A Unified End-to-End Communication Paradigm for Heterogeneous Networks

Vinaykumar Muralidharan

B.E. Electronics and Communications Engineering
Anna University, Chennai, India

Submitted to the Department of Electrical Engineering and
Computer Science and the Faculty of the Graduate School
of the University of Kansas in partial fulfillment of
the requirements for the degree of Master of Science

Thesis Committee:

_____

Dr. James P.G. Sterbenz: Chair

_____

Dr. Alexander M. Wyglinski: Co-Chair

_____

Dr. Victor S. Frost

_____

Date Defended

2008/03/24

The Thesis Committee for Vinaykumar Muralidharan certifies
That this is the approved version of the following thesis:

A Unified End-to-End Communication Paradigm for Heterogeneous Networks

Committee

_____

Chairperson

_____

Co-Chairperson

_____

_____

Date Approved

Abstract

The aim of this thesis research is to develop a unified communication paradigm that provides an end-to-end bursting model across heterogeneous realms. This model generates end-to-end bursts, thereby eliminating edge node burst assembly and its effect on TCP performance. Simulation models are developed in *ns-2* to validate this work by comparing it with edge burst assembly on OBS networks. Analysis shows improved end-to-end performance for a variety of burst sizes, time-outs, and other network parameters.

Acknowledgements

Before going through the research work, I would like to take a moment to thank one and all who were involved in bringing out this work.

Firstly, I would like to thank my advisor **Dr. James P.G. Sterbenz** for providing me with the right guidance towards my thesis direction and for his comments helping my work. I am glad to thank my co-advisor **Dr. Alexander M. Wyglinski** for his continuous support and encouragement without whom this work would not be possible.

I would like to thank my committee member **Dr. Victor S. Frost** for his advice and recommendations on my thesis document.

I sincerely thank Abdul Jabbar Mohammad for his advice and guidance with the simulation work of my thesis. I acknowledge Justin P. Rohrer, who spent precious time conducting weekly meetings to meet the requirements. I would also like to thank ITTC support staff for providing the necessary software requirements assisting my research work and solving my computer related issues in a timely manner.

I would like to thank my friends Jayanth, Prashanth, Kannan, Muthu, Bharat, Deepak, Payal, Madhuri, Aravind, Steve, Bala, Vijay, Sarvesh and others who were there for me during my difficult days.

Finally, I would like to thank my parents, who constantly encouraged me and supported me throughout my career.

Contents

List of Figures

List of Tables

x

Chapter 1

Introduction

1.1 Background

The demand for a high-speed networking has grown tremendously over the last decade, which has led to the research and deployment of high-speed optical backbone networks. At the same time, the demand for untethered access has lead to increasing research and deployment of wireless access networks. The wireless and optical network technologies have fundamentally different characteristics. The aim of this thesis is to develop a framework to integrate networks with heterogeneous technologies providing a unified paradigm for end-to-end communication. A new end-to-end bursting architecture is proposed to achieve this goal.

1.2 Motivation

The motivation for this thesis is the vision of an emerging network world, which has moved from the wired access domain to untethered wireless access.

One of the main problems that arises due to these differences is the impact on transport layer performance, which provides the incentive for developing a unified end-to-end communication paradigm across heterogeneous realms.

1.3 Research Contributions

This thesis aims at providing a unified end-to-end communication model, in which the transport layer at the end systems determine the end-to-end data transmission characteristics over wireless and optical realms. Listed below are the contributions of this thesis research.

- develop an end-to-end bursting model to match application layer requirements

- develop a multi-realm simulation model for a unified communication paradigm

- analyze transmission control protocol (TCP) performance over heterogeneous wireless and optical realms

The first and foremost work of this thesis is to implement a simulation model of a heterogeneous network [3–6] with wireless access and an optical backbone. The access and the core network constitute individual realms [7] with distinct infrastructure. The individual realms, when joined together to form a network, cause performance problems due to the differences in their underlying technologies. This research effort is to reduce the effects that arise due to these technology differences by employing an unified end-to-end communication framework.

