PhD Dissertation

International Doctorate School in Information and Communication Technologies

DIT - University of Trento

MAC Layer Investigation and Design for Gigabit Wireless LAN

Tao Chen

Advisor:
Prof. Imrich Chlamtac
Università degli Studi di Trento

Co-Advisor:
Dr. Hagen Woesner
European Center for Information and Communication Technologies

March 2007
Abstract

The age of the gigabit wireless LAN (WLAN) is coming. Evidences are witnessed from the increasing demand of emerging bandwidth intensive applications, the recent millimeter wave frequency band regulations around the world, the technical advances in the millimeter wave frontend design and signal processing, and the standardization activities on gigabit wireless personal area networks (WPAN). However, it is not a simple task to implement a reliable and efficient gigabit WLAN, in which three challenges are prominent: the choice of a proper radio frequency band with adequate spectrum capacity, the development of advanced air interfaces to fight channel impairments in high speed communications, and the design of a MAC protocol that can use the bandwidth efficiently. The 60 GHz unlicensed band, which is released by FCC in 2001, seems to meet the first challenge. Moreover, the recent research advances on low cost SiGe or CMOS based millimeter wave frontend pave the way for massive deployment of low cost 60 GHz radios. Nevertheless, the MAC protocols for gigabit WLANs are not well explored. It is the objective of this thesis to study the link and MAC layer issues of gigabit WLANs, under the assumption that future gigabit WLANs are based on the 60 GHz band.

To investigate these issues, the first question raised in the mind is to what extent the MAC protocols of current WLANs and WPANs can be employed into gigabit WLANs. To answer this, we identify the potential MAC candidates used in current WLANs and WPANs, including IEEE 802.11n, HiperLAN/2, and IEEE 802.15.3. Moreover, as the millimeter wave band is applied by IEEE 802.16 for high speed multi-point access, it is nature to consider its MAC protocol as a candidate as well. The functionalities of these MAC candidates are investigated. It concludes that from the functionality perspective each MAC candidate supports high speed communication
to a certain extent. However, the characteristics of the 60 GHz band significantly affect the gigabit WLAN design. The fact that 60 GHz signals do not penetrate solid materials very well demands a large number of BSs to cover a large building. To reduce the overall cost of the system, novel system architecture is needed. For this we propose a hybrid optical/wireless system architecture for 60 GHz based gigabit WLANs, which uses the WDM/TDM EPON as its fixed infrastructure and the 60 GHz radio for the wireless access. Based on the architecture, a MAC protocol is customized for this system. The issues of channel access, bandwidth allocation, error control, and handover are investigated and solutions are provided.

Since the bandwidth allocation is the core function of the proposed MAC protocol, we propose a multi-wavelength based dynamic bandwidth allocation (DBA) algorithm, which jointly considers the duplex modes of nodes, the cell plan, and the bandwidth requirement of nodes for allocation. Moreover, we study the approach to enable channel state aware scheduling in the proposed system architecture. A long-term channel state metric derived from the overhead of the error control scheme is proposed, and used in downstream scheduling algorithms.

Finally, the protocol efficiency of the MAC candidates is extensively studied. The result reveals that at the data rate of Gbps all MAC candidates but those of WIGEE, HiperLAN/2 and the data aggregate modes of 802.11n show low efficiency, which is resulted from the rate dependent overhead in those protocols. The overhead reduction therefore becomes the main consideration in the gigabit MAC protocol design.

**Keywords**

Wireless LAN (WLAN), Medium Access Control (MAC), Ethernet Passive Optical Network (EPON), Resource Allocation, BER Model
Acknowledgement

I would like to give my first thanks to my thesis advisor, Professor Imrich Chlamtac, for his continuous support during the entire course of my graduate studies. My second thanks are dedicated to Dr. Honggang Zhang, for his insightful advice, unreserved encourage and fruitful discussion during my Ph.D. studies. Moreover, I would like to thank Dr. Hagen Woesner, for his initial idea to start my thesis work and insightful discussion on my thesis, and Dr. Gian Mario Maggio, for his kind support during the last stage of my graduate studies.

I am very grateful to the members of my Ph.D. examination committee for their valuable time and advice.

I had the pleasure to share my work and life during my stay at the University of Trento with my friends Michele Merler, Mingmin Chi, Xiaofei Zhou, Jin Yu, Yabin Ye and Mingqing Hu.

The work would not be possible without the support of my family. I’d like to thank my parents and brother for their love and unconditional supports. My last thanks are reserved to my lovely wife Ping Zhang. Without her love and dedication, none of these is ever possible.
## Contents

1 Introduction 1

1.1 Structure of the Thesis ................................. 7

2 Gigabit Wireless MAC Candidates 11

2.1 Classification of Wireless MAC .......................... 12

2.2 State of the Art in MM-wave Based Multi-access Systems . 13

2.3 Challenges of Gigabit Wireless MAC Design ......... 16

2.3.1 Physical Constraints ............................... 16

2.3.2 Resource Allocation ................................. 19

2.3.3 QoS Provisioning ................................. 22

2.3.4 Handover ............................. 23

2.4 MAC Protocol Candidates for Gigabit Wireless LAN . 24

2.4.1 802.11n ........................................ 25

2.4.2 HiperLAN/2 ....................................... 32

2.4.3 802.15.3 ........................................ 36

2.4.4 IEEE 802.16 .................................... 41

2.4.5 Comparison of MAC Protocol Candidates ........ 46

2.5 Conclusion ........................................ 49

3 Wireless Gigabit Ethernet Extension 51

3.1 WiGEE System Architecture ............................. 53

3.1.1 Overall Architecture ............................... 54
3.1.2 OLT Design .............................................. 57
3.1.3 BS Design .............................................. 59
3.1.4 MT Design .............................................. 60
3.2 Scalability of the System ................................. 60
3.3 Medium Access Control of WiGEE ......................... 63
  3.3.1 Superframe Structure ................................. 63
  3.3.2 Link Layer Functions ................................. 66
3.4 Error Control in WiGEE ..................................... 73
  3.4.1 ARQ .................................................. 74
  3.4.2 Determine ARQ Parameters ......................... 75
  3.4.3 QoS Consideration ................................. 78
  3.4.4 Performance Analysis ............................. 78
3.5 Simulation Study of WiGEE ................................. 81
  3.5.1 Result Analysis ..................................... 83
3.6 Conclusion ................................................ 90

4 BER Models for WiGEE ........................................ 91
  4.1 Analytical Error Models ................................. 92
    4.1.1 BSC Model ....................................... 93
    4.1.2 Two State Markov Models ...................... 94
    4.1.3 N-state Models ................................. 95
    4.1.4 Other Models .................................. 96
  4.2 Trace File Analysis .................................... 96
  4.3 Suitability of Existing Models ..................... 102
  4.4 Conclusion ........................................... 106

5 Bandwidth Allocation Algorithm in WiGEE .................. 107
  5.1 DBA Algorithm ......................................... 109
    5.1.1 Compensate ARQ Overhead in DBA .............. 117
    5.1.2 Simulation Result .............................. 118
5.2 Downstream Scheduling Algorithm ........................................ 123
  5.2.1 Background .......................................................... 123
  5.2.2 System Model ........................................................ 125
  5.2.3 Channel Estimation .................................................. 127
  5.2.4 Scheduling Algorithms .............................................. 127
  5.2.5 Simulation Study ..................................................... 134
  5.3 Conclusion ............................................................... 139

6 Performance Study of MAC Protocols .................................... 141
  6.1 Performance Analysis of MAC Protocol Candidates ............... 141
    6.1.1 WiGEE ............................................................ 142
    6.1.2 IEEE 802.11n .................................................... 151
    6.1.3 HiperLAN/2 ....................................................... 173
    6.1.4 IEEE 802.15.3 .................................................... 176
    6.1.5 IEEE 802.16 ...................................................... 183
  6.2 Summary ................................................................. 183

7 Conclusion ........................................................................... 185
  7.1 Future Work ............................................................... 191

Bibliography ............................................................................. 193
## List of Tables

1.1 Bandwidth requirements of various applications ................ 2

2.1 Comparison of MAC protocol candidates ....................... 49

3.1 Notations and values to calculate link budget of optical in-
frastucture ............................................... 61

3.2 Extended types in MPCP control frame used for ARQ support 77

4.1 Parameter settings for 60 GHz channel models .................. 97

4.2 Statistics of trace files .................................... 99

4.3 Evaluation result based on area metric .......................... 105

5.1 Compensation algorithms in DBA ............................... 117

5.2 BER settings for different packet lengths in different simu-
lation scenarios ............................................. 135
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Evolution of WLAN</td>
<td>2</td>
</tr>
<tr>
<td>1.2</td>
<td>60 GHz band regulations around the world</td>
<td>4</td>
</tr>
<tr>
<td>2.1</td>
<td>The classification of wireless MAC protocols</td>
<td>13</td>
</tr>
<tr>
<td>2.2</td>
<td>MAC issues in using directional antennas</td>
<td>18</td>
</tr>
<tr>
<td>2.3</td>
<td>Contention free and contention period in 802.11</td>
<td>26</td>
</tr>
<tr>
<td>2.4</td>
<td>Differential services supported by AIFS in 802.11e</td>
<td>28</td>
</tr>
<tr>
<td>2.5</td>
<td>A-MSDU in 802.11n</td>
<td>29</td>
</tr>
<tr>
<td>2.6</td>
<td>A-MPDU in 802.11n</td>
<td>30</td>
</tr>
<tr>
<td>2.7</td>
<td>Block Acknowledgment in 802.11n</td>
<td>31</td>
</tr>
<tr>
<td>2.8</td>
<td>HiperLAN/2 frame structure</td>
<td>34</td>
</tr>
<tr>
<td>2.9</td>
<td>Piconet concept in 802.15.3</td>
<td>37</td>
</tr>
<tr>
<td>2.10</td>
<td>Superframe structure in 802.15.3</td>
<td>38</td>
</tr>
<tr>
<td>2.11</td>
<td>TDD access mode in 802.16</td>
<td>42</td>
</tr>
<tr>
<td>2.12</td>
<td>FDD access mode in 802.16</td>
<td>43</td>
</tr>
<tr>
<td>3.1</td>
<td>WiGEE network</td>
<td>52</td>
</tr>
<tr>
<td>3.2</td>
<td>Architecture of WiGEE</td>
<td>53</td>
</tr>
<tr>
<td>3.3</td>
<td>Alternative architecture of WiGEE</td>
<td>55</td>
</tr>
<tr>
<td>3.4</td>
<td>Baseband process in CS</td>
<td>58</td>
</tr>
<tr>
<td>3.5</td>
<td>Frontend process in BS</td>
<td>59</td>
</tr>
<tr>
<td>3.6</td>
<td>Structure of MT</td>
<td>61</td>
</tr>
<tr>
<td>3.7</td>
<td>Superframe structure of WiGEE</td>
<td>64</td>
</tr>
</tbody>
</table>
3.8 Statistical multiplex of downstream and upstream in WiGEE, observed at BSs .............................................. 66
3.9 Protocol stack of WiGEE ........................................ 68
3.10 Intra-wavelength and inter-wavelength handover. Intra-wavelength handover occurs in point 4 or 5; inter-wavelength handover occurs in point 2 or 3 ............................................. 72
3.11 Ethernet frame with extension to support ARQ ............ 77
3.12 MPCP control frame for ARQ control ...................... 78
3.13 IPACT-like DBA algorithm in full duplex mode .......... 83
3.14 Average throughput as a function of offered load per wavelength, no ARQ control, stationary scenario ............ 84
3.15 Average delay as a function of offered load per wavelength, no ARQ control, stationary scenario ..................... 85
3.16 Drop ratio as a function of offered load per wavelength, no ARQ control, stationary scenario ....................... 86
3.17 Average throughput as a function of offered load per wavelength, stationary and mobility scenario ............... 86
3.18 Average delay as a function of offered load per wavelength, stationary and mobility scenario ..................... 87
3.19 Drop ratio as a function of offered load per wavelength, stationary and mobility scenario ....................... 88
3.20 Average throughput as a function of offered load per wavelength, SR-ARQ, stationary scenario ..................... 89
3.21 Average delay as a function of offered load per wavelength, SR-ARQ, stationary scenario ....................... 89
4.1 CCDF of run lengths obtained from trace file 2 (Refer to Table 4.2) ....................................................... 98
4.2 CCDF of run lengths obtained from trace file 5, (Refer to Table 4.2) ................................. 100
4.3 ACF of run lengths in trace file 2, (Refer to Table 4.2) ... 100
4.4 ACF of run lengths in trace file 5, (Refer to Table 4.2) ... 101
4.5 CCDF of run lengths obtained from trace file 5 after corrected by RS(255,251,2) .................. 102
4.6 A set of distributions to fit CCDF of run lengths in trace file 2 2103
4.7 A set of distributions to fit CCDF of run lengths in trace file 5 103
4.8 A set of distributions to fit CCDF of run lengths in trace file 2 after RS code .......................... 104
4.9 A set of distributions to fit CCDF of run lengths in trace file 5 after RS code ...................... 105

5.1 Proposed DBA algorithm in half duplex mode ............ 111
5.2 Proposed DBA algorithm in full duplex mode ............ 112
5.3 Proposed DBA algorithm deals with overlapping area, in half duplex mode .......................... 115
5.4 Average throughput as a function of offered load per wavelength, SR-ARQ, BSC error model, frame size of 512 bytes, stationary scenario ................................. 119
5.5 Average throughput as a function of offered load per wavelength, SR-ARQ, BSC error model, various frame size, stationary scenario ................................. 120
5.6 Average delay as a function of offered load per wavelength under BSC model, SR-ARQ, various frame size, stationary scenario ................................. 120
5.7 Average throughput as a function of offered load per wavelength under different error models, SR-ARQ, various frame size, stationary scenario ................................. 121
5.8 Average delay as a function of offered load per wavelength under different error models, SR-ARQ, various frame size, stationary scenario ........................................ 121
5.9 Overall throughput under different algorithms, 4-node case 136
5.10 Overall throughput under different algorithms, 8-node case 136
5.11 Overall throughput under different algorithms, 20-node case 137
5.12 Throughput per node in 4 and 8-node cases ............... 138
5.13 Overall throughput under different algorithms, 4-node case, each node has different data rate ....................... 138

6.1 Absolute throughput under various data rates, half duplex mode, payload size of 512 bytes, and superframe length of 2ms .................................................. 146
6.2 Normalized throughput under various data rates, half duplex mode, payload size of 512 bytes, and superframe length of 2ms .................................................. 147
6.3 Normalized throughput under various data rates, half and full duplex mode, payload size of 512 bytes, and superframe length of 2ms ........................................ 148
6.4 Normalized throughput under various data rates, half and full duplex mode, four scheduled MTs per superframe, various payload sizes, and superframe length of 2ms ............ 149
6.5 Normalized throughput under various data rates, half and full duplex mode, four scheduled MTs per superframe, payload size of 512 bytes, various DS lengths, and superframe length of 2ms ........................................ 150
6.6 Normalized throughput under various data rates, half and full duplex mode, four scheduled MTs per superframe, payload size of 512 bytes, and various superframe lengths ...... 150
6.7 Common data frame structure of 802.11 . . . . . . . . . . 152
6.8 DCF access mode of 802.11n . . . . . . . . . . . . . . . . . 152
6.9 EDCA mode of 802.11n . . . . . . . . . . . . . . . . . . . 153
6.10 HCCA mode of 802.11n . . . . . . . . . . . . . . . . . . . 153
6.11 Data aggregate mode of 802.11n . . . . . . . . . . . . . . 154
6.12 A-MSDU frame in 802.11n . . . . . . . . . . . . . . . . . 154
6.13 A-MPDU frame in 802.11n . . . . . . . . . . . . . . . . . 155
6.14 Absolute throughput under 802.11n DCF, various data rates,
and various payload sizes . . . . . . . . . . . . . . . . . . . 160
6.15 Normalized throughput under 802.11n DCF, various data
rates, and various payload sizes . . . . . . . . . . . . . . . 161
6.16 Absolute throughput under 802.11n EDCA, various data
rates, various payload sizes, and TXOP length of 2ms . . . 162
6.17 Normalized throughput under 802.11n EDCA, various data
rates, various payload sizes, and TXOP length of 2ms . . . 163
6.18 Normalized throughput under 802.11n EDCA, various data
rates, normal ACK or BA, payload size of 512 bytes, various
TXOP lengths . . . . . . . . . . . . . . . . . . . . . . . . . 164
6.19 Normalized throughput under 802.11n HCCA, various data
rates, various payload sizes, and TXOP length of 2ms . . . 164
6.20 Normalized throughput under 802.11n HCCA, various data
rates, normal ACK or BA, payload size of 512 bytes, various
TXOP lengths . . . . . . . . . . . . . . . . . . . . . . . . . 165
6.21 Normalized throughput under 802.11n aggregate EDCA mode,
various data rates, various payload sizes, TXOP length of
10ms, A-MSDU=8191 bytes, and A-MPDU=23040 bytes . 166
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.22</td>
<td>Normalized throughput under 802.11n aggregate EDCA mode, various data rates, payload size of 512 bytes, TXOP length of 10ms, various A-MSDU lengths, and various A-MPDU lengths</td>
<td>167</td>
</tr>
<tr>
<td>6.23</td>
<td>Normalized throughput under 802.11n aggregate EDCA mode, various data rates, payload size of 512 bytes, various TXOP length, A-MSDU=8191 bytes, and A-MPDU=23040 bytes</td>
<td>167</td>
</tr>
<tr>
<td>6.24</td>
<td>Normalized throughput under 802.11n aggregate HCCA mode, various data rates, various payload sizes, TXOP length of 10ms, A-MSDU=8191 bytes, and A-MPDU=23040 bytes</td>
<td>168</td>
</tr>
<tr>
<td>6.25</td>
<td>Normalized throughput under 802.11n aggregate HCCA mode, various data rates, payload size of 512 bytes, TXOP length of 10ms, various A-MSDU lengths, and various A-MPDU lengths</td>
<td>169</td>
</tr>
<tr>
<td>6.26</td>
<td>Normalized throughput under 802.11n aggregate HCCA mode, various data rates, payload size of 512 bytes, various TXOP length, A-MSDU=8191 bytes, and A-MPDU=23040 bytes</td>
<td>169</td>
</tr>
<tr>
<td>6.27</td>
<td>Normalized throughput under A-MSDU of EDCA and HCCA mode in 802.11n, various data rates, payload size of 512 bytes, TXOP length of 10ms, A-MSDU=8191 bytes</td>
<td>170</td>
</tr>
<tr>
<td>6.28</td>
<td>Absolute saturated throughput under 802.11n DCF, various data rates, and various payload sizes</td>
<td>170</td>
</tr>
<tr>
<td>6.29</td>
<td>Normalized saturated throughput under 802.11n DCF, various data rates, and various payload sizes</td>
<td>171</td>
</tr>
<tr>
<td>6.30</td>
<td>Saturated throughput of 802.11n DCF under different access probabilities, various data rates, and payload size of 512 bytes</td>
<td>171</td>
</tr>
<tr>
<td>Section</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>6.31</td>
<td>Normalized throughput of different access modes in 802.11n, various data rates, payload size of 512 bytes, A-MPDU and A-MSDU take TXOP length of 10ms, A-MPDU takes length of 23040 bytes, A-MSDU takes length of 8191 bytes, EDCA and HCCA take TXOP length of 2ms.</td>
<td></td>
</tr>
<tr>
<td>6.32</td>
<td>Superframe structure of HiperLAN/2</td>
<td></td>
</tr>
<tr>
<td>6.33</td>
<td>Absolute throughput of HiperLAN/2 under various data rates, and different $N_{\text{down}}$ and $N_{\text{up}}$ sessions</td>
<td></td>
</tr>
<tr>
<td>6.34</td>
<td>Normalized throughput of HiperLAN/2 under various data rates, and different $N_{\text{down}}$ and $N_{\text{up}}$ sessions</td>
<td></td>
</tr>
<tr>
<td>6.35</td>
<td>Superframe structure of 802.15.3</td>
<td></td>
</tr>
<tr>
<td>6.36</td>
<td>ACK mechanisms of 802.15.3</td>
<td></td>
</tr>
<tr>
<td>6.37</td>
<td>Data frame structure of 802.15.3</td>
<td></td>
</tr>
<tr>
<td>6.38</td>
<td>Absolute throughput of 802.15.3 under Imm-ACK mode, various data rates, payload size of 512 bytes, and superframe length of 2ms</td>
<td></td>
</tr>
<tr>
<td>6.39</td>
<td>Normalized throughput of 802.15.3 under Imm-ACK mode, various data rates, payload size of 512 bytes, and superframe length of 2ms</td>
<td></td>
</tr>
<tr>
<td>6.40</td>
<td>Normalized throughput of 802.15.3 under No-ACK mode, various data rates, payload size of 512 bytes, and superframe length of 2ms</td>
<td></td>
</tr>
<tr>
<td>6.41</td>
<td>Normalized throughput of 802.15.3 under Dly-ACK mode, various data rates, payload size of 512 bytes, and superframe length of 2ms</td>
<td></td>
</tr>
<tr>
<td>6.42</td>
<td>Normalized throughput of 802.15.3 under different ACK modes, various data rates, payload size of 512 bytes, and superframe length of 2ms</td>
<td></td>
</tr>
</tbody>
</table>
6.43 Normalized throughput of 802.15.3 under all ACK modes, various data rates, payload size of 512 bytes, and various superframe length 

6.44 Normalized throughput of different systems under various data rates, payload size of 512 bytes. WiGEE in half duplex mode with 4 MTs; A-MPDU and A-MSDU of 802.11n have TXOP=10ms, and frame size of 23040 and 8191 bytes respectively; all other cases but the DCF access mode of 802.11n have superframe length or TXOP equal to 2ms.
Chapter 1

Introduction

The emergence of fiber optic technologies successfully pushes the bandwidth bottleneck problem from backbones to access networks. The FTTx\(^1\) technologies may eventually solve the access problems in the wired world, nevertheless, towards the wireless world, the gap is even enlarged. IEEE 802.11, a set of standards for wireless local area networks (WLAN), although successful in access networks due to flexibility and low cost, provides only limited effective bandwidth. For 802.11b the achievable data rate per cell is around 7 megabit/second (Mbps) and for 802.11a and -g around 25 Mbps. As seen from the Table 1.1, the capacities of current WLANs are far behind the bandwidth requirements of current and emerging bandwidth intensive applications. To support fully a variety of applications communicating over a relative short range, e.g. in indoor environments, a gigabit WLAN with a data rate up to one gigabit/second (Gbps) is desirable. The evolution of WLAN is shown in Fig. 1.1. The activities in frequency regulation, as well as the advances in frontend design and signal process, are all paving the way for gigabit WLANs. However, for gigabit WLANs to be practical, lot of researches have to be done. Two

\(^1\)FTTx refers to several different optical fiber architectures including: Fiber to the node/neighborhood (FTTN), Fiber to the exchange (FTTEx), Fiber to the cabinet (FTTCab), Fiber to the curb (FTTC), Fiber to the building (FTTB), Fiber to the home (FTTH) and Fiber to the premises (FTTP).
Table 1.1: Bandwidth requirements of various applications

<table>
<thead>
<tr>
<th>Application</th>
<th>Data rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>E-mail transmissions</td>
<td>2.4 to 9.6 kilobit/second (Kbps) or higher</td>
</tr>
<tr>
<td>Digitized voice phone call</td>
<td>64 Kbps</td>
</tr>
<tr>
<td>Digital audio</td>
<td>1 to 2 Mbps</td>
</tr>
<tr>
<td>Compressed video</td>
<td>2 to 10 Mbps</td>
</tr>
<tr>
<td>Medical transmissions</td>
<td>Up to 50 Mbps</td>
</tr>
<tr>
<td>Scientific imaging</td>
<td>Up to 1 Gbps</td>
</tr>
<tr>
<td>Full-motion video</td>
<td>1 to 2 Gbps</td>
</tr>
</tbody>
</table>

The main issues among them are: finding an applicable architecture that is able to support efficient communication at the data rate of multiple Gbps, and designing an efficient MAC protocol capable of fully exploring the capacity of the system. It is the objective of the thesis to investigate these issues.

Gigabit wireless systems have attracted extensive research attention in recent years. Several commercial products operated on the 60GHz band have already appeared on the market since 2003 [1]. While these products prove the maturity of wireless techniques to provide Gbps data transmission, to the best of our knowledge, gigabit WLANs are still in the early

Figure 1.1: Evolution of WLAN
Several approaches attempt to achieve gigabit wireless systems. One is the upgrade of the current IEEE 802.11 based systems. For instance, Extracom announced an interference free solution to extend the capacity of IEEE 802.11a and -g to 1Gbps and IEEE 802.11b to 99Mbps. The key of the solution is the tight channel re-use enabled by adaptive techniques. In December 2004, Siemens AG, IAF GmbH and HHI Heinrich-Hertz-Institute demonstrated 1Gbps data transmission using 5GHz band, which was achieved by the combination of multiple-input multiple-output antenna systems (MIMO) and orthogonal frequency division multiplexing (OFDM) modulation. Nevertheless, the technical and regulatory constraints in 2.4 and 5GHz band as well as the cost factor make them less attractive.

The ultra-wideband (UWB) technology is a promising technology to achieve multi-gigabit data rate for short range communications [2]. At such a data rate, however, the power emission constraint imposed by the regulation, i.e. the well known Federal Communications Commission (FCC) spectrum mask [3], makes it only suitable for wireless person area networks (WPAN).

The future of gigabit WLANs may exist in the millimeter wave (MM-wave) band. In 2001, the FCC set aside a continuous block of 7 GHz spectra between 57 and 64 GHz for license free wireless communications. Japan has released 59-66 GHz band for the same purpose. In Europe, the concern of allocating 54-66 GHz for the unlicensed usage has been shown by European Conference of Postal and Telecommunications Administrations (CEPT). The regulation of the 60 GHz band around the world can be found in Fig. 1.2.

Several characteristics make the 60 GHz band an ideal candidate for gigabit WLANs:

- Its 7 GHz spectra make the support of a data rate up to 10 Gbps
possible.

- Its low penetration property to most building materials confines the cell size to building rooms, which significantly improves the spectrum efficiency and security.

- Its low diffraction property makes it highly immune to co-channel interference.

- The high attenuation of 10-15dB/km due to oxygen absorption makes it very suitable for the indoor usage.

A variety of MM-wave WLAN prototype systems have been developed so far. Communications Research Laboratory in Japan has proposed two prototypes since 1998 [4][5]. In the first prototype an ATM based WLAN was developed to support multimedia transmission at a data rate of 51.84Mbps [4]. The second prototype is an IP based WLAN with a higher data rate of 64Mbps [5]. The IST project BROADWAY in Europe [6] proposed a hybrid frequency system which integrated HIPERLAN/2 operated at 5 GHz with an ad-hoc extension operated at 60 GHz. With the 60 GHz extension,
the system can support a data rate up to 100 Mbps. However, few research efforts has been invested on the MM-wave based gigabit WLAN for several reasons:

- The demand for high speed wireless systems, which is driven by emerging bandwidth intensive applications such as HDTV and multimedia applications, just appeared in last few years.

- A gigabit WLAN requires an appropriated frequency band with sufficient spectra to support Gbps transmission. However, the UWB and 60 GHz unlicensed bands have just been released by regulation bodies for few years.

- The traditional components to implement MM-wave frontends are III-V semiconductor materials, which are too expensive for massive deployment.

- The advances to solve technique problems in the 60 GHz band, e.g. the power consumption problem, are in progress.

Witnessing the intensive research efforts on low cost SiGe based MM-wave frontends [7] [8] and the standardization activities on the 60 GHz air interface [9], we believe the age of gigabit WLANs is coming. Before gigabit WLANs can be massively deployed, several challenges remain: the choice of a proper frequency band that is able to support adequate capacity, the design of an advanced physical layer that can fight channel impairments, and the development of an efficient MAC protocol that can full explore the capacity of the system. In this thesis, we assume the future gigabit WLANs are based on the 60 GHz band. The focus of the thesis is then concentrated on the MAC layer issues of gigabit WLANs.

Until now MAC protocols working for gigabit speed systems are only found in wired systems. CSMA/CD, the classical MAC protocol of Ethernet, is
used in half-duplex mode Gigabit Ethernet with a slot time extension [10]. However, to date, none of the vendors have plans to develop equipment based on half-duplex Gigabit Ethernet operation because of its low performance. CSMA/CD is also not suitable for the wireless link because in the wireless channel the signal strength sensed by the transmitter is not reliable enough to decide a collision considering the factors such as the receiving antenna position and inevitable interference in free space. Multi-point control protocol (MPCP) specified by IEEE 802.3ah [11] is another MAC protocol used in gigabit speed systems such as EPON. It is a centralized MAC protocol that uses a polling scheme for the uplink channel scheduling. The dynamic bandwidth allocation (DBA) algorithm in the MPCP provides system flexible on resource utilization and quality of service (QoS) support.

In WLAN, CSMA/CA and its variants are widely used MAC protocols. The CSMA/CA is a contention based MAC protocol used in IEEE 802.11 networks. It has been shown that the collision and contention overheads in CSMA/CA protocols significantly limit the system performance [12]. Therefore it is not a good candidate for gigabit WLANs. In addition to contention based MAC protocols, reservation based and hybrid protocols have been developed for different systems. The survey of those MAC protocols can be found in [13]. The reservation based and hybrid MAC protocols outperform the contention based protocols on the QoS support and protocol efficiency. According to the evolution of the WLAN MAC protocols, more efforts are put on these two aspects. Naturally, they are two important factors affecting the MAC protocol design of gigabit WLANs.

The main contribution of the thesis consists of three parts. In this first part, it investigates issues in gigabit WLANs, identifies the MAC candidates of gigabit WLAN from current WLAN, WPAN, and wireless metropolitan area network (WMAN) according to their performance under high data
rates [14]. Then in the second part, a novel system that is dedicated for indoor gigabit WLANs is proposed [15], which is a hybrid optical/wireless system that employs a wavelength division multiplexing/time division multiplexing passive optical network (WDM/TDM-PON) as the infrastructure and a 60 GHz radio for wireless access. A MAC protocol suitable for this system is developed and the performance is analyzed. The issues of channel access, error control, bandwidth allocation, and scheduling [16] are studied and the solutions are provided. At the final part of the thesis, the throughput performances of MAC candidates are extensively studied.

In addition to the traditional approaches to implement gigabit WLAN, the cognitive radio (CR) technologies [17] is also thinkable for gigabit WLANs. The basic idea of CR is to use the spectrum resources efficiently via implementing smart transceivers. It involves the cooperation of PHY, MAC and upper layers. In a typical CR scenario, the transmitter is able to sense the frequency bands, and choose a proper band for transmission. For coexistence, connectivity, and better QoS, it makes sense to use CR techniques in the context of MM-wave based gigabit WLAN systems. For instance, if a system can operate at the UWB band and 60 GHz band, it can use 60 GHz for high speed transmission, and UWB band to keep the connection alive when the blockage at 60 GHz is detected. Our initial efforts on this topic are reported in [18], [19], and [20].

