAP and UL_MAP during each frame, and it broadcast the scheduling decision in the DL/UL_MAP in the downlink sub-frame.

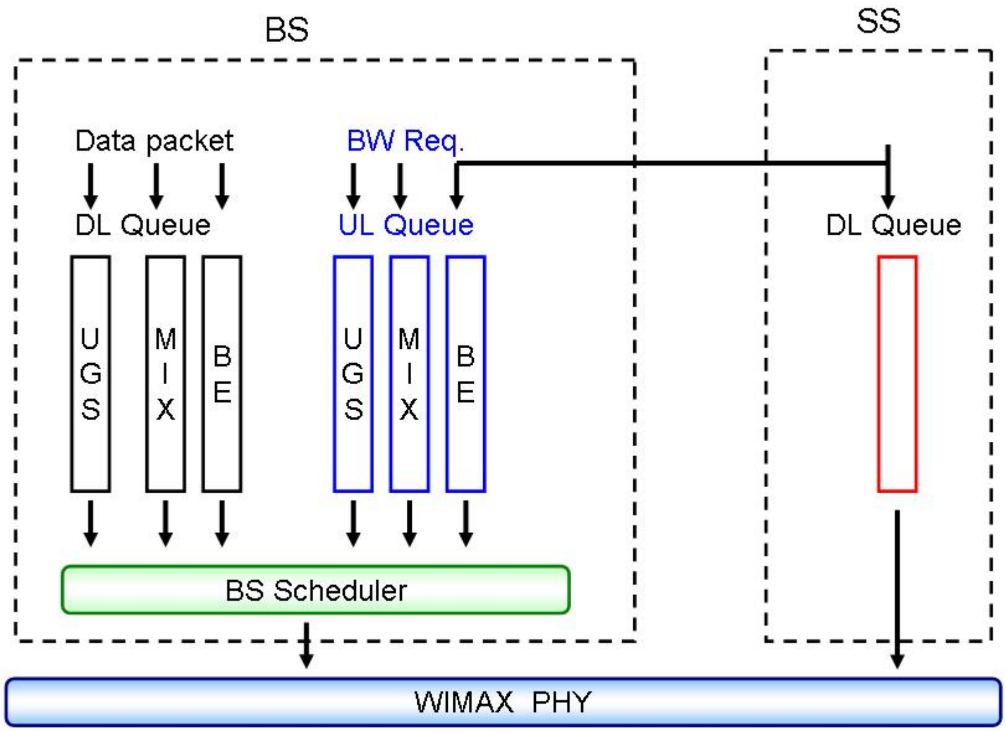

Figure 3.2 Proposed scheduling architecture.

Figure 3.2 presents the proposed scheduling architecture in WiMAX system. The proposed scheduler use two directional queues, i.e., downlink queue which deal with incoming data packets from each downlink connection, and uplink virtual queue which deal with bandwidth requests correspond to the uplink resource demands of SS. We define three kinds of queue in the two directions, referred as UGS queue, MIX queue, and BE queue. The UGS queue serves the bandwidth requests and data packets for UGS connections. The MIX queue holds the request and data packets sent by rtPS and nrtPS connections. The BE queue stores the requests and packets of the BE connections. In the uplink, we also need to stores periodic grants and unicast request opportunities for rtPS and nrtPS in the UGS queue.

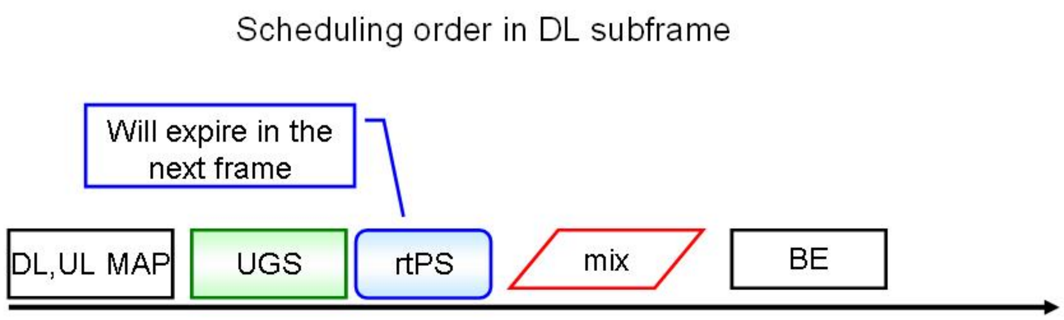

Figure 3.3 Scheduling order in DL sub-frame.

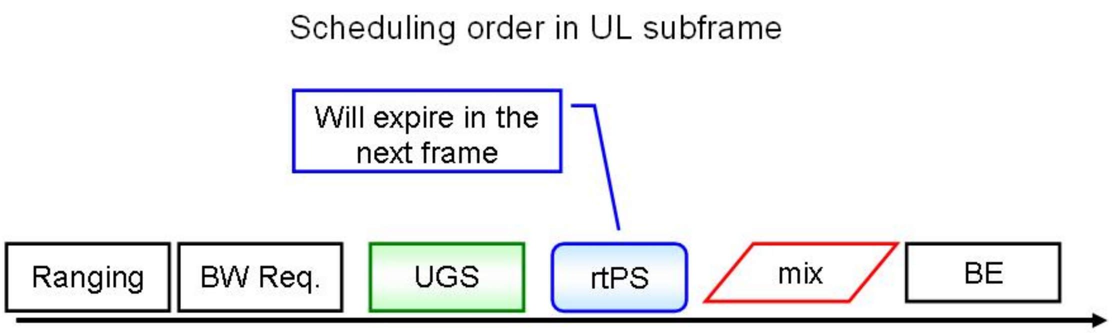

Figure 3.4 Scheduling order in UL sub-frame.

