

A Study of WiMax QoS Mechanisms

Masood KHOSROSHAHY

Vivien NGUYEN

RMOB Project

April 2006

Project supervisor: Prof. Philippe Godlewski

Properties of IEEE Standard 802.16

- Broad bandwidth–Up to 134 Mbit/s in 28 MHz channel (in 10-66 GHz air interface)
- Supports multiple services simultaneously with full QoS– Efficiently transport IPv4, IPv6, ATM, Ethernet, etc.
- Bandwidth on demand (frame by frame)
- •MAC designed for efficient use of spectrum
- Comprehensive, modern, and extensible security
- Supports multiple frequency allocations from 2-66 GHz– OFDM and OFDMA for non-line-of-sight applications
- TDD and FDD
- Link adaptation: Adaptive modulation and coding– Subscriber by subscriber, burst by burst, uplink and downlink
- Point-to-multipoint topology, with mesh extensions
- Support for adaptive antennas and space-time coding
- Extensions to mobility

802.16 MAC: Overview

- Point-to-Multipoint
- Metropolitan Area Network
- Connection-oriented
- Supports difficult user environments
 - High bandwidth, hundreds of users per channel
 - Continuous and burst traffic
 - Very efficient use of spectrum
- Protocol-Independent core (ATM, IP, Ethernet, ...)
- Balances between stability of contention-less and efficiency of contention-based operation
- Flexible QoS offerings
 CBR, rt-VBR, nrt-VBR, BE, with granularity within classes
- Supports multiple 802.16 PHYs

Subscriber Station (SS) initialization has several steps:

- -The SS scans for downlink channel and establishes synchronization with the BS.
- -It obtains transmit parameters.
- -It performs ranging and negotiating basic capabilities.
- -It is authorized by the BS and performs key exchange.
- -It performs the registration and IP connectivity establishment.
- -Time of day establishment and the transfer of operational parameters.
- -It sets up the connections.

PHY LAYER TECHNIQUES

OFDM OFDMA

OFDM

 orthogonal frequency-division multiplexing eliminates interference between channels no ISI (coherence band of channel) higher spectral efficiency high resistance to interference (channel code) adjustable rate (adaptive modulation)

OFDM Encoder



OFDMA

multiple user access
OFDM needs TDD or FDD
inter symbol interference immunity
frequency selective fading immunity

Studies of 2 papers

- Ian C. Wong, Zukang Shen, Brian L. Evans, and Jeffrey G. Andrews, "A Low Complexity Algorithm for Proportional Resource Allocation in OFDMA Systems", 2004. [WSEA-04]
- Mustafa Ergen, Sinem Coleri, and Pravin Varaiya "QoS Aware Adaptive Resource Allocation Techniques for Fair Scheduling in OFDMA Based Broadband Wireless Access Systems", IEEE Trans. Broadcast, vol. 49, no. 4, Dec. 2003. [ECV-03]

1- Study of

 Mustafa Ergen, Sinem Coleri, and Pravin Varaiya "QoS Aware Adaptive Resource Allocation Techniques for Fair Scheduling in OFDMA Based Broadband Wireless Access Systems", IEEE Trans. Broadcast, vol. 49, no. 4, Dec. 2003 [ECV-03]

Resource allocation [ECV-03]

 QoS fixed by the application
 goal : achieving a specified data transmission rate and BER for each user in each transmission

OFDMA System [ECV-03]



Orthogonal frequency division multiple access system.

Solution [ECV-03]

 subcarriers are distributed among users adaptive modulation in OFDMA systems optimal solution for ressouce allocation "Integer Programming" - data rate and BER vs total power

Control channels

• channel estimation

• ressouce allocation within coherence time

OFDMA System [ECV-03]



Orthogonal frequency division multiple access system.

Solution [ECV-03]

need of low complexity algorithms
need of solutions closed to the optimal one
subcarrier allocation (linear programming, hungarian algorithm)
adaptive modulation (bit loading algorithm)

 combination and improvements algo. (fair scheduling, greedy realising, horizontal and vertical swapping)

2- Study of

 Ian C. Wong, Zukang Shen, Brian L. Evans, and Jeffrey G. Andrews, "A Low Complexity Algorithm for Proportional Resource Allocation in OFDMA Systems", 2004 [WSEA-04]

Resource allocation [WSEA-04]

assigning subcarriers to SS
assigning power
BER and data rate contraints

Fixed data rate and BER [WSEA-04]

- fixed user data rates and BER requirements
- goal: minimize the total transmitted power
- problem : need of an adaptive modulation system

Rate-adaptive [WSEA-04]

maximize the total data rate (capacity) over all users
power and BER constraints
use of waterfilling
problem : no fairness
goal: minimum user's data rate

Resource allocation [WSEA-04]

 solution : maximization of the the total capacity subject to user rate constraints (fairness)

goal : low complexity algorithm

OFDMA system [WSEA-04]



[WSEA-04] OFDMA system block diagram for K users. Each user is allocated different set of subcarriers by the basestation.

Solution [WSEA-04]

- rate-adaptive resource allocation problem with "proportional rate constraints" for OFDMA systems
- subcarrier allocation scheme that achieves "approximate rate proportionality" while maximizing the total capacity
- direct algorithm with a much lower complexity versus an iterative algorithm
 higher total capacity

MAC LAYER QoS ARCHITECTURES

"Quality of Service Support in IEEE 802.16 Networks"

"Providing integrated QoS control for IEEE 802.16 broadband wireless access systems"

"Quality of Service Support in IEEE 802.16 Networks"

Analysis by simulating the performance of IEEE 802.16 in two application scenarios:
-residential subscribers
-SME (Small and Medium-sized Enterprises) subscribers

Analysis aimed at showing the effectiveness of the 802.16 MAC protocol in providing differentiated services to applications with different QoS requirements, such as VoIP and Web.