1.1 Structure of the Thesis

Chapter 2 identifies the characteristics of MM-wave bands that affect the MAC design, provides an overview of current MM-wave multi-access systems, investigates the MAC issues in MM-wave based WLAN/WPAN, and briefly introduces the MAC protocols of current WLAN, WPAN and WMAN, which can be candidate protocols for a MM-wave based gigabit
WLAN.
Since the MM-wave band has some unique features that demand a new system design for gigabit multi-access, in chapter 3 we introduce a hybrid optical/wireless system, named wireless gigabit Ethernet extension (WiGEE), as a candidate for MM-wave based gigabit WLANs. It takes advantage of the fiber optic infrastructure, 60 GHz radio, and centralized control scheme to meet the requirements of MM-wave based gigabit WLANs. The architecture of WiGEE is introduced first. Then the scalability of the fiber optic infrastructure in terms of the supported base station (BS) numbers is analyzed. Next, we focus on the issues of link and MAC layer. A MAC protocol suitable for the system architecture is proposed. The issues in the MAC protocol are identified and solutions are provided, which include the node discovery, bandwidth allocation, error control, and handover.

Chapter 4 derives the 60 GHz error models for the performance study of proposed MAC protocol. The motivation to derive the error models is described first. The state of the art of digital analytical error models is briefly introduced. Then, we focus the study on Gilbert-Elliot model and semi-Markov models. Based on the trace files obtained from the 60 GHz transceiver proposed in [21], the parameters of those models are extracted. The suitability of the models is analyzed based on the cumulative distribution functions of error models.

As the integral part of the MAC protocol, the bandwidth allocation and downstream scheduling issues of WiGEE are discussed in chapter 5. First, a DBA algorithm with the support on multiple wavelengths and different duplex mode of MTs is proposed to offer flexible upstream bandwidth allocation for WiGEE. Then a channel aware downstream scheduling scheme tailored for WiGEE is introduced, in which a long term channel state met-

\[^2\text{Unless state otherwise, the wavelength in this thesis denotes the optical wavelength.}\]
ric obtained from the error control scheme of the link layer is developed to exploit the multiuser diversity gain in the system. The performance of the scheduling algorithms is studied by simulation.

Chapter 6 studies the performance of MAC protocol candidates in terms of maximum throughput. The quantitative analysis results show the efficiency of candidates operating in different data rates and frame sizes. The main issues to utilize them in gigabit WLANs are identified.

Finally, the conclusion is drawn in chapter 7, where the future work is provided.
Chapter 2

Gigabit Wireless MAC Candidates

The MAC protocol is an essential part of every WLAN/WPANs. The function of MAC is to coordinate the channel access among competing nodes of a network in an orderly and efficient manner. It determines the behavior and performance of a network. On one hand, the MAC protocol of a gigabit WLAN should match the application scenario and satisfies general design rules such as efficiency, fairness, flexibility, and scalability. On the other hand, it should take into account the unique characteristics of operating frequency bands into account. In this chapter, the MM-wave band is assumed as the frequency band of the choice for Gigabit WLANs.

MM-wave bands have their own characteristics other than lower frequency bands. While signals at lower frequency bands can easily traverse through buildings, MM-wave signals do not penetrate solid materials very well. Consequently, MM-waves permit a dense packing of communication frequencies, thus providing very efficient spectrum utilization, and increasing security of transmissions. Moreover, like light waves, MM-wave signals result in low diffraction, but are subject to more shadowing and reflection. For Non-Line-of-Sight (NLOS) propagation, the greatest contribution at the receiver is the reflected power. Shorter wavelengths cause the reflecting material to appear relatively rougher, which results in greater diffusion.
of the signal and less direct reflection. Since diffusion provides less power at the receiver than directly reflected power, MM-wave systems usually rely on LOS communication condition. Directional antennas are normally required in these systems to achieve a reliable communication.

Apart from this, MM-wave systems share common features with other wireless systems. These include: 1) high error rate and bursty errors; 2) location-dependent and time-varying wireless link capacity; 3) half duplex communication; 4) user mobility; 5) power constraints of the mobile users. All of the above characteristics challenge the development of effective and efficient MAC protocols.

2.1 Classification of Wireless MAC

Wireless MAC protocols can be broadly classified into two categories: distributed and centralized protocols. In distributed protocols, competing nodes in the network contend for the medium access without any centralized coordination. On the other hand, in centralized protocols, there is a special node responsible for channel allocation. That node can be a dedicated node like BS, or a normal node elected from a group of nodes, like the Piconet Coordinator (PNC) in Bluetooth.

Based on the method of operation, the MAC protocols are further divided into three access modes: random access, guaranteed access, and hybrid access (see Fig.2.1). In the random access mode, all nodes contend for the medium access. When several nodes simultaneously access the channel, a collision occurs and leads to the failure of the transmission. A contention resolution algorithm (CRA) is adopted in the random access mode to resolve collisions. Three families of CRAs are widely used in wireless MAC protocols: the binary exponential backoff (BEB), the p-persistence backoff, and splitting algorithms like n-ary tree [22].
In the guaranteed access mode, competing nodes access the medium in an orderly manner, usually in a round robin fashion. There are two approaches to implement the guarantee access mode. One uses master-slave configuration, in which the master polls slaves for channel access. The other is to exchange token in a distributed way. Only the node holding the token is allowed to transmit. The token is passed to the next node after finishing the transmission. Because of its implementation complexity (measures have to be taken to re-create exactly one token after it was lost) it is rarely used in wireless systems.

The third access mode blends the benefits of above two modes. Most hybrid access protocols are based on request-grant mechanisms, i.e. each node sends a request to the BS indicating the bandwidth it requires on a contention basis. The BS then allocates upstream time slot for the actual data transmission and sends a grant to the requesting node indicating the allocation.

2.2 State of the Art in MM-wave Based Multi-access Systems

Among current wireless multi-access systems on the market few are operated at MM-wave bands. There are several reasons for this situation. Firstly, the front-end and baseband signal processing technologies capable
of operating at MM-wave bands are not mature enough for mass deployment. The III-V radio frequency (RF) components, which have typically been used for MM-wave front-end in the past, are too expensive for portable devices. Secondly, the demand for bandwidth is tightly related to applications. Truly bandwidth intensive applications such as HDTV emerged only in recent years. As long as legacy wireless systems are capable of providing adequate capacity for conventional applications, there is no a driving force to move to MM-wave bands. In the following, we briefly introduce the state of the art of MM-wave based multi-access systems.

LMDS (Local Multipoint Distribution System) is a fixed broadband wireless access system operating at the 28 GHz band in the United States and the 40 GHz band in Europe. Achievable data rates depend on distance and modulation format; typical figures are 40 Megabits/second (Mbps) for the downlink and 10 Mbps for the uplink for links of few kilometers. There are two specifications for LMDS: the Digital Video Broadcasting (DVB) specification from European broadcasting union, and the Digital Audio Video Council (DAVIC) specification from DAVIC, which is an international body formed by major network operators, service providers, and industry vendors. The DVB specification only focuses on broadcast services. The DAVIC defines an MAC protocol for uplink channel access. The DAVIC MAC is similar to 802.14 [22] and DOCSIS MAC. The latter two protocols are used in hybrid fiber coaxial systems. We can find LMDS, and 802.16 have a strong relationship with the DOCSIS standard. The DAVIC MAC uses the request and grant procedure for bandwidth allocation. The contention slots are used for registration, bandwidth request and short message exchange. The comparison of DAVIC, 802.14, and DOCSIS MAC can be found in [23].

In 2002, IEEE published 802.16.1 to provide high speed wireless Internet access between buildings with exterior antennas. It operates at 10 to
60 GHz bands. The MAC of 802.16.1 is based on a point to multi-point (PMP) topology. The uplink and downlink of the channel are structured into frames and the channel access is TDMA based. The BS governs the bandwidth allocation of both channels. The fine granularity QoS is guaranteed through a connection oriented MAC protocol. Due to the LOS propagation required in MM-wave bands, the 802.16.1 system has a coverage issue in urban areas. The IEEE then worked out a new PHY interface operating at 2-11 GHz. This new amendment is referred to as 802.16a. The current research on MM-wave bands is heavily focused on the 60 GHz band, with a goal to support multiple Gbps. The IEEE 802.15.3c working group is working on a 60 GHz PHY alternative for 802.15.3 WPAN. It is expected that the standardizing process will be completed in 2008. The objective of the 802.15.3c is to support bandwidth-intensive applications, and consumer device interconnection in WPAN. It is expected that its MAC will be developed based on 802.15.3 MAC. Due to the characteristics of the 60 GHz band, there is a need to support directional antenna or beamforming in its MAC protocol.

Several research projects have targeted on 60 GHz multi-access systems. The Europe IST BROADWAY project proposed the use of 60 GHz band as an ad-hoc extension for the HiperLAN/2 system [6]. It is a 5 GHz and 60 GHz hybrid system, where the 60 GHz sub-system provides a data rate of at least 100 Mbps for peer to peer connection. The MAC protocol is based on HiperLAN/2 but with modification to accommodate the 60 GHz ad-hoc extension. Currently, the Wireless Gigabit With Advanced Multimedia (WIGWAM) project has been setup up by 27 partners in Europe and coordinated by TU Dresden. It is aimed at designing a 1 Gbps system concept for home/office, public access and high velocity scenarios. A distributed MAC is used in home/office scenarios, and a centralized MAC is used in public access scenarios. The frequency bands to be used include
5 GHz, 17 GHz, 24 GHz, 38 GHz, and 60 GHz.

As we can see, currently there is no widely used MM-wave WLAN/WPAN system. LMDS and 802.16.1 are established products, but no multi-access systems, as there is one channel exclusively assigned to each node. All other systems are prototypes that arose from research projects. In the following section we will look into reasons for this situation.

2.3 Challenges of Gigabit Wireless MAC Design

According to the nature of multiple access, and the characteristics of MM-wave bands, the challenges of MM-wave based MAC exist in four aspects: physical constraints, resource allocation, QoS provisioning, and handover issues.

2.3.1 Physical Constraints

The Friis equation [24] indicates that the path loss of a radio signal is proportional to the square of its carrier frequency. For instance, with equal antenna gain, 60 GHz band has additional 21dB of path loss compared to 5 GHz band. As path loss combines with other channel impairments such as delay spread, there is a need to employ directional antennas in MM-wave systems in order to achieve reliable communications. A highly directive Horn antenna as used in [25] compensates for this loss entirely, but at the price of an opening angle of 7 degrees, thus basically prohibiting mobility. To enable user mobility, the beamforming process through the use of MIMO antennas seems to be a required process in MM-wave multi-access systems. The beamforming is achieved by adaptively adjusting the weight of each antenna branch in an antenna array. A typical beamforming process works as follows: the initiator sends a steering request to the responder, including a training sequence. The responder estimates the channel state and sends
back to the initiator the channel state information or computed steering matrix. Thereafter, the initiator uses the feedback to steering the beam of the MIMO antennas accordingly. The use of directional antenna brings several challenges on the MAC protocol:

1. Carrier Sensing

Most MAC protocols assume the use of the omni-directional antenna at the PHY layer for broadcasting control messages. If a directional antenna is used either at the transmitter or at the receiver, a new mechanism at MAC layer is required to detect neighbor nodes. As the data rate of a system increasing, the overhead of the MAC protocol becomes the bottleneck to the system throughput. As reported by Xiao [12], the upper bound throughput of 802.11a is only 75.24 Mbps when the data rate of the system becomes infinite. The MAC overhead in 802.11a includes a contention period, guard intervals between transmission operation, acknowledgement (ACK) process, and frame header for each packet. The most important parameter for carrier sensing based protocols is the Rx/Tx turnaround time, which is physically determined by the time to power up and down the antenna. This means that the number of changes between listening to the channel and transmitting for a given number of bits should be minimized. The block transmission and block acknowledgement are common approaches to reduce the MAC overhead, in which a transmitter is allowed to continuously transmit a number of packets with a single ACK message. To further reduce the overhead, the 802.11n introduces a concept of traffic aggregate, i.e. a number of packets from the upper layer can be aggregated into a single frame at the MAC layer. The guard interval and duplicated frame headers are therefore reduced. It is certain that large packets can improve the system throughput. However, there is a tradeoff between the packet size and channel condition.
The loss of large packets on the other hand degrades the performance. The challenge on this issue is to determine the optimum frame size according to the channel coherence time and collision rate, or adapt the maximum frame size to channel states.

2. Hidden Terminals

As shown in Fig.2.2A, the use of directional antennas creates a new hidden terminal problem. In contrast to omnidirectional transmission where hidden terminals are somewhat of an exception, with directed antennas the ongoing transmissions in the vicinity of a node are almost never detectable, as the narrow beam of the antenna is purposely suppressing signals from the side. This means that if there is a collision of two transmitters inside the receiver's reception beam/angle, it will be of the hidden terminal type. The exchange of RTS/CTS (Request-To-Send/Clear-To-Send) packets prior to transmission should help here.

3. Spatial Reuse

Directional antennas increase spatial reuse by means of narrowing the beamwidth. However, as shown in Fig.2.2B, with the same transmission power, the directional antennas may have a longer transmission range, which leads to a specific kind of exposed terminal problem (in
that the receivers and transmitters are exchanged compared to the usual sketch of the problem). Here, a potential transmitter node TX1 would hear an RTS from TX2, but not the CTS from RX2, as this is being directed towards TX2. According to the normal procedure, a node hearing an RTS but no CTS may conclude that its own transmission does not interfere with another ongoing transmission. Still, if immediate acknowledgements are to be sent out these may collide with ongoing data transmissions from the respective exposed terminal. Power control support at the MAC layer can alleviate this problem.

### 2.3.2 Resource Allocation

Main resources to be allocated in a wireless network include bandwidth, power and space. They are allocated through bandwidth allocation, rate control, and power control algorithms. Bandwidth allocation algorithms share bandwidth among multiple users. Rate control adapts bitrate to channel condition and the queue state of the nodes. Power control aims to prolong battery life, control interference level, and increase spatial reuse. Moreover, handover allows nodes to change the cell and use the shorter transmission links, which leads to a reduced area that is covered by the transmission and thus to an increased overall capacity.

**Bandwidth Allocation**  The challenge of a MM-wave system is to allocate the bandwidth according to the channel condition. The channel state dependent (CSD) scheduling exploits the time varying channel to achieve better system performance. It is highly suitable for MM-wave system since a channel in a MM-wave band experiences more fluctuation than a low frequency band due to fading and shadowing effects. Different CSD scheduling approaches are reported in [26], [27]. The essential idea of CSD scheduling is to schedule the node with good channel quality and defer the
one with bad channel quality; therefore the overall system performance improves. The challenges in CSD scheduling include efficient detection of the channel condition, and the fairness in scheduling algorithms.

**Power Control** The power consumption is a critical issue for MM-wave systems due to their high carrier frequency. A typical approach for energy saving is to shift the node into the sleep mode when it has no data to send \[28\], \[29\]. The sleeping node wakes up periodically to receive the data from access point (AP) or other nodes. Other approaches use power control scheme during transmission \[30\], \[31\]. In case of using MIMO, which is highly likely in MM-wave systems, a transceiver can further save the power by switching other antenna chains off and leaving only one antenna chain to monitor the network. The power saving mode needs the support of the MAC protocol.

The dilemma of power control is the following: When a transmitter observes there is high interference in the channel, it recognizes more power is needed to make a successful transmission. Therefore it would be better to back off, buffer the traffic and wait for the interference to lessen. However, the new traffic may arrive during the back off to fill up the buffer and delay raises, which pushes the transmitter to become power-aggressive in order to reduce its backlog. This problem has been studied by Bambos under the framework of dynamic programming \[32\].

Power control is also helpful in alleviating the hidden terminal problem in ad hoc networks \[33\], \[34\]. As seen from Fig.2.2.B, the transmission power can be effectively controlled so as to control the transmission range of a node. With channel knowledge from neighbor nodes, a hidden node can adjust its transmitting power in a way that the on-going transmission will not be disrupted. A joint optimization process to maximize the battery life, make the network stable, and improve the system performance and
2.3. CHALLENGES OF GIGABIT WIRELESS MAC DESIGN

capacity is a highly challenging task.

**Rate Control** The rate control is discussed on two levels. One is to control the data rate by means of different modulation and coding schemes (MCS). The motivation behind this is that current wireless systems usually support a set of MCS. The adaptive data rate control refers to the dynamic change of the transmission data rate according to the channel condition. The second refers to the adaptive coding rate control at the channel coding level, which adapts the data transmission rate with more or less redundancy to compensate channel fading and interference.

An adaptive data rate control schemes needs the estimation of the channel conditions. It can be done at transmitter or receiver [35], [36], however, the best place is to be done at the receiver. In a receiver initial scheme, the receiver feeds back the channel state or the rate of the choice to the transmitter. The transmitter chooses the appropriate data rate according to the feedback. For instance, Holland [36] proposes a Receiver Based AutoRate (RBAR) protocol based on RTS/CTS based protocols, in which the rate selection is performed on a packet by packet basis during the RTS/CTS exchange, just prior to packet transmission. The use of data rate control MAC scheme is reasonable in high data rate MM-wave systems. It provides the opportunity for the systems to adapt the complex channel conditions.

The power control scheme cannot solve the performance problem when the channel condition degrades to a certain level. The coding rate control is deployed in this case to further guarantee the performance. In a typical adaptive code rate scheme, the channel quality estimation is made at the receiver and sent back to the transmitter through the feedback message. To reduce the transmission overhead, incremental redundancy is usually applied, in which a special channel coding scheme is adopted such that
the high-rate code is the subset of the lower rate code. If the current
code cannot provide sufficient protection for decoding at the receiver, only
the redundant bits, which are different bits between two channel codes,
are transmitted. The example of incremental redundancy code is Rate
Compatible Punctured convolutional (RCPC) code [37]. In RCPC codes,
a high-rate code is generated by puncturing the lowest rate block of codes
bits. Since the encoder only needs to generate the code with lowest coding
rate, one encoder/decoder pair is enough to encode and decode the codes
with all coding rates. The incremental redundancy code scheme is one type
of type II hybrid automatic request repeat (ARQ) scheme.
It is noted that the adaptive data rate and code rate scheme both use the
channel estimation to determine the appropriate rate. It is an interesting
topic to combine two rate control schemes for better system performance.

2.3.3 QoS Provisioning

In wireless networks providing fairness becomes a complex problem. A
wireless link may turn to the error state when a flow is transmitted. To
maximize the overall system throughput, it is better to defer the transmis-
sion until the link changes to the good state. However, there is a tradeoff
between channel utilization and fairness. To ensure fairness, the flow or
node should be compensated for the loss when the link recovers. However,
the definition and objectives of fairness become ambiguous in wireless net-
works. The appropriate interpretation of fairness for wireless networks
depends on the service model, traffic type, and channel characteristics.
For ad-hoc networks, the challenges lie in the following factors:

- The location dependent contention of wireless channel;
- The tradeoff between the channel utilization and fairness;
- The inaccurate state and decentralized control.
The challenges of fairness in MM-wave based systems include the tradeoff between channel utilization and fairness, the guarantee of short term and long term fairness through scheduling algorithms.

Admission control is an essential component in MAC for strict QoS guarantee. It is difficult to do in a distributed MAC since there is no global information available. The research efforts based on measurement have been devoted to these issues [38], [39]. However, proposed distributed admission control schemes only address coarse QoS. The remaining open issues of distributed admission control include [40]:

- The support on heterogeneous traffic;
- The joint consideration of resources at network, link and physical layer;
- The cooperation with channel access methods;
- An efficient information exchange mechanism among nodes.

### 2.3.4 Handover

The use of cell based MM-wave systems makes the handover problem prominent. As aforementioned, the cell of a MM-wave based WLAN is largely confined to a room. The small cell size drastically increases the frequency of handover. As the data rate of the system increases, the overhead of handover becomes a critical performance issue. The fast handover at the MAC layer becomes a challenge in such a system.

Handover is broadly divided into two types: soft and hard handover. In soft handover the call is uninterrupted when the node is moving from one cell to another. It is supported in Universal Mobile Telecommunications System (UMTS). The hard handover tears down the old connection before setting up a new connection. IEEE 802.11F is an example of hard handover. It
is obviously that soft handover provides better performance; however, the internal procedure is more complex since the connection state of the handover node needs to be exchanged between the BSs involving the handover. The handover procedure consists of measurements, decision and execution of the handover. In cellular networks, the measurements are typically done via signal strength detection. The handover decision is made if a node detects the signal strength of the neighbor station outperforms the current one at a predefined level. In 802.11, the handover is done by mobiles. After losing the connection a mobile scans channels and tries to detect a new AP by using active probing or passively listening to beacon signal emitted periodically from APs. Research found the discovery (probe) phase has the dominant contribution in handover latency [41]. The total handover latency in 802.11b was shown to be between 75-350ms, which is far too long to the delay requirement of voice or video traffic (max 50ms). It is clearly that to enable fast handover, a cross layer approach with handover indication from the physical layer is necessary. The fast handover solution is also related to the MAC implementation, i.e. the connection oriented or connectionless.

2.4 MAC Protocol Candidates for Gigabit Wireless LAN

According to the challenges in designing MAC protocols for MM-wave systems, we identify several common features of MM-wave based MAC protocols:

- Be efficient and scalable to bear an ultra high data rate up to multiple Gbps;
- Be flexible to support QoS for a wide range of applications;
• Be capable of supporting directional antenna and advance antenna technologies like beamforming.

After investigating current and under developing WLAN/WPAN systems, we choose the MAC protocols of 802.11n, HiperLAN/2, and 802.15.3 as the candidates. Although the 802.16 is a standard for wireless metropolitan area networks (WMAN), its centralized architecture, operation mechanism, and support on QoS make it a competitive reference model for MM-wave based WLAN/WPAN systems.

Considering the unique characteristics of MM-wave bands, we propose a novel system for 60 GHz indoor WLAN, named wireless gigabit Ethernet extension (WiGEE), and design an MAC protocol that is suitable for pico-cell communications with a large number of cells. The WiGEE and corresponding MAC protocol are introduced in the next chapter.

In the rest of this section, we introduce the MAC protocols of 802.11n, HiperLAN/2, 802.15.3, and 802.16 on their channel access, bandwidth allocation, QoS, and error control scheme. At the end, a qualitative comparison is provided.

2.4.1 802.11n

IEEE 802.11n is the high throughput version of IEEE 802.11 standards, which is aimed at providing a minimum data rate of 100 Mbps, and a peak data rate as high as 600 Mbps. The first draft was approved by IEEE in March 2006. It is expected that the final standard will be published in October 2007. This new amendment boosts the system capacity primarily in three aspects: the use of MIMO to increase the peak data rate; the introduction of double width 40MHz channels for extra throughput; the enhancement of MAC to improve the MAC efficiency. The enhancements of 802.11n MAC can be summarized as follows: traffic aggregate, block
acknowledgement, link adaptation, and protection from legacy 802.11 services. The first two enhancements are aimed at reducing the MAC overhead, which is necessary when the data rate of the system is very high. The link adaptation functions provide 802.11n the capabilities like beamforming.

**Channel Access**

Like in legacy 802.11, two operation modes are defined in 802.11n draft: the infrastructure mode where an AP coordinates the communication between stations; and the ad hoc mode where stations directly communicate to each other without centralized coordination. The ad hoc mode provides the flexibility to form a network, however, the infrastructure mode is more popular since in most application scenarios stations need to access infrastructure networks.

The legacy 802.11 introduced two modes of channel access: Distribution Coordination Function (DCF) for random access and Point Coordination Function (PCF) based on DCF for coordinated access. The period operating in random access mode is called contention period (CP), and the period in coordinated access mode is called contention free period (CFP). The CP and CFP are announced in a special control sub-frame called Beacon. As shown in Fig.2.3, the Beacon, CP and CFP forms a structure called super-frame.

The DCF mode is the basic access mode that all other access modes rely on.

![Figure 2.3: Contention free and contention period in 802.11](image-url)
on. It uses Carrier Sensing Multiple Access/Collision Avoidance (CSMA/CA) mechanism for channel access. Truncated binary exponential backoff (BEB) is employed to resolve collisions.

In addition to DCF, 802.11n uses Hybrid Coordination Function (HCF) to support QoS in channel access. The HCF was first defined in 802.11e standard, which is the QoS amendment of legacy 802.11 standards. 802.11e extends the DCF to Enhanced Distributed Channel Access (EDCA), and the PCF to HCF Controlled Channel Access (HCCA) for better channel access control. Both modes need the participation of the AP to control access parameters on the fly. Therefore they can only operate in the infrastructure mode.

The EDCA enhances the DCF in two aspects: 1) An important concept, the transmission opportunity (TXOP), is introduced to limit the channel access time of each station. A TXOP is a time interval in which the station is permitted to deliver its MAC Service Data Units (MSDU). Once seizing the channel, a station is permitted to transmit as many as frames during its TXOP, but should not exceed the time interval defined by TXOP. Both EDCA and HCCA use TXOP to control channel access time of each station. For EDCA, the duration of TXOP is announced in the beacon. For HCCA, it is specified by the AP and sent in polling messages. The TXOP brings two benefits: The channel access time can be well controlled; the channel access overhead can be reduced. 2) To seize the channel in legacy 802.11, each node has to wait for a time interval called DCF Interframe space (DIFS) after sensing the channel to be idle. In EDCA, a new set of interframe spaces, named Arbitration Interframe Space (AIFS), are introduced to provide channel access priorities for differential services. As shown in Fig.2.4, a shorter AIFS gives a node higher priority to access the channel.

In the HCCA access mode, there is a Hybrid Coordinator (HC) residing
in the AP, fully controlling the HCCA operation. The HC uses polling based scheduling. It polls a station using a Contention Free Poll (CF-Poll) frame, or the combination of data and CF-Poll frame if it has data send to that station. The granted time for channel access is given by a form of TXOP specified in the CF-Poll frame. The polled station transmits its data frames and ACK during the granted TXOP. If the HC receives no response from a polled station after a predefined interval, it polls the next station or ends the CFP by broadcasting a CF-end frame. It is noted that although the EDCA is only permitted in CP, the HCCA can be used in both CP and CFP. In CFP, the HC uses shortest AIFS to seize the channel earlier than normal stations.

Frame Structure

802.11n specifies new frames and fields for high throughput operation, which includes frame structures for traffic aggregate, and control frames for block acknowledgement, MIMO manipulation and link adaptation. We only introduce the traffic aggregate since it can be generalized and used in MM-wave based MAC protocols. Using properly, it can significantly reduce MAC overhead.

The traffic aggregate is the opposite process of the fragmentation. The
purpose of traffic aggregate in 802.11n is to reduce the overhead of frame header and guard interval. Two kinds of traffic aggregate are defined in 802.11n draft: Aggregated MAC Service Data Unit (A-MSDU), which assembles multiple MSDUs into one MAC frame, and Aggregated MAC Protocol Data Unit (A-MPDU), which combines multiple MPDU into one Physical Layer Convergence Protocol (PLCP) frame.

The purpose of A-MSDU is to allow multiple MSDUs sent to a same receiver to be aggregated into a single MPDU. The efficiency of the MAC layer is improved particularly when there are lots of small MSDUs such as TCP acknowledgments. The structure of a Data MPDU containing an A-MSDU, which is a sequence of n MSDU sub-frames, is shown in Fig.2.5. Each sub-frame contains a sub-frame header followed by a MSDU and 0-3 bytes of padding. The sub-frame header consists of three fields: the Destination Address (DA), the Source Address (SA), and the length field which contains the length of the MSDU in bytes. It is noted that only those MSDUs transmitted to the same receiver can be aggregated. An A-MSDU is treated like any other MSDU. The maximum length of A-MSDU is limited to 8000 bytes.

The problem of A-MSDU is that a large frame size may degrade the system performance when the channel is in bad state. A-MPDU is proposed to alleviate this problem. The structure of A-MPDU is shown in Fig.2.6.
An A-MPDU consists of a number of MPDU delimiters each followed by a MPDU. The purpose of the MPDU delimiter is to robustly delimit the MPDUs within the aggregate even one or more MPDU delimiters are collapsed. A MPDU delimiter contains an 8 bits pattern field which provides a unique pattern to delimit the start of each MPDU. MPDUs are put into the aggregate without modification. Hence if one frame out of an aggregate is collapsed, it is possible to successfully decode all remaining frames. The maximum length of an A-MPDU is 65535 bytes.

![Figure 2.6: A-MPDU in 802.11n](image)

**Bandwidth Allocation and QoS**

As aforementioned, the 802.11 MAC is essentially a contention based protocol. However, 802.11e and 802.11n provide a centralized bandwidth allocation scheme in their controlled channel access mode that provides fine granularity QoS. In 802.11n, the HCCA mode is dedicated for centralized bandwidth allocation. A signaling mechanism is provided to let stations report their bandwidth requests to HC. Two kinds of requests are defined: the queue state indicator which reflects the queue state of station, or the traffic specification which defines the bandwidth and delay requirement of a traffic stream in a station. Requests are sent to HC in two ways: 1) Transmit during CFP period when using HCCA. 2) Contend during CP period when using EDCA. Once the HC admits the request, it issues a CF-Poll frame to the requesting station with channel access time and duration defined in a TXOP field.
In addition to support guaranteed QoS in the HCCA mode, 802.11n provides differential services in the EDCA mode. Four types of access categories (AC) are introduced and mapped to voice, video, best effort, and background traffic, respectively. The concept of AC was introduced in 802.11e. As shown in Fig.2.4, each AC has its own backoff entities and uses AIFS for channel access. The AIFS provides channel access priorities for each AC.

**Error Control Mechanism**

The legacy 802.11 MAC requires an explicit ACK for each unicast data frame. More flexible acknowledgement policies are provided in 802.11n to support differential services and improve throughput. 802.11n allows for two acknowledgement options: Normal Ack, and No Ack. The new No Ack option allows a recipient take no action upon receipt the frame. In addition, a block acknowledgement (BA) is introduced to further reduce transmission overhead. The block acknowledgement process is shown is Fig.2.7. Two kinds of BAs are used: N-Immediate BA that acknowledges the block of data right after the transmission, and N-delayed BA that delays the BA ACK to next TXOP.

![Figure 2.7: Block Acknowledgment in 802.11n](image-url)
2.4.2 HiperLAN/2

HiperLAN/2 is a WLAN standard developed by European Telecommunications Standards Institute (ETSI). Although it is technically superior to 802.11 standards in a variety of aspects such as throughput, and QoS support, it lost the competition with 802.11 in the marketplace. However, the design principles of HiperLAN/2 provide us valuable insights for the MAC protocol design of MM-wave systems.

HiperLAN/2 uses a TDMA based connection oriented MAC protocol. Similar to 802.11, it provides two operation modes: cellular access network configuration for business scenarios, and ad hoc LAN configuration for residential scenarios. The former uses AP to provide access over a certain area. An AP interconnects all mobile terminals (MT) associated with it, and all communication goes through the AP. APs are interconnected by external networks such as Ethernet to extend coverage. Ad hoc LAN configuration is operated in an ad hoc manner where MTs directly exchange data with each other. However, unlike the ad hoc mode in 802.11, a node is dynamically elected among MTs to provide control functions within its subnet. This special MT is called central controller (CC). Accordingly, HiperLAN/2 uses two access modes: centralized or direct mode. In the centralized mode all traffic goes through the AP. In the direct mode the traffic between MTs is exchanged directly without passing through the CC, while channel access still controlled by the CC.

HiperLAN/2 defines three layers in its protocol stack, from upper to lower listed as follows: the convergence layer to provide an interface between upper layer and data link layer; the data link control layer (DLC); and the physical layer. The DLC layer is the target of our interest. It consists of three main entities: the Radio Link Control (RLC), MAC, and Error Control (EC). The functions of DLC are divided into two categories:
user data plane functions are responsible for delivering data between upper layer and physical layer, and providing the error control mechanism; control plane functions reside in the RLC, providing the Radio Resource Control (RRC), Association Control Function (ACF), and DLC Connection Control (DCC). The RRC is responsible for detecting and efficiently utilizing available radio resources. It manages handover, frequency selection, power saving and power control functions. The ACF manages the association process for association, disassociation, authentication, key management and encryption. The DCC takes charge of connection setup, maintenance, and release. As we can see, the HiperLAN/2 uses a cross layer approach between PHY and DLC layer for efficiently utilizing radio resources.

