Efficient Routing and Resource Sharing Mechanisms for Hybrid Optical-Wireless Access Networks

by

Yousef Dashti

A Dissertation Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy

Approved March 2016 by the Graduate Supervisory Committee:

Martin Reisslein, Chair
Yanchao Zhang
John Fowler
Patrick Seeling

ARIZONA STATE UNIVERSITY
May 2016
ABSTRACT

The integration of passive optical networks (PONs) and wireless mesh networks (WMNs) into Fiber-Wireless (FiWi) networks has recently emerged as a promising strategy for providing flexible network services at relative high transmission rates. This work investigates the effectiveness of localized routing that prioritizes transmissions over the local gateway to the optical network and avoids wireless packet transmissions in radio zones that do not contain the packet source or destination. Existing routing schemes for FiWi networks consider mainly hop-count and delay metrics over a flat WMN node topology and do not specifically prioritize the local network structure. The combination of clustered and localized routing (CluLoR) performs better in terms of throughput-delay compared to routing schemes that are based on minimum hop-count which do not consider traffic localization. Subsequently, this work also investigates the packet delays when relatively low-rate traffic that has traversed a wireless network is mixed with conventional high-rate PON-only traffic. A range of different FiWi network architectures with different dynamic bandwidth allocation (DBA) mechanisms is considered. The grouping of the optical network units (ONUs) in the double-phase polling (DPP) DBA mechanism in long-range (order of 100 Km) FiWi networks is closely examined, and a novel grouping by cycle length (GCL) strategy that achieves favorable packet delay performance is introduced. At the end, this work proposes a novel backhaul network architecture based on a Smart Gateway (Sm-GW) between the small cell base stations (e.g., LTE eNBs) and the conventional backhaul gateways, e.g., LTE Servicing/Packet Gateway (S/P-GW). The Sm-GW accommodates flexible number of small cells while reducing the infrastructure requirements at the S-GW of LTE backhaul. In contrast to existing methods, the proposed Sm-GW incorporates the scheduling mechanisms to achieve the network fairness while sharing the resources among all the connected small cells base stations.
To my family
ACKNOWLEDGMENTS

I would like to thank my advisor Prof. Martin Reisslein for all his support, motivation, and guidance throughout my publications. My sincere thanks goes to my committee members, Dr. Yancho Zhang, Prof. John Fowler, Dr. Patrick Seeling, and Dr. Michael McGarry for their support and feedbacks. I would also like to thank my friends who were also my lab mates for several years, Anu Mercian, Akhilesh Thyagaturu, Ziyad Alharbi, Po-Yen Chen, Revak Tyagi, and Xing Wei, for their support and help. Lastly, but not the least, I would like to thank the Public Authority for Applied Education and Training (PAAET) in Kuwait who granted me the financial support through my graduate study.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>LIST OF TABLES</th>
<th>vii</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF FIGURES</td>
<td>viii</td>
</tr>
</tbody>
</table>

## CHAPTER

1. **INTRODUCTION**
   1.1 CluLoR: Clustered Localized Routing for FiWi Networks
   1.2 Grouping by Cycle Length (GCL) for Long-Range FiWi Networks
   1.3 Smart Gateways (Sm-GW) for LTE Femto-Cell Access Network: Resource Sharing

2. **CluLoR: CLUSTERED LOCALIZED ROUTING FOR FiWi NETWORKS**
   2.1 Related Work
   2.2 Principles of Clustered Localized Routing (CluLoR)
     2.2.1 Clustered Routing
     2.2.2 Localized Routing
   2.3 Simulation Setup
   2.4 Clustered Routing: Impact of Number of Cluster Heads
     2.4.1 Performance with Relay Routers
     2.4.2 Goodput for Delay Sensitive Traffic
   2.5 Localized Routing: Comparison with Unlocalized Minimum-Hop-Count Routing
   2.6 Evaluation for Highly Loaded Fiber Network

3. **GROUPING BY CYCLE LENGTH (GCL) FOR LONG-RANGE FiWi NETWORKS**
   3.1 Related Work
     3.1.1 FiWi Networks
<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1.2 Dynamic Bandwidth Allocation (DBA) Mechanisms</td>
<td>32</td>
</tr>
<tr>
<td>3.2 Long-Reach FiWi Network Model</td>
<td>34</td>
</tr>
<tr>
<td>3.2.1 Architecture</td>
<td>34</td>
</tr>
<tr>
<td>3.2.2 Dynamic Bandwidth Allocation (DBA) Mechanisms</td>
<td>35</td>
</tr>
<tr>
<td>3.3 ONU Grouping Strategies</td>
<td>39</td>
</tr>
<tr>
<td>3.3.1 Motivation</td>
<td>39</td>
</tr>
<tr>
<td>3.3.2 ONU Traffic Load Estimation</td>
<td>39</td>
</tr>
<tr>
<td>3.3.3 Distance Grouping (DG)</td>
<td>40</td>
</tr>
<tr>
<td>3.3.4 Distance Balancing (DB)</td>
<td>41</td>
</tr>
<tr>
<td>3.3.5 Load Grouping (LG)</td>
<td>41</td>
</tr>
<tr>
<td>3.3.6 Load Balancing (LB)</td>
<td>41</td>
</tr>
<tr>
<td>3.3.7 Grouping by Cycle Length (GCL)</td>
<td>41</td>
</tr>
<tr>
<td>3.4 Setup of Simulation Evaluation</td>
<td>43</td>
</tr>
<tr>
<td>3.4.1 FiWi Network Architecture</td>
<td>43</td>
</tr>
<tr>
<td>3.4.2 Network Traffic Scenarios</td>
<td>45</td>
</tr>
<tr>
<td>3.4.3 Network Protocols</td>
<td>47</td>
</tr>
<tr>
<td>3.4.4 Delay Metrics</td>
<td>48</td>
</tr>
<tr>
<td>3.5 Results for Mixing of Wireless (FiWi) and PON Traffic</td>
<td>48</td>
</tr>
<tr>
<td>3.5.1 Impact of DBA Mechanism</td>
<td>48</td>
</tr>
<tr>
<td>3.5.2 Impact of Long-Reach Propagation</td>
<td>53</td>
</tr>
<tr>
<td>3.6 Results for ONU Grouping in DPP with Excess Sharing</td>
<td>60</td>
</tr>
<tr>
<td>3.6.1 Traffic Ratio 1:10:40: Pronounced Mixed ONU to PON-only ONU Load Difference</td>
<td>60</td>
</tr>
<tr>
<td>3.6.2 Traffic Ratio 1:20:30: Mild Mixed ONU to PON-only ONU Load Difference</td>
<td>63</td>
</tr>
</tbody>
</table>
4 SMART GATEWAYS (Sm-GW) FOR LTE FEMTO-CELL ACCESS NETWORK:

RESOURCE SHARING .......................................................... 66

4.1 Related Work ............................................................... 66

4.2 Proposed Sm-GW Scheduler ........................................... 67

4.2.1 Equal Share Scheduling ............................................. 68

4.2.2 Excess Share Scheduling ........................................... 70

4.2.3 Fairness of the Schedulers ......................................... 73

4.2.4 Overhead Evaluation .................................................. 75

4.3 Evaluation of Sm-GW Scheduling ................................. 76

4.3.1 Simulation Setup ....................................................... 76

4.3.2 Sm-GW without Scheduler ........................................ 77

4.3.3 Sm-GW with Scheduler ............................................. 78

4.3.4 Delay Performance .................................................... 80

4.3.5 Fairness Index Performance ....................................... 81

5 CONCLUSION AND FUTURE DIRECTIONS ......................... 83

REFERENCES ........................................................................ 86
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Pseudo-Code Summary of Localized Routing Without Relays at Different Types of Network Nodes.</td>
<td>14</td>
</tr>
<tr>
<td>2.2 Quad Mode Payload Sizes</td>
<td>16</td>
</tr>
<tr>
<td>2.3 Pseudo-Code Summary of Localized Routing With Relay Routers for the Different Types of Network Nodes.</td>
<td>22</td>
</tr>
<tr>
<td>3.1 Summary of Dimensions of Design Space for Dynamic Bandwidth Allocation (DBA).</td>
<td>36</td>
</tr>
</tbody>
</table>
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Enterprise Deployment of Small (Femto) Cells: The Proposed Smart Gateway (Sm-GW) Is Inserted Between the Femto Cell Base Stations (LTE eNBs and Access Points) and the Conventional Backhaul Entities, Such As LTE MME and S-GW. The Sm-GW Flexibly Connects to a Multitude of Femto Cell Base Stations and Dynamically Allocates the Limited Backhaul Capacity to the Different Femto Cells.</td>
<td>6</td>
</tr>
<tr>
<td>2.1 FiWi Network Structure: Wireless Nodes Are Organized Into Different Zones That Operate on Different Radio Frequencies (Channels).</td>
<td>12</td>
</tr>
<tr>
<td>2.2 Illustration of Clustered Routing With Different Numbers of Cluster Heads in a Zone: Wireless Nodes Direct All Communication With Nodes Outside the Zone Through a Cluster Head. The Cluster Communicates With the Gateway Router, Which Is Connected to an ONU. The Configuration With 4 Heads Corresponds Effectively to an Unclustered Benchmark As All Nodes Communicate Directly (Without Going Through a Cluster Head) With the Gateway Router.</td>
<td>13</td>
</tr>
<tr>
<td>2.3 Clustered Routing: Mean End-to-End Packet Delay for Different Numbers of Cluster Heads in a Zone.</td>
<td>17</td>
</tr>
<tr>
<td>2.4 Clustered Routing: Normalized Throughput for Different Numbers of Heads.</td>
<td>20</td>
</tr>
<tr>
<td>2.5 FiWi Network Structure With Relays Operating at the Two Radio Frequencies of the Two Adjacent Zones. Packets Destined to an Adjacent Zone Are Routed Through the Relay Router, Bypassing the Cluster Heads and Gateway Router.</td>
<td>23</td>
</tr>
</tbody>
</table>
2.6 Clustered Routing With Relays: Mean Packet Delay As a Function of Traffic Load for Different Numbers of Cluster Heads in a Zone, With and Without Relays Between Adjacent Zones. ................................. 23

2.7 Clustered Routing With Relays: Normalized Goodput for Traffic With Delay Limit of 50 ms. ................................................................. 25

2.8 Illustration of Clustered Localized Routing (CluLoR) With Two Heads With Relays for a Scenario With a Regular Wireless Router As Source: CluLoR Avoids the Traversal of Zones That Do Not Contain the Source or Destination Node by Routing Through the Fiber Network. In Contrast, Unlocalized Routing Traverses Zones Without a Source/Destination (Using the Relays Operating at Both Radio Frequencies of Adjacent Zones) to Achieve the Minimum Hop-Count. ................................. 26

2.9 Localized Routing: Mean Packet Delay for CluLoR Vs. Unlocalized Minimum Hop-Count Routing. ......................................................... 27

2.10 Localized Routing: Normalized Goodput for 50 ms Delay Limit for CluLoR Vs. Unlocalized Minimum Hop-Count Routing. ................. 28

2.11 Mean End-to-End Packet Delay for PON With Fiber Background Traffic (32 ONUs With 2:1 Ratio of Background Traffic : Peer-to-Peer Wireless Station Traffic). ......................................................... 30

3.1 Illustration of Long-Reach (LR) FiWi Network Architecture: Distributed Optical Network Units (ONUs) Are Connected Via Optical Fiber to a Central Optical Line Terminal (OLT). Some of the ONUs Support Via Wireless Gateways the Communication of Distributed Wireless Stations. 34
<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2</td>
<td>Dedicated ONU FiWi Network Architecture Set-up: Four ONUs (ONUs 1–4) Are Dedicated to Wireless (FiWi) Traffic While Four ONUs (ONUs 5–8) Are Dedicated to Wired (PON) Traffic. The Figure Also Illustrates Cluster-LoR Routing: The Wireless Network Is Organized Into Zones, Each Operating at a Different Frequency. Wireless Stations Route Traffic Through a Cluster Head Towards the Gateway for Transmission Over the PON.</td>
</tr>
<tr>
<td>3.3</td>
<td>Mixed ONU FiWi Network Architecture Set-up: Four ONUs (ONUs 1–4) Serve a Mix of Wireless (FiWi) and Wired (PON) Traffic, While Four ONUs (ONUs 5–8) Serve Only Wired (PON) Traffic.</td>
</tr>
<tr>
<td>3.4</td>
<td>Mean Wireless (FiWi) and PON Packet Delays for Different DBA Mechanisms and Cycle Lengths. Fixed Parameters: 15–20 Km PON With Dedicated ONU Architecture, All-Server Traffic.</td>
</tr>
<tr>
<td>3.5</td>
<td>Mean Wireless (FiWi) and PON Packet Delays for Different DBA Mechanisms for CH-Server; STN-P2P Traffic. Fixed Parameters: 15–20 Km PON With Dedicated ONU Architecture, Cycle Length $Z = 2$ ms.</td>
</tr>
<tr>
<td>3.6</td>
<td>Mean FiWi Packet Traffic Delay for Different DBAs and Cycle Lengths for Long-Reach 90–100 Km PON. Fixed Parameters: Dedicated ONU Architecture, All-Server Traffic.</td>
</tr>
<tr>
<td>3.7</td>
<td>Mean PON Packet Traffic Delay for Different DBAs and Cycle Lengths for Long-Reach 90–100 Km PON. Fixed Parameters: Dedicated ONU Architecture, All-Server Traffic.</td>
</tr>
<tr>
<td>3.8</td>
<td>Mean FiWi Packet Delay for Different DBAs for Mixed ONU FiWi Architecture. Fixed Parameters: 90–100 Km Long-Range FiWi Network, All P2P Traffic, 1:20:30 Traffic Ratio.</td>
</tr>
</tbody>
</table>
3.9 Mean PON Packet Delay for Different DBAs for Mixed ONU FiWi Architecture. Fixed Parameters: 90–100 Km Long-Range FiWi Network, All P2P Traffic, 1:20:30 Traffic Ratio. .................................................... 59

3.10 Mean FiWi and PON Packet Delays for Different ONU Grouping Strategies. Fixed Parameters: Mixed ONU Architecture With 80–120 Km Long-Range Propagation, (dpp., exc.) DBA, \( Z = 4 \text{ ms} \) Cycle Length, All-P2P Traffic With Ratio 1:10:40. ................................................................. 61

3.11 Mean FiWi and PON Packet Delays for Distance Grouping (DG) and Distance Balancing (DB) ONU Grouping Strategies for (dpp., lim.) DBA Mechanism. Fixed parameters: Mixed ONU Architecture With 80–120 Km Long-Range Propagation, \( Z = 4 \text{ ms} \) Cycle Length, All-P2P Traffic With Ratio 1:10:40. ................................................................. 62

3.12 Mean FiWi and PON Packet Delays for Different ONU Grouping Strategies. Fixed Parameters: Mixed ONU Architecture With 80–120 Km Long-Range Propagation, (dpp., exc.) DBA, \( Z = 4 \text{ ms} \) Cycle Length, All-P2P Traffic With Ratio 1:20:30. ................................................................. 64

4.1 Illustration of Smart Gateway (Sm-GW) Function: (a) Sm-GW Receives Backhaul Transmission Bitrate Allocation \( G \) From Operator. Based on Uplink Transmission Bitrate Requests (Reports) \( \rho_n \) From Individual eNBs \( n, 1, \ldots, N \), Sm-GW Gives Uplink Transmission Grants \( \gamma_n \) to the eNBs. \( \delta \) Represents the Granting Delay. (b) The eNB Requests (Reports) \( \rho_n \) and Sm-GW Grants \( \gamma_n \) Are Followed by the eNB Uplink Data Transmissions (in Parallel) in a Fixed-Period Cycle. .............................................. 71

4.2 Overhead Analysis Due to Network Protocol Scheduler at the Sm-GW and Small Cell eNBs in the Indoor Environment (20m < \( d < 500 \text{m} \)). .......... 75
<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.3</td>
<td>78</td>
</tr>
<tr>
<td>4.4</td>
<td>79</td>
</tr>
<tr>
<td>4.5</td>
<td>80</td>
</tr>
<tr>
<td>4.6</td>
<td>81</td>
</tr>
</tbody>
</table>
Chapter 1

INTRODUCTION

Wireless and optical networking technologies at the early stages were deployed for different respective communication settings. Due to the fact that those technologies aim to solve different problems when they were initially developed, it is hard for one given technology to overcome many of the challenges arising in the access network area. The merging of optical access technologies with wireless access technologies by capitalizing on their respective advantages could lead to powerful solutions. Passive optical networks (PONs) connect several distributed optical network units (ONUs) at subscriber premises with a central optical line terminal (OLT) at high bandwidth of up to 10 Gbps [1–3] with reach extending over long distances [4–9]. We note that a plethora of studies has examined related TDM/WDM PONs, see e.g., [10–21]; however, they have high deployment costs. On the other hand, wireless mesh networks (WMNs) offer flexible communication and eliminate the need for a fiber drop to every user in the network, but offer only relatively low bandwidth, which is impacted by interference among ongoing wireless transmissions [22–34].