A realm is a group of nodes that share a common subnetwork technology and mechanism such as wired or wireless. The heterogenity at the realm boundary

may cause a number of performance problems. The aim of this research is that a smooth transition at heterogeneous realm boundaries with end-to-end bursting of transmission improves overall performance of the system.

The examples of realms that are considered in this thesis are optical burst swithed, and wireless (802.11). The first step in this approach is to model heterogeneous realms with appropriate and with edge gateways to communicate with other realms. Once each realm has been modeled, a new end-to-end inter-realm paradigm is designed and modeled that enables efficient data transition at the realm boundaries.

1.4 Organization of the Thesis

The rest of the thesis is organized as follows: Chapter 2 provides a broad review of heterogeneous networks, particularly of optical and wireless subnetwork realms and their interaction with the transport layer. Emphasis is placed on the performance of TCP over such networks under different scenarios. Chapter 3 considers the edge node burst assembly problem and its effects on TCP performance and proposes a new end-to-end bursting architecture and simulation model. Chapter 4 discusses simulation implementations for the various scenarios proposed in Chapter 3. Chapter 5 provides analysis and outlines the results obtained from these simulation models. Finally, Chapter 6 discusses the conclusions drawn from this work and suggests possible future directions of research.

Chapter 2

Background and Related Work

This chapter provides background on the related work that has been done in the field of heterogeneous optical and wireless networking, starting with Section 2.1 describing individual network characteristics and how they differ from one another and the problems in their integration. Section 2.2 explains the different optical switching paradigms with emphasis on the architecture design of optical burst switched networks (OBS). Section 2.3 discusses the characteristics of wireless access network architecture and provides a brief overview of IEEE 802.11. Section 2.4 summarizes the basic operation of TCP and its flow control and congestion control mechanisms. Section 2.5 explains the performance problems of TCP on OBS networks due to the burst assembly and disassembly process. Finally, the disadvantages of the current implementations are described, which provides the basis for this thesis.

2.1 Introduction to Heterogeneous Networks

The global internet consists of core and access networks. The core networks aggregate traffic and transfers data between the edges over high-speed switches. Access networks allows users to connect to the high-speed core. The core network consists mainly of optical network technology with link rates of 10-40 Gb/s per wavelength, using synchronous optical networking (SONET) and increasingly Metro Ethernet. The backbone links transmit data in the optical domain with optical-to-electrical (O/E) conversion at the edge. The access network consists of wired or wireless links including 802.3 Ethernet, 802.11 wireless LAN (local area network) [8], 802.16 wireless MAN (metropolitan area network) [9] links, which provide users with a means to connect to the high-speed backbone. The recent 802.11n standard supports data rates in excess of 100 Mb/s. The combination of wireless access connected to the high-speed optical core form the heterogeneous networks, that are the topic of this thesis.

2.1.1 Characteristics of Wireless and Optical Networks

There exist many challenges when we interconnect two heterogeneus subnetworks having entirely different properties. In particular, the optical and wireless network realms posses dissimilar properties so interconnecting them provides a rough boundary at the edge. Table 2.1 shows the difference between the two realms. A detailed overview of each characteristic is necessary to understand the boundary effects between the two technologies.

- *Switching paradigm*: The wireless access networks are generally packet switched

Table 2.1. Characteristics of wireless and optical realms

Characteristic	Wireless Realm	Optical Realm
switching paradigm	packet	circuit
bandwidth	limited	high
distance latency	low	high
queuing delay	high	low
bandwidth-$\times$-delay product	low	high
link reliability	weak/episodic	high
mobility	high	none
available power	limited	abundant

based on IP (Internet Protocol) [10]. On the other hand optical networks are circuit switched networks. Circuit switching is required due to the slow switching time of optical switch elements (e.g. micro-electro-mechanical systems [MEMS]).

- *Bandwidth*: Wireless link bandwidth is constrained due to the limited spectrum allocation as in the industrial, scientific and medical (ISM) radio bands. The bandwidth is further constrained due to the need for multiple users to share the medium. Wireless access rates typically vary from 2 Mb/s to 100 Mb/s depending on the technology such as 802.11 and 802.16. The access bandwidth is also limited by channel characteristics such as noise, interference and attenuation [11]. On the other hand, optical networks consist of a space division mesh consisting of point-to-point links with the capacity of the order of 10 Gb/s to 40 Gb/s per wavelength.