Frame Structure

Instead of using CSMA/CA model, HiperLAN/2 MAC uses a form of dynamic TDD and dynamic TDMA for channel access. The channel is structured into frames. A unique feature of HiperLAN/2 is that the duration of each frame is fixed to 2ms. Since the length of a frame header is fixed, as the data rate increases, a frame can accommodate more data payload and provide more throughputs. Xiao [12] showed that the MAC protocol of HiperLAN/2 outperforms that of 802.11 in terms of throughput. HiperLAN/2 introduces the concepts of logical channel and transport channel in DLC. They are used to classify control and user data messages originated from DLC or higher layer. Control and data messages are first mapped to proper logical channels based on their contents, and then passed to corresponding transport channels, where MAC frames are constructed and appropriate channel access is assigned. The logical channels are identified by message types. The transport channels are identified by message formats and channel access modes. The standard defines 10 logical channels and 6 transport channels.
As shown in Fig. 2.8, each frame starts with a Broadcast Channel (BCH) duration, followed by Frame Control Channel (FCH) duration, Access Feedback Channel (ACH) duration, data phase, and at least one Random Access Channel (RCH) duration. Let us define the transmission direction from MT to AP/CC as the uplink and the transmission direction from AP/CC to MT as the downlink. The BCH transport channel is used in downlink direction to carry the information about the entire cell, such as network identity (ID), AP ID, and transmission power for AP. The FCH is used in the downlink direction to carry the information describing the structure of the frame. The information specifies the transmission time slots, and the PHY mode used in each time slot. The RCH is used by MTs to send control information to AP/CC on a contention basis. A data phase is divided into several sub-phases. If transmissions occur between AP/CC and MTs, a data phase includes the Downlink (DL) phase and/or Uplink (UL) phase. If transmissions occur between MTs in the direct mode, a data phase consists of several Direct Link (DiL) phases.

**Figure 2.8: HiperLAN/2 frame structure**

Channel Access, Bandwidth Allocation, and QoS

HiperLAN/2 uses a request and grant channel access scheme, which is controlled by the AP/CC. A MT having pending data to send requests the bandwidth from AP/CC using the bandwidth allocation mechanism defined in the standard. The bandwidth allocation is performed per connection basis. Each connection is identified by the DLC User Connection.
ID (DUC ID). In the centralized mode, the DUC ID is comprised by DLC connection ID (DLCC ID), and MAC ID, which is assigned by the RLC of the AP during the association process. In the direct mode, the DUC ID is comprised by source MAC ID, destination MAC ID, and DLCC ID. The DUC ID plays as a tag for classification and enables per-flow QoS. The Resource Request (RR) message and Resource Grant (RG) message are used by MTs for bandwidth request and by AP/CC for bandwidth allocation on a connection basis, respectively. The RR contains the queue status of the requesting MT and is initiated by a MT in two ways: a response to a polling issued by AP/CC, or a standalone request sent in special transport channels. The RG is carried in the FCH transport channel, which specifies the location of the transmission/reception in the frame. HiperLAN/2 is capable of providing strict QoS guarantee due to its TDD/TDMA based channel access mode and centralized bandwidth allocation scheme. It is one of merits that outperform 802.11 MAC protocols.

**Error Control Mechanism**

The EC in the DLC layer manages the error detection and retransmission. It ensures in order delivery of packets. The selective repeat ARQ (SR-ARQ) mechanism is employed in the standard. Three modes are specified for different levels of transmission reliability: the acknowledge mode in which the EC retransmits acknowledgements from the receiver; the repetition mode in which the EC simply repeats the transmission without acknowledgement; and the unacknowledged mode for unreliable, low latency services without the need of retransmission. Unicast data are sent using either acknowledge or unacknowledged mode; Broadcast data are transmitted using either repetition or unacknowledged mode; Multicast data can be sent using unacknowledged mode.
2.4.3 802.15.3

The IEEE 802.15 working group focuses on developing standards for WPAN. A set of 802.15 standards have been developed so far. IEEE 802.15.1 adopts Bluetooth MAC and physical layer specifications for data and audio communications among portable devices. It operates at 2.4 GHz frequency band with a data rate up to 1 Mbps. IEEE 802.15.3 draft is aimed at providing high data rates in WPAN. The initial version provides data rates from 11 to 55 Mbps by means of a single carrier scheme. An amendment of 802.15.3, known as 802.15.3a, is on its process to develop an Ultra-wideband (UWB) (3.1-10.6 GHz) based PHY interface with a data rate up to 100 Mbps within 10 meter range and 400 Mbps within 5 meter range. Moreover, the IEEE 802.15.3c working group is working on a new PHY interface based on the 60 GHz unlicensed band. The initial objective of 802.15.3c is to provide a data rate up to 2 Gbps. The applications of WPANs include audio, digital video, HDTV, media rich interactive games, and data applications running on portable devices, laptops, home theater, and other consume electronics. Although air interfaces are different, MAC protocols of 802.15.3 standards are similar.

Architecture

The basic topology in 802.15.3 is piconet, a concept derived from Bluetooth. As shown in Fig.2.9, a piconet is formed in an ad hoc manner, and controlled by a Piconet Coordinator (PNC), who is a node dynamically elected among member devices (DEV). As its name suggesting, the piconet is confined to a small area typically covering a range from 10 meters to 70 meters. This is the range a MM-wave system usually covers in indoor environment. A device is allowed to join and leave the piconet with short association time. Once the piconet is formed, the PNC issues control mes-
sages in beacons periodically, through which the PNC maintains network synchronization, controls channel access, manages QoS provisioning, and performs admission control. To join a piconet, a device needs to associate with the PNC and follow the channel access information provided in beacons. A PNC can transfer its PNC role to other DEV when it leaves the piconet or needs to change its role.

To extend the coverage, increase the number of DEVs to be supported, or distribute control functions among DEVs, the concept of child piconet is introduced. The child piconet has the same topology as the parent piconet. The PNC of the child piconet fully controls its own piconet. However, the PNC of a child piconet is also the member of the parent piconet, and the parent PNC allocates time slots for the channel access inside the child piconet. DEVs in a child piconet are only allowed to transmit within the time slots assigned by the parent PNC.

In order to share frequency resources among different piconets, neighbor piconets are used to provide coexistence. Similar to a child piconet, a neighbor piconet associates with a parent piconet, and the channel access time of the neighbor piconet is allocated by the parent PNC. But unlike a child piconet, the PNC of a neighbor piconet is not a real member of...
the parent PNC. It is only permitted to exchange certain commands with the parent PNC, such as association/disassociation request, channel time request, and authentication. The topology of piconet, child piconet and neighbor piconet is shown in Fig.2.9.

Frame Structure

The 802.15.3 MAC is based on a time-slotted superframe structure that consists of three phases: beacon, Contention Access Period (CAP), and Channel Time Allocation Period (CTAP). The structure of a superframe is shown in Fig.2.10. At the beginning of each superframe, the PNC broadcasts a beacon, specifying time synchronization, control, resource allocation information of the piconet. It follows a CAP which is used to exchange small amount of data and commands such as bandwidth requests on a contention basis. The CTAP is a contention free period which consists of two types of time slots: management time slot (MTS) reserved for command exchange between PNC and DEVs, and guaranteed time slot (GTS) used for data transmission among DEVs. The length of each CAP and CTAP is specified in the beacon.

![Superframe structure in 802.15.3](image)

Figure 2.10: Superframe structure in 802.15.3

Channel Access

The channel access at CAP is on a contention basis. The collision is solved by a CSMA/CA mechanism similar to the one used in 802.11 MAC. Like 802.11, interframe spaces (IFS) are used to guarantee the protocol operation. The actual values of IFSs are PHY layer dependent. The traffic
allowed to be transmitted in a CAP period includes asynchronous data traffic and control commands. The isochronous traffic is only allowed in the CTAP.

The CTAP is based on TDMA in which the PNC specifies starting time and duration of time slot for each traffic stream. The PNC divides the CTAP into a number of channel time allocations (CTA). A CTA is specified by a CTA information element (IE) in the beacon. The CTA information includes the source and destination address of the DEV, the stream index that identifies the connection, and starting time and duration of reserved time slots. There are two types of time slots: GTS and MTS. GTS is used for asynchronous and isochronous traffic. Two kinds of GTSs are defined: dynamic GTS and pseudostatic GTS. The PNC can dynamically change the location of the dynamic GTS within the superframe on a superframe by superframe basis. Pseudostatic GTSs, which are only allocated to isochronous traffic, have relative fixed location within the CTAP. However, it can be changed by PNC on a long term basis in order to optimize channel utilization. MTSs are used for exchanging commands between PNC and DEVs. There are two ways to use MTSs: the direct uplink MTS for a dedicated DEV or the open MTS for multiple DEVs on a contention basis. For the latter, a slotted ALOHA protocol is used for contention resolution.

**Bandwidth Allocation**

Each DEV uses an explicit bandwidth allocation mechanism to request CTA in CTAP. A source DEV requiring time slots first sends a channel time request (CTR) command to the PNC, indicating the recurring duration and the number of required time slots. The parameters of CTR for asynchronous and isochronous traffic are different. The CTR for the former only contains the total amount of required time. For the latter, the
CTR includes the number of time slots or GTSs needed per superframe, and the minimum and desired duration of each time slot. Each GTS is actually a time slot reserved for a specified DEV. It is DEV’s duty to determine how to use the GTS, i.e. determining which command, stream and asynchronous traffic will be transmitted.

QoS

The CAP does not provide any service assurance. It only offers a best effort service. The QoS is guaranteed in CTAP, in which the channel access on a reservation basis. After receiving a CTR command, the PNC performs the admission control, and allocates GTSs to the requesting DEV when resources are available. For flexibility, the admission control and scheduling algorithms are open in the standard. The standard specifies stream index to identify the connection and traffic stream. A stream index is assigned by the PNC during the connection establishment procedure. There are three reserved stream indices: 0x00 for all asynchronous traffic, 0xFD for MCTA traffic, and 0xFE for unassigned streams. A stream index, except the reserved stream indices, is uniquely assigned for each isochronous stream in a piconet. Asynchronous traffic is assigned to one stream index. Therefore there is no service differentiation among asynchronous connections in a DEV.

Error Control Mechanism

Three ARQ policies are specified in 802.15.3: No-ACK, Immediate ACK and Delayed acknowledgement. If the ACK policy of a frame is set to No-ACK, upon the reception no ACK is sent by the intended recipient. The broadcast and multicast frames must use No-ACK policy upon transmission. If the immediate ACK policy is used, the intended recipient is required to send an Imm-ACK right after the reception. The delayed ac-
knowledge allows for a single ACK acknowledging a block of frames. It is only used for isochronous streams. The parameters of a delayed acknowledgement are negotiated between the source and destination DEV. The retransmission in CAP and CTAP are different. The former uses a backoff mechanism, i.e. the source DEV starts a backoff process to retransmit a frame when the expected ACK is not received in a given time interval. The maximum waiting time is limited by a predefined parameter. In CTAP, the source DEV waits the ACK for a period defined by Retransmission Interframe Space (RIFS) before starting the retransmission.

2.4.4 IEEE 802.16

IEEE develops IEEE 802.16 standards for WMAN. The objective of WMANs is to provide high speed wireless Internet access over a large geographic area, similar to wired access technologies such as digital subscriber line (DSL), Ethernet and fiber optic. The first standard of 802.16 was published in 2002, which is referred to as IEEE 802.16.1. It uses 10 to 60 GHz frequency band to provide fixed wireless access in a point to multi-point (PMP) topology. Because the LOS propagation condition limits the coverage of 802.16.1, especially in urban areas, IEEE developed a new air interface, known as 802.16a and operating at 2 to 11 GHz band. The 802.16a also adds optional capabilities at the MAC layer to support mesh networks so as to increase its coverage. Both 802.16.1 and 802.16a only support fixed wireless access. The mobility support is enabled in standard version 802.16 2005, which is published in December of 2005. The 802.16 standards are different in air interfaces but similar in MAC protocols, i.e. they are all connection oriented MAC protocols with the capability to support continuous or bursty traffic.
Frame Structure

802.16 uses a dynamic TDMA scheme for channel access, where the channels are structured into frames. The communication path from BS to SS is defined as downlink and the opposite direction as uplink. Both uplink and downlink can operate in Frequency Division Duplex (FDD) or Time Division Duplex (TDD) mode as shown in Fig.2.11, and Fig.2.12. In TDD, a frame is divided into uplink and downlink subframe, where the former follows the latter; In FDD, the uplink and downlink subframe are transmitted simultaneously over different frequency bands. Full duplex SS and half duplex SS are both supported in standards. Each subframe consists of a number of time slots. BS and SSs must synchronize and transmit data in predefined time slots.

It is the BS’s duty to assign time slots in both uplink and downlink subframe. The channel assignment information is issued at the head of each downlink subframe. The Downlink Map Message (DL-MAP) is used to specify the downlink channel usage in the current downlink subframe. The Uplink Map Message (UL-MAP) is used to allocate uplink channel to SSs. The standards support adaptive data burst profiling in which transmission parameters such as modulation and coding settings can be modified on a frame-by-frame basis in both uplink and downlink transmission. The data burst profiles are identified by a code named Interval Usage Code (IUC). The IUC for the downlink is called DIUC, and for the uplink is called

![Figure 2.11: TDD access mode in 802.16](image-url)
UIUC. The DL-MAP and UL-MAP uses the DIUC and UIUC to specify the data burst profiles used for each time slot, respectively.

Channel Access

The downlink channel allocation is simple in 802.16. The downlink subframes are different for FDD and TDD. The subframe in FDD is divided into TDM and TDMA portions as shown in Fig.2.12, where TDMA portions follow TDM portions. TDMA portions are separated by preambles. This design is used for half-duplex SS operating at the FDD mode. The half-duplex SS can transmit its data earlier in the subframe and synchronize back to the downlink using the preamble before receiving data. This allows a SS to decode a specific portion of the downlink without the need to decode the entire downlink subframe. The TDD downlink subframe only contains TDM portions.

The uplink subframe contains time slots for SS association, bandwidth request, and data transmission. As shown in Fig.2.12, it includes three periods: Initial Maintenance, Bandwidth (BW) Request Contention, and Data Grants. Different periods are identified by their UIUC. The BS announces these periods in UL-MAP, and can specify such periods in any order and
length. In the Initial Maintenance period, SSs send Initial Maintenance related messages, e.g. ranging requests for BS to determine the network delay. A new station may join the network in this period. In a BW Request Contention period, SSs request bandwidth based upon multicast and broadcast polls issued by the BS. The Data Grant period is dedicated for SSs to transmit their data. The Initial Maintenance period and BW Request Contention period are accessed on a contention basis. The truncated binary exponential backoff algorithm is used to solve collisions.

**Bandwidth Allocation**

Uplink bandwidth is always requested on a connection basis. The connection ID (CID) assigned by the BS is used to identify each connection. The BW request can be sent in an explicit packet or piggybacked on another packet. The requested bandwidth is expressed in two modes: either incremental, meaning how much additional bandwidth is required, or aggregate, meaning how much total bandwidth is needed. Both modes are allowed in explicit BW requests, while only incremental mode is allowed in piggyback BW requests. A BW request can be initiated directly by a connection of a SS or in response to a polling message given by the BS. A BW request is transmitted during the uplink subframe in one of three ways: contending in a BW request period, transmitted in a predefined time slot indicated by UL-MAP, or piggybacked on other packet.

BS issues unicast as well as multicast and broadcast polls. The polling process does not use an explicit poll message from the BS to a SS. Rather, the BS allocates bandwidth in UP-MAP for potential requests from SSs. The BS polls stations in multicast or broadcast method in case the BS finds there is no enough bandwidth to support unicast polling.

The BS grants uplink bandwidth to SSs based on one of two modes: Grant Per Subscriber Station (GPSS) or Grant Per Connection (GPC). In GPSS
mode, the BS allocates bandwidth for individual SSs. It is the SS’s duty to allocate bandwidth among its connections. In GPC mode, the BS allocates bandwidth for individual connections. The BS allocates bandwidth based on following facts: the amount of bandwidth requested by the connections, the QoS parameters of delay and bandwidth needed by current applications in the SS, and the available network resources.

QoS

The principal mechanism to enable QoS in 802.16 is to associate each connection with a service flow. A service flow is a unidirectional flow of packets provided with a particular QoS. A service flow is characterized by a set of QoS parameters such as latency, delay jitter, and throughput. Each network application is associated with a service flow by assigning a unique Service Flow ID (SFID). All packets must be tagged with SFID and CID in order for the network to provide appropriate QoS. Four types of service flows are defined in the standard: Unsolicited Grant Service (UGS), Real-Time Polling Service (rtPS), Non-Real-Time Polling Service (nrtPS), and Best Effort Service (BE). The UGS supports real-time service flows that generate fixed size periodic data packets. A connection with UGS service flow is prohibited from using any contention based BW request. The rtPS supports real-time service flows that generate periodic data with various sizes. The BS needs to poll a SS periodically in order to get unicast BW requests from the SS. A SS is prohibited from using contention based BW request in order to avoid unpredicted delay. The nrtPS supports non-real-time service flows that have data packets of various sizes. The BS needs to poll a SS on a regular basis and allows a SS to send its unicast BW request. A SS is allowed to use contention based BW requests for nrtPS services. Finally, the BE is used to support service flows that do not need QoS support. A SS sends its BW requests in the Request Contention period.
Error Control Mechanism

The ARQ mechanism is a part of the MAC, which is an option for implementation. When implemented, ARQ is enabled on a connection basis. A connection cannot have a mixture of ARQ and non-ARQ traffic. The standards use SR-ARQ as its ARQ mechanism. The transmission of data is based on blocks. A MSDU is logically partitioned into several blocks. Each block is identified by its block sequence number (BSN). Sets of blocks selected for transmission or retransmission are encapsulated into a PDU. To retransmit blocks in a PDU, there are two options: with or without rearrangement of blocks. The former retransmits the original PDU; the latter rearranges the blocks into different PDUs. The ARQ feedback information can be sent as a standalone message or piggybacked on an existing connection.

Two types of acknowledgements are defined: the accumulative ACK and selective ACK. The accumulative ACK acknowledges all blocks up to the BSN specified in the ACK. The selective ACK comes with a bitmap. The blocks with their BSN in the bitmap are acknowledged. A hybrid ACK that combines accumulative ACK and selective ACK is allowed in an ARQ feedback message.

2.4.5 Comparison of MAC Protocol Candidates

The design of an MAC protocol depends on network scenario, service requirements, and the number of nodes to be supported in a network. We identify three common features of MM-wave based MAC protocols: to support a very high data rate, to support a wide range of services, and to provide functionalities dedicated to MM-wave bands, e.g. directional antenna or beamforming support. Based on these features, we compare the MAC candidates on the aspects of channel access, data rate, service
support, MM-wave oriented functions, and other important aspects related to high speed WLAN/ WPAN. Table 2.1 gives the detailed comparison of these protocols.

As seen from Table 2.1, the centralized channel access mode is used by the majority of the candidates. It has great advantages to support fine granularity QoS, and to reduce channel access overhead. Considering emerging applications which are bandwidth intensive with varying requirements on QoS, a centralized access scheme is a better choice for high speed WLAN/ WPAN. Three of centralized access based candidates use hybrid access mode, in which the bandwidth request or other commands can be sent in a contention or contention free period. It increases the robustness of the system in time varying wireless channels. However, the distributed access mode can be considered to be more robust in the sense that it does not have a single point of failure.

Among the candidate systems, 802.15.3c aims at operating at data rates of multiple Gbps. 802.11n is going to support a data rate up to 600 Mbps. The remaining two support a maximum data rate less than 100 Mbps. However, the maximum data rate a system can support is limited by its physical layer, and is not equal to the maximum rate the MAC layer can support. For an MAC protocols, its efficiency determines the maximum throughput a system can afford. All candidates provide centralized access scheme to reduce channel access overhead. All of them provide block transmission and acknowledgement to reduce transmission overhead. In HiperLAN/2 and 802.16 2005, it is implemented by SR-ARQ. In 802.11n, and 802.15.3, it is realized by Block Acknowledgement mechanisms. Moreover, 802.11n provides traffic aggregate to further reduce the MAC layer overhead. In this sense, all candidates are suitable for high data rate wireless systems.

All candidates support QoS in different degrees. Centralized protocols
apparently outperform distributed protocols on QoS provisioning. The 802.11n supports differential services in the EDCA mode, and fine granularity services in the HCCA mode. Other candidates support fine granularity QoS by means of centralized control scheme, in which scheduling algorithms are all open for flexibility.

One unique feature of MM-wave based WLAN/WPAN is the support of directional antennas or beamforming in their MAC protocols. 802.11n is going to support MIMO and beamforming process in its MAC protocol. HiperLAN/2 is capable of using sector antennas at its AP to increase spatial reuse. The main problem of using directional antennas is the implementation of broadcasting function since most MAC protocols rely on broadcasting to realize their control and management functions. In HiperLAN/2 the broadcasting problem is solved by sending a copy of control messages on each sector antenna. In 802.11n, the beamforming process in only performed in the data transmission phase. In other phases, the antenna is working in the omni-directional mode. Therefore the broadcasting based control functions are not affected. For centralized schemes like 802.16, the sector antennas used in BS can be a simple solution. However, considering the flexibility, the beamforming is a better choice. For distributed system like 802.15.3, beamforming may be the only choice. MAC candidates need to be modified accordingly in order to accommodate the beamforming process.

All candidate systems support multiple MCS. It is easy for them to develop an adaptive data rate scheme with a minor modification on MAC protocols. Moreover, all candidates support power saving operation. It is an important feature for a MM-wave system since it has more power consumption compared to lower frequency based systems.

In addition, all candidates support a certain level of mobility. The 802.11n supports mobile initialized hard handover. HiperLAN/2 supports handover
Table 2.1: Comparison of MAC protocol candidates

<table>
<thead>
<tr>
<th></th>
<th>802.11n</th>
<th>HiperLAN/2</th>
<th>802.15.3c</th>
<th>802.16 2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>Channel Access</td>
<td>Distributed</td>
<td>Centralized/ Hybrid</td>
<td>Centralized/ Hybrid</td>
<td>Centralized/ Hybrid</td>
</tr>
<tr>
<td>Max Data Rate (Mbps)</td>
<td>600</td>
<td>54</td>
<td>2000</td>
<td>70</td>
</tr>
<tr>
<td>Frame Structure</td>
<td>Superframe</td>
<td>Superframe</td>
<td>Superframe</td>
<td>Downlink &amp; Uplink frame</td>
</tr>
<tr>
<td>QoS</td>
<td>Access Priority</td>
<td>Scheduling</td>
<td>Scheduling</td>
<td>Scheduling</td>
</tr>
<tr>
<td>Connection Oriented</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Error Control Scheme</td>
<td>BA</td>
<td>SR-ARQ</td>
<td>BA</td>
<td>SR-ARQ</td>
</tr>
<tr>
<td>CRA (^1) BEB(^2)</td>
<td>BEB</td>
<td>BEB</td>
<td>BEB</td>
<td>BEB</td>
</tr>
<tr>
<td>Duplex</td>
<td>TDD</td>
<td>TDD</td>
<td>TDD</td>
<td>TDD/FDD</td>
</tr>
<tr>
<td>Power Saving</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Mobility</td>
<td>Hard Handover</td>
<td>Handover</td>
<td>Partial</td>
<td>Hard Handover</td>
</tr>
<tr>
<td>Multiple MCS</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

in business scenarios. The handover is performed by the RRC unit in the RLC block. The 802.16 2005 supports hard handover. Even for 802.15.3c, the handover of PNC in a piconet guarantees a certain form of mobility.

2.5 Conclusion

The characteristics of MM-wave bands bring both challenges and opportunities for the MAC protocol design. In this chapter, we identify the characteristics of MM-wave bands that affect the MAC design, provide an overview of current MM-wave multi-access systems, investigate the MAC issues in MM-wave based WLAN/WPAN, and introduce the MAC protocols of current high data rate WLAN, WPAN and WMAN which have potential to be operated in MM-wave based multi-access systems. All candidate protocols are capable of working at high data rate with a decent
support on QoS. However, MM-wave systems are more susceptible to environmental changes than lower frequency based systems. Enhancements are needed to cope with the characteristics of MM-wave bands. To release the full power of a MM-wave system, the cross layer approaches jointly considering the power and rate control, beamforming, and channel state dependent scheduling between multiple layers of the network protocol stack are expected.
Chapter 3

Wireless Gigabit Ethernet Extension

In this chapter we follow the assumption that future indoor wireless LANs will use the MM-wave band. The 60 GHz band becomes the RF band of the choice. However, considering the penetration property of the 60 GHz band in indoor environment, a cell per room may be necessary in the system configuration, resulting in a large number of cells for an acceptable coverage. To support hundreds of cells with a maximum bursty data rate of 1 Gbps per cell, the fixed infrastructure will easily become the bottleneck. The passive optical networks (PON) are capable of solving the fixed infrastructure problem in an economical way. Two candidate solutions for PONs are currently competing: The Gigabit PON (GPON) allows a downstream data rate of 2.5 Gbps and an upstream rate of 1.25 Gbps [42]; Ethernet PON (EPON) provides an economical and reliable gigabit speed alternative for wired access networks [43]. Relaxed requirements for the optical layer as they are specified for the EPON standard allow for potentially lower cost equipment. An additional advantage for EPON is that the vast majority of the installed local area networks are based on Ethernet. For these reasons we will consider EPON as the more suitable candidate for the backbone of a gigabit WLAN.

An EPON system is served only by a single wavelength for downstream
CHAPTER 3. WIRELESS GIGABIT ETHERNET EXTENSION

Figure 3.1: WiGEE network

and upstream respectively, which limits its capacity. Wavelength Division Multiplexing (WDM) PONs are proposed to extend the system capacity [44] [45]. The concept of using WDM in access networks has been well established since the RITENET [46] and LARNET [47] beginnings in the early nineties, but the temperature dependency and the cost of the WDM components appeared as the basic technological problem. Recently a new architecture has been introduced that allows for potential low-cost components without temperature control [44]. Here, a broadband light source is spectrally sliced and then fed into Fabry-perot Laser Diodes (FP-LD) that are placed inside the ONUs. Therefore a cost efficient solution for ONUs in WDM PON is feasible.

We use the 60 GHz band to provide a data rate of 1 Gbps per cell, and Wavelength Division Multiplex/Time Division Multiplex (WDM/TDM) Ethernet Passive Optical Network (EPON) as the fixed infrastructure. The system is Ethernet based. We name it Wireless Gigabit Ethernet Extension (WiGEE) since the wireless link is used to replace a piece of fiber to reach mobile terminals (MT). A possible system layout is shown in Fig. 3.1. Combining with the fiber optic infrastructure, the 60 GHz radio,
and a centralized control scheme, the proposed system is able to support hundreds of cells with a maximum bursty data rate of 1 Gbps for each cell. According to its cost, capacity and capabilities, it will be very suitable for medium to large scale indoor scenarios, including office buildings, airports, hotels, bookstores, convention centers, schools and other public venues.

The rest of this chapter is organized as follows: In section 3.1 we develop the system architecture of the WiGEE system, where the basic building blocks are introduced. Then we analyze the scalability of the fiber optic infrastructure in section 3.2. The section 3.3 gives the detail description of the MAC protocol, in which the multi-wavelength (MW) multiple point control protocol (MPCP) is proposed. The simulation result of the performance is shown in section 3.5.

The material presented in this chapter is initially published in [15].

3.1 WiGEE System Architecture

There are several factors shaping the WiGEE design. The first factor is the cost of the system. To support hundreds of BS in WiGEE, a simple BS design becomes necessary. For this, we design a wavelength-selection-free BS with the remote light source provided by the optical carrier supply unit (OCSU) in the central station (CS). Moreover, the radio over fiber (ROF)
technology is employed to centralize the radio baseband and intermediate frequency (IF) band process at the CS. Without the link and upper layer logic, the BS is further simplified. The second factor is the ease of the maintenance. To achieve this, only passive optical components, i.e. optical spitter/combiner (OSC), and arrayed waveguide grating (AWG), are deployed in the optical transmission segment. Since WiGEE is an indoor WLAN, the use of passive optical components is sufficient for reachability. The third factor is the flexibility of the system. Our design provides flexibility at the CS in manifold:

- The shared radio baseband process located at the CS makes the upgrade of radio baseband process simple and easy. Moreover, advanced signal processing techniques such as MIMO process can be implemented at the CS for exploiting the spatial diversity of the wireless signals coming from multiple BSs.

- The embedded control channel coexisting with the data channel over the same wavelength provides the CS flexible to control BSs.

- The flexible configuration of OSC and AWG in the optical transmission segment provides considerable freedom for the cell plan.

3.1.1 Overall Architecture

The architecture is shown in Fig. 3.2, which consists of five basic building blocks: CS, optical Mutiplexer/Demultiplexer (MUX/DEMUX), BS, and MT. Analogous to EPON, the CS implements the OLT functionalities and the MT the ONU. The BS here can be considered as the wireless version of an OSC, which extends the optical signal to the air, and serves multiple MTs in its cell.
According to the position of the OSC and AWG in the optical transmission segment, two configurations of the system are available. The second configuration is simply shown in Fig. 3.3. They differ in the layout of the radio cells. The first configuration shown in Fig. 3.2 uses AWG plus OSC in the optical transmission segment as in [44]. The wavelengths from CS are first de-multiplexed and then each wavelength is fed into a single splitter. The BSs connecting to the same splitter share the same wavelength hence the same upstream and downstream bandwidth. Usually these BSs are close to each other. Therefore a virtual cell is formed if the same up/downstream RF is used by MTs in those cells. In this case the allocation of their upstream channel is performed on a virtual cell basis. Moreover, a spatial diversity can be achieved at the CS if multiple BSs belonging to the same branch of the AWG receive the copies of upstream radio signals from an MT. In the following, we use the terms *single color cells/BSs* to denote the cells/BSs served by the same wavelength, and *multi-color cells/BSs* to denote the cells/BSs served by different wavelengths.

The second configuration splits wavelengths before de-multiplexing them to different BSs. BSs connecting to the same AWG use different wavelengths. Correspondingly, adjacent BSs may use different wavelengths for their upstreams. In this case the upstream channels of adjacent cells can be

---

**Figure 3.3: Alternative architecture of WiGEE**
utilized simultaneously. It is suitable for the case where MTs are usually aggregated in particular areas of the building, e.g., offices or conference rooms. However, a spatial diversity is hardly to achieve in this configuration.

The system exhibits several interested characteristics: the light source of the BS is remotely provided by the CS; the radio over fiber (ROF) is partially used in the system; and the out-of-band control plane is overlapped with the data plane on the same wavelengths. The OCSU in the CS is responsible for generating optical carriers for each BS. Each BS receives two wavelengths from the downstream, one carrying downstream data and control messages, the other carrying carrier used for upstream. These two wavelengths are grouped at the CS by a coarse WDM (CWDM) multiplexer and traverse through the same port of the AWG in the optical transmission segment. A wavelength-seeded reflective semiconductor optical amplifier (R-SOA) is used at the BS for the optical modulation of upstream data received from MTs. This design results in a cost efficient BS with the wavelength-selection-free property. ROF has shown as a promise technology to distribute radio signals for cellular system [48]. However, it is difficult to directly apply to the 60 GHz band due to the technical challenge to modulate ultra-wide bandwidth signals over a wavelength. Alternatively, we only transmit IF signals over the fiber. According to the design, the radio baseband processes of BSs are moved to the CS. The BS simply performs the up/down-conversion of 60 GHz signals.