Figure 3.3 and Figure 3.4 represent the DL/UL scheduling order of the proposed scheduling solution. In our scheduling solution, we suppose that the priority of downlink queue is higher than the uplink queue. Therefore, the scheduling order for all kinds of queue is as follows: first, UGS DL, UGS UL, DL packets which will expire in the next frame for rtPS, UL bandwidth requests which will expire in the next frame for rtPS, mix DL, mix UL, BE DL, and BE UL. The proposed scheduling order can help us to adjust the DL/UL sub-frame ratio dynamically.

In the first, we should guarantee bandwidth for UGS service connections. Thus, the periodic data grants will be inserted into the UGS queue. The periodic data grant interval is determined by the BS at the establishment time. The polling intervals to rtPS and nrtPS are also defined by the BS during the network entry procedure.

There is one special QoS requirement to rtPS service namely the maximum latency. To suffice this requirement, the BS should assign deadline to the rtPS data packets (downlink) or the rtPS bandwidth request (uplink) in the MIX queue. As often as the BS scheduler works, the data packets/requests which will be expired in the next frame should be scheduled into the current frame. We have to know the arriving time of the packets/requests at the SS/BS queue if we want to calculate the deadline to each packet/request. In the uplink, the BS can not get this information, the worst case is considered, which corresponds to the arrival at the queue immediately after the connection sent the last bandwidth request [10]. Thus, in the uplink case, the deadline of an rtPS request is equal to the sum of the arriving time of the last request sent by the connection and its maximum latency. In the downlink case, the deadline of a data packet is equal to the sum of the arrival time of the data packet and the maximum latency for the connection.

We should also guarantee the minimum bandwidth requirement for rtPS nad nrtPS service connections in the MIX queue. As often as the BS scheduler is executed, it sorts the packets/requests in the MIX queue based on the granted bandwidth to each rtPS/nrtPS connection. If the connection gets more bandwidth than others, the priority of the connection is lower than others in the MIX queue. Finally, if there are some available symbols in the current frame duration, these symbols are divided equally among the BE connections.

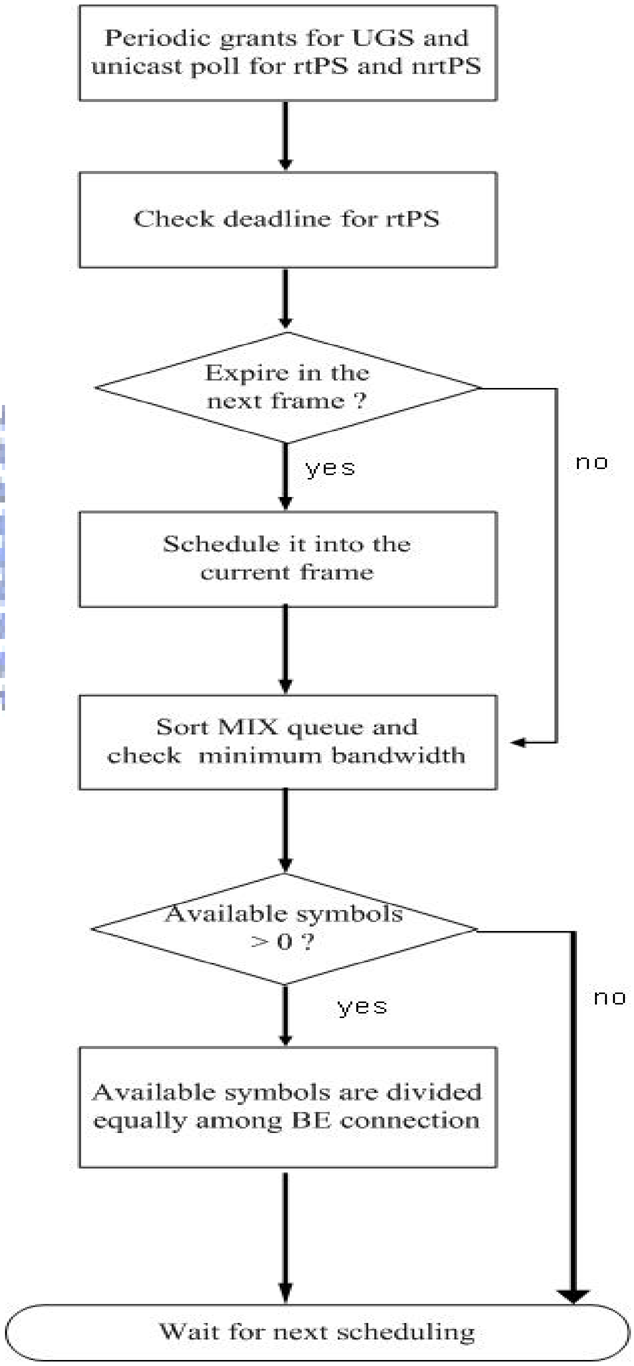

Figure 3.5 Scheduling algorithm

Check_deadline procedure:

For each packet/request i to rtPS connection in the MIX queue

$$check\_deadline[i] = \frac{deadline[i] - now}{frame\_duration}$$

if (1≤ check_deadlin[i] < 2) // will expire in the next frame

if(available_symbols ≥ (BW_req[i] in UL or pkt_size[i] in DL))

insert the packet/request into the current frame

granted_bit[CID] = granted_bit[CID] – (BW_req[i] in UL or pkt_size[i] in DL)

available_symbols = aviliabe_symbols – (BW_req[i] in UL or pkt_size[i] in DL)/bit_per_symbol(i)

else

allocate available_symbols to the connection CID

Check_minimum_bandwidth procedure:

For each connection CID and each packet/request i in the MIX queue

if granted_bit[CID] ≥ BW_min[CID]

priority[i] = 0;

else

priority[i] = BW_min[CID] – granted_bit[CID]

if(available_symbols ≥ (BW_req[i] in UL or Q_size[CID] in DL))

granted_bit[CID] = granted_bit[CID] + (BW_req[i] in UL or Q_size[CID] in DL)

available_symbols = aviliabe_symbols – (BW_req[i] in UL or Q_size[CID] in DL)/bit_per_symbol(i)

else

allocate available_symbols to the connection CID

sort the MIX queue base on priority

Figure 3.5 represents the simple flow chart for proposed scheduling algorithm. We suppose BW_req[i] is the ith bandwidth request in the MIX queue. Also suppose that Q_size[i] stands for queue size of the connection CID, the pkt_size[i] stands for the ith packet size in the MIX queue, and granted_bit[CID] stands for the granted bits to the connection CID. The procedure ***Check_deadline*** verifies for each rtPS packet/request whether it will expire in the next frame or not. If the packet/request will expire in the next frame and the amount of packet/request is less than available symbols, we allocate it in the current frame. In this case, we should update granted_bit[CID] and availialbe_symbols variables.

The procedure ***Check_minimum_bandwidth*** has to calculate priority[i] to each packet/request in the MIX queue. Next, the BS scheduler sorts the MIX queue according to the priority values. Like the ***Check_deadline*** procedure, we also update granted_bit[CID] and available_symbols variables.

3.4 Bandwidth allocation scheme

Several research works [5, 6, 7, 10], they propose scheduling algorithm and QoS architecture over WiMAX, but they do not take the modulation and coding into account in the bandwidth allocation procedure. However, one important characteristic of WiMAX PHY is adaptive modulation and coding. Thus, we proposed a modulation-aware bandwidth allocation scheme, so we have to translate the data byte request to symbols according to the modulation and coding scheme they used. In addition to modulation, we also guarantee the minimum QoS requirements (minimum reserved rate, maximum latency) for some specific service connections (UGS, rtPS, nrtPS). We need using expression (2) (3) (4) introduce in section 3.1.

$$bits\_per\_symbol = 192 * efficiency * coding\_rate - 8 \quad (2)$$

$$FPS = \frac{1}{frame\_duration} \quad (3)$$

$$T(i) = \frac{BW\_\min(i) / FPS}{bits\_per\_symbol(i)} \quad (4)$$

Suppose the $bits\_per\_symbol$ stands for the bits per symbol reflecting the modulation and coding, FPS stands for the number of frame per second, $BW\_req(i)$ stands for the bandwidth request for the i th connection (unit: bit), $BW\_\min(i)$ stands for the minimum reserved rate of the i th connection (unit: bps), $T(i)$ stands for the average required symbols per frame for the i th connection is calculated based on the minimum reserved rate (unit: symbol).

- **UGS queue**

In a UGS service, the BS should provide fixed-size data grants at periodic intervals. The periodic grant interval is decided by the BS at the establishment time. Unlike other service classes, UGS does not need bandwidth request in the uplink. In the downlink, the arriving time of data packets may not be periodic at the BS owing to the long transmission. Thus, we proposed the bandwidth allocation to i th connection for UGS downlink as expression (6).

$$allocated\_symbol(i) = \frac{\min\{Qsize(i), BW\_\min(i) - granted\_bit(i)\}}{bits\_per\_symbol(i)} \quad (6)$$

Suppose that $Qsize(i)$ stands for the queue size in unit of bit to the i th connection, $granted\_bit(i)$ stands for the number of the allocated bits to the i th connection. In the uplink, we can make sure that the SS will transmit the fixed-size data packets at the grant intervals periodically. So, the bandwidth allocation to i th connection for UGS is illustrated in expression (7).

$$allocated\_symbol(i) = \left(\frac{grant\_interval}{frame\_duration} \right) * T(i) \tag{7}$$

- **rtPS (packets/requests which will expire in the next frame)**

Next, we introduce the bandwidth allocation scheme for packet/request which will expire in the next frame. After we process the bandwidth allocation for UGS, we will choose the rtPS packets/requests which will expire in the next frame to allocate them bandwidth. We have to guarantee the maximum latency requirement for rtPS connections. In this case, the bandwidth allocation to i th connection for downlink is illustrated in expression (8). The bandwidth allocation to i th connection for downlink is illustrated in expression (9).

if ($granted\_bit(i) < BW\_\min(i)$)

$$allocated\_symbol(i) = \frac{\min\{\sum packet\_size, BW\_\min(i) - granted\_bit(i)\}}{bits\_per\_symbol(i)} \tag{8}$$

if ($granted\_bit(i) < BW\_\min(i)$)

$$allocated\_symbol(i) = \frac{\min\{\sum BW\_req(i), BW\_\min(i) - granted\_bit(i)\}}{bits\_per\_symbol(i)} \tag{9}$$

- **MIX queue**

For MIX queue, we guarantee the minimum reserved rate requirement for rtPS and nrtPS connections. After the BS sort the MIX queue based on the priority, we have to check whether the allocated bandwidth of each connection achieve the minimum reserved rate or not. In order to guarantee the minimum reserved rate for rtPS and nrtPS connections, we proposed bandwidth allocation scheme to i th connection for downlink as expression (10) and uplink as expression (11).