Fiber-Wireless (FiWi) access network combine wireless access networks with optical access networks. Wireless access networks can flexibly support distributed wireless users, while optical access networks provide high transmission bit rates through the optical fiber [35–38]. FiWi networks have begun to attract extensive research interest as they represent a promising approach for solving the problem of high-speed Internet access “over the last mile”. In particular, FiWi networks with a Passive Optical Network (PON) as the optical network have been intensely studied in the past few years as a PON can provide high transmission bit rates to support demand-
ing applications, e.g., multimedia applications [39–41], at relatively low maintenance cost [42–48].

We also focus on advances of the wireless access network based on FiWi access networks. Multitudes of wireless access technologies can be integrated in the front-end of the FiWi network mainly the WiFi, WiMaX, and the LTE [49, 50]. Recent advances in the field of FiWi has been the integrating of LTE access network into the FiWi [51]. Emerging applications for mobile devices of the 5G networks requires features such as increased capacity, bandwidth and ultra low latency. In order to meet the demands and requirements, access networks are undergoing transitions. LTE access network consists of entities such as P-GW, S-GW, eNBs, and signaling entities. Learned lessons in the early stages of FiWi networks can further help in improving LTE access networks. LTE access access network involves advanced efficient scheduling mechanisms between the eNBs and the end-users. Currently less attention is being paid to the scheduling mechanism between the eNBs and its associated S-GW. Advances in such scheduling mechanisms can further improve the performance of LTE access network to accommodate large number of eNBs.

1.1 CluLoR: Clustered Localized Routing for FiWi Networks

We focus on the problem of peer-to-peer communication within a given wireless mesh network (WMN). Integrating an optical access network with the wireless mesh network could possibly lead to higher throughput and lower end-to-end packet delays. Without an optical access network, all traffic has to go through the WMN, which results in high network interference and in turn limits the network throughput. By combining the optical access network and the WMN to an integrated fiber and wireless (FiWi) network, the traffic could be routed from the source node in the WMN over wireless hops to a nearby gateway wireless router where it could be routed via
the fiber network to a gateway wireless router near the destination node. This scenario would reduce interference in the wireless mesh network, and increase throughput between the two communicating peers.

As elaborated in Section 2.1, many FiWi network architectures and routing protocols have been explored in the past few years [35]. To the best of our knowledge, the existing FiWi routing approaches mainly consider a “flat” topology for the WMN, i.e., the existing approaches do not consider a hierarchical clustering structure of the WMN nodes. Moreover, the specific local network structure, i.e., the closest local gateway from the WMN to the PON, has not been prioritized over multi-hop transmissions through the WMN. Clustering has proven very beneficial in purely wireless networks [52, 53]. In this dissertation, we examine the combined effects of clustered localized routing [54]. We consider a common WMN setting where the wireless nodes are organized into zones that operate on different radio channels [55–57]. We allow wireless nodes to send traffic to each other directly only when they are in the same zone. Otherwise, all traffic has to go through an assigned cluster head which in turns routes the traffic to the assigned gateway router (which in turn routes the traffic to the destination zone, possibly utilizing the optical network).

1.2 Grouping by Cycle Length (GCL) for Long-Range FiWi Networks

FiWi network research based on PONs has mainly focused on normal-range PONs to date, as reviewed in detail in Section 3.1. A normal-range PON covers a distance of about 20 km between the central Optical Line Terminal (OLT) and the distributed Optical Network Units (ONUs), where users connect to the PON. In contrast, we investigate FiWi networks based on long-reach (LR) PONs covering on the order of 100 km [58]. LR PONs can amortize costs over a larger user population in the larger covered area, but pose unique challenges due to the long propagation delays [59–62].
Our first main contribution is an extensive simulation study of the mixing of two traffic types in the LR FiWi network: Wireless traffic that is generated at wireless stations and traverses the wireless network before transmission over the PON; this wireless traffic is true FiWi traffic as it traverses both the wireless and fiber network parts. Due to the transmission over the relatively lower-rate wireless network (compared to the fiber network), the wireless (FiWi) traffic has typically a relatively lower bit rate than conventional traffic that is transmitted only over the PON. This conventional PON-only traffic is our second considered traffic type. In our extensive simulations we examine the packet delays of wireless (FiWi) traffic and PON traffic for a wide range of FiWi network architectures and dynamic bandwidth allocation (DBA) mechanisms.

Our simulations revealed that the double-phase polling (DPP) DBA mechanism [63] achieves low packet delays for both traffic types. DPP relies on an assignment of the ONUs to two groups that take turns transmitting on the shared upstream wavelength channel and strive to mask each others’ idle times due to the long round-trip propagation delay of the PON polling control messages. The effects of this group assignment have to the best of our knowledge not yet been examined in detail. We compare several elementary grouping strategies and introduce a novel grouping by cycle length (GCL) strategy. The GCL strategy strives to balance the lengths of the polling cycles of the two ONU groups so that they can effectively mask each others’ idle times. We find that the GCL strategy, which is based on the OLT-to-ONU distances and the ONU load levels, significantly outperforms elementary grouping strategies that consider only OLT-to-ONU distances or ONU load levels.
1.3 Smart Gateways (Sm-GW) for LTE Femto-Cell Access Network: Resource Sharing

The expansion of the cellular industry in the recent past has presented a wide range of opportunities for technology development in the access networking domain. Cellular communication plays an important role in enabling emerging technologies for the Internet of Things (IoT) [64], cloud/mobile computing, and big data. The IoT paradigm connects virtually every electronic device which supports network connectivity to the Internet grid. Connecting such a high magnitude of devices to the Internet creates opportunities for seamless network access so as to monitor and control devices remotely. However, balanced progress has to be made across the entire technology chain of cellular wireless communication as well as the access (backhaul) networks and the core networks in order to benefit from the presented opportunities.

Emerging applications for mobile devices, such as artificial intelligence and virtual/augmented reality, require ultra-low latency and high data rates to offload the computationally intensive tasks at the mobile devices to the cloud [65]. Communicating with the cloud for computational offloading can incur large propagation delays. Therefore, the new paradigm of Edge/Fog networking [66-68], has been proposed to bring computational capabilities close to the end-devices. Although wireless technologies have offered various solutions [69, 70], such as carrier aggregation in LTE-Advanced, to support the advance applications, the wireless spectrum resources in the licensed bands are expensive and scarce. In contrast to having high aggregated spectrum bandwidth of to support the user requirements in a single cell, multiple small cells can be created to coexist with neighboring cells while sharing the same spectrum resources. Small cells offer the a potential solution to the limitation of present day of wireless protocols [71-77]. However, small cells challenges include
interference coordination [78], backhaul complexity [79, 80], and increased network infrastructure cost [81]. In this dissertation we propose a solution to reduce the backhaul infrastructure cost and the complexity of access networks supporting small cells.

Small cell networks are expected to be privatively owned [82]. Therefore it is important to enable usage flexibility and the freedom of investment in the new network entities (e.g., gateway, servers) and the network infrastructures (e.g., switches, optical fiber) by the private owners of small cells. The proposed Smart Gateways (Sm-GWs) enable the private owners of small cells to utilize the cellular gateways, e.g., LTE Serving Gateway (S-GW) and Packet Gateway (P-GW), based on different service level agreements (SLAs) possibly across multiple operators. In Figure 1.1, we illustrate a possible small cell deployment scenario in an enterprise building. We refer to small cells as *femto cells* in the context of LTE. Our proposed
(Smart-Gateway) Sm-GW at the enterprise building will enable the sharing of the access network resources by all the small cells in the building.

As the deployment of small cells grow rapidly, static allocations of the network resources in the backhaul entities for individual small cell would result in under-utilization of resources due to the bursty traffic nature of modern applications. Although with the proposal of several advance techniques for the management of eNB resources [83–87], very little attention has been given to the consequences of small cell deployments on the gateways [88]. In the present wireless network architectures, such as LTE, the main reasons for under-utilization are: 1) static (or non-flexible) network resource assignments between an eNB and the gateway (S-GW, P-GW), and 2) lack of traffic coordination between the eNBs and their gateways. For the same reasons, it is physically impossible to accommodate additional eNBs on a particular gateway (S/P-GW), e.g., to increase the coverage area, even when only very low traffic levels originate from the connected eNBs. Specifically, port exhaustion at the gateways commonly limits the deployment of additional eNBs. Additional eNBs would require a new gateway (S/P-GW), resulting in high expenditures. In contrast our proposed Sm-GW accommodate large numbers of eNBs by flexibly sharing network resources.

However, persistent over-subscriptions of the eNBs at a single Sm-GW under high load situation can affect the user satisfaction. In order to overcome such over-subscriptions, new Sm-GWs along with new connections to operator core gateways (S/P-GWs) have to be installed. Nevertheless, our proposed Sm-GW approach curtails the required infrastructure increase in the operator’s core (i.e., S/P-GW and MME for LTE) for the deployment of large numbers of small cells.

Typically, the aggregate service capacity (transmission bit rate) of all small cells within a building is much larger than the single connection service provided by
the cellular operators, thus creating a bottleneck. For instance, if 100 femto cells, each supporting 1 Gbps in the uplink are deployed in a building, two issues arise: 1) suppose one S-GW supports 16 connections, then 7 S-GWs are required, and 2) the aggregated traffic requirement from all the 7 S-GWs would result in 100 Gbps, causing a similar requirement at the P-GW. We emphasize that the discussed requirements are for a single building and there could be several building belonging to same organization within a small geographical location. We argue here that: 1) providing the 100 Gbps connectivity to every building would be very expensive, and 2) with sharing we can reduce the resource requirement to, say, 1 Gbps, and 3) we can curtail the infrastructure increase of S/P-GW, reducing the cost for the cellular operators.
Chapter 2

CluLoR: CLUSTERED LOCALIZED ROUTING FOR FiWi NETWORKS

2.1 Related Work

The recent survey [38] gives an overview of hybrid optical-wireless access networks. The Hybrid Wireless-Optical Broadband-Access Network (WOBAN) Architecture [37] is a pioneering FiWi network structure. The study [37] identified FiWi networking challenges with regard to network setup (placement of ONU, Base Stations (BSs), and OLT to minimize the cost), and efficient routing protocols. The FiWi network planning problem has been further studied in [89]. The studies [90–93] proposed FiWi architectures and reconfiguration algorithms in order to serve the needs of the hybrid access network users.

Some of the first studies that examined peer-to-peer communication in a FiWi network were by Zheng et al. [94, 95]. These studies noted the significance of integrating the optical networks with the mesh networks to achieve significant performance improvements in terms of overall throughput and average packet end-to-end delays. Also, a simple routing protocol was proposed based on minimum-hop-count, which includes the gateway routers to the fiber network as part of the hop count. Li et al. [96] also studied the problem of peer-to-peer communications. The main focus was on implementing a novel arrayed waveguide grating based WDM/TDM PON structure, including wavelength assignment for groups of ONU and a decentralized dynamic bandwidth allocation (DBA) algorithm, that supports direct communication between the ONU without the traffic going through the OLT which could lead to improved end-to-end delay and throughput. Similarly, studies [97, 98] focused on inter-ONU communications by deploying a star coupler (SC) at the remote node (RN) to broadcast the packets of one ONU to all other ONU, while [99] focused on the medium
access control problem in radio-over-fiber networks. A WDM EPON that supports inter-ONU communications in which the polling cycle is divided into two sub-cycles was proposed in [100]. In this study, which is focused on FiWi routing, we consider a TDM PON with interleaved polling with adaptive cycle time (IPACT) with gated service dynamic bandwidth allocation [101, 102].

Routing protocols and algorithms for FiWi access networks have been the main focus of several studies, whereby some focus on routing the packets in the wireless front-end only, or routing the packets through the wireless and optical domains combined to achieve better performance. Early work that focused on routing algorithms in FiWi access networks includes the Risk-And-Delay-Aware Routing (RADAR) [103], Delay-Aware Routing Algorithm (DARA) [104], Delay-Differentiated Routing Algorithm (DDRA) [105], and Capacity-and-Delay-Aware Routing (CaDAR) [106]. Other recent studies on routing techniques in hybrid wireless-optical access network have focused on energy efficient routing [107], and Availability-Aware routing [108] as well as analytical frameworks for capacity and delay evaluation [109]. Most of these studies approach the routing as an optimization problem in order to find the optimum solution. However, all of them considered a flat topology, without a cluster structure, in the WMN. In contrast, this study focuses on the effects of clustered localized routing in the WMN on FiWi network performance.

A number of other studies have focused mainly on load balancing and Transmission Control Protocol (TCP) related issues in FiWi networks. Shaw et al. [110] proposed an integrated routing algorithm that adapts to the changes of the traffic demands within different regions of the wireless network in order to achieve load balancing in the hybrid network. The route assignment is located in the central hub. A hybrid TDM/WDM network with a wavelength assignment scheme that focuses on
assigning a minimum number of wavelength to each group of ONUs while the maximum throughput at the ONUs is maintained was examined in [111]. The performance of multipath routing in FiWi and its effect on TCP performance due to out-of-order packets at the destination node was analyzed in [112, 113]. An integrated flow assignment and packet re-sequence approach that obtains the probabilities of sending along the different paths with the objective of reducing the arrived out-of-order packets at the OLT was explored in [112]. A DBA technique that gives higher priority to the flows that trigger TCP fast retransmissions was proposed in [113]. We do not specifically examine TCP traffic; instead, we focus on traffic transmitted with the User Datagram Protocol (UDP).

We note for completeness that recently energy efficiency in FiWi access network has begun to attract research interest, see e.g., [51, 114-117]. Survivability and protection techniques in FiWi access networks have been studied in [90, 118-124], while network coding in FiWi access network has been explored in [125, 126].

In summary, complementary to the existing FiWi networking literature, this study focuses on the effects of a combining (i) routing over cluster heads with (ii) prioritizing transmissions to be routed through the local WMN-PON gateway on overall FiWi network performance. While the existing FiWi routing literature has mainly considered a “flat” topology without a clustering structure of the wireless nodes, clustering techniques have been extensively studied in the area of purely wireless networking, see e.g., [52, 53]. To the best of our knowledge clustered routing in a FiWi network has so far only been studied in [127], which focused on the distribution of traffic in the downstream direction. The present study is the first to examine the benefits of clustered localized routing for peer-to-peer traffic involving both upstream and downstream PON transmissions in a FiWi network.
2.2 Principles of Clustered Localized Routing (CluLoR)

We focus on a setting where the wireless stations (nodes), which could be WiFi routers (e.g., IEEE 802.11g WiFi routers) are organized into different zones, as illustrated in Fig 2.1. Each zone operates on a different radio channel than its neighboring zones [55–57]. There is a single gateway router that serves the zone closest to it, e.g., zones 1–4 in Fig. 2.1 are served by gateway router G1. Each gateway router has an Ethernet interface that is connected directly to an ONU. Within this network setting, we examine the two principles of clustered and localized routing that are outlined in the next two subsections and combined to form the CluLoR scheme.

2.2.1 Clustered Routing

In each zone, there is a node that is assigned as a cluster head, as illustrated in the upper left illustration in Fig. 2.2. (It is possible to have multiple cluster heads for a zone, but for ease of exposition, we initially focus on the case of one cluster head per zone.) The cluster head of a zone is the node that is located closest to the gateway.

Figure 2.1: FiWi Network Structure: Wireless Nodes Are Organized Into Different Zones That Operate on Different Radio Frequencies (Channels).
Figure 2.2: Illustration of Clustered Routing With Different Numbers of Cluster Heads in a Zone: Wireless Nodes Direct All Communication With Nodes Outside the Zone Through a Cluster Head. The Cluster Communicates With the Gateway Router, Which Is Connected to an ONU. The Configuration With 4 Heads Corresponds Effectively to an Unclustered Benchmark As All Nodes Communicate Directly (Without Going Through a Cluster Head) With the Gateway Router.