It is worth noting that certain management tasks have to be performed in the BS. In order to avoid the optical noises produced by idle BSs bury the data signals transmitted in the upstream, the R-SOA of idle BSs should be shut temporally. To shut down specific BSs, the CS needs to know from which BS the upstream signals are transmitted. Hence it needs the identity of the BS and related information from the radio frontend of the
3.1. WIGEE SYSTEM ARCHITECTURE

BS. Moreover, the CS can benefit from the channel state information (CSI) measured at the BSs. With the reliable CSI from the BSs, the CS is able to control the R-SOAs of the BSs properly so that the optical interferences are minimized. In our design, the control channels are implemented over the wavelengths that serve the data channels on both downstream and upstream directions. Like the approach in [49], the control signals and data signals are modulated on different bands so that the co-channel interference is avoid.

3.1.2 OLT Design

The OLT consists of three main parts: the downstream, upstream, and control processes. In addition, it contains the necessary interface connecting to the backbone network. We describe these main processes in the following.

The downstream binary data stream coming from the backbone or WiGEE itself through the layer 2 (L2)/L3 switch is firstly injected into the radio process unit as shown in Fig. 3.4. The radio process unit is an array of radio process sub-units corresponding to each downstream and upstream wavelength, in which each sub-unit uses a radio baseband process and IF stage to perform modulation/demodulation process per downstream/upstream wavelength pair. To provide a flexible configuration to dealing with different wireless channel conditions, a set of modulation and coding schemes (MCS) can be implemented in each sub-unit. The adaptive MSC can be adopted in accordance with the CSI available from the radio frontend of the BS or from the MT. According to the suggestion of Cabric et al [50], the single carrier modulation scheme is a sound choice for the 60 GHz radio, due to its capability to deal with multipath impairments in 60 GHz channels. Grosskopf et al [25] has demonstrated a DBPSK-modulated transparent transmission of the 8B10B Gigabit Ethernet signals. Alternative modula-
tion candidate is the impulse based DS-UWB scheme [51], which is able to achieve multi-gigabit data rate in multipath environment. The modulated radio signals are transformed to optical signals by a FP-LD. The wavelength to carry the radio signals is predefined for each radio sub-unit, here letting $\lambda \alpha_1$ denote the wavelength for sub-unit 1. $\lambda \alpha_1$ is injected into an AWG after passing a CWDM MUX/DEMUX, and aggregated with other downstream wavelengths. The role of the CWDM MUX/DEMUX is to multiplex downstream and extract the upstream wavelength in bidirection. Similar to the CWDM MUX/DEMUX, the AWG works as a multiplex/demultiplexer for up to $N$ upstream/downstream wavelength pairs, where $N$ is the size of the AWG ports determined by the system configuration. Before transmitted over the optical transmission segment, the downstream wavelength goes through a back to back CWDM MUX/DEMUX pair, which is used to combine the carrier wavelengths generated by the OCSU to the downstream. Note that the CWDM MUX/DEMUX pair has different operating wavelength window than the CWDM MUX/DEMUX connected to the FP-LD.

The upstream data in the CS go through an opposite process as compared to the downstream data. The upstream optical signals, which carry the radio signals from a certain BS, are finally terminated at the corresponding photodiode (PD) and transformed to electronic signals. The electronic signals are in turn processed by the radio process unit shown in Fig. 3.4.

![Figure 3.4: Baseband process in CS](image-url)
Note that the single color BSs may receive multiple copies of the same radio signal. After appropriated exploited in the radio process unit, the spatial diversity gain can be obtained.

An out-of-band control plane is designed to coexist with the data plane. With this control plane, the CS is able to control the behaviors of BSs through dedicated control messages, which are multiplexed with the downstream data signals and traverse through the optical transmission segment to the corresponding BS. Two control functions are performed: the R-SOAs of the BSs can be shut down temporally in order to avoid optical interference produced by idle BSs; the radio transmitters of the BSs can be disabled temporally in order to support the half duplex operation mode of the BSs. The half duplex mode is detailed in section 3.3. Moreover, the BSs are able to report the CSI to the CS through the control plane.

### 3.1.3 BS Design

The structure of the BS is relatively simple. It contains a radio frontend, a control unit, and an optical transceiver. We only describe the frontend here since the other two have been introduced in the CS design. As shown in Fig. 3.5, the radio frontend simply up/down-converts radio signals which are received from or sent to the optical transceiver into 60 GHz signals. The local oscillator (LO) in the radio frontend generates 60 GHz carrier.
Using same LOs in the BSs, the carrier frequency is the same for all BSs. However, multiple channels on the 60 GHz band can be formed by using different IF bands on different downstream/upstream wavelengths. There are two reasons to put the LO at each BS instead of using ROF to convey 60 GHz modulated signal: to carry 60 GHz signals, the requirement on the optical components is high, resulting in high cost; due to the availability of SiGe or CMOS based 60 GHz frontend [8] [52], the cost to put an LO in the BS is acceptable.

### 3.1.4 MT Design

The simplified structure of the MT is shown in Fig. 3.6. It is a typical millimeter wave air interface whose design can be found in [53]. It is worth noting that for 60 GHz communication, directional antennas are usually required to achieve reasonable performance. However, directional antennas impose challenges in mobile communication. To solve this, the MIMO and beamforming process can be employed at both BS and MT sides [7]. Similar to the mechanism proposed in IEEE 802.11n draft standard [54], the communication between the BS and MT consists of two phases: In the initial phase, two peers start a beam steering training process through the omni-directional mode of the antenna; Then in the data exchange phase, the peers use the directional mode of the antenna for communication. Naturally, the beamforming process of the BS is controlled by the CS through the control plane.

### 3.2 Scalability of the System

The scalability of the fiber optic infrastructure in terms of BS numbers is analyzed in this section. The number of BSs is limited by the link’s power budget and available wavelengths. The link budget determines the link
reach range and the splitter ratio of the OSC. Due to the application scenarios of WiGEE, the link reach range is not the main concern here. The available wavelength numbers for the system determines the ports size of the AWG in the optical transmission segment, and in turn the number of supported BSs. Although the exact number of BSs depends on the characteristics of the optical components, such as the maximal transmission power, receiver sensitivity, and insertion loss, the approximated number can be estimated according to the common specifications of optical components as summarized in Table 3.1. The link budget of the downstream is

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Value</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_T$</td>
<td>Transmission power</td>
<td>5</td>
<td>dBm</td>
</tr>
<tr>
<td>$R_{scn}$</td>
<td>Receiver sensitivity</td>
<td>$-30$</td>
<td>dBm</td>
</tr>
<tr>
<td>$IL_{CW}$</td>
<td>Insertion loss: CWDM MUX/DEMUX</td>
<td>1.8</td>
<td>dB</td>
</tr>
<tr>
<td>$IL_{AWG}$</td>
<td>Insertion loss: AWG</td>
<td>5</td>
<td>dB</td>
</tr>
<tr>
<td>$IL_{CL}$</td>
<td>Insertion loss: Circulator</td>
<td>1</td>
<td>dB</td>
</tr>
<tr>
<td>$IL_{SA}$</td>
<td>Splicing and aging loss on the link</td>
<td>2</td>
<td>dB</td>
</tr>
<tr>
<td>$S$</td>
<td>Splitter ratio of OSC</td>
<td>4 – 16</td>
<td>none</td>
</tr>
<tr>
<td>$IL_{F}$</td>
<td>Fiber loss</td>
<td>0.2</td>
<td>dB/km</td>
</tr>
<tr>
<td>$G_{SOA}$</td>
<td>Gain of SOA</td>
<td>$10 \sim 25$</td>
<td>dB</td>
</tr>
</tbody>
</table>
expressed as the following:

$$P_T - 4 \times IL_{CW} - 2 \times IL_{AWG} - 3 \times \log_2 S - IL_F \times L - IL_{SA} \geq R_{sen}$$ (3.1)

where the L is the fiber length. Here we assume the maximum link reach range of a typical WiGEE system is 5km. The maximal splitter ratio of the OSC is then limited by

$$\log_2 S \leq \frac{14.8}{3} \Rightarrow S \leq 30$$ (3.2)

For the upstream, the link budget is obtained by

$$P_T - 2 \times (IL_{CL} + IL_{AWG} + 2 \times IL_{CW} + IL_F \times L + 3 \times \log_2 S) - 2 \times IL_{CW} - IL_{AWG} - IL_{SA} + G_{SOA} \geq R_{sen}$$ (3.3)

Taking $G_{SOA} = 25dB$, we get the maximally permitted splitter ratio of the upstream as

$$\log_2 S \leq \frac{14.1}{3} \Rightarrow S \leq 25$$ (3.4)

Considering the characteristics of the OSC, an OSC with a splitter ratio of 16 is the choice of our system to maximize the BS number. This conclusion remains holding when the AWG and OSC in the optical transmission segment exchanging their position.

As shown in Fig. 3.2, the number of available wavelengths is determined by the operation range of two back-to-back CWDM MUX/DEMUX in the CS. That number in turn limits the port size of AWGs in the optical transmission segment. According to the availability of commercial optical components, it is applicable to use up to 16 wavelengths for the downstream and upstream, respectively. Therefore the maximal port size of the AWGs is 16. Combining with the port size of the OSC, it is easy to get
that up to 256 BSs can be supported in WiGEE. Assume the line rate of each wavelength is 1 Gbps. In a full configuration, the peak data rate of each cell is 1 Gbps, and the average data rate per cell is 62.5 Mbps.

3.3 Medium Access Control of WiGEE

Regarding WiGEE as an extension of EPON by replacing a piece of fiber with 60 GHz wireless link to reach ONUs (MTs), it is nature to consider the MAC protocol of EPON, i.e. the MPCP, as the MAC candidate of WiGEE. However, several characteristics of WiGEE make the direct utilization of the MPCP inadequate:

- The movement of MTs between cells of WiGEE demands the inherent handover support in the MAC protocol. A mechanism capable of triggering and attaining fast handovers is desired.

- The above architecture involves two collision domains, i.e. the wireless and optical collision domain. To reduce the collisions and co-channel interferences in both domains, the BS control through the collaboration of MAC and out-of-band control plane of WiGEE is needed.

- Due to the error-prone nature of the wireless channel, the error control scheme has to be implemented in the MAC of WiGEE.

The MAC of WiGEE is described by first introducing its channel access mode, which is based on the superframe.

3.3.1 Superframe Structure

The concept of the superframe is introduced in WiGEE, with the goal to support efficient channel access and mobility. As shown in Fig. 3.7, the downstream and upstream channel access time is divided into a sequence
CHAPTER 3. WIRELESS GIGABIT ETHERNET EXTENSION

of non-overlapping continuous superframes. In general, a superframe consists of six parts: the beacon, MT presence, discovery, random access, and downstream as well as upstream data period. The arrangement of those periods depends on the duplex mode of MTs and the frequency plans of cells. Note that each wavelength has its independent superframe sequence, meanings that single color cells share same superframes and are allocated as a whole.

For simplicity, we assume a simple frequency plan, in which two radio frequency bands, $f_1$ and $f_2$, are used in the system, and $f_1$ is always used for downstream links. The radio frequency band used for the upstream link is determined by the duplex mode of the MT. A half duplex MT uses $f_1$, and a full duplex MT uses $f_2$, for the upstream link. Although different radio frequency pairs are allowed for the downstream and upstream of each wavelength, which results in different frequency plans, there are several reasons

Figure 3.7: Superframe structure of WiGEE
to give this simple frequency plan high priority to be considered: First, considering the vast bandwidth required to support reliable Gbps transmission, few channels are available in the 60 GHz unlicensed band. For instance, [8] used the spectral efficiency of 0.25bps/Hz to achieve 1 Gbps transmission. The occupied bandwidth in this case is 4 GHz. Second, the simple frequency plan raises interesting problems at the MAC layer. For instance, the coexistence problem when multiple cells are overlapped.

Accordingly, the cells of the system operate in three duplex modes: half-duplex, full-duplex, and hybrid duplex mode, which are shown in Fig. 3.7, respectively. The operating mode depends on the type of MTs in the cell. If all MTs are half-duplex nodes, it works in the half-duplex mode; if all MTs are full-duplex nodes, it works in the full-duplex mode; in all other cases, it works in the hybrid-duplex mode. The operating modes are different in the channel allocation. The full and hybrid duplex mode allow simultaneous transmission of downstream and upstream, while the half-duplex mode does not.

A superframe begins with a beacon period, which identifies the wavelength a cell gets served, and the start time and duration of the MT presence, discovery, and random access period in the superframes. To ease the handover process, the beacon periods of all cells are synchronized, which can be easily achieved by the CS. To avoid beacon collisions in case multi-color cells are overlapped, the CS arranges the beacons of those wavelengths in a way that their beacon periods are not overlapped, which is shown in Fig. 3.8.

Followed by the beacon period, the MT presence period is used for MTs reporting their presence in their cells. The MT presence period is divided into multiple time slots. An MT associated to a wavelength is assigned a fixed time slot in this period, used to report its presence in the cell. With the status report from control units in BSs, the CS is able to identify
in which cell(s) the MT locates, and thus uses this information for intra-wavelength handover and scheduling. The discovery period after the MT presence period is used for MT association and fast handover, which are described in section 3.3.2. The random access period right after the discovery period is used for MTs requesting bandwidth in a contention based way. The slotted ALOHA is employed to solve the collisions in this period. The interesting parts of the superframe is the upstream and downstream period. As shown in Fig. 3.7, the CS has the freedom to jointly allocate the downstream and upstream according to the duplex mode of the cell. Moreover, as shown in Fig. 3.8 and explained in section 3.3.2, the CS can achieve an efficient statistical multiplex of downstream and upstream in multiple cells of one wavelength through the use of the out-of-band control plane.

![Figure 3.8: Statistical multiplex of downstream and upstream in WiGEE, observed at BSs](image)

### 3.3.2 Link Layer Functions

Taking into account the requirements of WiGEE, we design the protocol stack of WiGEE as shown in Fig. 3.9. Since this section focus on the medium access control, the link layer becomes the layer of our interest.
As shown in Fig. 3.9, the link layer of the CS consists of three sub-layers: MAC, MPCP, and MAC client sub-layer. The MAC sub-layer is responsible for sending/receiving data to/from the corresponding wavelengths simultaneously. There is a MAC instance for each wavelength, respectively. Upon the MAC sub-layer is a unified MPCP sub-layer, which coordinates the channel access of MTs over all wavelengths. As similar to the MPCP in the EPON standard [11], the MPCP sub-layer provides the discovery, report and grant process for MT association, bandwidth request and allocation, respectively. Moreover, it provides control over multiple wavelengths instead of one wavelength in the current EPON standard. We call the new MPCP the multi-wavelength MPCP (MW-MPCP). The reason to have a unified MPCP sub-layer behind all wavelengths is to facilitate the inter-wavelength control. For instance, the fast handover function is implemented at this sub-layer. Moreover, the MPCP sub-layer interacts with the BS control unit of the CS to coordinate the BSs. The MAC client sub-layer is responsible for the error control and link layer queue management. In this sub-layer, an individual MAC client instance is created for independent error control and queue management of each MT. Separating error control and queue management from the MAC sub-layer provides the flexibility for management and control, e.g. the handover control.

The data link layer of the MT comprises same sub-layers as in the CS. However, the MAC and MAC client sub-layer only have one instance, respectively. For an MT, the MPCP sub-layer is responsible for report process, and upstream channel access according to assigned grants.

The protocol stack of the BS is relatively simple, in which no MAC layer is implemented. The main function of the BS is to connect optical and wireless domain through the O/E and E/O transformation. Moreover, BS performs certain control functions in the out-of-band control plane, which is discussed in subsection 3.3.2.
In the following, we discuss the main functions of the link layer. The error control of WiGEE is described in an individual section, i.e. in section 3.4.

**Discovery Process**

The discovery process basically works in the way similar to that specified in the EPON standard. An unregistered MT detects the presence of the network through beacon scanning. From the detected beacon the MT is able to synchronize with the network and obtain the cell information, including the location of the discovery period in the superframe. It then starts an association process in this period, negotiating the configuration with the CS in the following dedicated grant issued by the CS. Different from the registration process in EPON, the ranging is not performed in here since the distance between the BS and MT is negligible in WiGEE. Instead, the ranging is performed between the CS and BSs through the out-of-band channel plane.

Note that since the discovery period is shared in single color cells, the collision may occur in both wireless and optical domain. To trade off the collisions and control overhead of this period, the CS is allowed to adaptively adjust the length of this period.
3.3. MEDIUM ACCESS CONTROL OF WIGEE

Bandwidth Allocation

The proposed MW-MPCP is a reservation based MAC protocol, in which a request and grant mechanism is used for upstream bandwidth allocation. We use the corresponding MPCP control messages specified in the 802.3ah standard [11] for the bandwidth request and grant. The bandwidth request message is sent by an MT in two ways: contending in the random access period of a superframe, or sent in an upstream grant period of the MT. The reason to use a random access period for the bandwidth request is to provide MTs the opportunities to request upstream bandwidth in case the pervious bandwidth request or grant is missing due to the channel error. The grant message issues by the CS indicates the start time and period an MT is permitted to full utilize the upstream channel.

The request and grant mechanism provides WiGEE a flexible framework for upstream bandwidth allocation. The actual allocation is completed by a dynamic bandwidth allocation (DBA) algorithm, whose implementation relies on specific requirements of application scenarios. The DBA algorithms in WiGEE should support inter-wavelength allocation, which distinguishes it from the DBA algorithms in EPON.

BS Control

The prominent feature of WiGEE is the use of the out-of-band control plane to provide seamless integration of optical and wireless sub-systems. The control plane is designed to be bidirectional. The CS issues control messages to BSs in the downlink, while BSs report channel information to the CS in the uplink. Here we use the term downlink/uplink for the control channel and downstream/upstream for the data plane. For control purpose, a unique identity per BS is used to distinguish BSs in downlink control and uplink report messages.
The information the CS needs to know in order to efficiently control the BSs is the membership of MTs with regard to BSs. According to the design of WiGEE, both the CS and BS are not able to know directly from which BS the upstream data are transmitted since the OSC combines the data from multiple BSs. However, by our design the CSI of BSs can be obtained by the CS. Combining with the allocation information of the upstream, the CS is able to estimate in which cell(s) the MT is located. Note that if multiple cells are overlapped, the CS may find that an MT belongs to multiple BSs. In this case, it chooses the BS with the maximal received signal strength as its master BS for bandwidth allocation and BS control. Moreover, the CSI reported by BSs can be used to exploit the multiuser diversity gain in the DBA algorithm, where the channel aware scheduling algorithms are employed to improve the overall system capacity. One consideration here is the coherence time of channel as compared to the data propagation delay. It has been reported that the coherence time of 60 GHz channel in indoor environment is around 50ms [55]. Considering the propagation delay of 33 $\mu$s in a typical configuration of the WiGEE system, it is feasible to use channel aware scheduling algorithms in DBA.

Knowing the locations of MTs, the CS can control the BSs in two ways: temporarily shutting down the radio transmitter of a BS so as to support the half duplex mode, or temporarily shutting down the optical transmitter of a BS so as to guarantee the successful optical upstream transmission of the other BS. Without the BS control, the wavelength pair serving half duplex mode MTs has to operate as well in the half duplex mode. In this case half of the optical bandwidth is wasted. With the BS control, the CS can schedule simultaneous upstream and downstream transmission in different cells of the same wavelength pair. Therefore, the channel utilization of the optical transmission segment is improved.

Note that a BS can receive the downlink control messages at any time, but
the uplink control channel is not always available to the BS, for the reason that the optical transmitters of BSs need to be shut down frequently according to the upstream scheduling. Moreover, the collisions have to be solved in the uplink control channel since it is shared by multiple BSs. To solve these problems, the CS issues explicit grant messages to BSs, specifying the time and duration a BS is allowed to transmit report messages in the uplink.

**Handover**

A handover occurs when an MT roams from one cell to another. As shown in Fig. 3.10, there are two kinds of handovers in WiGEE: the intra-wavelength handover, and inter-wavelength handover. The former occurs when an MT moves between single color cells. The latter occurs when an MT moves to a different color cell. The reader may wonder the reason for the intra-wavelength handover. Actually, if the DBA algorithm takes the locations of MTs into account, the intra-wavelength handover process become necessary in order to update the location of roaming MTs promptly. In WiGEE, the MT presence period of a superframe is used to support intra-wavelength handover, which has been described in section 3.3.1. The beacon and discovery period are used to support inter-wavelength handover.

The identity of the working wavelength is contained in the beacon of the superframe. An MT detects an inter-wavelength handover if it receives a beacon tagged by a different wavelength identity. The potential handover is reported by the MT through sending a handover indicator in the coming discovery period of the new wavelength. Receiving the handover indicator, the CS negotiates with the MT to complete the handover process. In the next superframe, the traffic of the MT is directed to the new cell. In general, it takes two-superframe duration to complete a handover.
By choosing the duration of the superframe properly, the communication sessions of upper layers will not be interrupted.

The separation of the MAC client, MPCP, and MAC sub-layer reduces the state transitions of error control and queue management entities in an inter-wavelength handover. Note that each MT has a wavelength independent instance in the MAC client sub-layer of the CS. As an inter-wavelength handover occurs, the MPCP can simply redirect the MAC client instance of the corresponding MT to the related MAC sub-layer instance of the new wavelength. No additional process is needed to manipulate the error control and queues states maintained in the MAC client sub-layer.

Other functionalities

Other functionalities need to be supported by the MAC layer include the multiple physical data rate support, and the beamforming process of MIMO antennas for data transmission. Since they are tightly related to specific physical layer implementations, we only provide a brief introduction here.
It is reasonable for the system to accommodate different modulation and coding scheme at the PHY layer so as to provide reliable communications under different channel conditions. For instance, the control messages in a superframe, such as the beacon, grant, and report messages are transmitted at a base data rate, while the data are transmitted at higher data rate. As an impact to the MAC layer, the MW-MPCP should be multi-rate aware, and generates the grant accordingly.

The beamforming process is a useful approach to combat against multipath impairment in 60 GHz channels. It includes a training process between the CS, BS, and MT to negotiate the parameters of MIMO antennas at BS and MT. To support beamforming, special control frame are needed for training and data exchange at the MAC layer.

3.4 Error Control in WiGEE

Implementing gigabit wireless transmission in indoor environment is a challenge task due to the channel impairments. The Friis path loss equation [24] indicates that, for equal antenna gain, the pass loss is proportional to the square of the carrier frequency. A 60 GHz system has thus an additional 22 dB of path loss as compared to an equivalent 5 GHz system. Moreover, the delay spread caused by multipath fading is another obstacle to achieve high speed communications. The RMS delay spread of 60 GHz in indoor environment is typically up to tens of nanoseconds, which results in severe inter-symbol interference. Considering the path loss, the delay spread, the low penetration property of 60 GHz, the error control scheme in the link layer of WiGEE is necessary.

Two typical error control technologies, i.e. automatic repeat request (ARQ) and forward error correction (FEC), are used to provide reliable communication over an unreliable channel. The FEC corrects the transmission
error in real time through the use of redundant code (block or convolutional) attached in the transmitted message. The ARQ, however, drops the error frame and retransmits it until the correct one is received. Usually a wireless system adopts both to combat the error-prone channel. The FEC and ARQ can be employed independently without affecting the system performance. In this section, we assume the FEC is employed in the physical layer and focus our discussion on the ARQ on the link layer.

Although the collisions and errors may happen in the optical domain, the errors occurring in the optical domain in the order of $10^{-12}$ are negligible. For this, we only consider the error performance of WiGEE in the wireless domain. The error performance of WiGEE is determined by three factors: the channel behavior, the performance of physical layer implementation, and the link and upper layer error control schemes. This section studies the error control scheme at the link layer.

### 3.4.1 ARQ

The choice of the ARQ scheme is determined by the bandwidth delay product, which is the product of the data rate and propagation delay. A large bandwidth delay product with medium average frame size suggests the use of windows sliding ARQ, such as go back N (GBN) and SR-ARQ. It is easy to get that in a typical configured WiGEE with a fiber length of 5km, the bandwidth delay product is 24 kbytes, which is 47 times larger than the smallest Ethernet frame. Thereby, a sliding window based ARQ mechanism is needed.

The window sliding ARQ has been extensively investigated in data communication. GBN and SR-ARQ are two classic window sliding ARQ techniques. The main difference between them is the way to handle retransmission. The SR-ARQ is the ARQ of the choice for WiGEE, which only retransmits the problematic frames, and therefore reduces the amount of
retransmission.
To derive parameter and analyze performance, we assume WiGEE operates in the full duplex mode.

3.4.2 Determine ARQ Parameters

The parameters for SR-ARQ include the sequence space, sliding window size, retransmission timer and receiving buffer size. If the piggyback ACK is used, then the explicit ACK timer is required to issue an explicit ACK if no piggyback ACK is available in a given time. Most of the parameters are determined by the ACK time, which is the maximum time interval between the transmitter sends a frame and receives the ACK. Due to the asymmetric nature of downstream and upstream in WiGEE, different ARQ parameter settings are used in these two directions.

For the upstream, the ACK time can be expressed as

\[ T_{AT} = RTT + T_{proc} + T_{queue} \]  \hspace{1cm} (3.5)

where \( RTT = 2 \times T_{prop} + T_{trans} + T_{trans,ACK} \) is the round trip time, \( T_{prop} \) denotes the propagation delay, which is related to the fiber length, \( T_{trans} \) denotes the transmission delay of the data frame, \( T_{trans,ACK} \) denotes the transmission delay of the ACK frame, \( T_{proc} \) denotes the processing delay, \( T_{queue} \) denotes the queuing delay to send the ACK in the receiver. If piggyback ACK is used, the time waiting for piggyback ACK, denoted by \( T_{piggy} \), is added in the right hand of Eq. 3.5. The minimum window size for the upstream is obtained by \( W_{ul} = \lceil (R \times T_{AT})/L_{min} \rceil \), where \( R \) is the data rate of the system, \( L_{min} \) is the minimum Ethernet frame size in bit, and \( \lceil \cdot \rceil \) is the ceiling operation. The minimum retransmission timer is set as \( T_{RETXM} = T_{AT} + T_{guard} \), where \( T_{guard} \) is a variable used to reduce the unnecessary retransmission. For SR-ARQ, the sequence space must be two times greater than window size so as to avoid the old ACK acknowledges
the new frame. In general, a large window size provides SR-ARQ better performance in the high PER condition. However, a trade off should be taken between the infinite window size for an ideal error correction performance and a moderate window size for a practical system.

According to the upstream bandwidth allocation, the downstream has two sets of ACK time, i.e. the short ACK time when the MT is right in the grant period, or the long ACK time when the MT is not in its the grant. The ACK time for the former is approximated as

\[ T_{AT} = RTT + T_{proc} + T_{queue} + T_{piggy} \]

which is the same as the upstream case. For the latter,

\[ T_{AT} = RTT + T_{proc} + T_{queue} + T_{piggy} + T_{wait} \]

where \( T_{wait} \) is the maximum time between two consecutive allocations of any MT.

**Integrate ARQ Into Ethernet Frame**

The support of ARQ at the link layer needs additional ARQ control frames, and sequence number carried in each frame. The Ethernet frame, however, provides no field to support error control. Although the logical link control (LLC) layer in 802.2 provides sequence number fields, the length of these fields is not sufficient to accommodate sequence space required by WiGEE. Moreover, since the statistic information from the ARQ is used by the DBA algorithm for bandwidth allocation, from the WiGEE viewpoint, LLC is not the right place to process ARQ. It is reasonable to provide an acknowledge mechanism at the MAC layer of WiGEE.

To avoid modify the Ethernet frame, we put the sequence number of each frame at the beginning of the payload field. Five octets are used for the queue identity, sequence number of this frame, and piggyback acknowl-
edged sequence number for the queue identified by the queue identity, respectively. The MAC client sub-layer is responsible for removing the ARQ field and sending the pure payload to the upper layer. The Ethernet frame structure with additional ARQ control fields is shown in Fig. 3.11. Moreover, the MPCP control frames are extended to support ACK messages, which are shown in Table 3.2. The MPCP control frame used for ARQ control is shown in Fig. 3.12.

<table>
<thead>
<tr>
<th>Opcode</th>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>RR</td>
<td>9</td>
<td>Ready for Receipt, ACK sent by receiver</td>
</tr>
<tr>
<td>RREJ</td>
<td>10</td>
<td>Reject, NACK sent by receiver</td>
</tr>
<tr>
<td>SREJ</td>
<td>11</td>
<td>Selective Reject, NACK sent by receiver</td>
</tr>
</tbody>
</table>

Table 3.2: Extended types in MPCP control frame used for ARQ support
3.4.3 QoS Consideration

To support QoS for differential services, WiGEE supports priority queues at the link layer. The priorities of queues are mapped to the user priorities defined in IEEE 802.1Q [56]. The data frames stored in each queue are tagged with corresponding user priorities in the queue bit map field. The error control is performed on the queue basis, while the error control policy per queue is negotiated during the MT association process. For instance, a queue holding real time traffic may not need error control, while a queue holding data traffic should work with error control.

3.4.4 Performance Analysis

Let firstly consider the performance of SR-ARQ on a point-to-point (P2P) link where the guard time between consecutive frames is negligible, then we extend the analysis to the point to multi-point (PMP) case. Let $T_{prop}$ denote the propagation delay of the channel, $T_{trans}$ denote the transmission delay of one frame, where the frame length is assumed to be
a constant, and $\alpha$ denote $T_{\text{prop}}/T_{\text{trans}}$. For convenience, let us normalize frame transmission time to a value of one; thus, the propagation time is $\alpha$. Assuming frame processing time and the acknowledgement transmission time are ignored, it then takes at least time equal to $2\alpha + 1$ to send a frame and receive its acknowledgement. In an error-free environment, the normalized throughput \(^1\) for the sliding window ARQ can be expressed as:

$$S = \begin{cases} 
1 & W \geq 2\alpha + 1 \\
\frac{W}{2\alpha+1} & W < 2\alpha + 1 
\end{cases}$$

(3.6)

where $W$ denotes the ARQ window size in time. In the presence of errors, the above equation must be divided by average transmitted times per frame, denoted as $N_{\text{aver}}$. Let $P_e$ be the PER, thus in SR-ARQ,

$$N_{\text{aver}} = \sum_{i=1}^{\infty} i \times P_e^{i-1} \times (1 - P_e) = \frac{1}{1 - P_e}$$

(3.7)

Thus the throughput is:

$$S = \begin{cases} 
1 - P_e & W \geq 2\alpha + 1 \\
\frac{W(1-P_e)}{2\alpha+1} & W < 2\alpha + 1 
\end{cases}$$

(3.8)

The throughput we are interested in the PMP case is the aggregated throughput, which is the sum of the throughput from all MTs. Due to the guard time between the grants and the MPCP control messages, the overall throughput in the PMP case is less than that of the P2P case. In addition, the asymmetric nature of the downstream and upstream makes the ARQ implementation on each side different. In the following, we analyze the upstream and downstream separately.

**Upstream**

The upstream should take into account the overhead such as the guard time and control periods. Let define the sum of the guard time and control

\(^1\)Unless state otherwise, the throughput in this section means the normalized throughput
periods in a unit time as the system overhead, and the parameter $\beta$ as the ratio of the system overhead to channel capacity. Then the normalized effective channel capacity in the upstream becomes $1 - \beta$.