if ($granted\_bit(i) < BW\_\min(i)$)

$$allocated\_symbol(i) = \frac{\min\{Qsize(i)\ , BW\_\min(i) - granted\_bit(i)\}}{bits\_per\_symbol(i)} \tag{10}$$

if ($granted\_bit(i) < BW\_\min(i)$)

$$allocated\_symbol(i) = \frac{\min\{BW\_req(i)\ , BW\_\min(i) - granted\_bit(i)\}}{bits\_per\_symbol(i)} \tag{11}$$

- **BE queue**

After we process the MIX queue, if there are some available symbols in the current frame, these available symbols are divided equally among the BE connections. The BE service class has only the maximum sustained rate parameter, so we have to make sure that the number of allocated bandwidth must be less than the maximum sustained rate. We suppose n is the number of BE connections which the allocated bandwidth is less than the maximum sustained rate. We proposed bandwidth allocation scheme to i th connection for downlink as expression (12) and uplink as expression (13).

$$allocated\_symbol(i) = \min\{\frac{available\_symbols}{n}, \frac{Qsize(i)}{bits\_per\_symbol(i)}\} \tag{12}$$

$$allocated\_symbol(i) = \min\{\frac{available\_symbols}{n}, \frac{BW\_req(i)}{bits\_per\_symbol(i)}\} \tag{13}$$

3.5 Summary

In this chapter, we introduce the latency and modulation aware WiMAX scheduling solution for BS. It includes three primary schemes: admission control scheme, scheduling algorithm, and bandwidth allocation.

Because adaptive modulation and coding is one important characteristic in WiMAX PHY, and the channel quality is unstable due to the long-distance and electromagnetic interference, so we take modulation and coding into account in the proposed scheduling solution. In our admission control scheme, it can calculate the admit condition more precisely based on the modulation and coding scheme. Thus, we can allocate bandwidth for different kinds of service classes more efficiently. In addition to modulation and coding, the proposed scheduling solution can also guarantee the maximum latency for rtPS service, minimum reserved rate for some specific services such as UGS, rtPS, and nrtPS. In the bandwidth allocation scheme, we have different kinds of allocation schemes in different directions. Some properties of proposed scheduling solution are as follows: modulation and latency aware, easily implemented in BS, guarantee minimum QoS for some specific services, allocate bandwidth efficiently, precise admission control.

Chapter 4 Simulation and Numerical Results

4.1 Simulation environment

This section presents the simulation results for the proposed WiMAX scheduling solution. For testing, we use NS-2 (version 2.31) simulation tool [12] with WiMAX NIST module [13] which is originally designed and developed by the National Institute of Standards and Technology (NIST) in U.S.A. We implement the service flow, the service flow QoS, and the proposed scheduling solution in the NS-2 simulator.

The implemented PHY is OFDM, and we use the TDD duplexing mode in our simulation. Furthermore, our implementation also supports different modulation and coding schemes. In our simulation, the SSs may use different kinds of modulation and coding schemes. We present several simulation scenarios to analyze the proposed scheduling solution.

In the first scenario, we compare the general admission control scheme suppose all SS use the same modulation and coding scheme with the proposed admission control scheme. We will show that the proposed admission control scheme can guarantee the minimum reserved rate for the admitted connections, promote the throughput of BS, and utilize the frame space more efficiently. The second scenario will present a multi-service case. We use four kinds of service classes and increase the BE connections to verify whether the scheduler can guarantee the minimum reserved rate requirement for each connections.

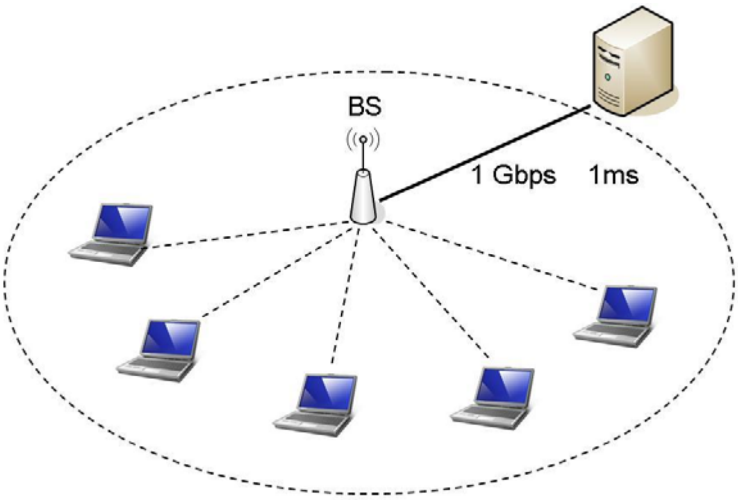

Figure 4.1 The network topology for simulation.

Each simulation scenario includes a wired node, a BS, and several SSs with different service flows. The wired node is connected to the BS with a link, whose bandwidth and delay are 1000 Mbps and 1ms, respectively. The bandwidth of this link is larger than the WiMAX network. Figure 4.1 depicts the network structure for the simulation. In each simulation scenario, the general parameters of the 802.16 network are the same, and they are presented in Table 4.1.