The cluster head is responsible for routing outbound packets from the regular wireless nodes in the zone on to the gateway router and for routing inbound packets from the gateway router on to the wireless nodes in the zone.

### 2.2.2 Localized Routing

The routing between the wireless nodes (peers) proceeds according to the following three rules, which are summarized in the pseudo-code in Table 2.1: (i) If the communicating peers are within the same zone, then the packet is directly wirelessly transmitted to the destination peer without going through a cluster head or gateway router. (ii) If the zone of the destination peer is serviced by the same gateway router as the zone of the source peer, then the packet is routed by the gateway router to the destination zone without going through the optical network. (iii) If the destination zone is not served by the same gateway router as the source zone, then the packet
Table 2.1: Pseudo-Code Summary of Localized Routing Without Relays at Different Types of Network Nodes.

**Wireless Station - not a Cluster Head:**
if (destination within the same Zone)
    Send packet directly to destination;
else
    Send packet to closest Cluster Head;

**Wireless Station - Cluster Head:**
if (destination within the same Zone)
    Send packet directly to destination;
else
    Send packet to Gateway Router;

**Gateway Router:**
if (destination is in a Zone associated with the same Gateway Router)
    Send packet to Cluster Head associated with the destination wireless station;
else
    Send packet to Optical Network;

is routed through the cluster head to the source-zone gateway router, then to the optical network.

The optical network broadcasts the packet in the downstream direction, where the ONU connected to the destination gateway (gateway router that is closest to the destination zone) accepts it while the other ONUs discard the packet. The destination gateway router then routes the packet via the cluster head to the destination peer in the destination zone. Localized routing ensures that a packet is never wirelessly transmitted in a zone that does not contain the source or destination of the packet.

2.3 Simulation Setup

In our simulations, we evaluate mean end-to-end packet delay and throughput of ChuLoR in a FiWi network. The simulations are conducted in OMNeT++ 4.2.2 [128] using INETMANET-2.0 modules [129]. Specifically, we initially simulate a FiWi
network with 64 wireless nodes and 4 gateway routers. The wireless nodes are placed uniformly in a 1600 m × 300 m region. The 64 wireless nodes are distributed evenly in 16 zones (4 nodes in each zone) resulting in each gateway managing 4 zones. Each of the 4 wireless gateway routers (IEEE 802.11g) is connected to its own ONU through an Ethernet cable with a transmission rate of 1 Gbps. All the ONUs are at a distance of 20 km from the OLT.

Each wireless node is equipped with a single radio interface. The gateway routers are equipped with four different radio interfaces (4 radio channels), whereby each channel is assigned to a single zone that operates on the given radio channel. There are 11 different radio channels possible, whereby a given channel is reused in the furthest zones in order to minimize interference. We employ a log-distance path loss channel model with a path loss alpha value of 2. The radio sensitivity is set to $-85$ dBm and the signal-to-noise ratio threshold is set to 4 dB, whereby the received packet is considered noise if it is below that value. The transmitting power for the wireless routers is set to 20 mW in order for the router that is located furthest in the zone to reach the gateway router. The transmission range is around 250 m. The physical data rate is 54 Mb/s. The retransmit limit for the wireless LAN is set to its default value 7. The buffer size for the wireless interface is set to 1000 packets regardless of the packet size.

We use a quad mode model of payload sizes at the UDP level in order to reach the quad mode of encapsulated packet sizes at the Ethernet level [130], see Table 2.2. The UDP level payload includes the UDP header of 8 bytes, the IP header of 20 bytes, and MAC level header of 18 bytes at the Ethernet layer. The maximum transmission unit (MTU) for the wireless interface is set so as to avoid fragmentation. We consider independent Poisson packet generation processes in the wireless nodes, whereby all
Table 2.2: Quad Mode Payload Sizes

<table>
<thead>
<tr>
<th>Ethernet encapsulated packet size</th>
<th>Payload size (UDP level)</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>64 bytes</td>
<td>18 bytes</td>
<td>60%</td>
</tr>
<tr>
<td>300 bytes</td>
<td>254 bytes</td>
<td>4%</td>
</tr>
<tr>
<td>580 bytes</td>
<td>534 bytes</td>
<td>11%</td>
</tr>
<tr>
<td>1518 bytes</td>
<td>1472 bytes</td>
<td>25%</td>
</tr>
</tbody>
</table>

the wireless nodes have the same mean packet generation rate. For each generated packet at a given wireless node, any of the other wireless nodes in the network is selected as destination with equal probability. All simulation are run until the 95% statistical confidence intervals of the performance measures are less than 5% of the sample means.

2.4 Clustered Routing: Impact of Number of Cluster Heads

In this section we examine the impact of the number of cluster heads in a given zone on the delay and throughput performance. As described in Section 2.3, the simulated wireless network is organized into different zones, whereby each zone has 4 wireless nodes. In each of the zones there is one wireless node (or multiple wireless nodes) that is (are) assigned as the cluster head (heads) of the zone and is (are) responsible for relaying the traffic from/to the gateway router. We examine the effects of having 1, 2, or 4 cluster heads (which we refer to as “heads” for brevity) in the zone, as illustrated in Fig. 2.2.

In the case of 1 head, the head is assigned as the wireless node that is located closest to the gateway router. In the case of 2 heads, the two wireless nodes in the zone that are closest to the gateway router are assigned as the heads. Ties in distance are broken through random selection. The outbound traffic from the other wireless nodes (that are not designated as heads) in a given zone is transmitted to the closest head; the head in turn transmits the traffic to the gateway router. Analogously, the
Figure 2.3: Clustered Routing: Mean End-to-End Packet Delay for Different Numbers of Cluster Heads in a Zone.

Inbound traffic is routed from the gateway router to the head that is closest to the destination node and then onwards by the head to the destination. In the case of 4 heads in a zone, all the wireless nodes in a zone are designated as heads and send their traffic directly to the gateway router. Note that the 4-head case is equivalent to unclustered routing in that all wireless nodes communicate directly with their gateway router, without a cluster hierarchy in the zone.

Figure 2.3 shows the mean end-to-end packet delay in the FiWi network for 1, 2, or 4 heads in a zone. (The 95% confidence intervals are too small to be visible and are omitted.) For each configuration of heads, the network traffic load is incremented until buffer overflows begin to occur. We observe from the figure that assigning 2-heads in the zone gives lower delays compared to the 1-head or 4-heads cases. In addition, we observe from Fig. 2.3 that at low loads, 1 head given lower mean delays than 4 heads. These performance characteristics are mainly due to a trade-off between mean hop-count and transmission distance. In particular, a smaller mean hop-count implies that a packet is transmitted on average fewer times on its way from source to destination. Clearly, fewer transmissions are generally preferable as each transmission
requires networking resources and incurs delay.

In the configuration with 4-heads in a zone (i.e., effectively the unclustered scenario), see Fig. 2.2 all four wireless nodes in a zone send directly packets to the gateway, i.e., all packets originating from the zone need only one hop to reach the gateway. Similarly, all packets arriving to the gateway for delivery to a node in the zone, reach their destination with one hop. Notice that the 4-heads configuration has the minimum mean hop-count among the three configurations illustrated in Fig. 2.2. As the number of heads decreases, the mean hop count increases. Specifically, the 1 head configuration requires one hop to reach the gateway from the head, but two hops to reach the gateway from any other node in the zone. Thus, to summarize, the 1 head configuration has the highest mean hop-count, the configuration with 2 heads has a moderate mean hop-count, and the 4-heads configuration has the lowest mean hop-count.

The transmission distance directly affects the received signal-to-interference and noise ratio (SINR), with transmissions propagating over longer distances being received with lower SINR. Among the considered configurations, see Fig. 2.2, the 4-heads configuration has the longest propagation distances, as all nodes in the zone transmit directly to and receive directly from the gateway router. Especially the propagation distance from the node in the upper left corner in the 4-heads illustration to the gateway router in Fig. 2.2 is the longest propagation distance among any of the three considered configurations. This long-distance transmission is particularly vulnerable to failure due to low SINR and requiring retransmissions. Notice from Fig. 2.2 that in comparison with the 4-heads configuration, the 1-head and 2-heads configurations both have moderate propagation distances, i.e., only moderate chances of a packet transmission being unsuccessful due to low SINR.
Returning to the interpretation of the results in Fig 2.3, we note that the 4-heads configuration incurs the highest mean packet delay mainly due to the long propagation distances and the resulting packet transmission failures due to low SINR and packet re-transmissions. The lower mean-hop count cannot overcome the disadvantage of the long propagation delays and results in relatively frequent packet failures and retransmissions, which dominate the delay characteristics.

In the configuration with 1 head, the propagation distances are shorter, reducing the probability of packet failure due to low SINR. Thus, mean packet delays are slightly reduced compared to the 4-heads configuration. But transmissions from/to 3 wireless nodes in the zone require two hops to reach/come from the gateway router.

The configuration with 2 heads strikes a good balance between low mean-hop count and short propagation distances (i.e., high SINR) achieving the lowest mean packet delays in Fig. 2.3. The 2-heads configuration has similarly short propagation distances for transmissions from/to the wireless nodes in the zone as the 1-head configuration. At the same time, the 2-heads configuration has a lower mean-hop count than the 1-head configuration, since the transmissions from/to one more node in the zone, i.e., the second head, require only one hop to reach/come in from the gateway router.

Figure 2.4 shows the 95 % confidence intervals of the normalized mean (long-run average) throughput in terms of traffic that reaches its final destination. The traffic load is incremented until buffer overflows occur; for each curve, the rightmost point corresponds to the highest traffic load without buffer overflows. The average throughput is measured based on the number packets with their corresponding numbers of bits that are received by the destination wireless nodes. The packets (bits) received by intermediate cluster heads and gateway routers are not taken into account.
Figure 2.4: Clustered Routing: Normalized Throughput for Different Numbers of Heads.

We observe from Fig 2.4 that the mean throughput is statistically the same for 1 head and 2 heads in the zone, whereby they both have higher throughput than the 4 heads. We further observe that the 2-heads case accommodates higher traffic loads, up to about 4.25 Gbps before buffer overflows occur, whereas the 1-head case avoids buffer overflows only up to a load of about 3.75 Gbps. The explanations for these behaviors are as follows. First, the 4-heads case has lower average hop-count than the other two cases; however, the long transmission distance from the wireless node farthest from the gateway router has lower SINR than any transmissions in the 1-head and 2-heads cases. Thus, the farthest-away node relatively frequently requires packet retransmissions and hence reaches the maximum retransmit limit relatively more often compared to the nodes in the 1-head and 2-heads configurations. As a result, more packets are dropped due to reaching the maximum retransmission limit in the 4-heads configuration compared to the 1-head and 2-heads configurations resulting in lower throughput for the 4-heads configuration at low to moderate traffic loads. Moreover, in the 4-heads configuration, the buffers fill up more due to more frequent packet retransmissions, leading to buffer overflows at lower traffic loads (3.5 Gbps);
whereas, the 1-head and 2-heads configurations avoid buffer overflows up to 3.75 and 4.25 Gbps, respectively.

At low loads, both the 1-head and 2-heads configurations achieve similar throughput levels. This is because the (very slightly) shorter transmission distances (i.e., higher SINRs) with the 1-head configuration largely counterbalance its higher hop-count. Similarly, the (very slightly) longer transmission distances (i.e., lower SINRs) largely counterbalance the lower hop-count for the 2-head configuration. As the traffic load grows high and buffer backlogs grow, the bottleneck in the 1 head leads to buffer overflows at a lower traffic rate compared to when 2 heads share the traffic load going wirelessly in and out of a zone. In fact, we have observed in our simulations that in the case of 2 heads, the buffer overflow first occurs at the gateway router as it wirelessly transmits all traffic destined into a zone to the two heads.

2.4.1 Performance with Relay Routers

Relay routers can be thought of as an extra cluster head in the zone. They are only used to relay the packets between neighboring zones, so that if the destination is in an adjacent zone, then the packet is directly transmitted to the relay router, which in turn sends the packet to the destination, as illustrated in Figure 2.5. The Pseudo-code for the routing algorithm is summarized in Table 2.3.

Relay routers are equipped with two different radio interfaces in which are configured to the two radio channels of the two adjacent zones. Relay routers relieve the cluster heads and the gateway routers from sending packets destined to a direct neighbor zone.

Figure 2.6 shows the mean packet delay with 22 relay routers added to the network configuration of Section 2.3 and without added relay routers for the different configurations of cluster heads in a zone. We first observe that the performance with
Table 2.3: Pseudo-Code Summary of Localized Routing With Relay Routers for the Different Types of Network Nodes.

**Wireless Station - not a Cluster Head:**

if (destination within the same Zone)  
   Send packet directly to destination;
else if (destination within adjacent  
   Zone & share a Relay Router)  
   Send packet to Relay Router;
else  
   Send packet to closest Cluster Head;

**Wireless Station - Cluster Head:**

if (destination within the same Zone)  
   Send packet directly to destination;
else if (destination within adjacent  
   Zone & share a Relay Router)  
   Send packet to Relay Router;
else  
   Send packet to Gateway Router;

**Gateway Router:**

if (destination is in a Zone associated  
   with the same Gateway Router)  
   Send packet to Cluster Head associated  
   with the destination wireless station;
else  
   Send packet to Optical Network;

added relay routers for the different numbers of cluster heads in the zone follows the same general pattern as for the network without relays. We also observe that adding relay routers results in substantially lower mean end-to-end packet delays, particularly for moderate to high traffic loads. These mean delay results illustrate the effects of bypassing the cluster heads and gateway routers, which lowers the mean hop-count. Also, the packets destined to adjacent zones avoid the queuing delays in the gateway and head routers.

Upon closer examination of Fig. 2.6, we notice that the relays have a slightly more pronounced effect for the 4-heads configuration compared to the 1-head and
Figure 2.5: FiWi Network Structure With Relays Operating at the Two Radio Frequencies of the Two Adjacent Zones. Packets Destined to an Adjacent Zone Are Routed Through the Relay Router, Bypassing the Cluster Heads and Gateway Router.

Figure 2.6: Clustered Routing With Relays: Mean Packet Delay As a Function of Traffic Load for Different Numbers of Cluster Heads in a Zone, With and Without Relays Between Adjacent Zones.

2-heads configurations. This is because the average propagation distance from the wireless nodes to the relay routers is lower than to the gateway router for the 4-heads configuration. On the other hand, for the 1-head and 2-heads configurations, the
propagation distances from the wireless nodes to the relay (without going through the head(s)) are somewhat longer than the distances to the cluster heads. Thus, the benefits of the relay routers are somewhat less pronounced with 1 or 2 cluster heads compared to the 4-heads configuration.

We observed from additional simulations for which we do not include plots to avoid clutter, that the throughput levels with relays are only very slightly elevated compared to the throughput levels without relay routers (see Fig. 2.4). However, the maximum traffic load that can be accommodated before buffer overflows occur is significantly increased by the relays; specifically for 1 head from 3.75 to 4 Gbps, for 2 heads from 4.25 to 4.75 Gbps, and for 4 heads from 3.25 to 3.75 Gbps.

2.4.2 Goodput for Delay Sensitive Traffic

To obtain deeper insights into the performance of the different cluster head and relay configurations, we simulated our FiWi network with a delay sensitive application. An example of delay sensitive application is online video gaming, for which packet delays should not exceed 50 ms. Higher delays disrupt the interactions between the players making the game impossible to play. In interactive video games, many of the participating players are located in the same geographic region and thus peer-to-peer traffic in a FiWi network is a reasonable model. Figure 2.7 shows the goodput, i.e., the portion of the normalized throughput that arrives within the 50 ms delay limit, for the different configurations. We observe from Fig. 2.7 that clearly the configuration with two heads in a zone combined with relays gives the highest goodput among the considered schemes. The goodput gains with relays are especially pronounced at high traffic loads. For a load of 4.25 Gbps, for instance, the relays increase the goodput by approximately 10 % for the 2-heads configuration.
2.5 Localized Routing: Comparison with Unlocalized Minimum-Hop-Count Routing

In this section we compare the performance of our proposed CluLoR with an unlocalized routing benchmark based on minimum-hop-count routing [94]. CluLoR transmits the traffic wirelessly only in zones that contain the source or the destination; while the traffic is routed through the fiber network from the source to the destination zone. In contrast, with unlocalized routing the traffic may be transmitted wirelessly in zones that contain neither the destination nor the source, i.e., the traffic may traverse some intermediate zones via wireless transmissions following minimum-hop-count routing. We consider in this section the best performing clustered routing configuration from Section 2.4, i.e., the configuration with two heads per zone and with relays.