In SR-ARQ, the aggregated throughput can be expressed by:

$$S = \begin{cases} 
(1 - \beta)(1 - P_e) & W \geq 2\alpha + 1 \\
\frac{W(1-\beta)(1-P_e)}{2\alpha+1} & W < 2\alpha + 1 
\end{cases} \quad (3.9)$$

**Downstream**

In an extreme case the OLT will take the time period $T_{\text{wait}} + 2\alpha + 1$ to wait an ACK back if a frame is sent at a time that the corresponding MT just uses out the grant and thus the ACK has to wait for the next grant. $T_{\text{wait}}$ is the time of two consecutive grants to an MT. Therefore the window size should take at least $T_{\text{wait}} + 2\alpha + 1$. To simplify the analysis, we assume it as a constant.

Let us first derive the throughput under the error-free operation. The window size $W$ is divided into three cases. When $W \geq T_{\text{wait}} + 2\alpha + 1$, the sender can always get the ACK back before the window is exhaust, and therefore the normalized throughput is 1. When $T_{\text{wait}} + 2\alpha + 1 > W \geq 2\alpha + 1$, in the grant period, the throughput is 1, but in other period, only $W$ can be sent, and thus the throughput is $W/T_{\text{wait}} \quad \text{2}$. The overall throughput then can be expressed by $S = (T_i + W)/(2\alpha + 1 + T_{\text{wait}} + T_i)$, where $T_i$ is the grant assigned to the MT. In the case when $W < 2\alpha + 1$, in $T_i$, the throughput becomes $W/(2\alpha + 1)$. The overall throughput turns to $(W \times T_i/(2\alpha + 1) + W)/(2\alpha + 1 + T_{\text{wait}} + T_i)$. In summary, the throughput

\footnote{We assume $W < T_{\text{wait}}$. This condition usually holds in a practical system.}
in error-free operation is:

\[
S = \begin{cases} 
1 & W \geq 2\alpha + 1 + T_{\text{wait}} \\
\frac{T_i + W}{2\alpha + 1 + T_{\text{wait}} + T_i} & 2\alpha + 1 < W \leq 2\alpha + 1 + T_{\text{wait}} \\
\frac{W \times T_i + W}{2\alpha + 1 + T_{\text{wait}} + T_i} & W < 2\alpha + 1
\end{cases}
\] (3.10)

The downstream window size does not change average transmitted times per frame. That is \(N_{\text{aveer}} = 1/(1 - P_e)\). Therefore the throughput in the presence of errors is:

\[
S = \begin{cases} 
1 - P_e & W \geq 2\alpha + 1 + T_{\text{wait}} \\
\frac{(T_i + W)(1 - P_e)}{2\alpha + 1 + T_{\text{wait}} + T_i} & 2\alpha + 1 < W \leq 2\alpha + 1 + T_{\text{wait}} \\
\frac{(W \times T_i + W)(1 - P_e)}{2\alpha + 1 + T_{\text{wait}} + T_i} & W < 2\alpha + 1
\end{cases}
\] (3.11)

### 3.5 Simulation Study of WiGEE

We use the simulator OMNET++ [57] to study the upstream performance of WiGEE under different ARQ policies. Two ARQ policies are used: one uses SR-ARQ to provide error free communication; the other does not use ARQ. The simulation mainly focuses on the throughput and delay performance under different bit error rate (BER), and the impact of user mobility on the system performance. We only simulate the first system architecture since the purpose of this simulation is the qualitative analysis of the system. From this viewpoint, the difference between the second architecture is small.

In the simulation, the system uses two wavelengths. Each wavelength serves two BSs with a line rate of 1 Gbps through a two-port splitter. In the optical segment, the distance between OLT and splitter is 5 km. For simplicity, the distance between the splitter and all BSs is identical to 1 km. The 60 GHz radio operates in the full duplex mode. The radio transmitted power is set to 10mw. The sensitivity of the receiver is -80dbm. According
to [58], the path loss coefficient for free space propagation is set to 2. For simplicity again, the reflection, diffraction, and scattering of radio signal are not considered in this simulation. The binary symmetric channel model is used as the error model. Moreover, BSs are separated enough to avoid co-channel interference. The MTs use waypoint mobility model.

The simulation consists of two scenarios. In the stationary scenario all MTs are stationary, while in the mobility scenario the MTs can move around in the playground with a speed of 3m/s, which is a normal walking speed of an adult. Each scenario includes three sub-scenarios, in which the number of MTs per cell at the initial stage is 2, 4, and 8 respectively. The first scenario focuses on the simulation of performance against different wireless BER settings. Three settings with the BER of 1e-4, 1e-6, and 1e-12 represent different indoor environments respectively. In the mobility scenario, we focus the simulation on the performance comparison with the stationary scenario. To isolate the problems caused by unreliable wireless channel, we only perform the simulation under low BER.

In this simulation, the length MT presence, discovery, and random access period are fixed in each superframe. The DBA algorithm is able to predict the start time and duration of those field, and therefore allocate the upstream bandwidth cross the superframe. As shown in Fig. 3.13, an IPACT-like DBA algorithm [59] is used in the simulation, in which the upstream is scheduled in a way that one grant right follows the previous grant when the load is sufficiently high.

The traffic mode used in the simulation is the on/off model with the exponential distribution on the state holding time. The frame length is fixed to 500 bytes. Since the retransmission in the ARQ control scheme introduces additional queue delay, to avoid extra long delay occurring in the high BER case, the MT under the SR-ARQ control uses smaller sending queue size than that without ARQ control. The former uses the queue size of 70, and
3.5. SIMULATION STUDY OF WIGEE

... the latter uses 1000.

![Diagram](image)

**Figure 3.13:** IPACT-like DBA algorithm in full duplex mode

### 3.5.1 Result Analysis

#### No Error Control Case

Fig. 3.14, 3.15 and 3.16 present the system performance under different BER settings in the stationary scenario. Note in this simulation no ARQ control is employed in the link layer. Fig. 3.14 shows the throughput as a function of the offered load per wavelength. It illustrates that the BER has a significant impact on system performance. In low BER cases, i.e. the BER of $1e^{-6}$ and $1e^{-12}$, the throughput approximately is equal to the load when the load is unsaturated. The slope of the throughput slows down when the load is large enough to cram the sending queue of MTs, leading to vast frame drop in MTs. The inflexion appears at the load of 800Mbps. It shows the system has good performance even under a very high load condition. It is worth noting that the system almost has the same performance under the BER of $1e^{-6}$ and $1e^{-12}$. In the high BER case, the gap between throughput and load becomes larger as the load increasing. The maximum throughput of the system only reaches half of the channel...
capacity. Moreover, the inflexion for unsaturated and saturated load appears at the load of 600Mbps, which is much less than that in low BER cases. As shown in Fig. 3.16, the high frame drop ratio due to the high channel error is the primary reason ruining the throughput. The solution to improve the performance is the use of FEC coding. The FEC scheme suggested by 802.3ah standard can be considered here.

Fig. 3.15 shows the average delay of data frames as a function of the offered load per wavelength. From it we can see that when the load is under 300Mbps, the average delay only has several milliseconds. It grows as the load increasing. The system with a lower BER has a lower average delay. When the load exceeds 600Mbps, the delay of the high BER case experiences a rapid growth, but after the load of 800Mbps, the growth rate decreases. However, in other two low BER cases, the delay increases slowly before the load approaches the inflexion of 800Mbps, whereas experiences
Figure 3.15: Average delay as a function of offered load per wavelength, no ARQ control, stationary scenario.

The queue length in MTs can explain the phenomena of the delay growth. The longer the queue length, the larger average delay the upstream frames experience. When the load is saturated, the queues in MTs become full. The average delay increases rapidly. Why the low BER cases outperform the high BER case can be explained by the fact that the probability the GRANT and REPORT messages being lost is higher in the latter.

Fig. 3.16 shows the data drop ratio as a function of the offered load per wavelength. In low BER cases, the drop ratio is almost equal to zero when the load is unsaturated. It increases slightly after the inflexion mentioned above. On the other hand, the drop ratio in the high BER case is higher than 30%, which is perfectly matched with the analytic result for a channel with the average frame length of 500 bytes under a BER of 1e-4 with a uniform distribution. After the inflexion, it experiences a fast increase.
Figure 3.16: Drop ratio as a function of offered load per wavelength, no ARQ control, stationary scenario

Figure 3.17: Average throughput as a function of offered load per wavelength, stationary and mobility scenario
than that of high BER cases. Before the inflexion, the BER is the main contribution for the frame dropping. It is a random process independent of the load. Therefore the drop ratio keeps stable. When the load exceeds the inflexion, MTs start dropping frames due to the limitation of queue length. As a result, the drop ratio increases.

The performance comparison between stationary and mobility scenario is illustrated in Fig. 3.17, 3.18 and 3.19. It can be seen that the stationary case outperforms the mobility case in each performance metric, due to the reason that the mobility of the MTs increases the co-channel interference, and reduces the efficiency of the channel utilization. However, when the load is low, the system can provide considerable performance even in the mobility case.

The simulation results show that the WiGEE system can provide considerable throughput and average delay even in the mobility scenario when
the system load is not saturated. It is foreseen that the mobility have a significant impact on the performance of the transport layer.

Error Control Case

Fig. 3.20 and Fig. 3.21 show the throughput and delay performance of WiGEE under the SR-ARQ control. Fig. 3.20 shows the throughput difference between the BER of $1e^{-6}$ and $1e^{-12}$ is minor, which is the same as the no ARQ control case. The main difference occurs at the high BER case. When the load is low, the throughput under the BER of $1e^{-4}$ is the same as the low BER cases. The inflexion appears at the load of 340 Mbps. As the load continues increase, the throughput keeps almost constant at 340 Mbps, which is less than the case under no ARQ control. It can be explained by the ARQ control overhead. When the load is low, the system is capable of accommodating the load plus the ARQ control overhead,
3.5. SIMULATION STUDY OF WIGEE

Figure 3.20: Average throughput as a function of offered load per wavelength, SR-ARQ, stationary scenario

Figure 3.21: Average delay as a function of offered load per wavelength, SR-ARQ, stationary scenario
leading to the throughput equal to the load. As the load goes behind the inflexion, the combination of the load plus ARQ control overhead makes the system saturated and thus limits the throughput.

Fig. 3.21 show the delay of the system under SR-ARQ control. Note that it is quite different from the no ARQ case, due to the different sending queue sizes in each case. As seen from the figure, since the ARQ control case uses small queue size, the delay is significantly smaller than that of the no ARQ control case. Under the SR-ARQ control, the BER case of $1e^{-4}$ has the delay significantly larger than the BER cases of $1e^{-6}$ and $1e^{-12}$, while the latter two have similar delay performance. The difference is resulted from the retransmission mechanism in SR-ARQ. Moreover, the delay increases as the number of nodes per cell increases. The delay difference between different node settings becomes larger when the load increases.

3.6 Conclusion

In this chapter we introduce a novel hybrid optical/wireless system as a candidate for the gigabit WLAN, where the WDM/TDM EPON is used as the fiber optic infrastructure, and 60 GHz radio for the wireless access. The unique feature of the system is the high data rate support among a large number of small cells. The architecture, medium access control, and error control issues are discussed in detail. The throughput and delay performance under a simple DBA algorithm is studied by simulation. Moreover, the scalability of the system in terms of the fiber optic infrastructure is studied. It show the proposed system has adequate capability to serve a large building.
Chapter 4

BER Models for WiGEE

The channel behavior and physical layer implementation greatly impact the performance of a communication system. It is particularly true for a wireless communications system. For instance, Zorzi et al [60] showed that a memory wireless channel, if exploited properly, can be used to improve the capacity of a system. For a digital communication system, the impact of the channel and physical layer on the performance is reflected in the error pattern of the receiving bit stream. If the error pattern can be replicated by some methods, e.g. the analytical error model or trace file, the performance of the link and upper layers can be evaluated even without the need of the input from a real physical layer. It is the objective of this chapter to develop suitable error models for WiGEE under the 60 GHz channel. The analytical error models are derived based on the trace files obtained from the 60 GHz transceiver proposed in [21].

An error model is a compact representation of error characteristics at physical or upper layers. Due to the intrinsically stochastic nature of communications, all relevant error models are random models, which can be classified into analog model and digital model. The analog models describe the behaviors of the received signal strength and/or the interference power; the digital models characterize the bit stream that the physical layer of
a receiver delivers to high protocol layers. The advantage of digital models is that they have already included all possible impediments including channel fading and synchronization problems from the receiver hardware. Therefore they are more suitable for the high layer performance evaluation. Unless stated otherwise, an error model in the remaining of the chapter is referred to a digital error model.

Based on different methodologies, the digital models are classified into analytic models or measurement based models. The Gilbert-Elliot model [61] [62] is a well known analytic model. The trace file model, which captures binary indicator sequences from the physical layer of a real or simulated system, is a measurement based model. The binary indicator sequence is a finite sequence of zeros and ones, with 0 indicating correct bits and 1 indicating erroneous bits [63]. The analytical model derived from a trace file characterizes the error pattern in the trace file, and provides a more compact and stochastic method to replicate error events. It is more desirable for the performance study.

The remaining of the chapter is organized as follows: In the next section, we give a brief introduction on existing analytical error models; Then the obtained trace files are analyzed and the parameters for various analytical models are derived; Finally, the suitability of analytical models is studied.

### 4.1 Analytical Error Models

The analytical error models are common in the sense that they are all used to decide whether a particular bit is erroneous or correct, resulting in a sequence of error bits. However, they are different in the complexity and accuracy to model the error patterns in trace files. Typically, the bit stream in a trace file is divided into runs and bursts, where a run is a stream of consecutively correct bits, and a burst is a stream of consecu-
tively erroneous bits. The lengths of runs extracted from a trace file form a run sequence, and the lengths of bursts form a burst sequence. The run sequence and burst sequence are usually treated as stochastic processes. The task of an analytical error model is to model the empirical distributions of run and burst sequence by appropriate distributions so that the simulated error bit stream matches the error pattern in the trace file. The simplest analytical model is the binary symmetric channel (BSC) model [64], in which only one parameter is required. The analytical models being widely used are time-homogeneous Markov models (discrete or continuous). The two-state Markov model, known as the Gilbert-Elliot model [61] [62], is the most widely used model. The problem of Gilbert-Elliot model is its limitation to catch the nature of complex error patterns. Semi-Markov and multi-state Markov models are thus used to model more complex distributions. To further model the heavy tail distribution of run lengths, advanced models such as bipartite model and chaotic map model can be employed.

4.1.1 BSC Model

In the BSC model, the Bernoulli experiment with the parameter $p_b \in (0, 1)$ is performed to determine the erroneous state of each bit independently, where $p_b$ is chosen to equal to the mean BER of the trace file. The run length distribution, therefore, follows a geometric distribution, i.e. $F_L(l) = 1 - (1 - p_b)^l$, where $l$ is the run length, and $F_L(l)$ is the cumulative distribution function (CDF) of $l$. It is easy to find that the complementary CDF (CCDF) of the run length is exponentially decaying. The advantage of this model is its simplicity. However, it is inadequate to capture the bursty nature of channel errors observed in wireless communications.
4.1.2 Two State Markov Models

The Gilbert-Elliot model is a two state Markov model that has a "good" channel state and a "bad" channel state. Two bit error probabilities $e_g$ and $e_b$ are associated with the good state and bad state individually, where $e_g << e_b$, and usually $e_b = 0$. The bits generated in each state use the corresponding error probability to produce error bits as in the BSC model. The transitions between states are determined by the transition matrix shown as follows:

$$P = \begin{pmatrix} p_{gg} & p_{gb} \\ p_{bg} & p_{bb} \end{pmatrix}$$

where $p_{mn}$ is the probability that the next state is $n$ given that the current state is $m$. The steady-state probability in the good state or bad state is

$$p_g = \frac{1 - p_{bb}}{2 - (p_{gg} + p_{bb})}, \quad p_b = \frac{1 - p_{gg}}{2 - (p_{gg} + p_{bb})}$$

respectively. The mean BER is $\bar{e} = p_g \cdot e_g + p_b \cdot e_b$.

In the Gilbert-Elliot model, the state holding time, which is the distribution of time the models stays in a certain state, characterizes the transition behavior of the model. For the good state, the mean state holding time is $1/(1 - p_{gg})$, and for the bad state, it is $1/(1 - p_{bb})$. Knowing the state holding time of two states is sufficient to determine the transition matrix $P$. The state holding time for a Markov model follows a geometric distribution. Accordingly, the run length and burst length distributions are exponentially decaying.

If either the distribution of the run length or burst length is not geometric, it is applicable to use other distributions, resulting in semi-Markov models. Candidate distributions includes binomial, Poisson distribution, or quantized versions of continuous distributions, such as log-normal, or gamma distributions. Semi-Markov models are more accurate to model error patterns. However, they are still inadequate to model heavy-tail distributions.
since they show short-term correlation prosperities for bit errors.

### 4.1.3 N-state Models

A widely used N-state model is the finite state Markov chain (FSMC) model proposed by Wang et al [65]. The Markov chain is derived from modeling the instantaneous signal-to-noise ratio (SNR) at the receiver under the Rayleigh fading assumption, where the transition is restricted to adjacent states. Physical parameters such as mean SNR and Doppler frequency are required by the model. It is not suitable for the trace file analysis in our case.

The bipartite model [63] is a multi-state Markov model that includes a number of “good” states and “bad” states and allows the state transition only from good states to bad states and vice versa (forming a bipartite graph). It allows the user to choose a proper tradeoff between model complexity and accuracy. However, a large number of parameters are required to achieve an acceptable fit with measured data. Moreover, the parameters are tightly coupled to the trace files they are generated from.

Fritchman model [66] is a Markov model with N states, which are subdivided into two state classes, namely error-free states (class A) and error states (class B). The transmitted symbols are error free in the class A and erroneous with probability one in the class B. In general, the possible state transitions are not restricted. Due to its complexity, we do not apply it to our trace files.

Other class of multiple states Markov models include Markov modulated process (MMP) model [67] and Hidden Markov model (HHM) [68]. MMP model uses MMP to approximate the first and second order statistics of the error rate measurement. It is not easy to apply to our trace file. For HHM, it uses only one state for the run sequence, leading to a prior independent of the run length. Moreover, it lacks a direct intuition between the channel
behavior and the underlying Markov chain.

4.1.4 Other Models

The main problem of aforementioned error models is the difficulty or complexity to model heavy-tail distributions of the run length, which occurs when the channel stays in the perfect state for long time. The chaotic map model [69] is proposed to model this kind of distribution. Similar to the Gilbert-Elliot model, the chaotic map model uses two states: ”good” and ”bad”. They are different, however, in deciding the erroneousness of a bit and in the rules to control the transition between two states. In the chaotic map model, the switching between states depends on the values of an auxiliary variable $x_t$, which is updated for each bit in a following way:

$$x_{t+1} = x_t + ux_t^z + \epsilon$$

The $u$, $z$ and $\epsilon$ are a set of parameters used to control the state transition between the good and bad state. When $x_{t+1} > 1$, the transition happens and a new number randomly chosen from $(0,1)$ is assigned to $x_{t+1}$. The advantage of the chaotic map model is its simplicity and flexibility. However, since a semi-Markov can properly model our trace files, we do not apply the chaotic map model in this chapter.

4.2 Trace File Analysis

The trace files are obtained from the simulation of a specific 60 GHz transceiver proposed in [21]. The channel models used in the simulation are given by the 802.15.3c channel modeling sub-group [70], which uses a modified Saleh-Valenzuela (S-V) statistical channel model [71] to model multipath channels in residential, office, library, desktop scenarios. Overall 8 channel models and 17 sub-models are defined by the 802.15.3c channel
modeling sub-group. Moreover, the directional antennas are taken into ac-
count by introducing the angle of arrival (AOA) information in the model.
We choose the residential LOS (channel model (CM) 1.1) and office LOS
scenario (CM 3.1) in our simulation. The parameters for the models are
listed in Table 4.1, where Λ is the cluster average arrival rate, λ is the pulse
average arrival rate, Γ is the power decay factor of clusters, γ is the power
decay factor of pulses within a cluster, σ_ξ and σ_ζ are the standard devia-
tion of the channel coefficients for clusters and pulses within each cluster,
respectively.

The error bit streams in trace files are recorded in the form of binary

<table>
<thead>
<tr>
<th>Channel</th>
<th>Λ (1/ns)</th>
<th>λ (1/ns)</th>
<th>Γ</th>
<th>γ</th>
<th>σ_ξ (dB)</th>
<th>σ_ζ (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>60 GHz CM1.1</td>
<td>0.19</td>
<td>1.22</td>
<td>4.46</td>
<td>6.25</td>
<td>6.28</td>
<td>13.0</td>
</tr>
<tr>
<td>60 GHz CM3.1</td>
<td>0.041</td>
<td>0.971</td>
<td>49.8</td>
<td>45.2</td>
<td>6.6</td>
<td>11.3</td>
</tr>
</tbody>
</table>

indicator sequences, in which the 0 denotes a correctly received bit, and
1 denotes an erroneous bit. The parameter k_0 is used to specify the runs
and bursts in the binary indicator sequence. A run of order k_0 is defined
as the maximum length of a consecutive all-zero subsequence with a length
of at least k_0; a burst of order k_0 is a subsequence that is at least one bit
long and with ones at its two ends. By this definition, a binary indicator
sequence b_1b_2...b_m of length m is segmented into n alternating runs and
bursts. The length of the j-th run is denoted as R_j, and the length of the
j-th burst is denoted as B_j, and L_j is the actual number of ones occurring
in the j-th burst. A length sequence of a binary indicator sequence is:

R_1, B_1, L_1, R_2, B_2, L_2...R_n, B_n, L_n

The run length sequence is defined as R_1, R_2...R_n, and the burst length
sequence is defined as B_1, B_2...B_n. Taking the binary indicator sequence
0010010100011000011 for example, the length sequences in order of $k_0 = 1$ and $k_0 = 2$ is:

\[
k_0 = 1 : \quad 2, 1, 1 \quad 2, 1, 1 \quad 1, 1, 1 \quad 3, 2, 2 \quad 4, 2, 2
\]

\[
k_0 = 2 : \quad 2, 1, 1 \quad 2, 3, 2 \quad 3, 2, 2 \quad 4, 2, 2
\]

From a length sequence, it is easy to get the simple statistics of binary indicator sequence, such as the mean error rate $\bar{e}$, and the mean error burst length $\bar{B}$:

\[
\bar{e} = \frac{\sum_{j=1}^{n} L_j}{\sum_{j=1}^{n} (R_j + B_j)}; \quad \bar{B} = \frac{1}{n} \sum_{j=1}^{n} B_j
\]

After obtaining the length sequences of the trace files, the empirical distri-

![Figure 4.1: CCDF of run lengths obtained from trace file 2 (Refer to Table 4.2)](image)

bution of a run/burst length sequence is computed by counting the number of observed runs/bursts that is smaller than a particular run/burst length. Let the random variable $L$ denote the length of a run/burst. The CDF function $F_L(l) = Pr(L \leq l)$ of $L$ expresses the probability that a run
length up to \( l \) occurs. We use the CCDF of \( F_L(l) \) for visualization purposes. The run length CCDF of two trace files are plotted in Fig. 4.1 and Fig. 4.2, respectively, in which the log-log plot is used to graphically emphasize the small probability values of the distribution function. As seen from Table 4.2, more than 90% of bursts have the length of one. For this sake, we focus our efforts on modeling the run length distribution.

In addition to the length distribution, the possible correlation between the run lengths, as expressed by their autocorrelation function (ACF), is an important metric to reveal the error pattern in a trace file. The autocorrelation function \( R(k) \) for lag \( k \in N \) of a discrete, (wide-sense) stationary stochastic process \( X_0, X_1, ... \) is defined as

\[
R(k) = \frac{E[(X_0 - \mu)(X_k - \mu)]}{\delta^2}
\]

where \( \mu \) and \( \delta^2 \) are mean and variance of \( X_k \), respectively.

The ACF of run length sequences in the trace file 2 and 5 are plotted in Fig. 4.3 and Fig. 4.4. For \( k \geq 1 \), the value of ACF is around 0.02, which is negligible. Therefore we do not model these patterns.

As seen from Table 4.2, even under a high \( E_b/N_0 \), the BER is quite large for a practical system. We use forward error correction (FEC) schemes to further reduce the BER, in which a block FEC scheme is applied.

<table>
<thead>
<tr>
<th>Trace file</th>
<th>CM</th>
<th>( Eb/N_0 ) (dB)</th>
<th>Length</th>
<th>Mean BER</th>
<th>Max run length</th>
<th>Max burst length</th>
<th>% of bursts with length 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12</td>
<td>2169304</td>
<td>0.0096</td>
<td>1035</td>
<td>4</td>
<td>91.83</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>15</td>
<td>2169304</td>
<td>0.0034</td>
<td>3394</td>
<td>4</td>
<td>92.3</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>18</td>
<td>2169304</td>
<td>0.0012</td>
<td>9654</td>
<td>4</td>
<td>91.77</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>3845016</td>
<td>0.0057</td>
<td>2071</td>
<td>4</td>
<td>96.36</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>15</td>
<td>3845016</td>
<td>0.001</td>
<td>10307</td>
<td>3</td>
<td>97.54</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>18</td>
<td>3845016</td>
<td>1.2e-004</td>
<td>46331</td>
<td>2</td>
<td>98.29</td>
<td></td>
</tr>
</tbody>
</table>

In a block FEC schemes [72], a block of \( k \) user bits is mapped onto \( n \) code
Figure 4.2: CCDF of run lengths obtained from trace file 5, (Refer to Table 4.2)

Figure 4.3: ACF of run lengths in trace file 2, (Refer to Table 4.2)
4.2. TRACE FILE ANALYSIS

Figure 4.4: ACF of run lengths in trace file 5, (Refer to Table 4.2)

bits, with \( n > k \). The ratio \( k/n \) denotes the code rate. The capability of a
general block code is measured by Hamming Bound [72], which states that
up to \( t \) errors can be corrected in a codeword of \( n \) bit length and \( k \) user
bits, only if the following relation holds:

\[
2^{n-k} \geq \sum_{i=0}^{t} \binom{n}{i}
\]

IEEE 802.3ah suggests the use of the Reed-Solomon (RS) code as the
optional FEC code [11]. The RS code [72] is a widely used block FEC
code that achieves the largest possible distance of any linear code. The
distance of two codewords is defined as the number of elements in which
two codewords differ. The linear codes is a class of codes with the property
that the linear combination of codewords is still in its code space. For the
RS code, the number of bits can be corrected by its codewords \((n,k,e)\) is
\( e = \frac{n-k}{2} \). We put the RS code on the top of the 60 GHz transceiver for
further BER performance improvement. Different user bit lengths are used
to obtain new trace files with improved mean BER. The CCDF of the trace file 5 after the correction of RS(255,251,2) is shown in Fig. 4.5, where the code length 255 is taken from the specification in IEEE 802.3ah.

### 4.3 Suitability of Existing Models

Various distributions are used to fit the empirical distributions of run lengths. Fig. 4.6, Fig. 4.7, Fig. 4.8, and Fig. 4.9 show the fitted CCDF curves of quantized exponential distribution, log-normal distribution, and gamma distribution as compared to the curve of the empirical distribution under different channel conditions and RS code settings. Fig. 4.6 and Fig. 4.7 show the exponential distribution fits the trace file 2 and 5 very well. However, in Fig. 4.8 and Fig. 4.9, as the mean BER becomes small, the heavy tails of the empirical distributions make the exponential distri-
4.3. SUITABILITY OF EXISTING MODELS

Figure 4.6: A set of distributions to fit CCDF of run lengths in trace file 2

Figure 4.7: A set of distributions to fit CCDF of run lengths in trace file 5
Figure 4.8: A set of distributions to fit CCDF of run lengths in trace file 2 after RS code

...distribution not suitable any more. The distribution of the choice goes to the
gamma distribution. It covers four trace files and provides good fit under
different BER conditions. Since the gamma distribution is a generalized
version of the exponential distribution. It is not surprise that the gamma
distribution fits the trace 2 and 5 very well.

To evaluate the suitability of models in a quantitative way, we use the
metric of area difference between two CDF functions, which compares two
distribution functions by the area between the functions. The area is ap-
proximated by the well known trapezium formula as follows:

\[ A = \frac{1}{N} \sum_{i=1}^{N} \left| \log(F_e^{-1}(i/N)) - \log(F_t^{-1}(i/N)) \right| - \frac{2N}{\left| \log(F_e^{-1}(1/N)) - \log(F_t^{-1}(1/N)) \right|} \]
where \( N \) determines the accuracy of the estimation. The reason to take logarithm in functions is to reduce the influence of long run lengths. We use the parameterized models to generate the CDF of run lengths, and compute the area difference between the CDF of the model and the trace. Table 4.3 summarizes the result of different analytical models with different trace files.

As shown in Table 4.3, a 2-state semi-Markov model with gamma distribution can be used for different BER conditions in different trace files. For the trace file without RS correction, the Gilbert-Elliot model can be deployed since the exponential distribution fits the empirical distribution well.
in those cases. For the trace files after RS correction, the log-normal distribution shows smallest area difference among three distributions. However, as shown from Fig. 4.8 and Fig. 4.9, the log-normal distribution does not fit the curve very well. Therefore, we do not use this distribution in our model.

As seen from Table. 4.2, the maximum burst length of all trace files is four, and 90% of burst lengths is below one. According to this, a simple method can be used to model the burst length distribution. Here we use a segment uniform distribution approach, in which a random variable \( r \in U(0, 1) \) is used to determine the state holding time in the bad state. For example, Eq. 4.1 shows the burst length distribution of the trace 5.

\[
\begin{align*}
0 &\leq Pr[r = 1] < 0.9 \\
0.9 &\leq Pr[r = 2] < 0.95 \\
0.95 &\leq Pr[r = 3] < 0.98 \\
0.98 &\leq Pr[r = 1] \leq 1
\end{align*}
\]

(4.1)

### 4.4 Conclusion

The objective of this chapter is to derive analytical error models of WiGEE. Based on a proposed 60 GHz transceiver [21], we generate trace files under multipath channel conditions. The run and burst length distributions in the trace files are analyzed, and corresponding analytical error models matching the empirical distributions of the run/burst sequences are derived. As a result, a semi-Markov is developed, which uses the exponential or gamma distribution to model the run length distribution, and segment uniform distribution to model the burst length distribution.
Chapter 5

Bandwidth Allocation Algorithm in WiGEE

This chapter discusses the upstream and downstream bandwidth allocation in WiGEE, where the former is dealt by DBA algorithms, and the latter is handled by downstream scheduling algorithms.

The DBA algorithm is the most flexible part of the system. It can be tailored to meet specific service requirements. For instance, to support real-time services, it is important to guarantee the QoS parameters such as bandwidth, delay and delay jitter of the traffic. It is the duty of the DBA algorithm to allocate the upstream bandwidth properly so as to satisfy those requirements.

In EPON, a variety of DBA algorithms have been proposed with the different emphases on the service support [73]. However, due to the significant difference between EPON and WiGEE, they cannot be directly utilized in WiGEE. The DBA algorithm of WiGEE should take the following factors into account:

- The allocation algorithm should jointly consider the downstream and upstream so as to support different duplex modes of MTs.

- The control of BS should be considered in the allocation algorithm in
order to improve the channel efficiency of the fiber optic infrastructure under different duplex modes of MTs.

- The allocation algorithm should avoid allocation collisions occurring at overlapping areas of multiple cells. An allocation collision is an event that occurs when an MT located in an overlapping area of multiple cells receives co-channel interference from other cells due to the uncoordinated allocation among multi-wavelengths.