Parameter	**Value**
PHY	OFDM
Duplexing mode	TDD
Frame duration	5ms
Frames per second	200
Available symbols per frame	70
DL/UL ratio	dynamic
Simulation start time (second)	10
Simulation end time (second)	30

Polling interval (rtPS) (ms)	20
Polling interval (nrtPS) (ms)	100

Table 4.1 General parameter of simulation scenario

Besides, we also set the minimum reserved rate and maximum sustained rate for all kinds of service classes (see Table 4.2). We have to guarantee these QoS requirements illustrated in Table 4.2 for each connection. To simplify the simulation, the SS only has one service flow connection. The UGS connection is simulated by the constant-rate data sent over the UDP protocol. According to the G.711 audio codec [11], we set the packet size for 160 bytes and the granted interval for 20ms. The rtPS connection is simulated variable-rate data sent over the UDP protocol. We set the rtPS packet size for 1378 bytes, the maximum sustained rate for 410880 bps, the minimum reserved rate for 205440bps, and maximum latency for 50ms. The nrtPS and BE traffic are generated by FTP application. For nrtPS connections, we set the packet size 1000 bytes and the maximum sustained rate for 200000 bps, the minimum reserved rate for 150000 bps. For BE connections, we set the packet size for 200 bytes, the maximum sustained rate for 120000 bps, and the minimum reserved rate for 0 bps.

Service class	Traffic type	Packet size (byte)	Bandwidth (bps)	
			Max	Min
UGS	UDP/CBR	160	80000	64000
rtPS	UDP/VBR	1378	410880	205440
nrtPS	TCP/FTP	1000	200000	150000
BE	TCP/FTP	200	120000	0

Table 4.2 Traffic type and QoS requirements.

4.2 Simulation scenario 1

In this scenario, we analyze the proposed admission control scheme, and general admission control scheme which suppose all SS use the same modulation and coding scheme. As we described in Section 3.2, the proposed admission control scheme can support different kinds of modulation and coding scheme, and the calculation for admission condition is more precise and efficient. However, if there are some SSs use different modulation and coding schemes, the general admission control may admit too many or too less SS.

We use several rtPS service connections to test the throughput of admission control scheme. Each SS has one rtPS connections to the BS. There are 20 64QAM_3_4 SSs, 15 16QAM_1_2 SSs, and 10 BPSK_1_2 SSs to random access the WiMAX network in our simulation scenario. The General_BPSK_1_2 admission control scheme means the general admission control scheme suppose all SS use the BPSK_1_2 modulation and coding scheme. The meaning of the General_16QAM_1_2 is similar to the General_BPSK_1_2. We use the proposed scheduling algorithm. The simulation result is shown in Table 4.3.

Admission control scheme	Admitted stations	Stations		
		64QAM_3_4	16QAM_1_2	BPSK_1_2
General_BPSK_1_2	5	3	1	1
Proposed	19	10	6	3
General_16QAM_1_2	25	12	8	5

Table 4.3 Simulation result of admission control

In Table 4.3, we can know the number of admitted stations for three kinds of admission control schemes. Next, we analyze the total throughput for the above-mentioned admission control schemes in Figure 4.2.

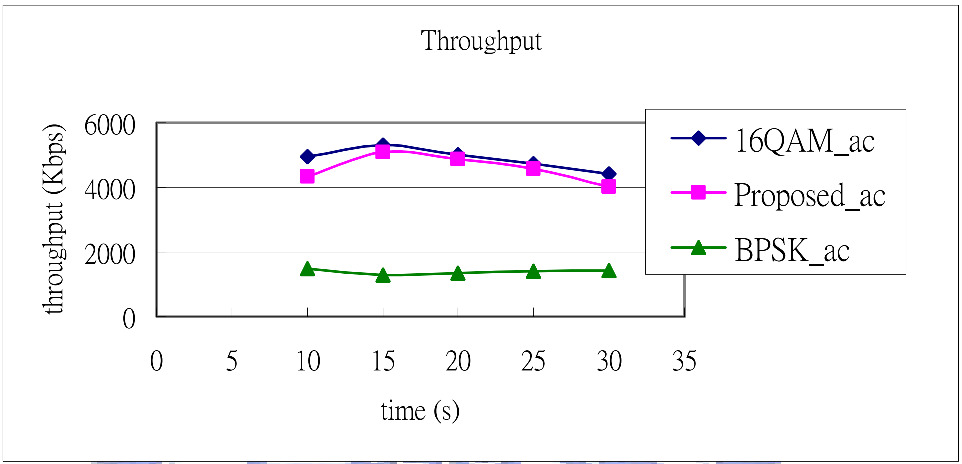

Figure 4.2 Total throughput of admission control.

In Figure 4.2, we can observe the total throughput of general_BPSK_1_2 admission control scheme is about 1400 Kbps. The total throughput of the proposed admission control scheme can achieve at least 4000 Kbps. The throughput of the general_16QAM_1_2 admission control scheme is higher than the proposed scheme slightly. However, the proposed scheme can guarantee the minimum reserved rate requirement for each admitted SS, but the general_16QAM_1_2 scheme can not. The throughput of each admitted stations for the proposed scheme is illustrated in Figure 4.3a and Figure 4.3b. In the two Figures, we can observe that the throughput of each station is higher than the minimum reserved rate of rtPS service (205.44 Kbps).