Figure 2.8 illustrates CluLoR and unlocalized minimum-hop-count routing for an illustrative example with one traffic source, namely a regular wireless station, and four possible destinations. With CluLoR, the traffic is routed through the cluster head (first hop) to the gateway router (second hop), from the gateway router G1 adjacent
Figure 2.8: Illustration of Clustered Localized Routing (ChLoR) With Two Heads With Relays for a Scenario With a Regular Wireless Router As Source: ChLoR Avoids the Traversal of Zones That Do Not Contain the Source or Destination Node by Routing Through the Fiber Network. In Contrast, Unlocalized Routing Traverses Zones Without a Source/Destination (Using the Relays Operating at Both Radio Frequencies of Adjacent Zones) to Achieve the Minimum Hop-Count.

...to the source zone through the fiber network to the gateway router G2 adjacent to the destination zone (third hop), to the cluster heads (fourth hop), and regular wireless station destinations (fifth hop). Clearly, with ChLoR, the traffic is only transmitted wirelessly in a zone that includes the destination; thus there is no wireless interference created in any other zones. In contrast, unlocalized routing based on the minimum hop-count routes the traffic from the source node to the relay router between zones 1 and 2 (first hop), then the relay router transmits the packet on the wireless channel of zone 2 to reach the relay router between zones 2 and 3 (second hop), and the packet...
Figure 2.9: Localized Routing: Mean Packet Delay for CluLoR Vs. Unlocalized Minimum Hop-Count Routing.

is then transmitted in turn by the relay router to reach the destinations in zone 3 (third hop).

Figure 2.9 compares the mean end-to-end packet delay for CluLoR with unlocalized minimum hop-count routing for the configuration with two heads per zone with relays. We observe from the figure that at lower traffic loads, the delays for both routing approaches are comparable. However, as the traffic load increases, CluLoR achieves lower mean packet delays than unlocalized minimum hop-count routing. For a traffic load of 18 Mbps, the mean packet delay with CluLoR is only about two thirds of the delay with unlocalized minimum hop-count routing. The higher delay with unlocalized routing is mainly due to “transit” traffic through zones that contain neither the source nor the destination. The wireless transmissions of this transit traffic increase the interference resulting in higher probability of packet transmissions failing due to low SINR as well as an increased chance of packet collisions. Consequently, more packet re-transmission are required, resulting in increased mean packet delays.

Both clustered localized routing and unlocalized minimum hop-count routing
achieve normalized throughput levels close to 100 %, we do therefore not include a throughput plot here to avoid clutter. The only noticeable difference between CluLoR and unlocalized minimum hop-count routing is that CluLoR accommodates traffic loads up to 19 Mbps without buffer overflows compared to 18 Mbps with unlocalized routing. This behavior is mainly due to the higher interference with unlocalized routing, which causes more packets to become backlogged due to the more frequent retransmissions; hence, increasing the chance of buffer overflows.

Figure 2.10 compares the goodput for a delay limit of 50 ms for CluLoR with unlocalized minimum hop-count routing. We observe that clustered localized routing achieves significantly higher throughput, particularly for high traffic loads. For a traffic load of 18 Mbps, the goodput is over 7 % higher with CluLoR compared to unlocalized routing.

2.6 Evaluation for Highly Loaded Fiber Network

So far our evaluation of CluLoR has focused on networking scenarios with only peer-to-peer traffic among the wireless stations. In this section, we add a high back-
ground traffic load that traverses only the fiber network and examine the impact on peer-to-peer traffic between the wireless stations that is routed following the ChuLoR approach. More specifically, we increase the ONU traffic load by adding a wired incoming traffic component. We also increase the number of ONUs to more heavily load the PON access network.

In particular, we simulate a PON network with 32 ONUs. The ONUs are divided into 8 groups; each group has 4 ONUs. One reason for dividing the ONUs into 8 groups is to minimize the interference between the zones. Having all the 32 ONUs in one region could significantly affect the performance because wireless routers in a far-away cluster could be within the sensing range of a given cluster. The wireless nodes are uniformly distributed in a given region with a distance between each node of 50 meters. As in the set-up in Section 2.3, each ONU handles 16 wireless nodes. The distance separation between each region (group) of ONUs is set to be larger than 1 km. The transmission power is set to 20 mW. Each ONU is associated with 4 zones and each zone is configured with 2 cluster heads. In this section, we do not consider relay routers as our focus is to examine the impact of background traffic in the fiber network on the peer-to-peer traffic of the wireless stations. Omitting relay routers forces more of the peer-to-peer traffic through the fiber network and thus gives a worst-case assessment of the impact of fiber network background traffic.

We maintain a ratio of the incoming ONU traffic to be 2:1 for fiber network background traffic: peer-to-peer wireless node traffic. All traffic follows independent Poisson packet generation processes. For the fiber network background traffic, the destination is considered to be an Internet destination outside the PON network. We measure the delay of fiber background traffic from the instant of packet generation to the instant that the packet is completely received by the OLT.
Figure 2.11 shows the delay performance for background traffic and peer-to-peer wireless station traffic. The x-axis represents the total aggregate incoming traffic load at the 32 ONUs. We observe from the figure that the delays for background and peer-to-peer wireless traffic follow the same curve shape at low loads. However, for moderate to high traffic loads, a pronounced gap opens up between the wireless traffic delay and the background traffic delay. This gap grows wider with increasing traffic load.

The delay results in Fig. 2.11 indicate that at low traffic loads, the delays in the optical network, which is the only network component traversed by the background traffic, dominate the wireless traffic delay. That is, the wireless transmissions to and from the gateway routers contribute relatively little to the delay experienced by the peer-to-peer wireless traffic; most of the delay comes from the PON transmissions (more specifically, the upstream transmissions, which require polling-based medium access control proceeding in polling cycles [131]). On the other hand, for high traffic loads, the probability of collisions in the random access of the wireless
channels increases, which causes retransmissions that in turn increase delays. These wireless transmission delays add quite significantly to the PON delays. For a traffic load close to 0.7 Gbps, the additional wireless transmission delay experienced by the peer-to-peer wireless traffic is approximately 10 ms on top of the roughly 13 ms of the PON delay.
Chapter 3

GROUPING BY CYCLE LENGTH (GCL) FOR LONG-RANGE FiWi NETWORKS

3.1 Related Work

3.1.1 FiWi Networks

The general challenges and benefits of FiWi networking have been discussed in [94, 109, 132, 133]. Architectures for FiWi networks have been examined in [37, 38]. FiWi networks appear particularly promising for the backhaul of wireless network traffic [134–136], and integrated control structures for low-delay transmission of mobile wireless traffic over PONs have been examined in [137, 138]. A wide range of specific issues have been examined for FiWi networks, such as the ONU placement [89, 139–142], energy efficient operation [143, 144], quality of service provisioning [145, 146], as well as survivability [147, 148].

The vast majority of the FiWi studies to date has considered normal-range FiWi networks with one-way PON propagation distances on the order of 20 km. To the best of our knowledge, only few studies have examined FiWi networks with long-range PONs covering on the order of 100 km. Specifically, the few prior studies on LR FiWi networks have mainly focused on energy-efficiency and fault tolerance [51, 120, 149, 150]. Complementary to these prior studies, we focus on the packet delay performance of dynamic bandwidth allocation (DBA) mechanisms in LR FiWi networks in this chapter.

3.1.2 Dynamic Bandwidth Allocation (DBA) Mechanisms

DBA mechanisms have been extensively studied for both normal-range and long-range PONs [151–156]. One branch of the DBA research has focused on PONs with multiple wavelength channels in each direction [17, 111, 157–164] or larger network
structures encompassing several PONs [165]. In contrast, we focus on a FiWi network with a single PON with a single wavelength channel in the upstream (ONUs to OLT) direction.

A wide variety of DBA enhancements have been developed in recent years to cope with the idle times due to the propagation delay of the polling control messages. The DBA enhancements typically stagger multiple polling processes over a basic polling cycle, so that the payload upstream transmissions of some polling process(es) mask the idle times of the other polling process(es) [7, 166–168]. Also, recent enhancements have sought to efficiently and fairly utilize the transmission resources within a given polling process [169–171] and to optimize the timing of the polling processes [172–175].

In the present study, we focus on the Double-Phase Polling (DPP) DBA mechanism [16, 63] as an example of an enhanced DBA mechanism with staggered multiple polling processed. DPP is simple and robust: the ONUs in a given PON are assigned to two distinct (non-overlapping) groups, and each group is then served with the elementary offline polling scheduling framework [16, 176]. DPP has also been found to give favorable performance [7, 16, 172, 177]. DPP has served as the basis for a number of recent DBA refinements, namely a predictive bandwidth DBA scheme with multiple QoS classes [178, 179] as well as a recent FiWi study with an long-term evolution LTE wireless component [180]. A similar approach to DPP has recently been proposed in [181] to use offline polling for low load and online polling for high load. Complementary to [181], our GCL approach exploits the load level to adapt the ONU grouping in DPP. To the best of our knowledge, the grouping of ONUs has not been previously examined in detail, neither in the general PON context, nor in the FiWi network context.
3.2 Long-Reach FiWi Network Model

This section gives an overview of the general features of the Fiber-Wireless (FiWi) network model considered in this study. We first present the general architectural structure of FiWi networks in Section 3.2.1. Then, we give an overview of the examined dynamic bandwidth allocation (DBA) mechanisms in the optical part of the FiWi network in Section 3.2.2. The novel grouping strategies for ONUs in the double-phase polling (DPP) DBA mechanisms are introduced in Section 3.3.

3.2.1 Architecture

As illustrated in Fig. 3.1, we consider a FiWi network architecture with $O$, $O > 1$, distributed Optical Network Units (ONUs). The ONUs are connected with a single shared upstream wavelength channel to the central Optical Line Terminal (OLT). We denote $\tau_o$, $o = 1, \ldots, O$, for the one-way propagation delay [in seconds] from ONU $o$ to the OLT, which we assume to be equal to the OLT-to-ONU $o$ one-way propagation delay. Some of the ONUs are connected with wires with wireless gateway routers.
The wireless gateway routers communicate wirelessly with distributed wireless stations. The other ONUs (without attached wireless gateway routers) support only conventional high-speed wired Internet access to homes and businesses.

### 3.2.2 Dynamic Bandwidth Allocation (DBA) Mechanisms

The ONU upstream transmissions on the shared upstream wavelength channel are coordinated by the standard (IEEE 802.3ah) polling-based Multi-Point Control Protocol (MPCP). In the MPCP protocol, REPORT messages that are included in the ONU upstream transmissions inform the OLT about the ONU queue occupancies. Based on the REPORT messages, the OLT dynamically allocates bandwidth in the form of upstream transmission windows (grants) to the ONUs. The OLT informs each ONU through a GRANT message about its allocated upstream transmission window; we denote $G_o$ for the duration [in seconds] allocated to ONU $o$ in a given polling cycle. We denote $t_G$ for the transmission time [in seconds] of a GRANT message on the downstream wavelength channel. Successive upstream transmissions from different ONUs are separated by a guard time $t_g$ [in seconds] on the upstream wavelength channel.

As summarized in Table 3.1, the design space of DBA mechanisms for the PON part [16] of the FiWi network consists of the dimensions:

- **Grant Scheduling Framework**: Decides when and for which ONUs the grants are sized and scheduled by the OLT
- **Grant Sizing**: Determines the amount of bandwidth (duration of upstream transmission window) allocated to an ONU
- **Grant Scheduling Policy**: Determines the ordering (sequence) of the ONU upstream transmissions on the upstream wavelength channel.
Table 3.1: Summary of Dimensions of Design Space for Dynamic Bandwidth Allocation (DBA).

<table>
<thead>
<tr>
<th>Grant Scheduling Framework</th>
<th>Grant Sizing Policy</th>
<th>Grant Scheduling Policy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Offline</td>
<td>Gated</td>
<td>SPD</td>
</tr>
<tr>
<td>Online</td>
<td>Limited</td>
<td></td>
</tr>
<tr>
<td>Double-Phase Polling (DPP)</td>
<td>Excess, Share</td>
<td></td>
</tr>
<tr>
<td>with Grouping Strategy</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The offline scheduling framework awaits REPORTs from all $O$ ONUs before sizing and scheduling grants to all $O$ ONUs [176]. In contrast, the online scheduling framework sizes and schedules a grant for an ONU $o$, $o = 1, \ldots, O$, immediately after receiving a REPORT from ONU $o$. The double-phase polling (DPP) scheduling framework partitions the set of $O$ ONUs into two groups; each group follows then the offline scheduling framework.

Different Grant sizing policies include:

- **Gated** – Granting the amount requested by the ONU. If $R(n, j)$ is the requested bandwidth in cycle $n$ by ONU $j$, then the Grant $G(n, j)$ for cycle $n$ will be:

  \[
  G(n, j) = R(n, j)
  \]  

  (3.1)

  The disadvantage of the Gated technique is that a single heavily loaded ONU may monopolize the upstream link and this would be unfair to other lightly loaded ONUs.

- **Limited** – This scheme overcomes the unfair distribution of grant by limiting the maximum grant size that could be given to an ONU. If the ONU requests more than the maximum grant size, then the OLT limits it to the maximum
grant size. If the maximum grant size is $G_{\text{max}}$:

$$G(n, j) = \begin{cases} 
R(n, j) & \text{for } R(n) \leq G_{\text{max}} \\
G_{\text{max}} & \text{for } R(n, j) > G_{\text{max}}.
\end{cases} \quad (3.2)$$

The disadvantages of Limited Grant Sizing are that the heavily loaded ONUs are restricted to send less than they request and that the entire cycle time may not be utilized.

- **Excess Bandwidth Grant Sizing** – This scheme overcomes the disadvantage of Limited by obtaining the excess bandwidth from the underloaded ONUs and using it for the overloaded ONUs. Therefore, if $R(n, j)$ is less than $G_{\text{max}}$, the ONU is given whatever it requests and the excess is saved into a pool to be used by heavily loaded ONUs. We use equitable excess bandwidth division for our setup, by which we divide the entire pool equally among all the overloaded ONUs.

$$E(n, j) = \sum_j G_{\text{max}} - R(n, j) \text{ if } R(n, j) < G_{\text{max}} \quad (3.3)$$

$$G(n, j) = \begin{cases} 
R(n, j) & \text{for } R(n, j) \leq G_{\text{max}} \\
\min\{R(n, j), G_{\text{max}} + \frac{E(n, j)}{G}\} & \text{for } R(n, j) > G_{\text{max}}.
\end{cases} \quad (3.4)$$

- **Excess Share Bandwidth Grant Sizing** – This scheme is exclusively used only with grouping or DPP grant scheduling framework as the sharing option is available when there are two or more groups. This scheme utilizes the advantage of excess bandwidth grant sizing by obtaining the excess bandwidth of the lightly loaded ONUs in the same group and the sharing option shares the excess bandwidth from the lightly loaded ONUs the other group as well. Along with the excess credits, we will add another excess share component as shown in the
equations below. For each group excess credits $E(n, j, k)$ is calculated, where $k$ is the group number 1 or 2.

$$E(n, j, k) = \sum_j G_{\text{max}} - R(n, j, k) \text{ if } R(n, j, k) < G_{\text{max}} \quad (3.5)$$

Based on DPP, Excess:Share [16], the total excess credit for groups 1 and 2 in each group can be given as the accumulated excess credits from its own group $E(n, j, 1)$ and the excess credits from the other group in the previous cycle $Sh(n-1, j, 2)$ correspondingly. We do not continuously accumulate the excess credit to avoid upper bounding.

$$E^*(n, j, 1) = E(n, j, 1) + Sh(n-1, j, 2) \quad (3.6)$$

$$E^*(n, j, 2) = E(n, j, 2) + Sh(n, j, 1) \quad (3.7)$$

The shared component $Sh(n, j, k)$ can be computed as the minimum of the unused excess credit from the other group or the present group itself.

$$Sh(n, j, k) = \min \{ E^*(n, j, k) - \sum_j G(j, n, k) - G_{\text{max}}, E(n, j, k) \} \quad (3.8)$$

After the computation of the Excess credits and Excess:Share credits, the grant size for this technique will be very similar to Eq. (3.4) but using the new component $E^*(n, j, k)$ which emphasizes the use of additional bandwidth from the previous group of the previous cycle. This is expressed by include the group number “$k$” in the Eq. (3.9).

$$G(n, j, k) = \begin{cases} R(n, j, k) \quad \text{for } R(n, j, k) \leq G_{\text{max}} \\ \min\{R(n, j, k), G_{\text{max}} + \frac{E^*(n, j, k)}{O} \} \quad \text{for } R(n, j, k) > G_{\text{max}}. \end{cases} \quad (3.9)$$
We consider throughout the shortest propagation delay (SPD) [182] grant
scheduling policy. For brevity, we use the terminology (on., gat.) for Online, Gated;
(on., lim.) for Online, Limited; (on., exc.) for Online, Excess; (off., lim.) for Offline,
Limited; (off., exc.) for Offline, Excess; (dpp., lim.) for DPP, Limited; and (dpp.,
exc.) for DPP, Excess with sharing.