Moreover, the error control scheme used in the link layer produces a certain amount of ARQ overhead according to the channel condition of each MT. For QoS to be guaranteed, it is better to consider ARQ overhead compensation in the DBA algorithm.

As discussed in chapter 2, the channel state information (CSI) can be used to improve the system performance of a wireless system. In a CSI aware scheduling algorithm, the scheduler gives the stations with better channel quality more priority to utilize the channel, while postponing the scheduling of the stations with worse channel quality. It is desirable to introduce the channel aware scheduling into WiGEE so as to increase its capacity. Usually, the CSI is obtained from the radio transceiver. However, in WiGEE, the CS, where the scheduler is located, cannot obtain the instantaneous CSI due to the separation of the CS and BSs. To enable the CSI scheduling in WiGEE, we develop another metric to present the long term CSI, which is derived from the overhead statistic of the SR-ARQ. A larger ARQ overhead in a given period indicates worse channel quality. Using this metric, the corresponding downstream scheduling algorithm and upstream DBA algorithm can be developed.

In this chapter, we first propose a DBA algorithm tailored for WiGEE, where the different duplex mode of MTs and the compensation for ARQ overhead are taken into account. Then in the second part, we develop the
metric to present the long term CSI per MT, and use it in the downstream scheduling algorithm. The performance of the proposed DBA algorithm and downstream scheduling algorithm are studied by simulations. The material presented in section 5.2 is also published in [16].

5.1 DBA Algorithm

A DBA algorithm tailored for WiGEE is proposed in this section. The algorithm is capable of allocating multiple wavelengths on the superframe basis. According to the duplex mode of MTs, we describe the operation of the DBA algorithm in two cases: the case where all MTs operate in the half duplex mode, and the case where all MTs operate in the full duplex mode. The difference between them is that in the former, the DBA algorithm should take the downstream into account as it shares the channel with the upstream. We do not extend the DBA algorithm into the hybrid duplex mode for two reasons: most likely the MTs will be operated in either the half duplex or full duplex mode; the algorithm for the hybrid duplex mode is the combination of the algorithm under the half and full duplex mode.

In designing DBA for WiGEE, we have two assumptions of the system. First, the CS knows which cell each MT belongs to. This information can be obtained from the collaboration of the MPCP sub-layer and out-of-band control plane. Moreover, if the MT is located in the overlapping area of multiple cells, the CS knows exactly which BSs serve the cells. Second, out-of-band control plane has perfect control on BSs. The interference due to the inaccurate control is avoided.

The proposed algorithm works as follows:

1. The bandwidth request message of an MT is issued either at the end of its granted upstream period, or in the random access period of the current superframe. A bandwidth request message includes the
lengths of the priority queues in the MT.

2. The bandwidth requests of MTs are collected at the end of each superframe period. Since the superframes of all wavelengths are synchronized, the CS is able to collect the bandwidth requests from all MTs before starting a new round of superframes.

3. Next, the algorithm allocates the superframes of wavelengths one by one until all wavelengths are allocated. Since the earlier processed wavelengths have more priority to utilize the channel, for fairness, in each round of allocation, the order to allocate wavelengths is randomly decided. The allocation process in a wavelength is called *intra-wavelength allocation.*

4. In the *intra-wavelength allocation*, the MTs are grouped based on the BSs they belong to. The allocation algorithm processes the groups one by one in order. Since the earlier processed groups has higher priority to access the channel, the order to process the group is randomly picked during each round.

Fig. 5.1 and Fig. 5.2 show the DBA algorithm works in half duplex and full duplex, respectively. Note that the propagation delay of the system is taken into account in the allocation. As illustrated in Fig. 5.1, the beacon is issued when the upstream transmission is not over.

As shown in Fig.5.1, after the MTs in a BS is processed, the holes of downstream or upstream are left in the superframe due to the allocation to half duplex MTs. The MTs of the next BS are permitted to utilize those holes. In this case, the BS should be well controlled so as to avoid allocation collisions. As seen from Fig.5.1, the gray area of the BS control denotes the downstream or upstream should be shut down accordingly. As a result, the statistic multiplex is achieved, which is shown in the downstream and upstream part in the CS.
Figure 5.1: Proposed DBA algorithm in half duplex mode
The CS calculates the grant to an MT in a whole according to the occupied states of its priority queues. To avoid an MT exhausts the resource, the maximum grant size is limited by a parameter called maximum transmission window (MTW).

In the following, we describe the bandwidth allocation for the MTs under two duplex modes, respectively.

In the full duplex mode, let $T_{sf}^j(l)$ denote the time duration of the su-

![Figure 5.2: Proposed DBA algorithm in full duplex mode](image)

perframe $j$ on the wavelength $\lambda_l$, excluding the beacon period; $T_{ctrl}^j(l)$
5.1. DBA ALGORITHM

denote the overall duration of upstream control periods in the super-frame \( j \), which include the MT presence, discovery, and random access period; \( T_{US}^j(l) \) denote the available upstream time for data transmission in the superframe \( j \); \( T_{BS}^{jBS(k)} \) denote the available upstream time for the BS \( k \) in the superframe \( j \); \( E_{jBS(k)}^j \) denote the time period in \( T_{US}^j(l) \) that cannot be assigned to the BS \( k \) for allocation collision avoidance; \( W_j(u) \) denote the amount of grant assigned to the MT \( u \) in bit; \( r_j(u) \) denote the overall requested bandwidth of the MT \( u \) in bit; \( c_j(u) \) denote the additional bandwidth assigned to the MT \( u \) in bit; \( W_{MTW} \) denote the maximum transmission window in bit; \( R \) denote the data rate of the system. Unless otherwise mentioned, the superframe round \( j \) in each notation is ignored in the remaining of the section.

In the full duplex mode, the amount of available upstream time in the superframe \( j \) of the wavelength \( \lambda_l \) is

\[
T_{US}(l) = T_{sf}(l) - T_{ctrl}(l)
\]

Assume the DBA algorithm now allocates the bandwidth for MTs belonging to the BS \( k \), where the BS \( k \) is served by the wavelength \( \lambda_l \). The time period \( E_{jBS(k)}^j \) that cannot be assigned to the BS \( k \) is

\[
E_{jBS(k)}^j = \frac{1}{R} \sum_{\lambda_m \in A} \sum_{MT_n \in B} W(n)
\]

where \( A \) is the set of allocated wavelengths in superframe \( j \), and \( B \) is the set of allocated MTs in superframe \( j \) that do not belong to the BS \( k \) but in the coverage area of the BS \( k \). The upstream time can be used by the BS \( k \) is

\[
T_{BS}^{jBS(k)} = T_{US}(l) - E_{jBS(k)}^j
\]

The order to allocate the MTs served by the BS \( k \) can be decided randomly or according to a certain rule, for instance, according to the
queue states of MTs or the QoS requirements of the queues. Assume it is the turn to allocate the MT $u$, the amount of bandwidth assigned to the MT $u$ is

$$W(u) = \min \left\{ \begin{array}{l}
\lfloor T_{BS}^{(k)}/R \rfloor \\
r(u) + c(u) \\
W_{MTW}
\end{array} \right. \quad (5.1)$$

where $\lfloor \cdot \rfloor$ is the flooring operation.

The start time and duration of the grant are determined accordingly. Note that $W(u)$ denotes the overall amount of grant assigned to the MT $u$, it may consist of several discontinuous time periods in the superframe due to the constraint of $E_{BS}^{(k)j}$, where may divide $T_{BS}^{(k)}$ into multiple segments. Fig. 5.3 shows the allocation constraint imposed by $E_{BS}^{(k)j}$. After the allocation, $T_{BS}^{(k)}$ is updated by:

$$T_{BS}^{(k)} = T_{BS}^{(k)} - W(u) \times R - T_g$$

where $T_g$ is the guard time, which takes a corresponding value according to the number of grant segments assigned to the MT $u$.

The allocation process continues until $T_{BS}^{(k)} = 0$ or all MTs belonging to the BS $k$ get assigned. Then it turns to the next unallocated BS served by the wavelength $\lambda_l$.

The DBA algorithm operating in the half duplex mode should takes the downstream time into account. We extend the aforementioned notations as following: let $\hat{T}_{jBS}^{(k)}$ denote the available downstream and upstream time for data transmission in the superframe $j$; $\hat{T}_{jBS}^{(k)}$ denote the downstream time assigned to the BS $k$ in the superframe $j$; $\hat{E}_{jBS}^{(k)}$ denote the time period in $\hat{T}_{jBS}^{(k)}$ that cannot be assigned to the BS $k$ for allocation collision avoidance; $T_{US}^{(k)j}$ denote the available upstream time for the BS $k$ in the superframe $j$; $D(u)$ denote the amount of downstream periods assigned to the MT $u$ in bit. Again,
Figure 5.3: Proposed DBA algorithm deals with overlapping area, in half duplex mode.
we ignore the $j$ in the notations in the following section.

The available downstream and upstream time in the superframe $j$ is

$$\hat{T}^{BS(k)} = T_{sf}(l) - T_{ctrl}(l) - \hat{E}^{BS(k)}$$

where

$$\hat{E}^{BS(k)} = \frac{1}{R} \sum_{\lambda_m \in A} \sum_{MT_n \in B} (W(n) + D(n))$$

in which $A$ is the set of allocated wavelengths in superframe $j$, and $B$ is the set of allocated MTs in superframe $j$ that do not belong to the BS $k$ but in the coverage area of the BS $k$. The available upstream time for the BS $k$ in the superframe $j$ is

$$T^{BS(k)}_{US} = \hat{T}^{BS(k)} - \hat{T}^{BS(k)}_{DS}$$

Assume it is the turn to allocate the MT $u$, the amount of bandwidth assigned to the MT $u$ is

$$W(u) = \min \left\{ \left\lfloor \frac{T^{BS(k)}_{US}}{R} \right\rfloor, r(u) + c(u), W_{MTW} \right\}$$

(5.2)

The start time and duration of the grant is determined accordingly. After the allocation, $T^{BS(k)}_{US}$ is updated by:

$$T^{BS(k)}_{US} = T^{BS(k)}_{US} - W(u) \times R - D(u) - T_g$$

The remaining process to complete the allocation is similar to that in the full duplex mode.

According to the specific implementations, $c_j(i)$ in Eq. 5.1 can take the values listed in Table. 5.1. At the end of intra-wavelength allocation, an allocation map for downstream and upstream, and corresponding control map of BSs for the next superframe are generated.
5.1. DBA ALGORITHM

Table 5.1: Compensation algorithms in DBA

<table>
<thead>
<tr>
<th>Formula</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_j(i) = c$</td>
<td>A constant value $c$, which could be zero.</td>
</tr>
<tr>
<td>$c_j(i) = \alpha \cdot O_{ARQ}^{(j-1)}(i)$</td>
<td>$O_{ARQ}^{(j-1)}(i)$ is the ARQ overhead of the previous round, $\alpha$ is a constant</td>
</tr>
<tr>
<td>$c_j(i) = \beta \cdot A_{CSI}(i)$</td>
<td>$A_{CSI}(i)$ is the award according to the historical channel condition of the MT $i$, and $\beta$ is a constant</td>
</tr>
<tr>
<td>$c_j(i) = c + \alpha \cdot O_{ARQ}^{(j-1)}(i) + \beta \cdot A_{CSI}(i)$</td>
<td>The combination of above three.</td>
</tr>
</tbody>
</table>

5. After one wavelength is allocated, next wavelength is randomly picked from unallocated wavelengths for allocation. The process continues until all wavelengths are allocated. At the end, the allocation maps of the downstream and upstream, and the control maps of BSs over all wavelengths are generated. The allocation of the next superframe is performed based on those maps. The grants of the upstream are sent at the first downstream period of the next superframe round.

6. When an MT moves to a new cell belonging to a different wavelength, the DBA algorithm reallocates the grant to that MT in the next superframe of the new cell. If the reallocation is permitted, the upstream and downstream allocated to that MT in the old cell can be revoked and reallocated to other MTs.

5.1.1 Compensate ARQ Overhead in DBA

Since the ARQ overhead in WiGEE consumes partial of grant period, it is desirable to compensate this part of overhead in the DBA algorithm so as to guarantee the QoS.

In the simplest compensation algorithm, a fixed amount of bandwidth is assigned in each grant. However, due to the randomness of the errors, the ARQ overhead in each grant is unexpected. It is reasonable to compensate the amount of grant proportional to the ARQ overhead observed in
the previous grants. We propose the use of the following compensation algorithm: Let \( N_{o}^{j-1}(u) \) denote the ARQ overhead of the MT \( u \) observed at the previous grant, \( N_{c}^{j-1}(u) \) denote the compensated value given to the previous grant \( j-1 \) of the MT \( u \), and \( N_{c}^{j}(u) \) denote the compensated value given to the following grant of the MT \( u \), we have:

\[
N_{c}^{j}(u) = \beta \cdot ((1 - \alpha)N_{o}^{j-1}(u) + \alpha N_{c}^{j-1}(u))
\] (5.3)

where \( \alpha (0 \leq \alpha \leq 1) \) is the smoothing factor that determines the weight of the old compensated value, and \( \beta (0 \leq \beta \leq 1) \) is the weight that determines the actual compensated amount. The choice of the parameter \( \alpha \) is determined by the channel conditions. In a slow fading channel, a big \( \alpha \) is used to smooth the fluctuation of the compensated amount. For the parameter \( \beta \), a big value can provide extra grant for compensation, however, in the cost of wasting bandwidth and leading to unfairness.

5.1.2 Simulation Result

We study the performance of the DBA algorithm operating in the full duplex mode. The simulation setup is similar to that in chapter 3, however, with several differences. Firstly, the error models proposed in chapter 4, which include the Gilbert-Elliot model and semi-Markov model with gamma distribution for run lengths, are employed for performance comparison with the BSC error model. Note that we only use the Gilbert-Elliot model and semi-Markov model to simulate the high BER conditions. For the BER greater than \( 1e^{-6} \), the errors generated by different models have limited impact on the system performance. Therefore only the BSC model is utilized for the low BER simulation. We use the trace file 6 in chapter 4 to derive the parameters of proposed error models.

Secondly, in addition to fixed frame size, various frame size with a length distribution matching the Internet traffic [74] is applied in the simulation.
5.1. DBA ALGORITHM

Fig. 5.4 shows the throughput in the fixed frame size case, in which the

![Figure 5.4: Average throughput as a function of offered load per wavelength, SR-ARQ, BSC error model, frame size of 512 bytes, stationary scenario](image)

BSC model is employed. Firstly, it shows that the number of the nodes has minor impact on the throughput when the node size is medium. The node number 2, 4, and 8 per cell have similar throughput performance. Secondly, under the BER conditions of $1e^{-6}$ and $1e^{-9}$, the throughput experiences a linear increase with the amount approaching the load, which shows the DBA algorithm achieves good performance at the good channel quality. As the BER degrades to $1e^{-4}$, the throughput under the load lower than a certain value is not affected. However, as the load continues increasing, the throughput keeps almost constantly. That constant value is related to the BER condition and the packet length. The bandwidth in this case is wasted by the error packets and ARQ control overhead.

Fig. 5.5 shows the throughput in the various frame size case. It illustrates the DBA algorithm achieves good throughput performance as similar to
Chapter 5. Bandwidth Allocation Algorithm in WIGEE

Figure 5.5: Average throughput as a function of offered load per wavelength, SR-ARQ, BSC error model, various frame size, stationary scenario

Figure 5.6: Average delay as a function of offered load per wavelength under BSC model, SR-ARQ, various frame size, stationary scenario
5.1. DBA ALGORITHM

Figure 5.7: Average throughput as a function of offered load per wavelength under different error models, SR-ARQ, various frame size, stationary scenario

Figure 5.8: Average delay as a function of offered load per wavelength under different error models, SR-ARQ, various frame size, stationary scenario
the load when the channel quality is good. Under a high BER condition, it exhibits the same result as shown in Fig. 5.4. The normalized throughput is less than 20% under the BER of $1e^{-4}$, suggesting the exploiting of multiuser diversity to increase the overall system throughput.

Fig. 5.6 shows the average delay in the various frame size case. From the figure we can see the average delay under the BER of $1e^{-6}$ and $1e^{-9}$ is less than 2 ms. It shows the DBA algorithm has good delay performance under low BER conditions small. As the BER becomes $1e^{-4}$, the delay increases significantly. Moreover, a higher node number per cell produces higher delay. The impact of node number on the average delay is prominent when the BER is high.

Fig. 5.7 shows the throughput performance under different error models. When the mean BER is $1.24e^{-4}$, which is the mean BER of trace file 6 in chapter 4, the DBA algorithm operating under different error models shows the similar behavior as in Fig. 5.5, i.e. the throughput at a high load is limited to a certain value. However, the achieved throughput under the Gilbert-Elliot model is slightly higher than that under the semi-Markov model, and the throughput under the BSC model is slightly higher than that under the Gilbert-Elliot model. Fig. 5.8 reveals the influence of error models on the average delay of the DBA algorithm. In line with Fig. 5.7, the BSC model has better delay performance than the Gilbert-Elliot model, and the latter has better delay performance than the semi-Markov model. These two figures exhibit the influences of error models on the system performance. Since the semi-Markov model is more accurate to model the trace file, it suggests the use of the semi-Markov model to study the performance of the DBA algorithm under bad channel conditions.
5.2 Downstream Scheduling Algorithm

In the following section, we develop two channel aware downstream scheduling algorithms that are suitable for the architecture of WiGEE. The proposed algorithms use the long term channel state information estimated from the error control scheme of the link layer as the metric to achieve the multi-user diversity gain in WiGEE. The first algorithm jointly considers the channel quality and queue states for throughput optimization, and the second algorithm tries to achieve long term throughput fairness among MTs.

Note the same metric can be used in the DBA algorithm of the upstream, e.g. to determine the $c(u)$ in Eq. 5.1 and Eq. 5.2. However, an additional mechanism is needed to report the channel state information from the MT to the CS, which needs an extension in the MPCP protocol. For this sake, we only discuss the downstream scheduling algorithm in this section.

5.2.1 Background

The channel aware scheduling problem in multi-access wireless networks has been studied extensively in last decade [75]. The motivation behind the channel aware scheduling is the asynchronous fading experienced by terminals of a network due to their location and movement. The basic idea is to postpone the transmission of the users with bad channel conditions, while giving the priority to the users with good channel conditions [26]. It has been proven that this approach can significantly improve the system capacity [76].

Normally, channel aware scheduling algorithms exploit the instantaneous channel state information (CSI) estimated from the radio receiver of the physical layer to achieve multiuser diversity gain. The channel state dependent packet (CSDP) scheduling [26] is one of the first papers addressing
channel aware scheduling issues, in which the link with better channel quality is always scheduled with higher priority. The fairness issue arises in the CSDP scheduling since nodes with bad channel conditions receive less bandwidth. A variety of compensation techniques are introduced to solve this issue [77], [78], in which an error-free reference model is running at the background to count the amount of services each node should receive in an error free condition. If a node with a bad channel condition lags behind its reserved service rate, the scheduler compensates the node the lagged amount of bandwidth when its link is recovered from the bad state. Usually the generalized processor sharing (GPS) is employed as the reference model [79].

Early algorithms simply divide the channel quality into the good and bad state, and thus are inadequate to fully exploit the multiuser diversity gain [26], [77], [78]. The recent developed channel-aware scheduling algorithms use the achievable rate of the wireless link to represent the channel quality [80], [81], [82], [27]. Among them, the proportional fair (PF) algorithm [80] is the most studied algorithm, which achieves a balance between the maximum overall throughput and the fairness. The proportional fairness is defined as the fairness achieved by an allocation so that any change to that allocation leads to a negative average change on the bandwidth share [83]. The PF algorithm is known to be not stable [84]. To solve this, the queue length of the node is considered in the scheduling. The most widely studied algorithms includes Max-Weight type algorithms [85], [86], which always serve the user that maximizes the product of the achievable rate and the queue length, and the exponential rule algorithm [82], which provides a more complex approach to avoid queue overflow. Moreover, maintaining service rates in desired proportions among the users is studied in [87]. Scheduling algorithms usually need to meet certain resource constraints, for instance, the minimum rate constraint required by an application, or
the maximum rate constraint imposed by a system operator. The minimum and maximum rate constraint issues are studied in [88], and the time fraction constraint is considered in [27]. Normally, the optimization problem of the channel-aware scheduling is expressed by utility functions [27], [89], [88]. The linear utility function is studied in [27] and concave utility function is studied in [88].

Most channel-aware scheduling algorithms rely on the instantaneous CSI obtained from the physical layer. Due to the design constraints of some wireless systems [15], this kind of information may not be available to the scheduler. Instead of looking for the instantaneous CSI, the long term CSI extracted from the error control scheme of the link layer can be used for the same purpose. In this section, we propose a metric to denote the long term CSI, and utilize it in downlink scheduling algorithms. Note that our approach to estimate the channel is different from that of [90], in which a single transmission event may conclude the channel in the good or bad state. Due to the time varying nature of wireless channel, the estimation in [90] may not reflect the real state of the channel.

5.2.2 System Model

To quantize the assignment of bandwidth, the downstream channel time is divided into time slots of fixed duration. A time slot is the minimum time unit assigned to an MT. One or several variable size data packets are allowed to be transmitted in one time slot, while the number is determined by the length of time slot and packet sizes.

The CS in WiGEE obtains the long term stationary channel quality of MTs from the error control scheme at the link layer. We define the metric of the long term channel quality as the ratio of throughput to load in a given period, denoted as \( \frac{\text{throughput}}{\text{load}} \) (TLR). Note in the fixed length packet case, the \( 1 - \text{TLR} \) is equal to the packet error rate (PER),
however, in other cases they are different. Consider a case where two MTs have same PER of 0.2 but different packet length distributions. Assume MT1 receives 1000 packets with same length of 64 bytes, among which 200 packets are corrupted; MT2 receives 1000 packets with 500 packets of 64 bytes and 500 packets of 1500 bytes, among which 200 longer packets are corrupted. The TLR for MT1 is 0.2, but for MT2 is 0.38. Obviously these two MTs should be treated differently by the scheduler.

Under the control of SR-ARQ, the CS detects downlink transmission errors in two ways: by selective reject messages or by the retransmission timer. To guarantee orderly delivery of packets at the link layer, each downlink packet is stamped by a sequence number. In an error free transmission condition, an MT receives data packets in order. The receiving two successive packets with disjointed sequence numbers imply the occurrence of transmission errors. In case the sequence numbers of error packets can be inferred by the ARQ entity in the MT, explicit selective reject messages are sent to the CS, indicating the sequence number of the error packets. If the MT fails to identify all errors, the remaining errors are detected by the retransmission timer at the CS. The CS sets a retransmission timer for each outstanding packet. The retransmission timer is revoked if the acknowledgement of a packet is received in time. Otherwise, the packet is regarded being lost and retransmitted by the transmitter.

To use the proposed mechanism, the stationary scenario is assumed, in which the MT is fixed or moves in a relatively slow speed. Moreover, the channel experiences slow fading. Therefore, the channel state between BS and each MT keeps relatively stable in a long period. However, different MTs may own different channel states, which are related to the locations of the MTs.

Note that the metric can be used to schedule real time services which can tolerate a certain level of errors and may not require ARQ at the link layer.
5.2. DOWNSTREAM SCHEDULING ALGORITHM

to provide reliable transmission.

5.2.3 Channel Estimation

The channel state of each MT is estimated independently at the CS by tracking the average downlink TLR of the MT, which is obtained through counting the amount of transmitted packets and error packets in bit periodically. The counting period is a variable time interval during which the CS transmits a given number of packets, denoted as $N_{TXM}$, to an MT. The light load to an MT may lead to a long counting period. To avoid this, TLR estimated interval, denoted as $T_{est}$, is applied to update TLR immediately when the counting period exceeds $T_{est}$. If no data transmitted to an MT in $T_{est}$, the CS uses the previous TLR of the MT for scheduling. Let $A_{txm}(i)$ denote the amount of transmitted packets in bit, and $A_{err}(i)$ denote the amount of error packets in bit estimated for MT $i$ in a given counting period. The TLR of MT $i$ in this period, denoted as $TLR(i)$, is $TLR(i) = 1 - A_{err}(i)/A_{txm}(i)$. Due to the time varying nature of a wireless channel, the value of estimated TLR is expected to be fluctuating over time. The exponential moving average technique is employed to smooth the TLR. The average TLR for MT $i$ is obtained by

$$TLR_k(i) = \alpha \times TLR_{k-1}(i) + (1 - \alpha) \times TLR_k(i)$$

where $k$ is the counting period index, $\alpha$ is the smoothing factor, and $TLR_0(i) = 1$. Unless state otherwise, we use TLR($i$) to denote $TLR_k(i)$. Note the choice of $\alpha$ determines the weight of estimated TLR on the average TLR.

5.2.4 Scheduling Algorithms

The channel-aware scheduling algorithms based on instantaneous CSI usually set the optimization goals as overall throughput maximization or fair-
ness. In this section, we show the long term CSI can be used to achieve same optimization goals.

We propose two scheduling algorithms with different optimization goals. The throughput optimal (TO) algorithm jointly considers the channel quality and the queue length at the CS for throughput maximization and queue stabilization. Under similar queue lengths, the CS schedules lower TLR MTs with higher priorities. However, the queue length becomes the dominant factor to determine the scheduling priorities when the queue length of an MT is significantly larger than other MTs. The queue overflow issue is ameliorated thereafter.

The second algorithm is the throughput fairness (TF) algorithm, which provides throughput fairness among MTs in the sense that similar throughput is achieved at the receivers if the loads at the transmitters are similar. It is realized by compensating the retransmission overhead of lower TLR MTs with extra bandwidth. The compensation is achieved through a stochastic approximation scheme. The basic idea is the following: The throughput fairness criteria are transformed into a set of target probabilities that reflect the chance an MT wins a downlink time slot. The proportion of the downlink slots that an MT has received is compared with the corresponding target probability. The MT with the largest difference wins the slot.

**TO Algorithm**

The TO algorithm trades the overall downlink throughput with the downlink queue stability. Under the same level of queue occupancies, the algorithm schedules the MT with highest potential throughput, which is defined as the product of $TLR(i)$ and the link rate $R_i$ of the MT $i$. It is likely an MT with a high link rate but low TLR frequently wins time slots. As a result, the system achieves maximum throughput, but mean-
5.2. DOWNSTREAM SCHEDULING ALGORITHM

while produces plenty of retransmissions. To solve this, we set a threshold $T L R_{tsh}$ to avoid scheduling low TLR MTs.

Another important concern in the TO algorithm is to achieve the queue stability in the CS. For the physical constraint and cost reason, the CS usually has limited amount of queue length for each MT. Considering the fact that data traffic usually arrives in burst, it is highly likely the queues of some MTs is occupied more than others. If the scheduler pays no attention on the queue occupancy, the continuous arrival of traffic may eventually overflow the queue. To solve this, we bring the exponential scheduling rule [82] into the algorithm and give nearly exhausted queues higher priorities to be scheduled.

For each time slot, the scheduling priority of each MT is computed as follows:

$$\varphi_i = w(i) \times TLR(i) \times \frac{R_i}{R_N} \times \exp \left( \frac{Q_i(t)}{\mu + Q(t)} \right)$$  \hspace{1cm} (5.4)

where $w(i)$ is the weight assigned to the MT $i$, $R_i$ is the line rate of the MT $i$, $R_N$ is the maximum line rate of the system, $\mu$ is a positive constant, $Q_i(t)$ is the queue length of the MT $i$ at the scheduling time, $Q(t) = \sum_i Q_i(t)/M$ is the average queue length of all MTs, and $M$ is the number of MTs in the system. The scheduling algorithm always schedules the MT with the highest priority, i.e.

$$i = \arg \max_j \{ \varphi_j \}$$

Ties are broken randomly. Once determined the qualified MT, the whole downlink frame is assigned to that MT. In [82], Shakkottai et al have proven the queue stability can be achieved in a scheduling rule similar to Eq. 5.4.

A potential starving problem may occur in above algorithm. When the TLR of an MT is lower than $T L R_{tsh}$, the MT loses the chance to win time slots and keeps the MT unserved. To solve this, the aforementioned
TLR estimation algorithm is slightly modified. If $TLR(i) \leq TLR_{tsh}$, the $TLR(i)$ of MT $i$ is increased by

$$TLR(i) = \begin{cases} \lambda \times TLR(i) & \text{if } \lambda \times TLR(i) \leq 1 \\ 1 & \text{if } \lambda \times TLR(i) > 1 \end{cases}$$

at each $T_{est}$, where the $\lambda \in (1, 2]$ is a positive constant. $\lambda$ determines the delay the MT $i$ will experience for scheduling. A small $\lambda$ provides the system a backoff mechanism that prevents an MT with bad channel from wasting the downlink bandwidth.

The scheduling algorithm is summarized as the following:

1. Update the TLR of each MT independently according to the algorithm.

2. At the beginning of the downlink frame, obtain $R_i, Q_i(t)$, and calculate $\overline{Q(t)}$.

3. Calculate the $\varphi$ of each MT according to Eq. 5.4.

4. Schedule the MT $i$ where $i = \arg \max \{ \varphi_i \}$ with the whole downlink frame.

**TF Algorithm**

The idea of the TF algorithm is to schedule smaller TLR MTs more often so as to compensate their retransmission overhead. The objective of this algorithm is to make the throughput of an MT proportional to its load without respect to channel conditions. The specific feature of the algorithm is that it provides fairness at the goodput level. The reason is straightforward since the TLR itself indicators the normalized goodput. Moreover, the TF algorithm takes into account the situation where MTs have different data rates, which are treated in two ways: the algorithm
provides same throughout to MTs no matter which data rates they adopt, or the level of throughput is proportional to the data rate of the MT. For the latter it means after normalized by their data rate the throughput of MTs is similar. Let us denote the former as TF1 and the latter as TF2. Note that in order to provide error free transmission, the MT with a small TLR may produce a large number of retransmission overhead. The same approach applied in TO algorithm is used to suppress the scheduling of low TLR MTs.

At the beginning of each downlink frame, the MTs satisfying that $TLR(i) > TLR_{tsh}$ and their queues in the CS are not empty are selected into the scheduling set $S$. The time slots of downlink frames are allocated to the MTs according to a stochastic scheduling policy similar in [27],

$$Pr\{U(S) = i\} = \gamma_i$$  \hspace{1cm} (5.5)

where $i = 1, ..., M$ is the index of MT in the set $S$, $M$ is the number of MTs in the set $S$, $\gamma_i$ is the probability that the MT $i$ should be scheduled, and $U(S) = i$ is the scheduling policy that chooses the MT $i$ from the set $S$ for scheduling. $Pr\{U(S) = i\}$ is a parameter maintained in the CS for all MTs, which reflects the historical scheduling information of the MT $i$. To get $\gamma_i$, we first calculate the weight $\beta$ of each MT, i.e

$$\begin{cases}
\beta_i w(i) TLR(i) R_i = \beta_j w(j) TLR(j) R_j & \text{in TF1} \\
\beta_i w(i) TLR(i) = \beta_j w(j) TLR(j) & \text{in TF2}
\end{cases}$$  \hspace{1cm} (5.6)

for all $i, j \in S$. $w(i)$ in Eq. 5.6 is the weight assigned to the MT $i$. $\gamma_i$ is obtained by

$$\gamma_i = \frac{\beta_i}{\sum_{j \in S} \beta_j}$$

$\gamma_i$ is inversely proportional to $TLR(i)$. A lower $TLR(i)$ implies the higher probability an MT will win a time slot.
To map Eq. 5.5 to a deterministic scheduling algorithm, we use the parameter $c_i$ for MT $i$ to track the historical scheduling information. The initial value of each $c$ is set to zero. $c_i$ is updated according to certain rules, i.e.

$$
\begin{cases}
  c_i = c_i / 2 & \text{if } Pr\{U(S) = i\} > \gamma_i \\
  c_i = c_i + \gamma_i & \text{if } Pr\{U(S) = i\} \leq \gamma_i
\end{cases}
$$

(5.7)

Then the MT $i$ with $i = \arg\max_{j \in S}(c_j)$ is scheduled, i.e. the MT with maximum $c$ in the set $S$ is scheduled. Ties are broken randomly.