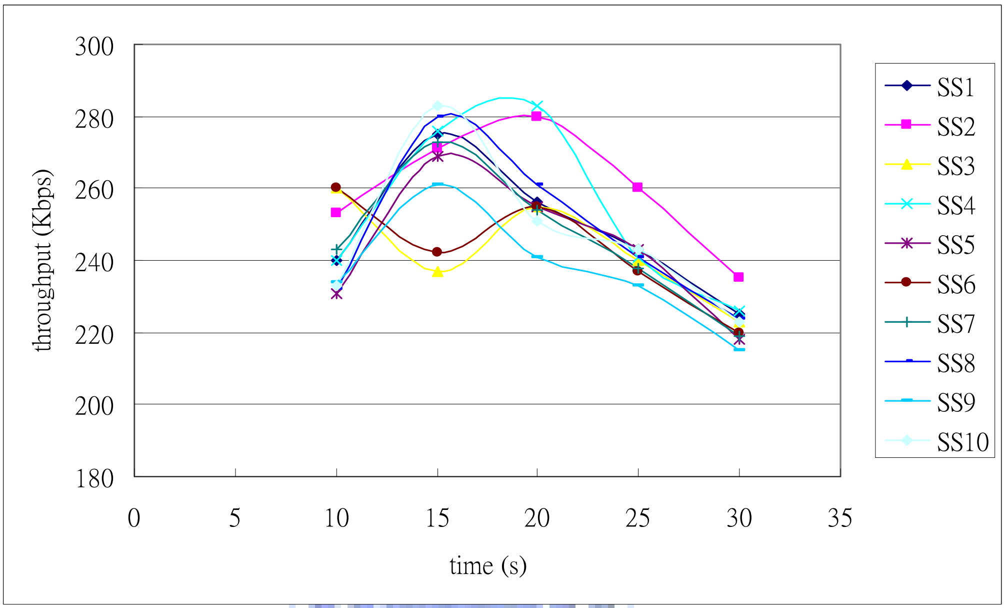

Figure 4.3a Throughput of each SS for the proposed scheme.

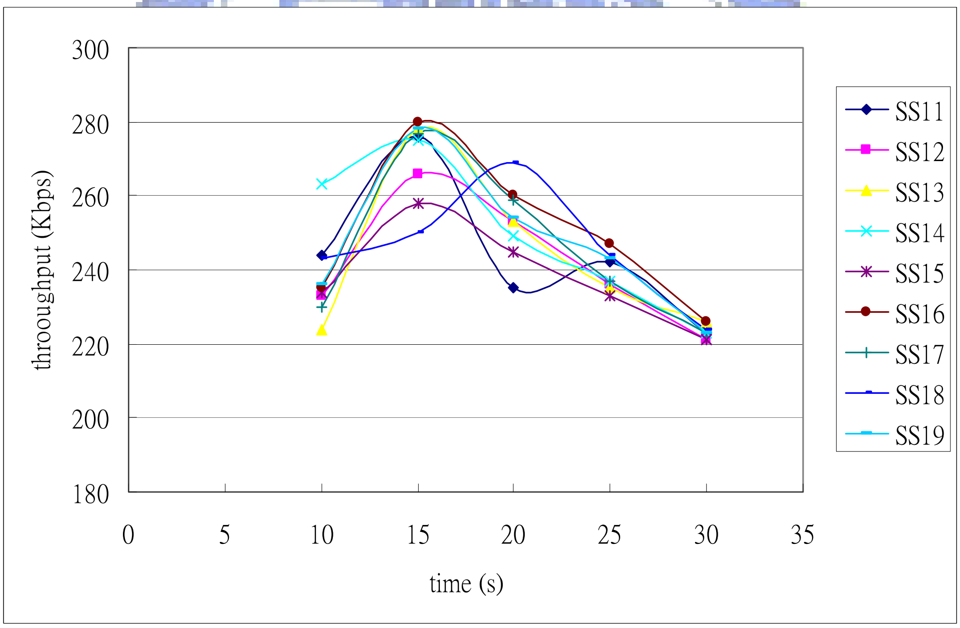

Figure 4.3b Throughput of each SS for the proposed scheme.

We show the throughput of each admitted SS for the general_16QAM_1_2 scheme in Figure 4.4a, Figure 4.4b, and Figure 4.4c. We can observe that the throughput of SS2, SS3, SS6, SS15, and SS18 is zero Kbps, so the general_16QAM_1_2 scheme can not guarantee the minimum reserved rate requirement of each admitted SS. It does not take the modulation and coding scheme into account, so it admits too many rtPS connections into network. Thus, the admission control scheme will affect the QoS guarantee in scheduling deeply.

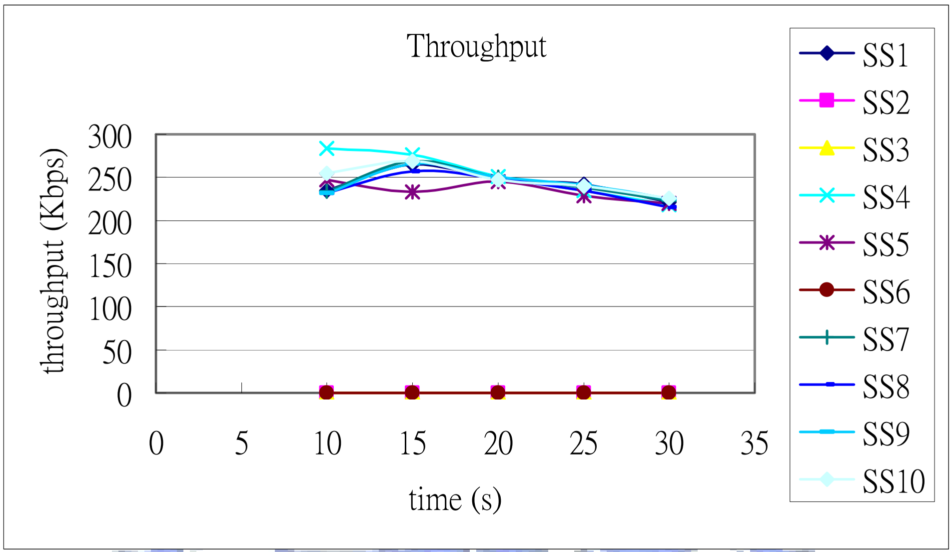

Figure 4.4a Throughput of each SS for the general_16QAM_1_2 scheme.