Blueprint of the functional entities for QoS support



- Each downlink connection has a packet queue at the BS.
- Uplink connection queues reside at SSs. Since the BS controls the access to the medium in the uplink direction, bandwidth is granted to SSs on demand.
- Bandwidth requests are used on the BS for estimating the residual backlog of uplink connections.
- Allocates future uplink grants according to the respective set of QoS parameters and the virtual status of the queues.

Implemented Algorithms:

Since minimum reserved rate is the basic QoS parameter negotiated by a connection, so: the class of *latency-rate scheduling algorithms*

Deficit round robin (DRR) as the downlink scheduler at the BS:
 the ability of providing fair queuing in the presence of variable length packets
 the simplicity of implementation.

-DRR assumes that the size of the head-of-line packet is known at each packet queue: cannot be used by the BS to schedule transmissions in the uplink direction.

 Weighted round robin (WRR) is selected as the uplink scheduler. Like DRR, WRR belongs to the class of ratelatency scheduling algorithms.

-DRR is implemented as the SS scheduler, because the SS knows the sizes of the head-of-line packets of its queues.

The performance of an IEEE 802.16 system has been assessed under two traffic scenarios:

The first one (residential scenario) dealt with data (non-QoS) traffic only: BE scheduling service:

-The average delay of the uplink traffic is higher than that of the downlink traffic. And the former increases more sharply than the latter with the offered load. This behavior can be explained by means of both the bandwidth-request mechanism and the overhead introduced by physical preambles.

The second scenario (SME scenario), they have shown the service differentiation, in terms of delay, between data (served via BE) and multimedia traffic (served via rtPS). This is achieved because scheduling in 802.16 is controlled by the BS in both the downlink and uplink directions. Therefore, it is possible to employ scheduling algorithms which are able to provide QoS guarantees.

"Providing integrated QoS control for IEEE 802.16 broadband wireless access systems" In this study, the authors propose a new integrated QoS architecture for IEEE 802.16 Broadband Wireless MAN in TDD mode. A mapping rule for providing DiffServ between IP layer and MAC layer is given and a fast signaling mechanism(IntServ) is designed to provide cross layer integrated QoS for Point to Multi-Point (PMP) mode.

IntServ:

 Applications requiring guaranteed service or controlled-load service must set up the paths and reserve resources before transmitting their data.
 The QoS provision procedure will consist of the following two part:

-The secondary management connection will be used for RSVP to provide the layer 3 QoS

-The primary management connection will be used for DSA/DSC/DSD to provide the layer 2 QoS.

Result > The whole QoS provision will be rather slow.

They propose an integrated QoS control architecture as shown in the figure, which implements a cross layer trafficbased prioritization mechanism in a comprehensive way.



- Step 1 and 2, when a new service flow arrives in IP layer, it will be firstly parsed according to the definition in PATH message (for InteServ) or Differentiated Services Code Point (DSCP for DiffServ); then classified and mapped into one of four types of services (UGS, rtPS, nrtPS or BE).
- In step 3, the dynamic service model in SS will send request message to the BS, then the admission control in BS will determine whether this request will be approved or not.
- If not, the service module will inform upper layer to deny this service in step 4.
- If yes, admission control will notify scheduling module to make a provision in its basis scheduling parameter according to the value shown in the request message and the accepted service will transfer into traffic grooming module in step 5.

 In step 6, according to the traffic grooming result, SS will send Bandwidth Request message to BS.

 The scheduling module in BS will retrieve the requests (step 7) and generate UL-MAP and DL-MAP message (step 8) following the bandwidth allocation results.
 In steps 9-10, the SS will package SDUs

from IP layer into PDUs and upload them in its uplink slot to BS.

Admission Control and Scheduling in BS



- It will collect all the DSA/DSC/DSD requests and update the estimated available bandwidth based on bandwidth change. The hierarchical structure of the bandwidth allocation in BS is shown in the figure.
- Two-layer scheduling is deployed. Six queues are defined according to their direction (uplink or downlink) and service classes (rtPS, nrtPS and BE). Since service of UGS will be allocated fixed bandwidth (or fixed time duration) in transmission, these bandwidths will be cut directly before each scheduling.

The algorithm of the first layer scheduling is called Deficit Fair Priority Queue (DFPQ) which is basically based on priority queue.
In the second layer scheduling, three different algorithms are assigned to three classes of service to match its requirements:

Earliest deadline first (EDF) for rtPS
Weight fair queue (WFQ) for nrtPS.
BE connection by round robin (RR).

Comparison between two ways of RSVP

- The negotiation of QoS parameters for one traffic will be processed two times:
- For the first time, the parameters are carried in RSVP messages and transmitted through the Secondary Management connection.
- For the second time, the same parameters are mapped in MAC message and transmitted through the Primary Management Connection.
- With the utilization of the new mapping rule, the RSVP signaling messages are mapped directly into the MAC messages, and then transmitted through the Primary Management Connection. In this way, the messages are transmitted only once, reducing the delay.

Other studies:

- "A QoS Architecture for the MAC Protocol of IEEE 802.16 BWA System"
- "A Quality of Service Architecture for IEEE 802.16 Standards"
- "Quality of service scheduling in cable and broadband wireless access systems"
- "Exploiting MAC flexibility in WiMAX for media streaming"
- "Algorithms for routing and centralized scheduling to provide QoS in IEEE 802.16 mesh networks"
- "Modeling and performance analysis of the distributed scheduler in IEEE 802.16 mesh mode"

Thank you for your attention ...