After scheduling, $Pr\{U(S) = i\}$ is updated by

$$
Pr\{U(S) = i\} = \eta Pr\{U(S) = i\} + (1 - \eta)1_{U(S)}
$$

(5.8)

where $1_{U(S)}$ is an indicator function that

$$
1_E = \begin{cases}
  1 & \text{If E occurs} \\
  0 & \text{Otherwise}
\end{cases}
$$

and $\eta$ is a smooth factor.

In the following, we prove that the throughput fairness can be achieved by the TF algorithms. Here we assume the channel quality and data rate of each MT is fixed during the lifetime of the MT in the network. Accordingly, each MT has fixed but may different TLR. $TLR(i)$ of the MT $i$ can be expressed as $TLR(i) = Thpt(i)/L(i)$ where $Thpt(i)$ is the average throughput of the MT $i$, and $L(i)$ is the average load corresponding to $Thpt(i)$. Since $w(i)$ is a constant for all MTs, for simplicity, we set $w(i)$ to 1 for all MTs. Let $\beta_i TLR(i) R_i = 1$ for $i = 1 \ldots M$, we have

$$
\beta_i = \frac{1}{TLR(i) R_i}
$$

which is also fixed for each MT. As a result, we get that $\gamma$ is fixed for each MT, and

$$
\gamma_i = \frac{1}{\sum_{j \in S} \beta_j TLR(i) R_i} \frac{L(i)}{R_i}
$$

(5.9)
According to the proposed TF algorithm, after a long run, $Pr \{U(S) = i\} \approx \gamma_i$ can be achieved. $Pr \{U(S) = i\}$ can be expressed as

$$Pr \{U(S) = i\} = \frac{T(i)}{T_a}$$

(5.10)

where $T(i)$ is the slotted time assigned to the MT $i$, and $T_a$ is the overall slotted time in a long observation window. Combining Eq. 5.9 and Eq. 5.10, we get

$$Thpt(i) \approx \frac{1}{\sum_{j \in S} \beta_j R_i T(i)} T_a L(i)$$

Since $R_i T(i) = L(i)$, each MT has the same level of throughput.

For TF2, we have

$$Thpt(i) \approx \frac{1}{\sum_{j \in S} \beta_j} T_a L(i)$$

$$Thpt_N(i) = \frac{Thpt(i)}{R_i} \approx \frac{T_a}{\sum_{j \in S} \beta_j}$$

which is the same of all MTs. As a conclusion, the TF1 and TF2 algorithms are throughput fairness.

As the same problem in the TO algorithm, in case $TLR(i) \leq TLR_{tsh}$, the MT $i$ will lose the chance to be scheduled and keep starving. We use the same approach in TO algorithm to solve this problem.

The scheduling algorithm is summarized as the following:

1. At the beginning of each downlink frame, the CS obtains the set $S$.

2. Get $\gamma_i$ for all MT in the set $S$.

3. Update $\{c_i\}$ according to Eq. 5.7.

4. Schedule the MT $i$ one time slot if $i = \arg\max_{j \in S} (c_j)$.
5. Update $Pr\{U(S) = i\}$ according to Eq. 5.8.

6. Repeat step 3) until all time slots of the downlink frame are scheduled

### 5.2.5 Simulation Study

The simulation is developed based on the OMNET++ 3.1 simulation platform [57]. Besides the proposed algorithms, we develop a round robin algorithm for the performance comparison, which simply schedules the MTs in a round robin fashion. Simulations are run under a single data rate or multiple data rates. The overall throughput of the system and the individual throughput of each MT under different algorithms are compared.

#### Channel Model

The channels between the CS and each MT are assumed independent of each other. The channel is modeled by binary symmetric channel model [64]. We assume all channels are relatively stable, and a predefined Bit Error Rate (BER) is assigned to each channel according to the location of each MT.

#### Parameter Setting

Fixed packet length is used in the simulation. We perform the algorithms under three packet sizes: 128, 512, and 1024 bytes. The duration of the downlink frame is set to 1 ms. For three packet sizes, the time slots in a downlink frame is 56, 14, and 7, respectively. If one packet occupies a time slot, the maximum load of the system is 57 Mbps, which is similar to the data rate of the IEEE 802.11 a/b/g systems. The multiple rate support of the system is implemented by allowing multiple packets being transmitted in a time slot. Different MTs may have different data rates, however, their data rates are fixed during the simulation.
The number of MTs in different simulation runs varies from 4 to 20. The channel condition of each MT is summarized in Table 5.2. In the single rate case, only one packet is transmitted in each time slot of the downlink frame. Therefore the maximum load of the system is 57 Mbps. In the multiple data rate case, we only simulate the system with 4 MTs and under the packet size of 512 bytes. The channel condition of each MT in this case is listed in Table 5.2. The rates of 4 nodes are set according to the BERs of MTs. From the low BER to the high BER, the number of packets transmitted in a time slot is 4, 3, 2, and 1, respectively.

The traffic arrival process for each MT is a Bernoulli distribution characterized by two parameters: the average data rate $r$ and probability $p$. In each cycle, the arrival rate for a queue is zero with probability $p$, and $r/(1-p)$ with probability $1-p$. The parameters of the arrival process are chosen so that the saturated throughput of the system is achieved.
Result and Discussions

![Diagram](image)

Figure 5.9: Overall throughput under different algorithms, 4-node case

![Diagram](image)

Figure 5.10: Overall throughput under different algorithms, 8-node case

Fig. 5.9, Fig. 5.10 and Fig. 5.11 show the overall throughput of the system under three scheduling algorithms, in which we can see all algorithms provide high saturated throughput under different packet sizes and different node numbers. The TO algorithm provides the maximum throughput among the three, and the round robin algorithm provides the throughput slightly higher than the TF1 algorithm. It is evidence that the proposed
algorithms can take advantage of the proposed metric for performance improvement.

Fig. 5.12 shows the throughput of each MT under TF1 and TO algorithms. In the TF1 algorithm, each MT has similar throughput without regard to the channel quality. We can see the MTs with poor channel qualities are compensated with extra bandwidth for the throughout fairness. The TO algorithm, on the other hand, provides MTs with better channel qualities higher scheduling priorities for overall throughput maximization. As a result, the MTs with poor channel qualities have significantly less throughput than the MTs with good qualities. Fig. 5.12 explains why the TF algorithm has less overall throughput than the TO algorithm. The compensation mechanism in the TF algorithm leads to that difference.

Fig. 5.13 shows the performance of algorithms in the multiple data rate case. As seen from the figure, the TF1 algorithm achieves similar absolute throughput among 4 MTs. The overall throughput under this algorithm is significant lower than other two algorithms. The TO algorithm provides the MTs with higher product of TLR and the data rate more chance to win time slots, and hence produces the highest throughput among three algorithms.

Figure 5.11: Overall throughput under different algorithms, 20-node case
Figure 5.12: Throughput per node in 4 and 8-node cases

Figure 5.13: Overall throughput under different algorithms, 4-node case, each node has different data rate
5.3 Conclusion

In this chapter, we study the bandwidth allocation issues for both downstream and upstream, which are important for WiGEE to meet different service requirements. For the upstream, a DBA algorithm that jointly considers multiple wavelengths is proposed. The solution is provided in the algorithm to avoid the co-channel interference among multiple cells. Moreover, two duplex modes of MTs are taken into account in the algorithm. The main difference of the half duplex mode DBA and full duplex mode DBA is that the former should jointly consider the downstream bandwidth allocation since the downstream and upstream share the wireless channel.

In the second part of this chapter, the downstream scheduling algorithm suitable for the architecture of WiGEE is studied. A long term channel state metric obtained from the error control scheme of the link layer is used to enable channel aware scheduling in WiGEE. Two algorithms based on this metric are proposed. One is the throughput optimal algorithm, which jointly considers the channel quality and the queue length at the CS for throughput optimization and system stabilization, the other is the throughput fairness algorithm, which provides throughput fairness among MTs in the sense that similar throughput achieved at the receivers if the loads at the transmitters are similar. The performances of the proposed algorithms are studied in simulations.
Chapter 6

Performance Study of MAC Protocols

In this chapter, we study the performance of MAC protocol candidates described in chapter 2 and chapter 3, with an emphasis on the metrics of throughput and overall control overhead. The throughput of each protocol as a function of different physical layer data rates and frame lengths is analyzed. The performance of each protocol under a high data rate up to Gbps is compared.

6.1 Performance Analysis of MAC Protocol Candidates

For any network, four parameters, the bandwidth, delay, delay jitter, and packet loss ratio, are tightly related to the QoS of the applications, and are of special interest for the performance study. Among them, the bandwidth and delay are usually tightly coupled, while the delay jitter and packet loss ratio are more related to the specific scheduling algorithm and queue management scheme at the link and upper layers. Considering this, it is nature to choose the throughput of the MAC protocol as the performance metric. Here we use the maximum throughput, which is defined as the
overall maximum throughput of the MAC protocol achieved under three conditions:

- The channel is error free.
- No collision occurs during the communication. This condition holds for the contention period of any MAC protocol.
- Perfect scheduling among nodes.

It is the metric that indicates the overall amount of data can be delivered by the MAC of the system in unit time. The normalized throughput, which is the ratio of the time used to transmit data payload to the sum of the time needed to complete the transmission, reflects the efficiency of a MAC protocol, and can be used for MAC performance comparison.

In the following of this section, we focus the study on the theoretic maximum throughput of MAC protocol candidates.

6.1.1 WiGEE

Since multiple wavelengths in WiGEE are basically independent in operation, we only study the performance of WiGEE on a single wavelength. Depending on the duplex modes of nodes, each wavelength supports three operation modes: half duplex, full duplex and hybrid duplex mode. Note that the half duplex mode is similar to the time division duplex (TDD) mode in IEEE 801.16 systems, and the full duplex and hybrid duplex mode is similar to the frequency division duplex (FDD) mode in IEEE 801.16 systems. We study the maximum throughput of the half duplex and full duplex mode. Obviously, the hybrid duplex mode has the throughput performance between these two.
6.1. PERFORMANCE ANALYSIS OF MAC PROTOCOL CANDIDATES

Maximum Throughput of WiGEE

We use the duration of the superframe, denoted by $T_{sf}$, as the measured time period to compute the maximum throughput of WiGEE. The maximum throughput is obtained by counting the maximum amount of payload transmitted in a superframe, and dividing it by the length of a superframe. To get the maximum amount of payload in a superframe, we need to know the minimum control overhead required in a superframe. Here for simplicity, we assume the length of a superframe and data frames are fixed.

The overheads in the half duplex and full duplex mode are slightly different. In the half duplex mode there is radio turnover time, denoted by $T_{\text{turn}}$, between the downstream period and upstream period. Other overheads are basically common in two modes.

The common overhead are divided into the control overhead and the overhead of data frames. The control overhead of the downstream includes the beacon period, and the grant messages issued to MTs. The control overhead of the upstream includes the MT presence period, discovery period, random access period, report messages sent by MTs, and the guard time between two consecutive grants. Moreover, there is guard time between two consecutive superframes. Let $T_b$, $T_{\text{MTP}}$, $T_{\text{disc}}$, and $T_{\text{rnd}}$ denote the length of them, respectively. For simplicity, we assume the guard time between grants and superframes is same, which is denoted by $T_g$. Referred to the chapter 3 and [11], the length of grant and report message are 72 bytes if the preamble and start-of-frame delimiter (SFD) fields are included [10]. They are rate dependent and rely on the number of MTs allocated in the upstream of the superframe. An average number, denoted by $N_{MT}$, is used in the overhead calculation, which is determined by the length of superframe and the grant threshold assigned to an MT. The grant threshold
is a system parameter that specifies the maximum grant allowed to assign to an MT in a superframe.

As described in chapter 3, the data frame of WiGEE is basically an Ethernet data frame with extensions to support error control at the link layer. According to the specification, the overhead of a data frame is 26 bytes in no ARQ control scheme, and 31 bytes in ARQ control scheme. Let $L_{Of}$ denote this overhead. The data payload takes the length of 46 to 1500 bytes. For simplicity, we use fixed length payload in the maximum throughput analysis, which is denoted by $L_{payload}$.

The absolute throughput of the system is obtained by:

$$ T_{hpt} = \frac{(N_{f_{ds}} + N_{f_{us}}) \times L_{payload}}{T_{sf}} \quad (6.1) $$

where $N_{f_{ds}}$ is the number of data frames transmitted in the downstream of the superframe, and $N_{f_{us}}$ is the number of data frames transmitted in the upstream of the superframe. The difference of the half duplex mode and full duplex mode in Eq. 6.1 exists in calculating the $N_{f_{ds}}$ and $N_{f_{us}}$. Assume the data rate of the system is $R$, the grant threshold is $L_g$. In the half duplex mode, the control overhead is

$$ T_{Oc} = T_b + T_{MTP} + T_{disc} + T_{rnd} + 2N_{MT} \times 72 \times 8/R + N_{MT} \times T_g + T_{turn} \quad (6.2) $$

where $N_{MT}$ is obtained by solving

$$ N_{MT} = \left\lfloor \frac{(T_{sf} - T_{ds} - T_{Oc}) \times R/L_g} {T_{sf}} \right\rfloor \quad (6.3) $$

jointly with Eq. 6.2. $T_{ds}$ in Eq. 6.3 is the time assigned to the downstream in the superframe, and $\lfloor \cdot \rfloor$ is the floor operation. Note that $T_{sf}$ and $L_g$ in Eq. 6.2 and Eq. 6.3 are properly chosen so that the overhead is minimized. The number of frames transmitted in an upstream grant is $N_{fg} = \lfloor L_g/(L_{payload} + L_{Of}) \rfloor$, and the $N_{f_{us}} = N_{fg} \times N_{MT}$. The frame overhead of the upstream is $T_{Of_{us}} = N_{MT} \times N_{fg} \times L_{Of}$. For the downstream, the number of transmitted frames is $N_{f_{ds}} = \lfloor (T_{ds} \times R - N_{MT} \times
72 \times 8)/(L_{payload} + L_{Of})\right].

In the full duplex mode, the downstream period becomes \( T_{ds} = T_{sf} - T_b - T_g \).

The control overhead of the upstream is

\[ T_{Oc.us} = T_b + T_{MTP} + T_{disc} + T_{rnd} + N_{MT} \times 72 \times 8/R + N_{MT} \times T_g \]  

(6.4)

The \( N_{MT} \) is obtained by solving

\[ N_{MT} = \left\lfloor \left( T_{sf} - T_{Oc.us} \right) \times R/L_g \right\rfloor \]

jointly with Eq. 6.4. The number of frames transmitted in the downstream is \( N_{f.ds} = \left\lfloor (T_{ds} \times R - N_{MT} \times 72 \times 8)/(L_{payload} + L_{Of}) \right\rfloor \).

The number of frames transmitted in the upstream is \( N_{f.us} = N_{fg} \times N_{MT} \).

Note that the \( N_{MT} \) in half duplex and full duplex mode take different values.

The normalized throughput in the half duplex mode is

\[ Thpt_{N_{hf}} = \frac{(N_{f.ds} + N_{f.us}) \times L_{payload}}{T_{sf} \times R} \]  

(6.5)

The normalized throughput in the full duplex mode is

\[ Thpt_{N_{fl}} = \frac{(N_{f.ds} + N_{f.us}) \times L_{payload}}{2 \times T_{sf} \times R} \]  

(6.6)

where the \( N_{f.ds}, N_{f.us} \) take the corresponding values in half and full duplex mode.

**Numerical Result**

To evaluate the efficiency of the WiGEE MAC protocol, we take a quantitative analysis on the maximum throughput of WiGEE in the following. Referring to the specifications in IEEE 802.15.3, we set \( T_g = 2\mu s \) and \( T_{turn} = 8\mu s \). The throughput is analyzed under different data rates, i.e. 54 Mbps, 120 Mbps, 480 Mbps, and 1 Gbps. The length of a superframe is set to 2ms, 10ms, 30ms, and 100ms, respectively. The grant threshold for
each MT is set with regard to the data rate, the length of the superframe, and the average number of MTs allocated in the superframe accordingly. Fig. 6.1 shows the absolute throughput of the half duplex mode under four data rate sets. The curves show the resultant throughput with different MTs scheduled on the upstream of a superframe. From the figure, two results are obtained: 1) The absolute throughput experiences linear increase with regard to the data rate. As the number of allocated MTs is small or medium, the throughput approaches to the data rate. 2) Even the number of allocated MTs is increased to 32, the throughput drops a little with regard to the data rate. However, the throughput is continuously deteriorated as the number allocated of MTs increases, due to the increasing control overhead in allocation.

The normalized throughput of the half duplex mode is presented in Figure 6.1: Absolute throughput under various data rates, half duplex mode, payload size of 512 bytes, and superframe length of 2ms

Figure 6.1: Absolute throughput under various data rates, half duplex mode, payload size of 512 bytes, and superframe length of 2ms

Fig. 6.2. It shows that the smaller the number of MTs is allocated on the upstream of the superframe, the better the normalized throughput.
Even the number of allocated MTs becomes 64, the normalized throughput is as high as 80% at the data rate of 1 Gbps. It thus suggests that the proposed MAC protocol of WiGEE is suitable for high speed communication. Moreover, the number of allocated MTs has more impact at the low data rate. As aforementioned, the amount of MPCP control messages and guard time grows as the number of allocated MTs increases, resulting in higher control overhead. Since the overhead contributed by control messages is rate dependent, the throughput at the lower data rate is more vulnerable to the node number.

Fig. 6.3 shows the performance difference of the half duplex and full duplex mode under various data rates and MT configurations. From the figure, it is clear that the full duplex mode outperforms the half duplex mode under all MT configurations. The result matches our expectation since the full duplex mode has less overhead than the half duplex mode. It is also shown from the figure that the difference of two modes is related
large when the data rate is low. As the data rate increases to Gbps, the difference under the same MT configurations is less than 6%.

The impact of the payload size on the throughput performance is shown in Fig. 6.4. The interesting result is that the normalized throughput of two duplex modes is significantly low when the payload size takes the minimum permitted size, which is 46 bytes. As seen from the figure, the normalized throughput is around 60% even when the data rate boosts to 1 Gbps. The main contributor of the overhead comes from the data frame, which has 30 bytes header and tail length per frame in our analysis. Such amount is comparative to the payload size of 46 bytes. As the payload size increases, the normalized throughput becomes better, which suggests the use of large payload in the system. However, this analysis assumes the error free channel condition. The result may not hold when taking channel errors into account.

In the half duplex mode, the downstream and upstream share the same
channel. It is worth to know whether the proportion of downstream period in a superframe affects the throughput performance. Fig. 6.5 shows the normalized throughput when the downstream occupies different proportions of the superframe. From the figure, we can see the influence of that proportion is minor, especially when the data rate is high. It is a reasonable result since no additional control overhead is produced due to the change of that proportion.

The last figure studies the impact of the superframe length to the throughput. As seen from the Fig. 6.6, a larger superframe length produces higher normalized throughput. The conclusion holds for two duplex modes. The superframe length has more influence on the low data rate than the high data rate. As the data rate boosts to 1 Gbps, that difference becomes minor. Since a long superframe can hold more data frames, the percentage of the overhead is reduced as compared to the increased data load. As the superframe increases to a certain length, the improvement on the perfor-
Figure 6.5: Normalized throughput under various data rates, half and full duplex mode, four scheduled MTs per superframe, payload size of 512 bytes, various DS lengths, and superframe length of 2ms

Figure 6.6: Normalized throughput under various data rates, half and full duplex mode, four scheduled MTs per superframe, payload size of 512 bytes, and various superframe lengths
mance becomes minor. In this case, the main contributor to the overhead is the frame overhead, which is independent from the superframe length.

### 6.1.2 IEEE 802.11n

The performance of IEEE 802.11 networks has been studied extensively in the last decade [91] [92] [93]. Based on the developed methodologies, we study the throughput performance of the 802.11n under different data rate conditions. Here we assume the reader has adequate knowledge on 802.11 based networks. If no, please refer to the chapter 2 or [94] for more information.

Note that the 802.11n standard is still under development at the time of this writing. The analysis is therefore based on the draft standard P802.11n-D1.0 [54]. Firstly, we develop the analytical framework of the maximum throughput under the DCF access mode, in which the data aggregate function is not considered. Next, we put the data aggregate mechanism into the analytical framework and study the performance improvement. Finally, we study the saturated throughput of 802.11n. The difference between maximum throughput and saturated throughput is that in the later the collision is inevitable due to the saturated load in the system.

The maximum throughput of 802.11n is analyzed under three access modes: DCF, EDCA, and HCCA mode. Under the DCF and EDCA mode, no collision occurs at the contention window. The DCF mode is the basic access mode specified in the first 802.11 standard; The EDCA and HCCA modes were developed in 802.11e and will be employed in 802.11n.

In addition to enhance the access mode, 802.11e extents the acknowledgement (ACK) mechanisms to further improve the throughput. Three ACK schemes, i.e. immediate block ACK (BA), delayed BA, and No ACK, were introduced. In 802.11n, the first two become N-immediate BA and
N-delayed BA, which are extensions to those in 802.11e with the support on data aggregate. We only analyze the case with N-immediate BA. It is expected that the performance difference between N-immediate BA and N-delay BA is minor since they use the similar mechanism. The No ACK is not analyzed here because in wireless data communication the error control scheme is normally required.

Note that the standard supports RTS/CTS mechanism to solve the hidden terminal problem. For simplicity, our analysis does not take the RTS/CTS mechanism into account. Moreover, the fixed frame length is used in the analysis.

The notations for the analysis are listed in the following. Fig. 6.7 shows the relationship between MSDU and MPDU. For convenience, we use the terms data frame and MPDU interchangeable. As referring to Fig.6.7, Fig.6.8, Fig.6.9, Fig.6.10, Fig.6.11 let $L_{\text{payload}}$, $L_{\text{MAC_HDR}}$ and $L_{\text{FCS}}$ denote the length of payload, MAC header and frame check sequence (FCS).
in bit, $T_d$ and $T_{ACK}$ denote the time to transmit a data frame and an ACK, $T_{DIFS}$, $T_{SIFS}$, $T_{RIFS}$ and $T_{AIFS}$ denote the duration of DIFS, SIFS, RIFS and AIFS, $\tau$ denote the propagation delay of two nodes, and $\overline{CW}$ denote the average contention window (CW) length in time, respectively. Since in the maximum throughput analysis no collision is considered, it is easy to get $\overline{CW} = (CW_0 - 1) \times T_{slot}/2$ where $CW_0$ is the initial CW size, and $T_{slot}$ is the slot time in CW.

In the no data aggregate case, $T_d$ of each data frame contains the following
parts:

\[ T_d = T_{pre} + T_{PHY\_HDR} + T_{MAC\_HDR} + T_{payload} + T_{FCS} \]

where \( T_{pre} \) is the PHY preamble duration, \( T_{PHY\_HDR} \) is the PHY header duration, \( T_{MAC\_HDR} \) is the MAC header duration, \( T_{payload} \) is the payload duration, and \( T_{FCS} \) is the FCS duration. Usually the physical preamble and physical header are transmitted with the base data rate. Therefore the period of \( T_{pre} \) and \( T_{PHY\_HDR} \) are fixed in the analysis. \( T_{MAC\_HDR} \) and \( T_{FCS\_HDR} \) are rate dependent. In 802.11, \( L_{MAC\_HDR} + L_{FCS} \) is equal to 28 bytes. Assuming the data rate of system is \( R \), \( T_{MAC\_HDR} + T_{FCS} = 28 \times 8/R \).

\( T_d \) in the data aggregate case relies on the aggregate mode. As shown
in Fig. 6.12 and Fig. 6.13, there are two aggregate modes: A-MPDU, and A-MSDU. According to the standard [54], the maximum length of an A-MPDU is 65535 bytes, and the maximum length of an A-MSDU is limited by the maximum size of the MSDU accepted by the physical layer. In 802.11, 802.11a, and 802.11b, the maximum size of MSDU is 4095 bytes in the frequency hopping spread spectrum (FHSS) mode, and 8191 bytes in the direct sequence spread spectrum (DSSS) mode. Let the $L_{AMPDU}$ and $L_{AMSDU}$ denote the length of an A-MPDU and A-MSDU frame in bit, respectively. $T_d$ for an A-MPDU and A-MSDU frame are

$$T_{d_{AMPDU}} = T_{pre} + T_{PHY_HDR} + L_{AMPDU}/R$$

$$T_{d_{AMSDU}} = T_{pre} + T_{PHY_HDR} + L_{AMSDU}/R$$

The number of payloads transmitted in an A-MPDU and A-MSDU are

$$N_{AMSDU} = \left\lfloor (L_{AMSDU} - L_{MAC_HDR} - L_{FCS})/(14 \times 8 + L_{payload}) \right\rfloor$$

$$N_{AMPDU} = \left\lfloor (L_{AMPDU})/(32 + L_{MAC_HDR} + L_{FCS} + L_{payload}) \right\rfloor$$

**Maximum Throughput In DCF Access Mode**

As shown in Fig. 6.8, the DCF access mode uses the normal ACK, in which each data frame needs an immediate positive ACK. The maximum throughput of 802.11n in the DCF access mode is expressed as:

$$Thpt_{DCF} = \frac{L_{payload}}{T_d + T_{ACK} + T_{DIFS} + T_{SIFS} + 2 \times \tau + CW}$$

(6.7)
$T_{ACK}$ is calculated by $T_{ACK} = T_{pre} + T_{PHY} + L_{ACK}/R$, where $L_{ACK} = 112$ bits is the length of the ACK frame. Note that the overhead of the beacon period is not taken into account. Since its duration is significant small as compared to the data period, we ignore this part of overhead in the following analysis.

Maximum Throughput In EDCA mode

The EDCA and HCCA mode introduce the concept of transmission opportunity (TXOP) for overhead reduction and fine transmission control. A node can exclusively use its TXOP period to transmit multiple frames without the need to contend the medium. Let $T_{TXOP}$ denote the duration of a TXOP.

As shown in Fig. 6.9A, in the normal ACK mode, the number of data frames transmitted in a TXOP is:

$$N_{frm} = \left\lfloor \frac{T_{TXOP}}{T_d + 2 \times T_{SIFS} + T_{ACK}} \right\rfloor$$

(6.8)

The maximum throughput in the non-aggregate data frame case is

$$Thpt = \frac{N_{frm} \times L_{payload}}{T_{TXOP} + T_{AIFS} + CW}$$

(6.9)

For simplicity, we set $T_{AIFS} = T_{DIFS}$.

As seen from Fig. 6.9A, the way to analyze the throughput of the data aggregate case is similar to that in the normal ACK mode. The number of data frames transmitted in a TXOP can be calculated by

$$N_{frm} = \left\lfloor \frac{T_{TXOP}}{T_d + 2 \times T_{SIFS} + T_{BA}} \right\rfloor$$

(6.10)

where $T_d$ is the duration to transmit an A-MPDU or A-MSDU, and $T_{BA}$ is the duration to transmit a BA frame, whose length is 36 bytes. The
throughput in the data aggregate case it is
\[ Thpt = \frac{N_{frm} \times N_{Ag} \times L_{payload}}{T_{TXOP} + T_{AIFS} + CW} \] (6.11)

According to the aggregate mode, \( N_{Ag} \) is equal to \( N_{AMPDU} \) or \( N_{AMSDU} \), respectively.

In the BA mode, the number of data frames transmitted in a TXOP becomes:
\[ N_{frm} = \left\lfloor \frac{T_{TXOP} - (T_{BAR} + 2 \times T_{SIFS} + T_{BA})}{T_d + T_{RIFS}} \right\rfloor \] (6.12)

where \( T_{BAR} \) is the duration to transmit a BA request frame, whose length is 24 bytes. The throughput can be calculated by Eq. 6.9 for the non-aggregate data frame case.

**Maximum Throughput In HCCA mode**

The HCCA mode uses the polling approach to control the channel access, which is governed by a node called HC. The contention period is therefore avoided in this mode. The HCCA mode is similar to the half duplex mode of WiGEE.

The throughput calculation of this mode is quite similar to that in the EDCA mode, except that the contention overhead is not counted (Fig. 6.10). In the normal ACK mode, the number of data frames transmitted in a TXOP can be obtained by Eq.6.8.

The maximum throughput becomes
\[ Thpt = \frac{N_{frm} \times L_{payload}}{T_{TXOP} + T_{SIFS} + \tau + T_{CF\_Poll}} \] (6.13)

where \( T_{CF\_Poll} \) is the duration to transmit a CP_Poll frame, whose length is 20 bytes.

As the same in the EDCA mode, the throughput calculation of the data aggregate case is similar to that of the normal ACK mode in the HCCA
mode. The number of data frames transmitted in a TXOP can be obtained by Eq. 6.10. The throughput for the data aggregate case it is

\[
Thpt = \frac{N_{frm} \times N_{Ag} \times L_{\text{payload}}}{T_{TXOP} + T_{\text{SIFS}} + \tau + T_{\text{CF POLL}}}
\]  

(6.14)

According to the aggregate mode, \( N_{Ag} \) is equal to \( N_{\text{AMPDU}} \) or \( N_{\text{AMSDU}} \), respectively.

In the BA mode, the number of data frames transmitted in a TXOP can be calculated by Eq. 6.12. The throughput can be calculated by Eq. 6.13 for the non-aggregate data frame case.

Saturated Throughput

We only analyze the saturated throughput under the DCF access mode. The same analytical framework can be applied to the EDCA mode. Since the TXOP in the EDCA model allows for transmitting multiple frames once a time, it has better saturated throughput than the DCF access mode. For the HCCA mode, since no contention is involved, the saturated throughput is equal to its maximum throughput.

The saturated throughput of 802.11n is expressed as [93]:

\[
Thpt_S = \frac{P_s \times L_{\text{payload}}}{(1 - P_b) \times T_{\text{slot}} + P_s \times T_s + (P_b - P_s) \times T_c}
\]  

(6.15)

where \( P_s \) the probability of a successful transmission occurring in a slot time, \( P_b \) is the probability that the channel is busy, \( T_s \) is the average time the channel is sensed busy due to a successful transmission, and \( T_c \) is the average time that the channel suffers a collision. Let \( p \) be the probability that a station start a transmission in a randomly chosen slot time. \( P_b \) expressed by \( p \) is

\[
P_b = 1 - (1 - p)^n
\]

where \( n \) is the number of nodes in the system. \( P_s \) is derived from [93]

\[
P_s = n \times p \times (1 - p)^{n-1}
\]
6.1. PERFORMANCE ANALYSIS OF MAC PROTOCOL CANDIDATES

$T_s$ and $T_c$ are obtained by:

$$T_s = T_d + T_{ACK} + T_{DIFS} + T_{SIFS} + 2 \times \tau$$

$$T_c = T_d + T_{DIFS} + \tau$$

To maximize the saturated throughput, $p$ should satisfy that

$$p = \frac{1}{n \sqrt{T_c/(2 \times T_{slot})}}$$

(6.16)

According to Eq. 6.15 and Eq. 6.16, we analyze the saturated throughput of the 802.11n system under different node number and transmission probability $p$.