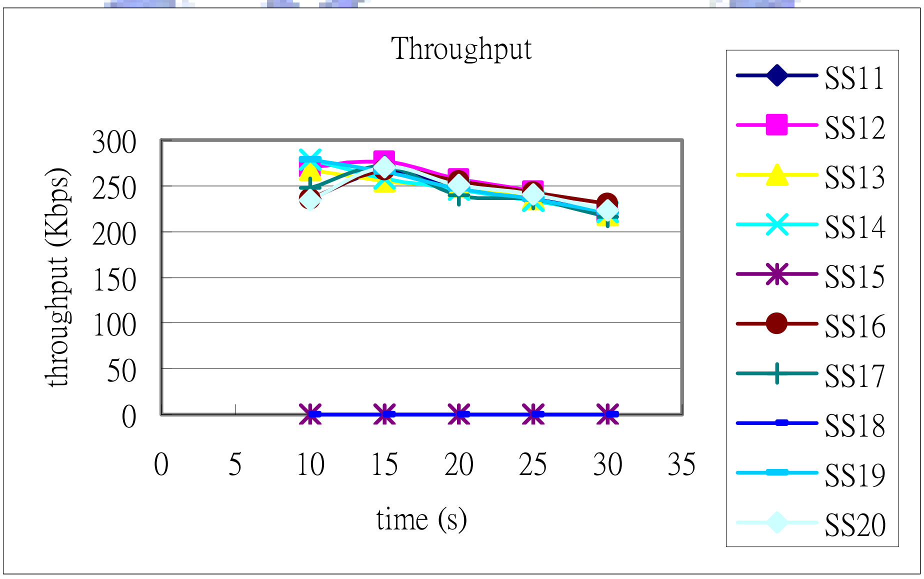

Figure 4.4b Throughput of each SS for the general_16QAM_1_2 scheme.

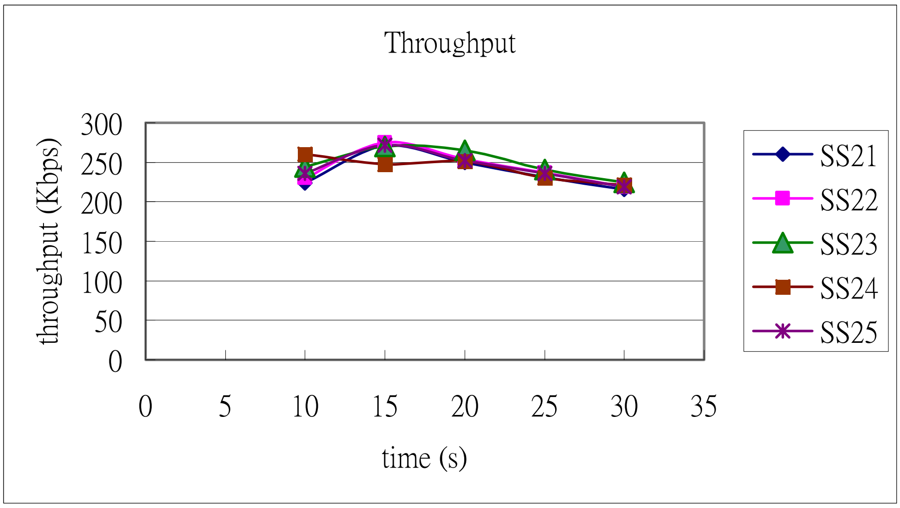

Figure 4.4c Throughput of each SS for the general_16QAM_1_2 scheme.

The above-mentioned results show that the proposed admission control scheme and the proposed scheduling algorithm can utilize the bandwidth more efficiently and guarantee the minimum reserved rate requirement of each admitted SSs.

4.3 Simulation scenario 2

The aim of scenario 2 is to investigate whether the increase of the BE service traffic load influences or not the QoS level of service which higher priority. For this purpose, the simulated scenario includes one BS and 32 SSs. There are 4 UGS connections, 8 rtPS connections, 8 nrtPS connections, and the number of BE connections varies from 2 to 12. The UGS, rtPS, and nrtPS connections use 64QAM_3_4 modulation and coding scheme. The BE connections use 16QAM_1_2 modulation and coding scheme.

As can be seen in Figure 4.5, the average throughput of the rtPS and nrtPS connections decreased slightly as the BE connection increased. The average throughput of the UGS is always 64 Kbps. The proposed scheduling solution can guarantee the minimum reserved rate (64, 205.44, 150 Kbps) requirement to UGS, rtPS, nrtPS connections even though we increase the BE traffic load. Furthermore, we also guarantee the maximum sustained rate (120 Kbps) requirement to the BE connections. Since the BE connection has no minimum reserved rate requirement, the average throughput decreased as the traffic load of whole system increased.