3.3 ONU Grouping Strategies

3.3.1 Motivation

A key principle of efficient polling-based medium access in PONs with long propa-
gation delays is the masking of idle times arising from control message propagation.
In particular, the delay between the ONU transmission of a REPORT message and
the arrival of the corresponding grant message leads to idle times on the upstream
channel, unless transmissions by other ONUs mask the idle time. Therefore, the prin-
cipal strategy of double-phase polling (DPP) is that the upstream transmissions of
one ONU group mask the idle times between transmissions of the other ONU group.
We consider four elementary grouping strategies based on ONU traffic load and prop-
agation distance well as a grouping strategy based on the polling cycle durations
of the ONU groups.

3.3.2 ONU Traffic Load Estimation

The propagation delays $\tau_o$, $o = 1, \ldots, O$, are constants available from the registration
of the ONUs with the OLT. (Inaccuracies in the propagation delays can be compen-
sated with the approaches in [183].) The ONU traffic loads are typically variable
quantities. By combining historic traffic patterns with traffic measurements and esti-
mations following the strategies in [184–187], the ONU traffic loads can be periodically
updated with strategies similar to [106, 188]. We denote $\hat{R}_o$, $o = 1, \ldots, O$, for the
ONU traffic load long-run estimates expressed in terms of the requested bandwidth.
per ONU REPORT. These long-run ONU traffic loads vary typically slowly, e.g., with a diurnal pattern.

In DPP, the polling cycles of the two ONU groups are interleaved. Thus, in order to update the ONU grouping, e.g., according to new ONU load estimates, the operation of the PON needs to be briefly interrupted. After the new ONU groups have been formed, the interleaved DPP polling cycles are launched anew. The long-run ONU traffic load estimates vary typically on the time scale of hours while the interruption due to re-grouping is on the time scale of a polling cycle (usually a few milliseconds in duration). Thus, the service disruption due the re-grouping should typically be minimal.

We denote $\hat{G}_o, o = 1, \ldots, O$, for the corresponding estimates of the durations of the ONU upstream transmission windows that are obtained according to the employed grant sizing policy from the traffic load estimates $\hat{R}_o, o = 1, \ldots, O$.

### 3.3.3 Distance Grouping (DG)

Distance Grouping (DG) orders the $O$ ONUs in increasing one-way propagation distance from the OLT. In particular, with $(o), o = 1, \ldots, O$, denoting the ordered position, e.g., $(1)$ denotes the first ONU in the ordering, the ordered ONUs satisfy $\tau(1) \leq \tau(2) \leq \cdots \leq \tau(O)$. We initially assume that the number of ONUs $O$ is an even number; which is common since the splitting ratio of optical splitters is typically a power of two. DG assigns the first half of the ONUs, i.e., the ONUs $(1), (2), \ldots, (O/2)$ with relatively short propagation delays, to group 1, while the second half, i.e., the ONUs $(O/2 + 1), \ldots, (O)$ with the relatively long propagation delays, are assigned to group 2. If the number of ONUs $O$ is an odd number, then one (arbitrarily selected) group is assigned one less ONU than the other group; the impact of this uneven assignment is minimal for typical ONU numbers.
3.3.4 Distance Balancing (DB)

Distance Balancing (DB) is also based on the ordering of the ONUs in increasing propagation delays. DB assigns the ONUs to the two groups in a round robin fashion, i.e., ONU (1) to group 1, ONU (2) to group 2, ONU (3) to group 1, and so on.

3.3.5 Load Grouping (LG)

Load Grouping (LG) is analogous to DG, but is based on the estimated ONU loads $\hat{R}_o$. The ONUs are sorted in increasing load, i.e., the ordered ONUs satisfy $\hat{R}_{(1)} \leq \hat{R}_{(2)} \leq \cdots \leq \hat{R}_{(O)}$. Then, LG assigns the relatively lightly loaded half, ONUs (1), (2), ..., (O/2), to group 1, and the relatively heavier loaded half, ONUs $(O/2 + 1), \ldots, (O)$, to group 2.

3.3.6 Load Balancing (LB)

Load Balancing (LB) is analogous to DB and assigns ONUs, which are ordered in terms of increasing load, in round-robin fashion to the two groups.

3.3.7 Grouping by Cycle Length (GCL)

GCL strives to balance the durations of the offline polling cycles of the two ONU groups. As a basis for GCL, we first analyze the duration of the offline polling cycle of a given ONU group. Suppose that there are $\gamma$, $\gamma > 1$, ONUs in the group that have been sorted in increasing order of the one-way propagations delay, i.e., $\tau_{(1)} \leq \tau_{(2)} \cdots \leq \tau_{(\gamma)}$.

We initially consider only the first ONU, i.e., ONU (1). The cycle duration is commonly defined as the time period from the instant when the OLT begins to transmit the GATE message to the instant when then end of the corresponding upstream transmission arrives at the OLT. The components of the cycle duration $\Gamma_1$ for one ONU are: the transmission time $t_G$ of the GATE message, the round-trip propagation
delay $2\tau(1)$ for the GATE message to propagate from the OLT to the ONU and for the upstream transmission to propagate from ONU to OLT, as well as the upstream transmission time (transmission window duration) $G(1)$ (which includes the REPORT message). Thus,

$$\Gamma_1 = t_G + 2\tau(1) + G(1).$$

(3.10)

We proceed to consider the first two ONUs, i.e., ONUs (1) and (2). There are two cases: First, the propagation delay $\tau(2)$ is sufficiently short and the transmission window $G(1)$ sufficiently long such that the upstream transmission of ONU (2) can immediately (with a guard distance) follow after the upstream transmission of ONU (1); resulting in the overall cycle duration $t_G + 2\tau(1) + t_g + G(2)$. Second, the propagation delay $\tau(2)$ is sufficiently long and the transmission window $G(1)$ sufficiently short such that the upstream transmission of ONU (1) is finished before the upstream transmission of ONU (2) can start; thus, ONU (2) governs the overall cycle duration with $2t_G + 2\tau(2) + G(2)$ (the second $t_G$ accounts for the GATE message transmission to ONU (1), which precedes the ONU (2) GATE message according to the SPD scheduling). Thus, the overall cycle duration for first two ONUs is analogously to [182, Eqn. (3)]:

$$\Gamma_{1,2} = \max\{t_G + 2\tau(1) + t_g + G(2), 2t_G + 2\tau(2) + G(2)\}$$

(3.11)

$$= \max\{\Gamma_1 + t_g + G(2), 2t_G + 2\tau(2) + G(2)\}$$

(3.12)

$$= \max\{\Gamma_1 + t_g, 2t_G + 2\tau(2)\} + G(2).$$

(3.13)

Analogously to the reasoning leading to Eq. (3.13), we complete the induction step to the first three ONUs:

$$\Gamma_{1,2,3} = \max\{\Gamma_{1,2} + t_g + G(3), 3t_G + 2\tau(3) + G(3)\}.$$  

(3.14)
We obtain in general for \( \gamma \) ONUs:

\[
\Gamma_{1,2,...,\gamma} = \max\{\Gamma_{1,2,...,\gamma-1} + t_g, \; \gamma t_G + 2\tau(\gamma)\} + G(\gamma).
\] (3.15)

Based on the cycle duration in Eq. (3.15), GCL assigns the ONUs as follows to the two groups: ONU (1) is assigned to group 1 and ONU (2) is assigned to group 2. We compare the resulting cycle durations evaluated with Eq. (3.10). Then, we add ONU (3) to the group with the shorter cycle duration and update the cycle duration with Eq. (3.15). Then, we repeat comparing the cycle durations and adding the next ONU to the group with the shorter cycle duration until all \( O \) ONUs have been placed in a group.

3.4 Setup of Simulation Evaluation

We conducted our simulation evaluations with the OMNet++4.2.2 [128] simulator framework. Within OMNet++, we employed the INETMANET-2.2 modules [129] and integrated a self-built optical network simulator with the INETMANET modules.

3.4.1 FiWi Network Architecture
3.4.1.1 Overall FiWi Architecture

We consider two FiWi architectures: a dedicated ONU architecture illustrated in Fig 3.2 and a mixed ONU architecture illustrated in Fig 3.3.

3.4.1.2 Optical Network

The one-way distances from the OLT to the ONUs are uniformly randomly distributed: for the normal-reach PON in the range from 5 km to 20 km and for the long-reach (LR) PON in the ranges from 90 km to 100 km or 80 km to 120 km. Each ONU, including the ONUs in the mixed architecture serving FiWi and PON traffic, has one queue serving the traffic in first-come-first-served order.
Figure 3.2: Dedicated ONU FiWi Network Architecture Set-up: Four ONUs (ONUs 1–4) Are Dedicated to Wireless (FiWi) Traffic While Four ONUs (ONUs 5–8) Are Dedicated to Wired (PON) Traffic. The Figure Also Illustrates ChuLoR Routing: The Wireless Network Is Organized Into Zones, Each Operating at a Different Frequency. Wireless Stations Route Traffic Through a Cluster Head Towards the Gateway for Transmission Over the PON.

Figure 3.3: Mixed ONU FiWi Network Architecture Set-up: Four ONUs (ONUs 1–4) Serve a Mix of Wireless (FiWi) and Wired (PON) Traffic, While Four ONUs (ONUs 5–8) Serve Only Wired (PON) Traffic.

3.4.1.3 Wireless Network

The wireless network supported by the FiWi network has a total of 64 wireless stations (regular wireless source and destination nodes) and four gateway routers (one for
each ONU supporting the FiWi network). The 64 wireless stations are uniformly distributed in an area of 1000 m \times 1600 m. More specifically, the wireless stations are arranged into 16 zones, each containing four wireless stations, whereby any two wireless stations are 100 m apart. Each zone operates on a different radio frequency channel compared to its neighboring zones. We employ the eleven radio channels of the IEEE802.11g standard and reuse some of the radio channels in distant zones in order to minimize interference. Each gateway router operates on four radio channels to serve four zones. A 1 Gbps cable connects a given gateway router to the corresponding ONU. In addition, there are 22 relay routers, each operating on two radio channels to serve two adjacent zones. The two wireless stations closest to the gateway router in a zone are designated to serve as cluster heads in the wireless routing protocol, see Section 3.4.3.2.

We employ a path loss wireless channel model with an alpha value of 2 and a signal-to-noise ratio of 4 dB. Received packets that are below 4 dB are considered noise. The radio sensitivity is set to −85 dBm and the transmission power of the wireless stations is 20 mW, which permits the most distant wireless station in a zone to reach the gateway router. The transmission range is around 250 m. The physical transmission rate for all wireless stations is 54 Mbps. Each wireless station has a buffer size of 1000 packets for each radio channel interface. The queues in the wireless stations follow the drop tail queueing policy.

3.4.2 Network Traffic Scenarios
3.4.2.1 Packet Level Traffic Characteristics

We consider UDP packet traffic with packet sizes based on quad mode distribution: 60 % 64 byte packets, 25 % 1518 byte packets, 11 % 580 byte packets, and 4 % 300 byte packets. These packet sizes include the payload as well as 8 bytes of UDP header, 20 bytes of IP header, and 18 bytes of MAC (Ethernet) header. The maximum
transmission unit (MTU) for the wireless domain is set to 1500 bytes to avoid packet fragmentation.

Packets are generated following independent Poisson processes. All wireless stations have the same wireless packet traffic generation rate, while all wired PON traffic generators have the same PON traffic generation rate.

3.4.2.2 Flow Level Traffic Characteristics (Source-Destination Traffic Matrix)

PON traffic is generated by the PON traffic generators attached (wired) to the ONUs and is always destined upstream to the server (sink) node, which is directly attached (wired) to the OLT.

We consider three traffic matrices (scenarios) for FiWi traffic:

**All-Server Scenario** The FiWi traffic generated at all wireless stations, including cluster heads, is destined to the server attached to the OLT.

**CH-Server; STN-P2P Scenario** The FiWi traffic generated at the wireless nodes that are cluster heads is destined to the server. The FiWi traffic generated at the other (non cluster head) wireless stations is peer-to-peer (P2P) traffic that is uniformly randomly destined to any other wireless node (including the cluster heads); whereby for each generated FiWi packet, a new random destination is drawn.

**All-P2P Scenario** The FiWi traffic generated at all wireless stations (including the cluster heads) is P2P traffic destined to any other uniformly randomly drawn wireless station.

3.4.2.3 Traffic Ratios

In the dedicated ONU architecture, we set the ratio of FiWi traffic:PON traffic to 1:30, that is the aggregate wired (PON) packet traffic generation rate is 30 times
higher than the aggregate FiWi (wireless) packet traffic generation rate in the overall
FiWi network.

In the mixed ONU architecture, we prescribe FiWi traffic:PON Traffic at mixed
ONUs:PON Traffic at PON-only ONUs traffic ratios of 1:10:40 or 1:20:30. With
the 1:10:40 ratio, the packet generation rate of the PON traffic generator at a given
mixed ONU (that serves FiWi and PON traffic) is ten times higher than the aggregate
packet generation rate of the 16 wireless stations associated with the ONU. Moreover,
a traffic generator at an ONU that serves only PON traffic has a four times higher
packet generation rate than the PON traffic generator at a mixed ONU.

3.4.3 Network Protocols
3.4.3.1 Optical Network

In the optical network, we examine the dynamic bandwidth allocation (DBA) mech-
anisms outlined in Section 3.2.2.

3.4.3.2 Wireless Network

We consider clustered localized routing (CluLoR) [54] for the wireless (FiWi) traffic
with two cluster heads in each zone. CluLoR routes traffic from wireless stations
in a zone through the two cluster heads in the zone to the gateway router. FiWi
traffic destined to the server is then forwarded to the corresponding ONU for PON
upstream transmission. FiWi traffic destined to a non-adjacent zone is forwarded to
the ONU, and then transmitted downstream on the PON to the ONU associated with
the zone of the destination wireless station (and then onwards via a cluster head to
the destination). FiWi traffic destined to an adjacent wireless zone is forwarded by
the relay router between the two adjacent zones. The wireless network follows the
IEEE802.11g MAC protocol with a retransmit limit of seven.
3.4.4 Delay Metrics

Throughout, we evaluate the mean end-to-end packet delays, i.e., the mean time periods from the instant of packet generation to complete delivery to the destination. Specifically, we evaluate the mean PON packet delay, i.e., the mean end-to-end delay for PON packet traffic from the instant of packet generation at a PON packet generator to the instant of complete packet delivery to the server (sink). We also evaluate the mean FiWi packet traffic delay, i.e., the mean end-to-end delay for FiWi traffic from the instant of packet generation at a wireless station to the complete packet delivery to the server or destination wireless station. The delay samples are collected with a batch means method from the simulation runs until the 95% confidence intervals of all delay metrics are less than 5% of the corresponding sample means. These confidence intervals are too small to be visible in the plots.

3.5 Results for Mixing of Wireless (FiWi) and PON Traffic

In this section we investigate the effects of mixing traffic from the wireless stations, i.e., FiWi traffic, with conventional PON traffic. In particular, we initially focus on the dedicated ONU network architecture, see Fig. 3.2, in order to bring out the fundamental effects due to mixing FiWi and PON-only traffic from distinct ONUs. We investigate the impact of the DBA mechanisms and the PON propagation distance on the delays experienced by wireless (FiWi) traffic and PON-only traffic.