**Numerical Result**

For the analysis, the following values of parameters are used [91] [95]: $T_{pre} = 16 \mu s$, $T_{PHY} = 4 \mu s$, $T_{DIFS} = 34 \mu s$, $T_{SIFS} = 16 \mu s$, $T_{RIFS} = 2 \mu s$, $T_{slot} = 9 \mu s$, $C W_0 = 15$, and $\tau = 1 \mu s$. Note that some of parameters are related to specific physical layer implementations. However, the use of different parameter sets does not alter the analysis result since the difference between them is minor. The rate dependent parameters are not listed here. They are calculated in the analysis according to specific data rates. Four data rate sets, 54 Mbps, 120 Mbps, 480 Mbps, and 1 Gbps, are used, which are related to the data rate of 802.11a/b/g, 802.11n, 802.15.3a, and WiGEE, respectively.

Fig. 6.14 shows the absolute throughput of the 802.11n DCF mode under different data rates and payload sizes. The first observation from the figure is that as the data rate increases, the absolute throughput increases under most payload size cases. The increasing amount is proportional to the payload size. The maximum throughput achieved at the data rate of 1 Gbps is 106 Mbps under the payload size of 2304 bytes, which is quite low as
compared to the data rate. As seen from Eq. 6.7, the reason for the low throughput in the DCF mode is the fixed amount of the rate independent overheads in each packet transmission. This part of overheads has to be reduced for the system operating at a high data rate.

The other observation is that the throughput under the payload size of 46 bytes keeps almost constant at 2.36 Mbps under different data rates. Beside the rate independent overhead, the frame overhead, which is comparative to the payload size, is contributed to this result.

As a conclusion, the DCF mode or 802.11n is not suitable for gigabit WLANs.

The normalized throughput of the DCF mode is shown in Fig. 6.15, from which two observations are obtained. Firstly, as the data rate increases, the normalized throughput, or in other words, the efficiency of the MAC protocol, decreases rapidly, and approaches to a low limit. The low limit for different payload size can be obtained from Eq. 6.7 by setting $T_{payload}$
in $T_d$ to zero. Secondly, larger payload has better normalized throughput than smaller payload. Nevertheless, at a high data rate, the protocol efficiency under a large payload is still too low. The maximum normalized throughput at the data rate of 1 Gbps is only 10.6%. Fig. 6.15 supports the conclusion drawn from Fig. 6.14.

Fig. 6.16 shows the absolute throughput under normal ACK and BA mechanisms of the 802.11n EDCA mode. Fig. 6.17 is the normalized throughput version of Fig. 6.16. Here we use the term *normal EDCA mode* to denote the EDCA mode without using data aggregate, and *aggregate EDCA mode* otherwise. The same notation is used for the HCCA mode. From Fig. 6.16 and Fig. 6.17, we first notice that the EDCA mode provides better throughput performance than the DCF mode. The maximum throughput provided by the normal ACK and BA mode at the data rate of 1 Gbps is 412 Mbps and 194 Mbps, respectively. However, that fact that the maximum normalized throughput at the data rate of 1 Gbps is less than 0.5 suggests that
this mode is insufficient to support gigabit transmission. The second observation from the figures is the BA has better performance than the normal ACK, which matches our expectation since the BA has less overhead than the normal ACK mechanism. The third observation is that the normalized throughput goes down as the data rate increases. When the payload size is as small as 46 bytes, the normalized throughput keeps almost constant, like the case in the DCF access mode. The reason is the same as in the DCF mode.

Fig. 6.18 shows the EDCA mode under different TXOP time, in which the TXOP time is ranging from 2\text{ms} to 100\text{ms}. It shows that the length of TXOP length has minor impact on the throughput performance when the TXOP is greater than a proper size. The reason is the following: From the data exchange sequence in the EDCA mode, we can easily find that the main overhead comes from the interframe time following each data frame, which is independent from the TXOP. Combining with the results obtained
from Fig. 6.16 and Fig. 6.17, we get the conclusion that the normal EDCA mode is not suitable for gigabit WLANs.

The main difference between the EDCA and HCCA mode is that the EDCA needs a contention process to seize the channel. In the maximum throughput analysis, the overhead contributed by the contention process is minor, therefore it is the performance of the EDCA and HCCA mode in our analysis is similar.

Fig. 6.19 shows the normalized throughput of the HCCA mode under different payload size. Fig. 6.20 shows the normalized throughput of the HCCA mode under different TXOP length. Except that HCCA yields a throughput slightly higher than EDCA, the curves in these two figures are rather similar to the outcome of EDCA shown in Fig. 6.17 and Fig. 6.18. Therefore, the same conclusion holds: the normal HCCA mode is not suitable for gigabit WLANs.

Then we turn to the study of the data aggregate mode in 802.11n.
Figure 6.18: Normalized throughput under 802.11n EDCA, various data rates, normal ACK or BA, payload size of 512 bytes, various TXOP lengths

Figure 6.19: Normalized throughput under 802.11n HCCA, various data rates, various payload sizes, and TXOP length of 2ms
shows the normalized throughput of two data aggregate modes under different data payload sizes. The data payload here means the data payload in the individual sub-frame of an aggregate frame. From the figure, we can find the throughput under the data aggregate mode is better than the normal EDCA mode, due to the overhead reduction. The A-MPDU mode outperforms the A-MSDU mode when the data payload size is identical. At the data rate of 1Gbps, the A-MPDU mode achieves the maximum throughput of 753 Mbps, while the A-MSDU mode achieves 504 Mbps. The reason for this is that the size of an A-MPDU is much larger than that of an A-MSDU in our setting. The figure shows the throughput decreases as the data rate increases, however, the decreasing rate is slower than that of the normal EDCA and HCCA mode. Moreover, we find an interesting phenomenon in the figure, i.e. instead of 2304 bytes, the maximum throughput occurs at the payload size of 1500 bytes. In the analysis, we use the same aggregate size for all payload cases. In this case, a larger
payload may waste more bits in an aggregate frame if the payload size and aggregate frame are mismatched.

The normalized throughput of aggregate EDCA mode under different aggregate frame sizes is shown in Fig. 6.22. From the figure, we can see the throughput increases as the size increases. The maximum normalized throughput at the data rate of 1 Gbps is 87.2% under the A-MPDU mode, and 48.2% under the A-MSDU mode, which are higher than the maximum normalized throughput of 42.8% in the normal HCCA mode. The high throughput achieved by the data aggregate mode suggests the transmission of large data frames in high speed systems. However, the trade off should be taken since a larger frame is more vulnerable to errors.

Fig. 6.23 shows the impact of the TXOP length on the performance of the system. For the TXOP length greater than 10ms, the performance difference is insignificant.
6.1. PERFORMANCE ANALYSIS OF MAC PROTOCOL CANDIDATES

Figure 6.22: Normalized throughput under 802.11n aggregate EDCA mode, various data rates, payload size of 512 bytes, TXOP length of 10ms, various A-MSDU lengths, and various A-MPDU lengths.

Figure 6.23: Normalized throughput under 802.11n aggregate EDCA mode, various data rates, payload size of 512 bytes, various TXOP length, A-MSDU=8191 bytes, and A-MPDU=23040 bytes.
Fig. 6.24, Fig. 6.25, and Fig. 6.26 show the throughput performance of data aggregate under the HCCA mode. The throughput under the HCCA mode is slightly higher than that of the EDCA mode, due to the overhead reduction in the HCCA mode. Except this, the throughput performance is quite similar in these two modes.

Fig. 6.27 compares the performance of the A-MSDU frame of different sizes under the EDCA and HCCA mode. It further confirms the conclusion that the throughput performance of A-MSDU under EDCA and HCCA are similar.

The absolute and normalized maximum saturated throughput of 802.11 under the DCF access mode are shown in Fig 6.28 and Fig. 6.29, respectively. The saturated throughput performance is quite similar to that of the DCF access mode in the maximum throughput analysis. However, the saturated throughput has slightly higher value than the maximum through-
Figure 6.25: Normalized throughput under 802.11n aggregate HCCA mode, various data rates, payload size of 512 bytes, TXOP length of 10ms, various A-MSDU lengths, and various A-MPDU lengths

Figure 6.26: Normalized throughput under 802.11n aggregate HCCA mode, various data rates, payload size of 512 bytes, various TXOP length, A-MSDU=8191 bytes, and A-MPDU=23040 bytes
Figure 6.27: Normalized throughput under A-MSDU of EDCA and HCCA mode in 802.11n, various data rates, payload size of 512 bytes, TXOP length of 10ms, A-MSDU=8191 bytes.

Figure 6.28: Absolute saturated throughput under 802.11n DCF, various data rates, and various payload sizes.
6.1. PERFORMANCE ANALYSIS OF MAC PROTOCOL CANDIDATES

Figure 6.29: Normalized saturated throughput under 802.11n DCF, various data rates, and various payload sizes

Figure 6.30: Saturated throughput of 802.11n DCF under different access probabilities, various data rates, and payload size of 512 bytes
Figure 6.31: Normalized throughput of different access modes in 802.11n, various data rates, payload size of 512 bytes, A-MPDU and A-MSDU take TXOP length of 10ms, A-MPDU takes length of 23040 bytes, A-MSDU takes length of 8191 bytes, EDCA and HCCA take TXOP length of 2ms.

At the data rate of 1 Gbps, the throughput of the former under the payload size of 2304 bytes is 123 Mbps, while the latter is 106 Mbps. The discrepancy is resulted by the different ways the rate independent overhead is calculated. Note that Fig 6.28 and Fig. 6.29 show the maximum saturated throughput. In such a case, the overhead is slightly lower than the maximum throughput case. As the parameter $p$ in Eq. 6.15 takes sub-optimal values, which is shown in Fig. 6.30, the saturated throughput is deteriorated. The result of saturated throughput shows again that the DCF access mode of 802.11 is not suitable for gigabit WLANs.

We use a single plot to compare the normalized throughput of different access modes in 802.11n, which is shown in Fig. 6.31. From the figure, it shown that at the data rate of 54 Mbps, except the DCF access mode, all other modes can provide the maximum throughput approaching or greater
than half of the data rate. However, when the data rate becomes 1 Gbps, only data aggregate functions can provide acceptable throughput performance. Note that the throughput in this analysis is achieved in the error free condition. In realistic, the throughput should be worse than the maximum throughput. This is especially true for the data aggregate case since the large frame size is more vulnerable to the errors.

6.1.3 HiperLAN/2

The unique feature of the HiperLAN/2 is the fixed length of the superframe equal to 2ms. As the control overhead of the superframe is almost fixed in time, a high data rate leads to a high throughput. In the following, we analyze the maximum throughput of HiperLAN/2.

A typical superframe structure of HiperLAN/2 in the centralized access mode is shown in Fig.6.32, in which one BCH, one FCH, one ACH, zero or several downlink (DL) phases, zero or several uplink (UL) phases, and one or several RCHs are included. These fields are divided into control fields and data fields. The DL phase and UL phase are data fields, while remaining fields are control fields. Note that HiperLAN/2 also supports peer to peer transmission between two nodes, which uses the DiL phase in a superframe. However, for the MAC protocol comparison, we only consider the centralized access mode of HiperLAN/2.

Figure 6.32: Superframe structure of HiperLAN/2
To compute the maximum throughput of HiperLAN/2, we need to know the overall overhead in a superframe, which is comprised of control overhead and data overhead. The MAC protocol of HiperLAN/2 is connection oriented, where the transmission in DL and UL is organized by sessions. The data overhead is the preamble inserted before each transmission session. We use $N_{UP}$ to denote the number of sessions in the UL phase of a superframe, and $N_{DOWN}$ as the number of sessions in the DL phase. The overall overhead of a superframe expressed in time is:

$$T_o = T_{BCH+Pre} + T_{FCH} + T_{ACH} + T_{RCH+Pre} + T_{CtrlDL} + T_{CtrlUL} + T_{DLPre} + T_{ULPre} + T_{GT} + T_{Turn}$$

(6.17)

where $T_{BCH+Pre}$, $T_{FCH}$, $T_{ACH}$ and $T_{RCH+Pre}$ are the durations for BCH plus preamble, FCH, ACH, and RCH plus preamble, $T_{CtrlDL}$ and $T_{CtrlUL}$ are the durations for transmitting control signal in DL (an SCH for ACK per session) and UL (SCH for ACK and resource request per session), $T_{DLPre}$ and $T_{ULPre}$ are the durations for preambles in DLs and ULs, $T_{GT}$ is the guard time in UL, and $T_{Turn}$ is the radio turnover time, respectively. Among the overheads, $T_{FCH}$, $T_{CtrlDL}$, $T_{CtrlUL}$, $T_{DLPre}$, $T_{ULPre}$ and $T_{GT}$ are related to $N_{UP}$ and $N_{DOWN}$. Others are only related to the base data rate, which is 6 Mbps under the BPSK modulation scheme. The length of each overhead is listed as follows: $T_{BCH+Pre} = 36\mu s$, $T_{FCH} = ((N_{DOWN} + N_{UP})/2) \times 36\mu s$, $T_{ACH} = 12\mu s$, $T_{RCH+Pre} = 28\mu s$, $T_{CtrlDL} = N_{DOWN} \times T_{SCH}$ where $T_{SCH}$ is the duration for an SCH, $T_{CtrlUL} = 2 \times N_{UL} \times T_{SCH}$, $T_{DLPre} = N_{DOWN} \times 8\mu s$, $T_{ULPre} = N_{UP} \times 12\mu s$, $T_{GT} = N_{UP} \times 2\mu s$, and $T_{Turn} = 12\mu s$. The data payload is transmitted in LCH, whose duration for transmission is denoted by $T_{LCH}$. Note that $T_{SCH}$ and $T_{LCH}$ are related to the data rate used to transmit data payload. The length of an SCH and LCH is 9 bytes and 54 bytes, respectively. In an LCH, 48 bytes are used to transmit data payload.
After put the values of the overheads into Eq. 6.17, we get:

$$T_o = 88 + 26 \times N_{DOWN} + 32 \times N_{UP} + (N_{DOWN} + 2 \times N_{UP}) \times T_{SCH} \mu s$$ (6.18)

The throughput of the system is

$$Thpt = \frac{(0.002 - T_o) / T_{LCH}}{0.002} \times 48 \times 8$$ (6.19)

in which the time unit of $T_o$ is second.

**Numerical Result**

![Figure 6.33: Absolute throughput of HiperLAN/2 under various data rates, and different $N_{down}$ and $N_{up}$ sessions](image)

The absolute throughput of HiperLAN/2 under different data rates is shown in Fig. 6.33. The interesting feature of HiperLAN/2 is that the throughput undergoes linear increase under a fixed configuration. As shown in Fig. 6.34, the normalized throughput under different data rates are almost constant. It is benefited from its special MAC design, i.e. the
amount of overhead is fixed in the fixed length superframe. Therefore, the high throughput can be achieved under proper configuration.

However, as shown in Fig. 6.33, the throughput performance of Hiper-

![Normalized throughput of HiperLAN/2 under various data rates, and different N_{down} and N_{up} sessions](image)

LAN/2 is significantly affected the number of downlink and uplink sessions. The $N_{up}$ has more impact than the $N_{down}$. As $N_{down}$ and $N_{up}$ becomes large, the overhead increases substantially, greatly reducing the normalized throughput.

As a conclusion, the MAC protocol of HiperLAN/2 is suitable for gigabit WLANs, however, only under small or medium size of transmission sessions.

### 6.1.4 IEEE 802.15.3

The superframe structure of 802.15.3 is shown in Fig. 6.35, which consists of a Beacon period, an optional CAP, and a mandatory CTAP. The CAP
is a random access period used to exchange small amount of data or report bandwidth requests on the contention basis. The CTAP is used to transmit two types of data: the MTCA for management messages, and the CTA for normal data. Three types of acknowledgement mechanisms are defined for CTAs: the No-ACK, Imm-ACK, and Dly-ACK, which are illustrated in Fig. 6.36, respectively. In the following, we analyze the maximum throughput of 802.15.3 under those acknowledgement mechanisms. To simplify the analysis, we do not consider CAP and MTCA in the superframe. Moreover, we only include one CTA in the CTAP. In this case, the guard time between multiple CTAs is avoided.

Assuming the length of a superframe is $T_{sf}$, the time for the CTA is

$$T_{CTA} = T_{sf} - T_o$$

where $T_o = T_b + T_g$, $T_b$ is the beacon time, and $T_g$ is the guard time. Note that the control overhead inside $T_{CTA}$ is not counted in $T_o$. The length of beacon varies according to the number of contained information elements (IE). Here we assume the beacon only include three IEs: a BSID IE, a CTA IE, and a CTA status IE. Under this setting the total length of the beacon is 88 bytes. Considering the base data rate of $R_b = 22$ Mbps in 802.15.3, we get $T_b = 32\mu s$.

From Fig.6.37, we get the time to transmit a frame is:

$$T_d = T_{pre} + T_{PHY\_HDR} + T_{MAC\_HDR} + T_{payload} + T_{FCS}$$

Note that since the PHY preamble, PHY header, MAC header and tail are transmitted in the base data rate, the duration of them are fixed, in which
The general form of the maximum throughput is

\[ T_{hpt} = \frac{N \times L_{\text{payload}}}{T_{sf}} \] (6.23)

where \( N \) denotes the number of data frames transmitted in a superframe, and \( L_{\text{payload}} \) is the payload size in bit.
According to the standard [96] and [95], the setting of the parameters is following: $T_{MISF} = 5\mu s$, $T_{SISF} = 8\mu s$, $T_{ACK} = L_{ACK}/R_b = 3.7\mu s$, $T_{Dly\_ACK} = L_{Dly\_ACK}/R_b = (136 + 16 \times N)/R_b$, $T_g = 10\mu s$, where $N$ is the number of data frames transmitted in the CTA, $L_{ACK}$ and $L_{Dly\_ACK}$ are the length of ACK and Dly-ACK frame in bit, respectively. Moreover, the maximum length of a superframe is 65535\(\mu s\), and the maximum length of a data frame is 2048 bytes.

Numerical Result

![Figure 6.38: Absolute throughput of 802.15.3 under Imm-ACK mode, various data rates, payload size of 512 bytes, and superframe length of 2ms](image)
The absolute throughput and normalized throughput of 802.15.3 under Imm-ACK mode are shown in Fig. 6.38 and Fig. 6.39, respectively. The throughput is similar to that of 802.11n under the BA of the EDCA mode, which is shown in Fig. 6.17. However, its throughput is lower than of 802.11n. At the data of rate of 1 Gbps and payload size of 1500 bytes, 802.15.3 has a throughput of 249 Mbps, while 802.11n has a throughput of 316 Mbps. As shown from Fig. 6.39, the normalized throughput under 1 Gbps is less than 31%, which is not suitable for gigabit WLANs.

Fig. 6.40 and Fig. 6.41 show the normalized throughput of 802.15.3 under No-ACK and Dly-ACK mode. These two modes exhibit almost the same throughput performance. It is not surprise since in the maximum throughput analysis the overhead difference of these two modes is negligible. From Fig. 6.39 and Fig. 6.41, we can see the Dly-ACK mode has better throughput performance than the Imm-ACK mode. The difference increases as the data rate increases. At the data rate of 1 Gbps, the normalized through-
6.1. PERFORMANCE ANALYSIS OF MAC PROTOCOL CANDIDATES

Figure 6.40: Normalized throughput of 802.15.3 under No-ACK mode, various data rates, payload size of 512 bytes, and superframe length of 2ms

Figure 6.41: Normalized throughput of 802.15.3 under Dly-ACK mode, various data rates, payload size of 512 bytes, and superframe length of 2ms
put of the Dly-ACK mode is 36.2%, while that of the Imm-ACK mode is 24.9%. However, even for the Dly-ACK mode, the maximum normalized throughput is 43.4% when the payload size is 2034 bytes, which is the maximum allowed size. It shows that to achieve a high throughput at the data rate of 1 Gbps, the same data aggregate approach in 802.11n may need in 802.15.3. The throughput comparison of 802.15.3 under different ACK modes is shown in Fig. 6.42, in which the Dly-ACK mode has almost the same throughput performance as the No-ACK mode, while the Imm-ACK has worse throughput performance than other two.

Fig. 6.43 shows the impact of the superframe length on the throughput performance. Under the configurations of all ACK modes, different superframe lengths only have limited impact on the throughput performance. It can conclude that the superframe length has minor influence on the maximum throughput of 802.15.3 when the superframe length takes reasonable values.
6.2 SUMMARY

Figure 6.43: Normalized throughput of 802.15.3 under all ACK modes, various data rates, payload size of 512 bytes, and various superframe length

6.1.5 IEEE 802.16

Note that the TDD and FDD mode in 802.16 are quite similar to the half duplex and full duplex mode of WiGEE. In this sense, after minor modification, the analytical framework built for WiGEE can be used for 802.16 systems. As expecting the main results derived from the same analytical framework is similar, we do not provide the maximum throughput analysis for 802.16 networks.

6.2 Summary

Fig. 6.44 show the normalized throughput comparison of different systems. As seen from the figure, except WiGEE and HiperLAN/2, the operating modes in all other systems share a common property, i.e. the normalized throughput decreases as the data rate increases. The rate independent
Figure 6.44: Normalized throughput of different systems under various data rates, payload size of 512 bytes. WiGEE in half duplex mode with 4 MTs; A-MPDU and A-MSDU of 802.11n have TXOP=10ms, and frame size of 23040 and 8191 bytes respectively; all other cases but the DCF access mode of 802.11n have superframe length or TXOP equal to 2ms.

overhead in those systems is contributed to this result. As a conclusion, from the throughput perspective, the overhead becomes the main obstacle to keep the current wireless WLAN and WPAN MAC protocols from being utilized in gigabit WLANs. There are two approaches to solve this problem: improving the efficiency of current wireless WLAN and WPAN MAC protocols by reducing the rate dependent overhead in the protocols, or design a new MAC protocol that is suitable for gigabit WLANs.
Chapter 7

Conclusion

We study the link and MAC layer issues of gigabit WLANs in this thesis. To investigate these issues, the first question raised in the mind is whether the MAC protocols of current WLAN and WPAN can be directly used in gigabit WLAN. To answer this, we identify the potential MAC protocols utilized in current WLAN and WPAN, including IEEE 802.11n, HiperLAN/2, and IEEE 802.15.3. As the MM-wave band is employed in IEEE 802.16 for high speed multi-user access, its MAC protocol is as well considered as a candidate. The functionalities of the MAC candidates are studied in four aspects: the channel access, bandwidth allocation, QoS, and error control. The results of the qualitative analysis is following:

- The centralized channel access mode is the majority used mode among the candidates. It has great advantages to support fine granularity QoS and to reduce control overhead. Considering the fact that most emerging applications are bandwidth intensive with varying requirements on QoS, a centralized access scheme is a better choice for high speed WLAN/WPAN.

- All of them provide block acknowledgement mechanism to reduce transmission overhead. In HiperLAN/2 and IEEE 802.16, it is implemented by SR-ARQ. In IEEE 802.11n, and 802.15.3, it is realized
by block acknowledgement mechanisms. Moreover, 802.11n provides data aggregate to further reduce the MAC layer overhead. In this sense, all candidates are suitable for high speed wireless systems.

• All candidates support QoS in different degrees. Centralized protocols apparently outperform distributed protocols on QoS provisioning. The 802.11n supports differential services in the EDCA mode, and fine granularity services in the HCCA mode. Other candidates support fine granularity QoS by means of centralized control scheme.

• The MAC protocol for future WLAN should be able to support beamforming process in order to provide reliable communication in gigabit WLANs. In candidate protocols, only 802.11n provides optional beamforming process. The extension on beamforming support are needed in other candidate protocols.

• As the BER performance of the system is more sensitive to the link quality under a high data rate, an adaptive data rate scheme will be necessary in gigabit WLANs. All candidate systems support multiple MCS. It is easy for them to develop an adaptive data rate scheme with a minor modification on MAC protocols.

• In addition, all candidates support a certain level of mobility.

Usually, the characteristics of the system should be taken into account in the MAC protocol design so as to achieve reliability, scalability, and best performance. The 60 GHz band has a unique feature that sharps the WLAN design, which in turn affects the design of the MAC protocol. Since MM-wave signals do not penetrate solid materials very well, the cell size is usually equal to the room size. Consequently, a large number of BSs may need to provide a considerable coverage. To reduce the cost and provide efficient communication, a system capable of supporting a large number of
BSs is demanded. For this we propose innovative system architecture for the 60 GHz based gigabit WLAN.

The proposed system is a hybrid optical/wireless system, in which the WDM/TDM EPON is used as the fiber optic infrastructure, and the 60 GHz radio for the wireless access. From the link and MAC layer perspective, the system has following features:

- The system is based on a tree topology, in which the CS at the root provides the centralized control on the whole system. As the system architecture is similar to that of EPON except the wireless medium replacing a piece of fiber to reach the ONU(MT), a centralized MAC protocol capable of jointly solving the collisions in the optical and wireless domain is more suitable for the system. The MPCP in the EPON standard 802.3ah can be the candidate.

- Different from the EPON, WDM technologies have to be employed so as to provide sufficient capacity. The MPCP has to be extended to support multiple wavelengths, especially for the handover between wavelengths.

- No MAC functionalities are implemented in the BS. It can be considered as a simple relay node connecting the optical and wireless segment of the system. The error control scheme, which is usually implemented between the BS and MT, has to be extended to the CS. Considering this, SR-ARQ is the error control scheme of the choice for the system.

- To avoid co-channel interference, and provide better channel resource utilization, an out-of-band control plane is implemented between the CS and BS, from which the on/off state of optical transmitter and 60GHz radio transmitter in the BS can be well controlled by the CS.
The MAC protocol, therefore, needs to cooperate with the out-of-band control plane for better access control.

- The concept of the superframe is introduced in the channel access in order to implement coordination functions, such as synchronization and handover, in the system.

The primary function of the MAC protocol in the proposed system is the upstream bandwidth allocation, which is achieved by the DBA algorithm. After studying the bandwidth allocation issues related to the characteristics of the proposed system, we propose a DBA algorithm, which has following features:

- The DBA algorithm uses a request and grant mechanism for bandwidth allocation. An MT reports its queue length, which includes the lengths of multiple queues, to the CS through the MPCP control message. The CS allocates the bandwidth to an MT according to its request, and the resource utilization state of the system. The compensation for ARQ overhead is taken into account in the DBA algorithm.

- The allocation algorithm jointly considers the downstream and upstream so as to support different duplex modes of MTs. In the half duplex or hybrid mode of the system, the downstream and upstream of different BSs served by the same color are jointly allocated so that the full duplex is achieved in the optical segment.

- The control of BS is considered in the allocation algorithm in order to improve the channel efficiency of the fiber optic infrastructure under different duplex modes of BSs.

- The allocation algorithm provides the mechanism to avoid co-channel
interference at overlapping areas of multiple cells caused by uncoordi-
nated allocation among multi-wavelengths.

Since a gigabit wireless system needs more link budget for reliable com-
munication, it is more sensitive to the multipath fading than a low data rate wireless system. Moreover, according to the low diffraction property of the 60 GHz band, the channel quality of the link is highly related to the location of the MT. It is likely that the MT served by the same BS experience significantly different channel quality. Therefore the channel aware scheduling in the proposed system is highly suggested for better system performance. However, the proposed system architecture prevents from the utilization of traditional channel aware scheduling schemes, which utilize instantaneous CSI. To provide CSI at the CS, we proposed a long term channel state metric that is obtained from the error control scheme of the link layer. The assumption to use this metric is that the channel quality of the MT is stable in long run, which is reasonable in most cases of indoor communication. Two algorithms based on this metric are proposed for the downstream scheduling. One is the throughput optimal algorithm, which jointly considers the channel quality and the queue length at the CS for throughput optimization and system stabilization, the other is the throughput fairness algorithm, which provides throughput fairness among MTs in the sense that similar throughput achieved at the receivers if the loads at the transmitters are similar. Providing a proper feedback mechanism, the metric can be used in the DBA algorithm for the upstream bandwidth allocation.

Taking into account the MAC protocol of the proposed system, we analyze on the maximum achievable throughput of the MAC candidates in the last part of the thesis. For each candidate, the maximum throughput under different data rates is obtained and normalized by the date rate. The normalized throughput provides a good indicator to compare the efficiency of
the different MAC protocols. The following results are obtained:

- From the maximum throughput analysis, the MAC protocol of WIGEE, HiperLAN/2 and the data aggregate modes of 802.11n have the potential to operate at the data rate of Gbps. Other protocols and other modes of 802.11n show low efficiency at the data rate of Gbps.

- All protocols but WiGEE and HiperLAN/2 share a common property, i.e. the normalized throughput decreases as the data rate increases. The rate independent overhead is contributed to the throughput decrease. This part of overhead becomes the main obstacle to directly utilize current wireless WLAN and WPAN MAC protocols in gigabit WLANs.

- A large frame size provides better throughput performance than a small frame size in an error free or low BER conditions. It suggests the use of a large payload size for high speed communication under those channel conditions.

- In addition to using large payload sizes, the block ACK and data aggregation mechanism can significantly reduce the control overhead during the data transmission, and thus increase the protocol efficiency.

- Most protocols use superframes to group the channel access. The length of superframe has limited impact on the throughput performance when it takes the value greater than a certain threshold, which is related to the average frame size.

Moreover, we study the saturated throughput of the 802.11 system under the DCF access mode, which is a typical contention based access mode in WLANs. The result shows the contention based protocol has significant low throughput efficiency at the data rate of Gbps. The throughput of the system is highly related to the number of collisions occurring during
the contention process. As a result, the overall throughput of the system varies according to the number of nodes in the system and the load of each node. To achieve a stable throughput performance, a guaranteed access based MAC protocol is preferred in gigabit WLANs.

7.1 Future Work

Although it is technically feasible to provide Gbps transmission in WLAN, lot of research remains before gigabit WLANs can be massively deployed. Here we focus the issues on the link and MAC layer.

Firstly, since the gigabit WLAN is more sensitive to the channel condition, the cross layer design that introduces the physical layer information such as CSI into the MAC layer is desirable for better access control, bandwidth allocation, and QoS guarantee. The CSI can be used to achieve an adaptive data rate scheme at the MAC layer, which adapts the data rate to the channel condition of the link. The mechanism to determine the data rate wisely is required. Moreover, it is meaningful to introduce the CSI into the scheduling or bandwidth allocation algorithm at the MAC layer. The scheduling or bandwidth allocation algorithm that jointly considers the multiple data rate and channel condition will be a promising approach to achieve high performance in gigabit WLANs.

Secondly, the use of directional antennas will be necessary in gigabit WLANs. A more advanced option is beamforming technique via MIMO antennas. The advanced antenna techniques provide gigabit WLANs extra gain to combat against the multipath fading, thus increasing the data rate and the communication range. Those techniques require inherent support in the MAC layer for negotiating the operating parameters. This is an especially interesting topic for distributed systems, where the MAC needs to be combined with the beamforming respectively the antenna selection.
It will therefore hold tables with antenna parameters for each destination node and thus perform routing for each packet, leading to the integration of layer 1 and 3 into the MAC.

Finally, it has been proven that block ACK and data aggregation are effective approaches to improve the system throughput. However, a large frame is more susceptible to channel errors. Therefore a trade off should be taken on the frame size according to the channel condition. The ideal approach is to develop an algorithm that adapts the frame or block size to the channel conditions.

For WiGEE, the future work includes implementing multicast function in the system, introducing the VLAN concept into the system for service isolation, and developing better scheduling and bandwidth allocation schemes for different application scenarios.
Bibliography


[94] “Information technology - Telecommunications and information exchange between systems - Local and metropolitan area networks specific requirements- Part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications,” IEEE 802.11, 1999.