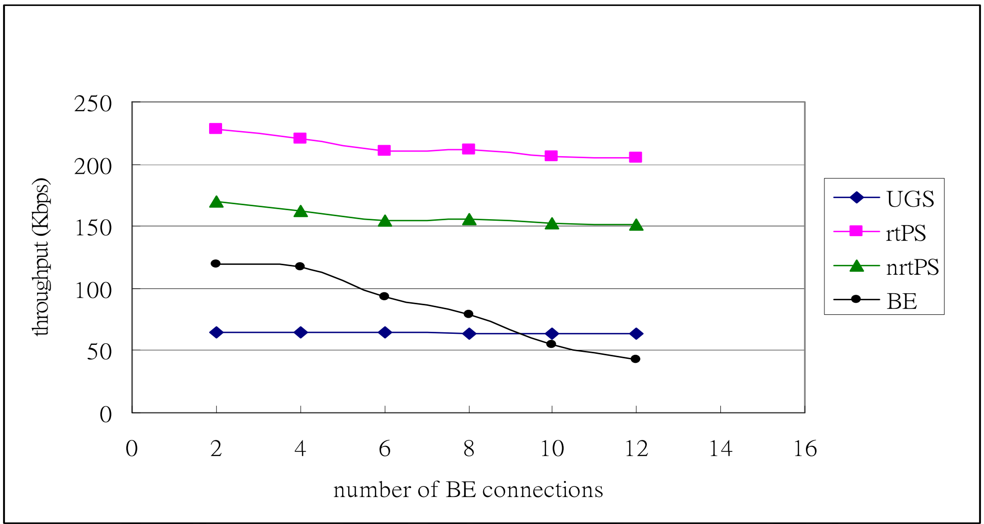

Figure 4.5 Average throughput of four kinds of service connections

Chapter 5 Conclusion and Future Works

In this paper, we take latency and modulation and coding scheme into account in our scheduling solution. The proposed scheduling solution contains three primary schemes: admission control scheme, scheduling algorithm, and bandwidth allocation scheme. Some properties of proposed scheduling solution are as follows: modulation and latency aware, easily implemented in BS, guarantee minimum QoS requirements for some specific services (UGS, rtPS, nrtPS), and allocate bandwidth efficiently.

A precise admission control is very important to the scheduling algorithm as the entire traffic load is heavy. The proposed admission control scheme can calculate the admit condition more precise in the situation which all SSs use different kinds of modulation and coding schemes. The simulation scenario 1 shows that the proposed admission control scheme can admit the suitable number of SSs to help the scheduler guarantee the minimum reserved rate to each admitted connection. If the admission control scheme is not precise, it will influence the guarantee of QoS in the scheduling deeply. In the serious situation, the admission control which does not consider the modulation and coding scheme can results in the failure of QoS guarantee to each admitted SS in the scheduling.

The proposed scheduling algorithm and bandwidth allocation scheme can guarantee the QoS requirements of all kinds of service connections. We allocate the precise and suitable symbols for each connection based on the modulation and coding scheme. Thus, we can utilize the finite bandwidth more efficiently, and improve the total throughput of the entire network. The simulation scenario 2 shows that our scheduling solution can guarantee the minimum reserved rate requirement of some specific connections (UGS, rtPS, nrtPS), and maximum sustained rate requirement of BE connection as the BE traffic load increased.

We use NS-2 (version 2.31) simulation tool [12] with WiMAX NIST module [13] which is originally designed and developed by the National Institute of Standards and Technology (NIST) in U.S.A. This module still has many functions need to be implemented, and we implemented the necessary functions such as the service flow, the service flow QoS, the admission control scheme, and the scheduler algorithm to accommodate our task. Based on the simulation results, the proposed solution has a precise admission control scheme to help the scheduler to guarantee the QoS requirements of each connection when all SS use different modulation and coding schemes. However, our scheduling solution still has some defective situations and we can improve it in future, such as that we could not avoid the starvation of BE connections as the other service traffic load is too heavy. Because the BE service has no minimum reserved rate requirement, we do not reserve some bandwidth for BE service connection in advance.

In our simulation solution, we just consider some parameters such as maximum latency, maximum sustained rate, minimum reserved rate, and modulation and coding scheme. We will take more parameters into account and modify our scheduling solution to solve the starvation of BE connections when other service traffic load is too heavy.

References

[1] IEEE Std 802.16-2004 (Revision of IEEE Std 802.16-2001)

[2] IEEE Std 802.16e-2005 and IEEE Std 802.16-2004. Cor 1-2005(Amendment and Corrigendum to IEEE Std 802.16-2004)

[3] Loutfi Nuaymi, “WiMAX: Technology for Broadband Wireless Access”, John Wiley & Sons, Ltd, 2007

[4] Jheng-Yue, Liou, “An adaptive bandwidth request scheme for 802.16e polling services”, 2007

[5] J. Chen, W. Jiao, H. Whang, “A service flow management strategy for IEEE 802.16 broadband wireless access systems in TDD mode”, IEEE International Conference on Communications, vol. 5, 2005

[6] K. Wongthavarawat, “IEEE 802.16 based last mile broadband wireless military networks with quality of service support”, IEEE Military Communications Conference, vol. 2, 2005

[7] G. Chu, D. Whang, and S. Mei, “A QoS Architecture for the MAC Protocol of IEEE802.16 BWA System”, IEEE International Conference on Conmunications, 2002

[8] M. Hawa and D. W. petr, “Quality of Service Scheduling in Cable and Broadband Wireless Access Systems”, the 10th IEEE International Workshop on Quality of Service, 2002

[9] C. Cicconatti, L. Lenzini, E. Mingozi, “Quality of service support in IEEE 802.16 networks”, IEEE Networks 20 (2), 2006

[10] J. Freitag and Nelson L. S. Fonseca, “Uplink Scheduling with Quality of Service in IEEE 802.16 Networks”, IEEE GLOBECOM, 2007

[11] Pulse code modulation (PCM) of voice frequencies, ITU-T recommendation G.711, 1988

[12] http://www.isi.edu/nsnam/ns/

[13] http://www.antd.nist.gov/seamlessandsecure/download.html