3.5.1 Impact of DBA Mechanism

We focus initially on the All-Server traffic scenario, see Section 3.4.2.2, to observe the effects of all the FiWi traffic and the PON-only traffic competing on the upstream wavelength channel (A preliminary version of Section 3.5.1 appeared in [189]). We consider the CH-Server; STN-P2P traffic scenario, which includes a P2P traffic com-
Figure 3.4: Mean Wireless (FiWi) and PON Packet Delays for Different DBA Mechanisms and Cycle Lengths. Fixed Parameters: 15–20 km PON With Dedicated ONU Architecture, All-Server Traffic.
3.5.1.1 Online, Gated

We observe from Fig. 3.4 that the (on., gat.) DBA achieves the lowest mean delays for the PON packet traffic across the entire range of traffic loads. On the other hand, (on., gat.) DBA gives the highest delays for wireless (FiWi) packet traffic at high traffic load levels, e.g., for traffic loads above 0.5 Gbps in Fig. 3.4(a). The gated grant sizing allocates to each ONU an upstream transmission window corresponding to its full request. This is beneficial for the conventional high-rate PON-only packet traffic which dominates the network for the considered FiWi:PON traffic ratio of 1:30. Even as the traffic load grows very high, the (on., gat.) DBA allocates the PON-only ONUs their full requests, leading to very long cycles. The long cycles result in relatively long mean wait times for the lower-rate FiWi traffic that arrives over the wireless network and gateway router to the FiWi ONUs; i.e., the FiWi ONUs have to wait relatively long for their turn on the upstream wavelength channel and then only occupy it for a relatively short time.

The (on., gat.) DBA mechanism does not depend on a prescribed cycle length and hence has been repeated in Fig. 3.4(a), (b), and (c) as reference for the other DBA mechanisms.

3.5.1.2 Offline, Limited

We observe from Fig. 3.4 that the (off., lim) DBA mechanism works in favor of FiWi traffic, compared to the (on., gat.) DBA. Limiting the cycle length ensures that the relatively lightly loaded FiWi ONUs can transmit more frequently and do not have to wait until the PON traffic ONUs transmit their entire queues. However, limiting the allocations to the heavily loaded PON traffic ONUs results in growing queues, and eventually buffer overflow, for increasing traffic load. The load point where buffer
overflows for PON traffic occur indicate the stability limit of the network. Average packet delays grow very high near and beyond the stability limit.

We observe a “knee point” of the FiWi traffic delay near the load level corresponding to the stability limit for the PON traffic. Once the PON traffic queues fill up and the PON traffic completely utilizes its limited share of the cycle length, no further increases in the carried upstream load are possible for PON traffic. Instead, further increases in the carried upstream traffic load are due to FiWi traffic only, which follows the fixed 1:30 FiWi:PON traffic ratio. That is, only 1/30th of a given increase in the total traffic load contributes to the actual increase of the carried upstream traffic load. This “switch” from all generated traffic contributing to the carried upstream traffic load to only 1/30th of the generated traffic load contributing to the carried upstream traffic load results in the substantially lower slope of delay increases with increasing generated traffic load, i.e., the observed “knee point”.

3.5.1.3 Online, Limited

The (on., lim.) DBA follows the same performance trend as of the (off., lim) DBA, while performing slightly better than the (off., lim.) DBA. The offline scheduling framework [16, 176] waits for all the REPORTs from all ONUs before commencing the sizing and scheduling of the grants. The delay difference is small because only eight ONUs are considered. As the number of ONUs increases, the delay difference would also increase.

3.5.1.4 Offline, Excess

For the (off. exc.) DBA, we observe from Fig. 3.4 that for low loads, the PON traffic delay is the same as for the (off., lim.) DBA. This is because, all ONUs have typically requests below the limit at low loads and do not require the excess feature. However, as the traffic load increases, the delay for the (off., exc.) DBA is lower than for the
(off., lim.) DBA. Also, the (off., exc.) DBA reaches higher stability limits than the (on., lim.) DBA due to the re-allocation of unused portions of the grant limit to ONUs with presently large requests. Comparing Figs. 3.4(a) and (c), we observe that the relative increase of the stability limit is especially pronounced for short cycle length limits. Specifically, the increase is approximately 40% for $Z = 1$ ms in Figs. 3.4(a) compared to about 15% for $Z = 4$ ms in Figs. 3.4(c). Longer cycle length limits are less restrictive, thus re-allocations of excess “slack” are relatively less effective for long cycle length limits.

For FiWi traffic, we observe that the (off., exc.) delay is higher than the (off., lim.) delay. This is due to the increase in the average cycle length as the re-allocation of unused portion of the grant limits leads to longer mean cycle lengths. The re-allocation benefits mainly the heavily loaded PON ONUs. On the other hand, the FiWi ONUs need to wait on average longer for the next grant, while typically not enlarging their grants with the re-allocation.

3.5.1.5 Impact of Cycle Length

We observe from Fig. 3.4 that increasing the cycle length limit $Z$ benefits PON traffic through increasing stability limits and lowered mean packet delays. Longer cycle length limits allow the heavily loaded PON ONUs to transmit more traffic in each cycle, thus fewer cycles and cycle overhead (e.g., guard times) are incurred to transmit a given PON traffic amount.

On the other hand, FiWi traffic benefits from short cycle length limits. Shorter cycles allow the FiWi ONUs to transmit their relatively small upstream grants more frequently, incurring shorter FiWi packet traffic delays. We also observe that the “knee points” in the FiWi delay curves move to lower traffic load levels for decreasing cycle lengths. This is because the PON traffic queues fill up at lower traffic loads,
preventing the PON ONUs from transmitting more traffic (beyond their respective stability limits) already at relatively low traffic loads.

3.5.1.6 DBA Impact for CH-Server; STN-P2P Traffic

We observe from Fig 3.5 for the CH-Server; STN-P2P traffic scenario similar trends as for the All-Server traffic scenario with $Z = 2$ ms in Fig. 3.4(b). However, we observe that the FiWi packet delays tend to be higher in Fig 3.5 compared to Fig. 3.4(b). The CH-Server; STN-P2P traffic scenario has wireless packet traffic entering the zones to reach the P2P traffic destinations. Thus, more interference and collisions are introduced in the zones, causing the wireless nodes to resend the traffic more frequently than in the All-Server traffic scenario. The retransmissions cause wireless (FiWi) packets to queue up longer and experience longer delays as they traverse the wireless network.

3.5.2 Impact of Long-Reach Propagation

In this section we focus on the impact of the long propagation delays of long reach PONs with 90 km to 100 km between the OLT and the ONUs. For ease of com-
parison with the normal-range PON, we consider initially the same dedicated ONU architecture as in Section 3.5.1. In order to comprehensively examine the impact of the long-range propagation, we consider then in Section 3.5.2.2 the mixed ONU FiWi network architecture.

3.5.2.1 Dedicated ONU FiWi Network Architecture

We consider initially the same scenario as in Section 3.5.1, however, we doubled the considered cycle length limits from 1, 2, and 4 ms in Fig. 3.4 to 2, 4, and 8 ms in Figs. 3.6 and 3.7. Long cycle length limits are important for good utilization of the long-range PON so as to ensure that the long propagation delays and resulting idle times are kept small relative to the durations of the upstream transmission windows (grants). In particular, we observe from Fig. 3.7(a) that with a $Z = 2$ ms cycle length, the (off., lim.) DBA reaches the stability limits around 0.173 Gbps, compared to approximately 0.46 Gbps for the normal-range PON in Fig. 3.4(b).

Generally, we observe from Figs. 3.6 and 3.7 that the cycle length limit $Z$ affects the stability limits and the knee points of the different DBA mechanisms in similar manner as for the normal reach PON in Fig. 3.4. Specifically, the stability limits for the PON traffic and the knee points of the FiWi traffic curves are pushed to higher loads as the cycle length increases. However, for the PON traffic, the trend for the performance impact of the DBA mechanism for the long-range PON is different from the normal-range PON: For the normal-range PON in Fig. 3.4 the stability limit increase achieved by the sophisticated (off., exc.) DBA compared to the simple (off., lim.) DBA is approximately 63 % for $Z = 1$ ms and 18 % for $Z = 4$ ms; in contrast the corresponding increases for the long-range PON are 57 % for $Z = 2$ ms and 33 % for $Z = 8$ ms. Thus, in comparison to the normal-range PON, the impact of the DBA mechanism remains relatively stronger for increasing cycle length in the long-range

54
Figure 3.6: Mean FiWi Packet Traffic Delay for Different DBAs and Cycle Lengths for Long-Reach 90-100 Km PON. Fixed Parameters: Dedicated ONU Architecture, All-Server Traffic.
a) Cycle length $Z = 2$ ms

b) Cycle length $Z = 4$ ms

c) Cycle length $Z = 8$ ms

Figure 3.7: Mean PON Packet Traffic Delay for Different DBAs and Cycle Lengths for Long-Reach 90-100 Km PON. Fixed Parameters: Dedicated ONU Architecture, All-Server Traffic.

For FiWi traffic, we observe from Fig. 3.6 that lower cycle length limits benefit the FiWi traffic in the long-range FiWi network. In particular, reducing the cycle
length from 8 to 2 ms reduces the mean FiWi (on., lim.) packet delay by approximately 29\% in Fig. 3.6, which is equivalent to the corresponding delay reduction in the normal-range FiWi network.

3.5.2.2 Mixed ONU FiWi Network Architecture

In this section we expand the investigation of the impact of the long-range propagation delay in FiWi networks by considering the mixed ONU network architecture, see Section 3.4.1.1. In the considered mixed ONU architecture, four ONUs serve wireless (FiWi) and PON traffic, while the other four ONUs serve only PON traffic. In the mixed ONU architecture, the ONUs serving FiWi and PON traffic have already a large upstream traffic component from the PON traffic; thus, we consider the All-P2P traffic scenario for the wireless FiWi traffic, specifically with the 1:20:30 traffic ratio, see Section 3.4.2.

**Online and Offline DBA Mechanisms** We observe from Figs. 3.9 and 3.8 for the mixed ONU architecture generally similar behaviors for the PON traffic and FiWi traffic as for the dedicated ONU architecture in Section 3.5.2.1. However, we observe that the stability limits and delays are generally slightly higher for the mixed architecture compared to the dedicated architecture. The PON traffic at the mixed ONUs leads to longer upstream transmission grants of the mixed ONUs serving both FiWi and PON traffic. These longer grants increase the utilization of the upstream wavelength channel relative to the idle times (and overheads), resulting in increased stability limits.

On the other hand, FiWi traffic suffers substantially higher delays in the mixed architecture, see Fig. 3.8, compared to the dedicated architecture, see Fig. 3.6. The wide load range with very slowly increasing FiWi packet delays from the “knee point” onwards towards high loads in Fig. 3.6 is replaced by a narrow load range between the
knee point and the load point indicating the stability limit for the FiWi traffic (where the FiWi packet delays shoot up sharply) in Fig. 3.8. In the mixed ONU architecture, wireless (FiWi) traffic is mixed with high-rate conventional PON traffic. Thus, for increasing traffic load, the queues in the mixed ONUs grow very large, causing high FiWi packet delays. In order to preserve the wide load range of slowly increasing FiWi packet delays for traffic loads above the knee point, QoS mechanisms would be needed to protect the FiWi traffic from the PON-only traffic.
Double-Phase Polling (DPP)  Following the indication of the strong effect of sophisticated DBA mechanisms for long-range PONs in Section 3.5.2.1 and Fig. 3.9, we included the sophisticated DPP DBA mechanism (with DG ONU grouping, see Section 3.3.3) in the evaluation of the long-range mixed ONU architecture. The results in Figs. 3.9 and 3.8 demonstrate the superiority of the DPP DBAs over the online and offline scheduling framework DBAs. With DPP, the upstream transmissions of one ONU group can mask the idle times of the other ONU group. Reduced idle times increase the utilization of the upstream transmission wavelength for payload transmissions.
We observe from Figs. 3.9 and 3.8 that DPP achieves lower PON and FiWi packet delays than the other DBA mechanisms, except for (on., gat.), which however has the drawback that a single ONU can monopolize the upstream bandwidth usage for extensive time periods [176]. We also observe that DPP with sharing of the excess bandwidth (dpp, exc.) gives significant performance improvements over simple limited grant sizing (dpp, lim.). Given the favorable performance of DPP with excess sharing (dpp, exc.), we proceed to examine the ONU grouping strategies for this DBA mechanism in the next section.

3.6 Results for ONU Grouping in DPP with Excess Sharing

In this section we examine the different ONU grouping techniques introduced in Section 3.3 in the context of the double-phase polling (DPP) DBA mechanism with excess bandwidth sharing. As in the preceding Section 3.5.2.2, we continue to consider the mixed ONU FiWi network architecture with long-range propagation and with the All-P2P traffic scenario. The ONUs within a group continue to be scheduled based on SPD. We focus on the $Z = 4$ ms cycle length in this section and initially consider the 1:10:40 traffic ratio, followed by the 1:20:30 traffic ratio.

3.6.1 Traffic Ratio 1:10:40: Pronounced Mixed ONU to PON-only ONU Load Difference

For PON packet traffic, we observe from Fig. 3.10(b) that load balancing (LB) gives lower delays than load grouping (LG). LG assigns the lightly loaded ONUs to one group and the highly loaded ONUs to the other group. The lightly loaded ONUs have typically shorter upstream transmission windows than the highly loaded ONUs. Consequently, the lightly loaded ONU group has shorter mean cycle lengths than the heavily loaded ONU group. With shorter cycles there is a smaller probability that the upstream transmissions of the lightly loaded ONU group mask the idle time
Figure 3.10: Mean FiWi and PON Packet Delays for Different ONU Grouping Strategies. Fixed Parameters: Mixed ONU Architecture With 80-120 Km Long-Range Propagation, (dpp., exc.) DBA, $Z = 4$ ms Cycle Length, All-P2P Traffic With Ratio 1:10:40.

between the upstream transmissions of the highly loaded ONU group in successive cycles. In contrast, with LB grouping, there is a higher chance that the upstream transmissions of each group are sufficiently long to mask the idle times between the successive upstream transmission cycles of the other group. Improved masking of idle times increases the utilization of the upstream wavelength channel and lowers the packet delays.

We further observe from Fig. 3.10(b) that for low to moderate loads up to around 0.4 Gbps, distance balancing (DB) gives lower mean PON packet delays than
Figure 3.11: Mean FiWi and PON Packet Delays for Distance Grouping (DG) and Distance Balancing (DB) ONU Grouping Strategies for (dpp., lim.) DBA Mechanism. Fixed parameters: Mixed ONU Architecture With 80–120 Km Long-Range Propagation, $Z = 4$ ms Cycle Length, All-P2P Traffic With Ratio 1:10:40.

distance grouping (DG). This result is due to the analogous masking effect as for the LB vs. LG result, i.e., the two ONU groups can mask each others’ idle times better if the propagation delays are balanced. However, for loads above 0.4 Gbps, DB gives higher PON packet delays than DG. In order to further investigate the DB and DG behaviors, we plot the FiWi and PON packet delays for DB and DG with the simple (dpp., lim.) DBA mechanism in Fig. 3.11. We observe from Fig. 3.11 the expected reduction of the mean packets delays achieved with DB in comparison to DG. In contrast, with the complex (dpp., exc.) DBA mechanism, the sharing of the excess bandwidth among the groups, which do not have balanced loads in the DB and DG strategies may counter the effects of the distance balancing and lead to unexpected results. This indicates that it is crucial to consider both ONU traffic load and propagation distance, i.e., the two main variables that govern the durations of the upstream transmission windows and the round-trip idle times, and thus the cycle lengths (see Eq. (3.10)).

Grouping by Cycle Length (GCL), as derived in Section 3.3.7, considers both
the ONU traffic loads and the propagation delays to balance the cycle durations of the two DPP groups. We observe from Fig. 3.10(b) that GCL indeed achieves the lowest delays and highest stability limit among the considered grouping strategies.

For FiWi packet traffic, we observe from Fig. 3.10(a) that LG achieves significantly lower delays than GCL (while the other grouping strategies give very similar delays). LG assigns the lightly loaded ONUs, i.e., the four ONUs with FiWi traffic, to one group. Since these four ONUs are all lightly loaded, they typically do not need to share any excess bandwidth amongst themselves; rather they typically provide excess bandwidth to be shared among the four highly loaded ONUs in their next cycle. Since the four FiWi ONUs are grouped together, the (typically short) FiWi ONU upstream transmissions (within a given cycle) follow each other on the upstream wavelength channel. In contrast, with GCL, lightly loaded ONUs with FiWi traffic are typically grouped with heavily loaded PON-only ONUs. Thus, there is typically some excess bandwidth sharing within a given GCL group. Consequently, the (typically short) upstream transmissions of the FiWi ONUs (within a given cycle) are interspersed with (typically long) upstream transmissions of heavily loaded PON-only ONUs. This interspersing of (long) PON-only ONU transmissions among the (short) FiWi ONU transmissions within the upstream transmission of a given group in a given cycle tends to increase the mean FiWi packet delay.