- *Latency due to distance*: Wireless links generally have low latency due to the short distance in access network. Optical backbone links deployed in long-haul wide area networks have high latency due to the speed of light delay.

- *Queuing delay*: Optical networks generally have low queuing delay as they are circuit switched networks without significant internal queuing. The wireless access networks may have high queuing delay that is dependent on the traffic in the network and the channel conditions.

- *Bandwidth-×-delay product*: The bandwidth-×-delay product [12] is the product of link capacity and edge-to-edge delay, which is the number of bits in flight in the network. Long-haul optical networks generally have high bandwidth-×-delay product and are referred to as long fat networks (LFN). This product is an important performance metric for TCP performance.

- *Link reliability*: Wireless network connectivity may be weak and episodic, as there are many factors that affect the performance of the links such as noise, interference and mobility. Optical links are engineered to be highly reliable with forward error correction (FEC) using amplifiers and regenerators, and are relatively immune to noise and interference.

- *Mobility*: Wireless nodes may be mobile in nature to allow untethered use. This relative mobility results in dynamic topology and traffic conditions, and can impact connectivity. Optical fiber links are stationary resulting in a static topology.

- *Available Power*: Optical links have abundant power as they are connected directly to the power grid. Untethered wireless nodes require a battery that provides limited power supply.

2.2 Optical Core Networks

Optical networks evolved from research on the possibility of transmitting light through an optical waveguide. The first waveguide was a silica-based fiber made as a multimode fiber [1]. The multi-mode fibers suffer intermodal dispersion due to the delay in arrival time of the different reflection modes, requiring regenerators at frequent intervals. Single-mode fibers were later introduced permitting higher data rates over longer distances. Optical networks consist of high bandwidth point-to-point links, but at the edge of the link the signals are converted from optical to electrical for switching.

Much of the processing time and power consumption on these networks are due to the optical-electrical O/E converters at the interface to electronic switches, resulting in research leading to the introduction of all-optical networks [1]. Optical switching provides such a solution by directing the traffic to the appropriate egress switch without optical-electrical-optical (OEO) converters.

2.2.1 Switching Paradigms

Current optical networks use optical circuit switching (OCS). Recently, optical burst switching (OBS) has been proposed, with some researchers pursuing optical packet switching (OPS).

2.2.1.1 Optical Circuit Switching

Optical circuit switching is a paradigm that evolved from the optical-electrical-optical (OEO) based networks with electronic switches, to all-optical networks with wavelength routed networks. This switching operates on a three state basis which involves circuit setup, data transfer and circuit teardown.

All optical circuit switching reserves a light path along the transmission path through all the switches involved in transmission. Light path circuits are permanently provisioned, but may also be signaled with setup and teardown messages.

2.2.1.2 Optical Packet Switching:

Optical packet switching (OPS) is a potential futuristic per packet switching technology. The major advantages of OPS would be bandwidth efficiency, flexibility and functionality [1]. Wavelength routed networks provides a granularity of one wavelength, so data streams smaller than this results in inefficient link utilization which can be offset avoided by aggregating multiple flows.

OPS would solve this problem by providing greater granularity by using smaller size packets transmitted optically with headers that are processed either optically or electronically. The main problem arises here are due to the nature of optics which does not provide any random access memory (RAM) equivalent to buffer the data in optical domain. The queuing in optical domain are done using FDL (fiber delay lines) by delaying the packet on a fiber loop to provide time to process the header. Secondly, OPS would require switching rates much faster than possible with current switching elements. These two fundamental problems makes OPS a dream which would become reality only with significant innovations in optical

hardware technology.

2.2.1.3 Optical Burst Switching

Optical burst switching (OBS) [13–17] is a paradigm that evolved as a intermediate bridging technology between OCS and OPS with some of the advantages of each. OBS has the ability to transmit bursts of packets across the network, which makes it a viable option for future Internet transport. This switching paradigm is still evolving and has not yet seen commercial deployment. Some of the characteristic properties of OBS include:

- *Bandwidth utilization:* Bandwidth utilization is better in OBS than OCS networks as they multiplex bursts from multiple flows on each lightpath, which is not the case in OCS which requires dedicated light path.