3.6.2 Traffic Ratio 1:20:30: Mild Mixed ONU to PON-only ONU Load Difference

We note that with the 1:10:40 traffic ratio considered in Section 3.6.1, the load differences between the mixed ONUs serving FiWi and PON traffic and the ONUs serving only PON traffic are relatively pronounced. We next reduce this load difference by considering the 1:20:30 traffic ratio in Fig. 3.12.

We observe from Fig. 3.12 that for the 1:20:30 traffic ratio the FiWi and
Figure 3.12: Mean FiWi and PON Packet Delays for Different ONU Grouping Strategies. Fixed Parameters: Mixed ONU Architecture With 80–120 Km Long-Range Propagation, (dpp., exc.) DBA, $Z = 4$ ms Cycle Length, All-P2P Traffic With Ratio 1:20:30.

PON packet delays exhibit generally similar trends as for the 1:10:40 traffic ratio in Fig. 3.11. GCL gives still the lowest PON packet traffic delays while LG still gives the lowest FiWi packet delays. However, we observe that for FiWi packets, the delay difference between LG and GCL is reduced in Fig. 3.12(a) compared to Fig. 3.11(a). For the mild load difference between the lightly loaded mixed (FiWi and PON traffic) ONUs and the heavily loaded PON-only ONUs, the delay effect of interspersing long PON-only ONU upstream transmissions among the short FiWi ONU upstream transmissions within a given group and cycle is reduced.
We also observe from the comparisons of Figs. 3.12(a) and 3.11(a) that FiWi traffic has lower stability limits for the 1:20:30 traffic ratio compared to the 1:10:40 traffic ratio. With the 1:20:30 traffic ratio, the mixed ONUs are relatively higher loaded (21/51 proportion of the total traffic) compared to the 1:10:40 ratio (11/51 of total traffic). Thus, with the 1:20:30 ratio, the mixed ONUs saturate for a lower total traffic load, thus reducing the maximum throughput levels that FiWi traffic can achieve.

Overall, the results for the ONU grouping strategies indicate that GCL has favorable performance characteristics. GCL achieved the lowest PON packet traffic delays among the considered five grouping strategies. LG achieved lower FiWi packet delays than GCL; however, for FiWi networks with only mild load differences between the ONUs serving FiWi traffic and the ONUs serving only PON traffic, the delay reduction with LG is relatively small. Thus, GCL appears overall to be a promising strategy ONU grouping strategy for FiWi networks.
Chapter 4

SMART GATEWAYS (Sm-GW) FOR LTE FEMTO-CELL ACCESS NETWORK: RESOURCE SHARING

4.1 Related Work

Schedulers implemented through wireless networking protocols share the wireless resources at the eNB among all the connected user equipment (UE) nodes. For example, the LTE medium access control (MAC) protocol [190] is responsible for scheduling the wireless resources between a single eNB and multiple UEs in LTE. Likewise to sharing of resources in the wireless protocols such as the air-interface of LTE, we propose the network protocols to share the network resources between the Sm-GW gateway and the eNBs to dynamically allocate the network resources based on the requirements and availability from the cellular operators. Schedulers for wireless resource sharing have been extensively researched. Proportional fair scheduling (PFS) of the bandwidth allocations in wireless networks [191, 192] can support high resource utilization while maintaining good fairness among the network flows. An algorithm to support quality of service (QoS) aware uplink scheduling and resource allocation at the small cell eNB in the LTE network has been examined in [193]. However, most studies to date have been limited to the sharing of the wireless resources.

Similar research, namely a mechanism for network sharing of resources among small cell base stations has been conducted in [194, 195]. Specifically, the H-infinity scheduler for limited capacity backhaul links in [194] schedules the traffic in the downlink from a centralized scheduler focused on buffer size requirements at the base stations in the small cell networks. In contrast, we focus on the uplink traffic from the eNBs to the Sm-GW. To the best of our knowledge, we propose the first network protocol scheduler for the uplink transmissions from the eNBs to the gateways in the
context of LTE small cells.

For completeness, we note that resource allocation in cellular networks has been studied from a number of complementary perspectives. Destounis et al. [196] discusses the power allocation of the interfering neighboring base stations based on the queue length for QoS provisioning. D2D resource allocation by negotiations for offloading of traffic to small cell networks has been studied in [197], which can be easily supported by our proposed Sm-GW. Coordinated scheduling algorithm in the context of small cells with dynamic cell muting to mitigate the interference has been discussed in [198], cell muting technique can be further benefited from our approach of traffic scheduling to eNBs based on the requirements.

We present a new access networking framework for supporting small cell deployments based on the sharing of network resources. We introduce a new network entity, the Smart GateWay (Sm-GW) as shown in Figure 1.1. The Sm-GW flexibly accommodate the eNB connections at the gateway and dynamically assigns uplink transmission resources to the eNBs. The main contribution in this article is a Sm-GW scheduling framework to share the limited backhaul network resources among all the connected small cell eNBs.

4.2 Proposed Sm-GW Scheduler

The main purpose of a scheduler is to maximize the utilization of the network resources, and to ensure the fairness to all the eNBs connected to an Sm-GW for the uplink transmissions. In case of no scheduling, highly loaded eNBs can impair the service of lightly loaded eNBs connected to the same Sm-GW. When many eNBs are flexibly connected to an Sm-GW, traffic bursts from heavily loaded eNBs can overwhelming the queue of an Sm-GW, resulting in excessive packet drops and high delays, even for lightly loaded eNBs. The absence of a scheduling mechanism in such
a situation would affect the fairness of all eNBs. On the other hand, with scheduling mechanism, a large number of eNBs can be flexibly connected (making the network scalable) to the Sm-GW while ensuring prescribed QoS and fairness levels. Each eNB can possibly have a different service level agreement inside the enterprise building (Fig. 1.1). The dynamic resource allocation in the Sm-GW scheduler ensures fairness across all eNBs, while providing flexible connectivity.

4.2.1 Equal Share Scheduling

Equal share scheduling shares the available backhaul (transmission bitrate) capacity equally among all eNBs connected to an Sm-GW. A configuration message is sent to all eNBs in the event of a change in connectivity at the Sm-GW, i.e., addition of new eNB or disconnection of existing eNB. Configuration messages include the maximum allowed bandwidth for the eNB uplink transmissions. We propose to dynamically compute the maximum allowed transmission amount (in Bytes) for eNB \( n \), such that \( n = 1, 2, 3, \ldots, N \), where \( N \) is the number of eNBs, within the transmission window of duration \( W \) [seconds] assigned by the the Sm-GW.

More specifically, the maximum allowed transmissions \( \Gamma \) [Byte] during a cycle (reporting window) of duration \( W \) [seconds] are dynamically evaluated by the Sm-GW based on the allocated backhaul link capacity \( G \) [bit/s] from a given operator and the number of eNBs \( N \) for each eNB \( n \) which connected to the Sm-GW. \( \Gamma \) [Byte] is computed as:

\[
\Gamma = \frac{G}{N} \times W. \tag{4.1}
\]

The traffic amount limit (maximum grant size) \( \Gamma \) and window duration \( W \) are sent to the eNBs as a part of initial configuration message. Each eNB schedules the uplink transmission such that no more than \( \Gamma \) [Byte] of traffic are sent in cycle
Algorithm 1: Equal Share Scheduler

1. Initialization
(a) Sm-GW calculates $\Gamma$ bytes based on number of eNBs $N$, link capacity $G$ and window duration $W$;

2. Configuration Packet
(a) Sm-GW generates configuration packet and send it to each eNB;

3. Configuration Update
if Change in connectivity at the Sm-GW then
    $\Gamma \leftarrow \frac{G}{N} \times W$;
    Send configuration packet to all eNBs;
end

Equal Share Algorithm

Step 1: Based on equation (4.1), if there are $N$ eNBs connected to an Sm-GW, maximum allowed bandwidth for each eNB is calculated such that, $\Gamma_n \in \{\Gamma_1, \Gamma_2, \Gamma_3, \ldots, \Gamma_N\}$. However for an equal share scheduling, the grants are allocated in equal to all the eNBs connected, therefore, $\{\Gamma_1 = \Gamma_2 = \Gamma_3 = \cdots = \Gamma_N\}$. In addition, duration of transmission windows $\{W_1, W_2, W_3, \ldots, W_N\}$ required for the bandwidth evaluation at each eNB are selected by the scheduler based on the operator configuration. For an equal share scheduling the transmission window should be same for the all the eNBs $\{W_1 = W_2 = W_3, \ldots, W_N\}$.

Step 2: Parameters $\Gamma_n$ and $W_n$ are sent to the corresponding eNB $n$ as a part of initial configuration message.

Step 3: eNB $n$ upon the reception of configuration message with the parameter $\Gamma_n$ and $W_n$, network protocol identifies the non signaling traffic and schedules the transmissions.
Step 4: Each eNB \( n \) schedules the transmissions \( \gamma_n \in \{ \gamma \mid \gamma \leq \Gamma_n \} \) such that total transmissions do not exceed the assigned maximum bandwidth from the Sm-GW within the transmission window duration of \( W \) seconds.

The simple equal share scheduler can flexibly accommodate large numbers \( N \) of eNBs. However, the equal bandwidth assignments by the simple equal share scheduler to the eNBs under-utilize the network resources when some eNBs have very little traffic while other eNBs have high traffic rates (above \( \Gamma / W \)). We next present an excess share scheduling technique that utilizes the unused equal share portions of the lightly loaded eNBs for highly loaded eNBs.

4.2.2 Excess Share Scheduling

Equal share scheduler results in under-utilization of network resources when the traffic requirements at the eNBs are highly asymmetric i.e., very high requirements at some eNBs and very low at others. In order to effectively utilize the network resources we propose the mechanism of excess share scheduling at the Sm-GW such that excess bandwidth from the lightly loaded eNBs are allocated to highly loaded eNBs.

Excess share scheduling operates with a cyclical reporting-granting-transmission cycle of duration \( W \) [seconds] illustrated in Fig. 4.1. At the start of the cycle, each eNB \( n, n = 1, 2, \ldots, N \), sends a report (uplink transmission bitrate request) \( \rho_n \) [in units of traffic amount in Byte to be transmitted over a cycle of duration \( W \)] to Sm-GW. Once all reports have been received (i.e., following the principles of the offline scheduling framework [16, 176]) i.e., \( \forall n, \rho_n \), grant \( \gamma_n \) is estimated at the Sm-GW and sent to the eNB \( n \) such that \( \sum \gamma_n \leq \Gamma + \epsilon \), where \( \epsilon \) is the slack variable. Algorithm (2) presents the excess share scheduler.

Lightly traffic eNBs are granted with the resources based on Eq. 4.2 which is equal to the requested bandwidth \( \rho_n \), while Eq. 4.3 accumulates in an excess pool the
Figure 4.1: Illustration of Smart Gateway (Sm-GW) Function: (a) Sm-GW Receives Backhaul Transmission Bitrate Allocation $G$ From Operator. Based on Uplink Transmission Bitrate Requests (Reports) $\rho_n$ From Individual eNBs $n$, 1, ..., $N$, Sm-GW Gives Uplink Transmission Grants $\gamma_n$ to the eNBs. $\delta$ Represents the Granting Delay. (b) The eNB Requests (Reports) $\rho_n$ and Sm-GW Grants $\gamma_n$ Are Followed by the eNB Uplink Data Transmissions (in Parallel) in a Fixed-Period Cycle.

unused (excess) portion of the equal share allocation to the eNB$_n$, $n \in (1, 2, ..., |L|)$, where $|L|$ is the number of lightly loaded eNBs in the Sm-GW.

$$\forall \rho_n \leq \Gamma, \quad \gamma_n = \rho_n. \quad (4.2)$$

$$\xi = \sum_{\forall \rho_n \leq \Gamma} \Gamma - \rho_n. \quad (4.3)$$

Following the principles of controlled equitable excess allocation [199, 200], highly loaded eNB$_n$, $n \in (1, 2, ..., |H|)$ are allocated an equal share of the excess up
Algorithm 2: Excess Share Scheduler

1. Sm-GW: Initialization
   (a) Evaluate $\Gamma$
   (b) $\{\xi, \mathcal{H}\} \leftarrow 0.$

2. eNB: Report $\rho$ Evaluation
   (a) Send $\rho_n$ to Sm-GW based on the requirement

3. Sm-GW: Grant $\gamma$ Evaluation
   if $\text{Count}(\rho) = N$ then
     
     for $(n \leftarrow 0; \ n \leq N; \ n + +)$ do
       if $\rho_n \leq \Gamma$ then
         $\gamma_n \leftarrow \rho_n$;
         Accumulate $\xi$;
       else
         $\mathcal{H} \leftarrow \mathcal{H} \cup \{\rho_n\}$;
       end
     end
     if $\xi = 0$ then
       foreach $\rho_n \in \mathcal{H}$ do
         $\gamma_n \leftarrow \Gamma$;
       end
     end
     else
       
       while $(\xi \neq 0 \land |\mathcal{H}| \neq 0)$ do
         foreach $\rho_n \in \mathcal{H}$ do
           $\gamma_n \leftarrow \min\left(\rho_n, \ \Gamma + \frac{\xi}{|\mathcal{H}|}\right)$;
           Accumulate $\hat{\xi}$;
         end
         Reevaluate $|\mathcal{H}|$;
         $\Gamma \leftarrow \Gamma + \frac{\hat{\xi}}{|\mathcal{H}|}$;
         $\xi \leftarrow \hat{\xi}$;
       end
   end
   foreach $n \in N$ do
     Send $\gamma_n$ to eNB$_n$;
   end

   to their request. That is, with $|\mathcal{H}|$ highly loaded eNBs, the grants are

   $$\forall \rho_n > \Gamma, \quad \gamma_n = \min\left(\rho_n, \ \Gamma + \frac{\xi}{|\mathcal{H}|}\right). \quad (4.4)$$
4.2.3 Fairness of the Schedulers

Within the context of our proposed Sm-GW, fairness is the measure of network accessibility of all NeNBs connected to the Sm-GW based on individual throughput requirement at each eNB. We denote $T$ for the long-run average throughput level [bit/s] of uplink traffic generated at eNB $n$, $n = 1, 2, \ldots, N$, at Sm-GW. The throughput level $T$ can for instance be obtained through low-pass filtering of the requests $\rho$ over successive cycles (windows) $w$. In the case when traffic requirements for all eNBs is less than or equal to the $\Gamma$, i.e., if the incoming traffic onto the Sm-GW does not exceed the outgoing link capacity $G$, even if no scheduling mechanism is employed at the Sm-GW, eNBs always get the fair share of the network resources due to availability. But, a scheduling mechanism has to be employed, if the system is overloaded i.e., when there are large number of eNBs trying to send very large traffic and only few eNBs try to send very low traffic to Sm-GW, and the aggregated total traffic from all the eNBs exceeds the outgoing total link capacity at the Sm-GW.

Under overload conditions in the network of an Sm-GW (i.e., aggregated traffic from all eNBs is larger than the outgoing link bandwidth $G$ at the Sm-GW set by the operator), suppose if there are highly loaded eNBs $h \in \mathcal{H}$ with throughput levels $T_h > \Gamma/W$ and lightly loaded eNBs $l \in \mathcal{L}$ with throughput levels $T_l < \Gamma/W$, the optimal throughput $\Omega_h$ for the highly loaded eNBs is given as:

$$\Omega_h = \begin{cases} G - \sum_{l \in \mathcal{L}} T_l, & \text{if } \sum_{h \in \mathcal{H}} T_h + \sum_{l \in \mathcal{L}} T_l > G \\ T_h, & \text{otherwise.} \end{cases} \quad (4.5)$$

The highly loaded eNBs should be able to transmit traffic up to an equitable share of the transmission bitrate (capacity) not used by the lightly loaded eNBs.
Similarly, the optimal throughput for the lightly loaded eNBs is given by:

$$\Omega_l = T_l. \quad (4.6)$$

We define the normalized distance $\mathcal{E}_n$ of the actually achieved (observed) throughput $\tau_n$ and the optimal throughput $\Omega_n$, i.e.,

$$\mathcal{E}_n = \tau_n - \Omega_n. \quad (4.7)$$

Based on the optimal fair throughput criteria described by the Eqs. 4.5 and 4.6 along with the achieved (observed) throughput at the Sm-GW, the Fairness Index $\mathcal{F}_T$ for the throughput is evaluated based on the Raj Jain Fairness based on normalized distance [201] and is defined as:

$$\mathcal{F}_T = \frac{\sqrt{\sum_{n \in (1, 2, \ldots N)} \mathcal{E}_n^2}}{\sqrt{\sum_{n \in (1, 2, \ldots N)} \Omega_n^2}}. \quad (4.8)$$

A fairness index $\mathcal{F}_T$ close to zero indicates fair Sm-GW scheduling. When the traffic received by the Sm-GW from all eNBs is not exceeding the outgoing link capacity $G$ and all eNBs transmitting at $T < \Gamma/W$, the $\mathcal{F}_T$ is expected to be close to zero regardless if there is a scheduling mechanism or not due to the availability of network resources. However, under highly asymmetrical load conditions, $\mathcal{F}_T$ of the equal share scheduler can be larger (i.e., $\mathcal{F}_T >> 0$) even when the aggregated traffic from all the eNB at the Sm-GW is less than the link capacity $G$. eNBs in the equal share mechanism are granted fixed value of $\Gamma$ resource, resulting in the starvation of the highly loaded eNBs even with the excess availability of resources at the lightly loaded eNBs.
4.2.4 Overhead Evaluation

Excess share scheduler implementation requires each eNB $n$ to send a report $\rho_n$ to Sm-GW every cycle of duration $W$ seconds. Upon reception of the reports from all $N$ eNBs, the Sm-GW evaluates and sends the grants $\gamma_n$ to the respective eNBs. Therefore, the factors affecting the effective throughout due to the scheduling mechanisms are transmission delay of the report $\rho$ and grant $\gamma$, processing delay at the Sm-GW, and the round-trip propagation delay. However in an equal share scheduler, a configuration packet is sent to the eNB. Typically, eNB could be connected to the Sm-GW with a simple plug-in feature. The connection of new eNB triggers the evaluation of new maximum bandwidth $\Gamma$ as discussed in the Section 4.2.1 and recomputed $\Gamma$ value is sent as a reconfiguration message to all the eNBs. Although the reconfiguration messages are sent to the eNBs at times when an eNB changes its status to i.e., either a connect or disconnect, the overhead of sending a reconfiguration message to all the eNBs is negligible compared to the timescale of reports and grants sent every cycle duration $W$ in the excess share scheduler. As noted, the overhead for the excess share

![Overhead Analysis due to Network Protocol Scheduler at Sm-GW and Small Cell eNBs in the Indoor Environment (20m < d < 500m).](image-url)
scheduling mechanism can be significant. Although there exists many scheduling algorithms for optical links which consider negligible processing delay [131], it could reasonably assume the processing delay in the order of micro seconds to account for the LTE protocol stack. Figure 4.2 describes the behavior of the percentage overhead as a function of the propagation distance \( d \) [meters] between the eNB and the Sm-GW, and the cycle (reporting window) duration \( W \) [seconds]. We assume a short frame in the Ethernet link for the report packet size \( \sigma_r \) and grant packet size \( \sigma_g \) of 70 Bytes each, and the regular Ethernet link capacity \( \beta \) of 1 Gbps for the overhead evaluation. The overhead due to the scheduler mechanism in the excess share scheduler can be evaluated as,

\[
\text{Overhead \%} = \frac{1}{W} \left( 2t_{\text{prop}} + t_{\text{proc}} + \left( \frac{\sigma_r + \sigma_g}{\beta} \right) \right).
\] (4.9)

For typical parameter settings, such as up to 500 m eNB-to-Sm-GW propagation distance, and \( W = 1 \) ms cycle duration, the overhead is less than half a percent. It can be observed that as the cycle duration decreases, overhead percentage increases due to more frequent exchanges of reports and grants between the eNB and Sm-GW accounting for the overhead. However, as the propagation distance between Sm-GW and the eNB increases, overhead percentage also increases due to the increased propagation delay in the report and grant cycles. Therefore, based on the trade-offs in the complexity, overhead and the delay, the operator can choose the value of cycle duration \( W \) and maximum distance \( d \) for operation of eNBs within network of Sm-GW.

4.3 Evaluation of Sm-GW Scheduling

4.3.1 Simulation Setup

We evaluate the performance of the Sm-GW scheduling with the discrete event simulator OMNeT++ 4.6 [128]. We consider a given Sm-GW with a backhaul transmission bitrate (capacity) of \( G = 1 \) Gbps. A typical contemporary S-GW of an operator can
support $N = 8$ eNB connections. For our evaluation of the flexible small cell eNB support, we consider $N = 20$ eNBs connected to the Sm-GW. The LTE access network typically requires the packet delay to be less than 50 ms [202]. Therefore, we set the Sm-GW queue size to 20 MBytes, which is equivalent to a maximum queuing delay of 20 ms over the $G = 1$ Gbps link. Without any specific scheduling, the Sm-GW operates in first-come-first-served mode with taildrop.

We simulate the typical bursty eNB traffic generation pattern, with two eNB traffic rate states: low and heavy. The sojourn time in a given traffic rate state is randomly drawn from a uniform distribution over 1 ms to 4 ms. At the end of the sojourn time, a switch to another state occurs with a probability of 70% in the low traffic state and 30% in the heavy traffic state. The traffic bitrate ratio between the heavy and low traffic states is $4:1$. Within a given traffic rate state, data packets are randomly generated according to independent Poisson processes.

We consider $|\mathcal{L}| = 10$ lightly loaded eNBs and $|\mathcal{H}| = 10$ highly loaded eNBs. Each eNB, irrespective of whether it is lightly or highly loaded, generates traffic according to the two traffic rate state (low and heavy) model. The low and heavy traffic rates are set such that the long-run average generated traffic rate corresponds to a prescribed required throughput (load) level $T_L < G/N = 50$ Mbps for a lightly loaded eNB and a prescribed required throughput (load) level $T_H > G/N = 50$ Mbps for a highly loaded eNB. For all simulations, the 95% confidence intervals are less than 5% of the corresponding sample mean.

### 4.3.2 Sm-GW without Scheduler

Figures 4.3 and 4.4 show the performance of the Sm-GW with and without scheduling for the lightly and highly loaded eNBs respectively. It shows the actual (achieved, observed) throughput $\tau$ as a function of the required (generated) throughput level
Figure 4.3: Observed Average Throughput of Lightly Loaded eNBs.

We observe from Fig. 4.3 that without scheduling, the lightly loaded eNB suffer reductions in the achieved throughput, that are especially pronounced (over 30 %) for the high $T_H = 200$ Mbps load of the highly loaded eNBs. At the same time, we observe from Fig. 4.4 that without scheduling, the highly loaded eNBs achieve more than their fair throughput share. For instance, for the highly loaded eNB throughput requirement (load) $T_H = 140$ Mbps, and $T_L = 30$ Mbps, the observed throughout of the highly loaded eNBs is $\tau_H = 76$ Mbps, which is significantly higher than the fair share of $(G - |L|T_L)/|H| = 70$ Mbps. As the average level of lightly loaded traffic is increased, the effect of subdue in accessing the Sm-GW network by the highly loaded eNBs can be observed.

### 4.3.3 Sm-GW with Scheduler

Figures 4.3 and 4.4 also illustrates the behavior of Sm-GW with equal share and excess share schedulers. The highly loaded eNBs are restricted with the resources so as to yield for the lightly loaded eNBs allowing to achieve the throughput requirements. In contrast to the behavior with no-scheduling, we observe in Fig. 4.3 that
Highly loaded eNBs benefit from scheduling mechanisms in that they get the full share of their optimally (fair) throughput. However, we observe from Fig. 4.4 that the behavior of the highly loaded eNBs are yielded based on the scheduler implementation. For equal share scheduler, highly loaded eNBs achieve only a throughput of \( G/(|\mathcal{L}|+|\mathcal{H}|) = 50 \text{ Mbps} \) (lower than its optimal throughput) as equal share Sm-GW scheduling assigns a static allocation of equal shares of the limited backhaul capacity \( G \) to all eNBs irrespective of their traffic generation rates. For instance, observed throughput is 50 Mbps when \( T_H \) is 100 Mbps and \( T_L \) is 30 Mbps, while optimal throughput is 70 Mbps. But the excess share scheduler restricts the highly loaded eNBs such that lightly loaded eNBs achieve their throughput requirements. Therefore we observe the throughput of approximately 70 Mbps for the excess share scheduler.

Performance comparisons of schedulers along with no-scheduling for different mean traffic loads of \( T_H \) and \( T_L \) is also shown in Figs. 4.3 and 4.4.
4.3.4 Delay Performance

Delay experienced by the packets transmitted by the lightly loaded eNBs are shown in Fig. 4.5. In a situation when large number of eNBs are connected to the Sm-GW, it is likely for the overload conditions to occur during which the queues at the Sm-GW would overflow incurring the large delays along with the dropping of packets. We observe a sharp delay increase at $T_L = 20$ Mbps, when the total traffic load $|\mathcal{L}|T_L + |\mathcal{H}|T_H$ approaches the backhaul capacity $G$. Therefore all the additional packets which cannot be queued should be dropped. However, with the scheduler implementation due to the yielding mechanism of highly loaded eNBs, lightly loaded eNBs are allowed to send the traffic with the minimal delay. When lightly loaded eNBs reach 50 Mbps scheduler assumes all eNBs are highly loaded and hence allocates all the available resources along with a slack quantity such that queue is also utilized at the Sm-GW. Therefore, the delay of 8.5ms is experienced by the packets when $T_L$ is 50 Mbps.
4.3.5 Fairness Index Performance

Figure 4.6 emphasizes the importance of scheduler mechanism in the presence of large number of eNBs when connected to the Sm-GW. Fairness Index, $F_T$ described in the section 4.2.3 is used to quantify fairness of the system based on the normalized distance from the optimal value. As noted earlier, closer the value of $F_T$ to zero, the system is fairer. Figure 4.6 clearly illustrates that excess share scheduler has the best fairness index performance throughout the varying traffic load of $T_L$ while equal share scheduler performs the worst for $T_L$ traffic loads lower than 36 Mbps. Although no scheduling appears to be fairer with the values of $F_T$ at the lower values of $T_L$, due to the overload conditions, packets would incur large delay and hence not suitable for the cellular applications which require dedicated QoS. Having $T_H = 200$ Mbps leads to buffer overflow at the Sm-GW as it exceeds the backhaul capacity $G$. In no-scheduling case, we observe an increasing fairness index $F_T$ as the lightly loaded eNBs generate more traffic (i.e. for increasing $T_L$). That is, as the lightly loaded eNBs try to transmit more traffic, their achieved throughput falls more and more.
below their fair share [see growing divergence between the no scheduling curves and straight lines for scheduling in Fig. 4.3], leading to increasing unfair treatment of the lightly loaded eNBs. Equal share and excess share schedulers provide protection to lightly loaded eNBs as its achieved throughput matches its fair optimal throughput. In equal share scheduler, a high fairness index $\mathcal{F}_T$ for low traffic loads of the lightly loaded eNBs, as the highly loaded eNBs receive only unfairly small shares of the backhaul capacity. Equal share scheduler limits the achieved throughput of the highly loaded eNBs to $\Gamma/W$ even there are unused resources by the lightly loaded eNBs. With excess share scheduler, it grants the excess resources from lightly loaded eNBs to highly loaded eNBs resulting in better overall fairness index performance when compared to other scenarios. Moreover, in excess share scheduler, the gap between the achieved throughput and the optimal throughput for the highly loaded eNBs is narrower when compared to equal share scheduler.
Chapter 5

CONCLUSION AND FUTURE DIRECTIONS

We have examined the combined effects of clustered and localized routing (CluLoR) in fiber-wireless (FiWi) networks. CluLoR is a simple routing strategy that does not require route discovery and maintenance between distant regions of the wireless mesh network (WMN) of a FiWi network. Instead, WMN nodes require only local routes to and from their nearby cluster heads, while in turn the cluster heads require only routes to their nearby gateway routers that interface the WMN with the fiber network.

Our evaluations for CluLoR in a FiWi network organized into zones operating on different radio channels revealed that the clustered routing strategy where regular wireless nodes communicate via cluster heads with the gateway router improves the throughput-delay performance compared to unclustered routing where wireless nodes directly communicate with the gateway router. Our evaluation of the localized routing strategy indicated substantial throughput-delay improvements over an unlocalized minimum hop-count routing strategy.

We have also examined fiber-wireless (FiWi) networks with long-range propagation on the order of 100 km in the fiber-based passive optical network (PON) part of the overall FiWi network. We have conducted extensive simulations to investigate the mixing of low-rate wireless (FiWi) traffic that first traverses the wireless traffic and then the PON with high-rate PON traffic that traverses only the PON. For a FiWi network architecture with dedicated ONUs for wireless (FiWi) traffic and dedicated ONUs for PON-only traffic, we found that strategies that lower the PON traffic delay, generally increase the FiWi traffic delay. That is, enhancements to the dynamic bandwidth allocation (DBA) mechanisms and extended cycle lengths reduce
the PON packet traffic delay, but increase the FiWi packet traffic delay in the dedicated ONU FiWi network. In contrast, for a mixed ONU FiWi network architecture where some ONUs serve both low-rate wireless (FiWi) and high-rate PON traffic, the enhanced DBA mechanisms and extended cycle lengths benefit both FiWi and PON packet traffic. We also found that the double-phase polling (DPP) DBA mechanism gives the best performance among a wide range of compared DBA mechanisms.

We closely examined the ONU grouping in DPP and introduced a novel grouping by cycle length (GCL) strategy. The GCL strategy considers both the OLT-to-ONU propagation distances and the ONU load levels and strives to balance the polling cycle durations of the two ONU groups in DPP. We found that DPP achieves the lowest PON packet traffic delays among a range of considered grouping strategies and gives relatively low delays for FiWi traffic. For FiWi networks with a pronounced disparity between the traffic load levels of the ONUs serving FiWi traffic and the ONUs serving only PON traffic, ONU grouping by load levels can achieve somewhat lower FiWi delay than GCL.

We have created a new backhaul architecture by inserting a novel Smart Gateway (Sm-GW) between the wireless base stations (eNBs) and the conventional gateways, e.g., LTE S/P-GW. The Sm-GW enables flexible support for large numbers of small cell base stations. In particular, the Sm-GW adaptively allocates (schedules) uplink (backhaul) transmission grants to the individual eNBs on a fast (typically millisecond) timescale. Simulation results have demonstrated that the scheduling of eNB grants by the Sm-GW can greatly improve the fairness of the backhaul service over conventional static backhaul capacity allocations.

There are many directions for important future research on long-range FiWi networks. One direction is to examine quality of service differentiation for differ-
ent classes of wireless and PON-only traffic in long-range FiWi networks. Another important direction is to examine the interactions with specific wireless networking protocols and standards, such as long-term evolution (LTE) in the wireless part of the long-range FiWi network. Also, the internetworking of the PON part of the long-range FiWi network with metropolitan area networks [203–209] leading to the backbone of the Internet is an interesting direction for future research. Yet another emerging important research direction is the incorporation of energy efficiency mechanisms, such as [210–215], into low-delay long-range FiWi networking. Other interesting research directions for small cells in LTE access network is applying Software Defined Networking (SDN) concept to manage the small cells backhaul [216–228]. SDN can facilitate network optimization techniques that enable higher utilization of the network resources.
REFERENCES


[50] R. R. Tyagi, F. Aurzada, K.-D. Lee, and M. Reisslein, “Connection establish-
ment in LTE-A networks: Justification of poisson process modeling,” IEEE


[54] Y. Dashti and M. Reisslein, “CluLoR: Clustered localized routing for FiWi

novel routing metric for multi-radio wireless mesh networks,” in Proc. IEEE

[56] F. Theoleyre, B. Darties, and A. Duda, “Assignment of roles and channels for
a multichannel MAC in wireless mesh networks,” in Proc. IEEE ICCCN, 2009,
pp. 1–6.

238–247.

[58] Y. Dashti, A. Mercian, and M. Reisslein, “Grouping by cycle length (GCL) for

network with a novel excess bandwidth distribution scheme,” Opt. Fiber Tech-

[60] M. D. Andrade, A. Buttaboni, M. Tornatore, P. Boffi, P. Martelli, and A. Pat-

[61] A. Helmy, H. Fathallah, and H. Mouftah, “Interleaved polling versus multi-
thread polling for bandwidth allocation in long-reach PONs,” IEEE/OSA J.

“Dynamic bandwidth allocation for long-reach PON: overcoming performance