- *One way fast reservation:* The reservation of channels is made by transmitting a control packet which does not generate an ACK for reservation. The offset time used in separating the control packet from the subsequent data burst eliminates the need for buffering.

- *Setup latency:* The setup latency is less in OBS than OCS since it has a one-way fast reservation and does not require the setup and tear down of light path as in OCS.

- *Switching granularity:* OBS provides a granularity in between that of OPS and OCS by transfering data in the form of bursts, which are a set of data packets combined to produce a frame transmitted over the optical links.

- *Separating data and control information:* Control information is transmitted out-of-band or in-band with respect to wavelength depending on the OBS signalling type but separated from the data.

- *Traffic adaptability:* OBS can adapt relatively well to bursty traffic as it can multiplex the incoming traffic at the ingress switch.

2.2.2 Optical Burst Switching Architecture

The OBS [13] network consists of a group of *edge* and *core* nodes. The edge nodes perform routing, burst assembly and sorting of data packets. The core nodes switch the bursts to the receiving edge nodes as shown in Figure 2.1. The edge node that accepts traffic from other subnetworks is called the ingress edge and the edge node that delivers traffic to other subnetworks is called the egress edge. The architecture of the ingress edge switch is depicted in Figure 2.2 containing
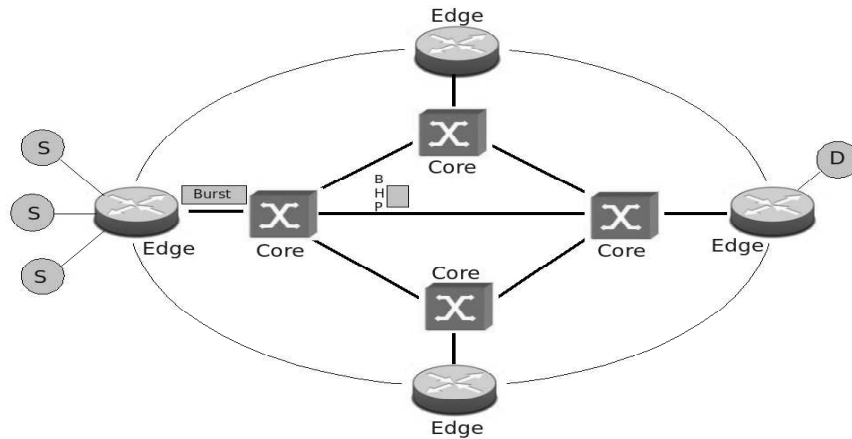

Figure 2.1. OBS network architecture

three essential modules: the burst assembler [18], routing module, and scheduler. The routing module selects the specific burst assembler module depending on the destination of the packet. Once the burst assembler is chosen, it frames the packets into bursts as shown in Figure 2.3. Each packet class has a distinct queue in the assembly unit in which the packet gets queued to be formed as a burst.
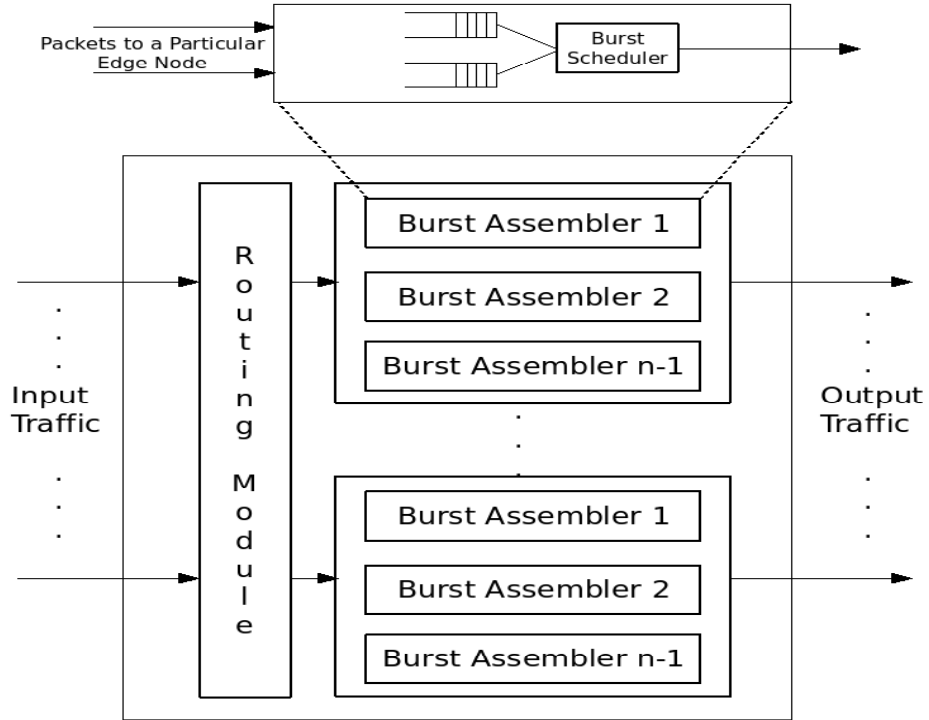
Figure 2.2. Architecture of edge node [1]

The ingress edge is also responsible for sending the burst header packet (BHP) that provides the advance control information about the burst. This control packet is initiated by the ingress edge node and passed on to the receiving edge for wavelength reservation. The ingress edge then transmits the burst after an offset time without waiting for an acknowledgement.

The receiving edge node receives the burst and sends it to the burst disassembler, which disassembles the burst into packets that are sent to the receiving

13
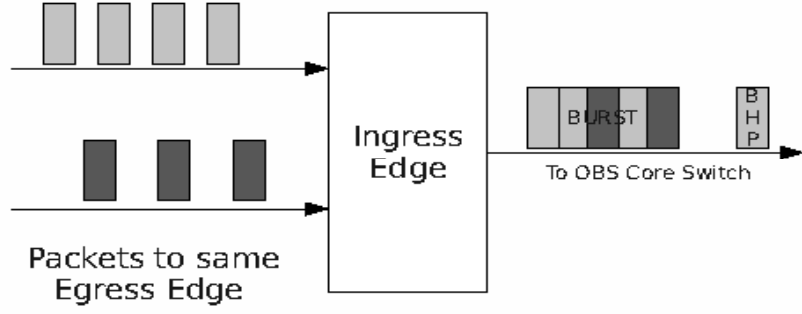
Figure 2.3.   Burst aggregation at edge node

subnetwork. The burst disassembly process is depicted in Figure 2.4.
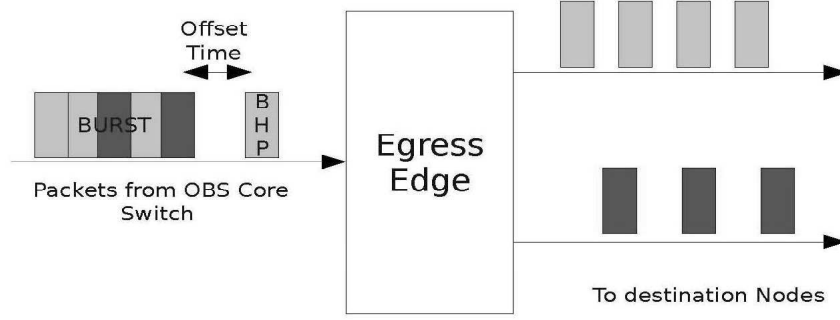
Figure 2.4.   Burst segregation at receiving edge node

The core nodes of the OBS network consist of two planes, the control plane and the data plane. The control plane processes and forwards signaling control information. The data plane is responsible for transferring data from the input ports to the output ports of the switching fabric. The architecture of the core node is illustrated in Figure 2.5. When the burst reaches a core node, the node demultiplexes the control information and sends it to the control plane where the information in the control packets are processed by converting it into an electrical signal through an O/E converter and sent to the data channel scheduler.

14

Depending upon the nature of channel requirement of the burst, the scheduler assigns a channel to the burst and the control information is sent back in optical format with new headers, which may differ from older header information that was in the packet at the ingress port, such as the offset time that depends on the contention at the node depending on the current traffic.
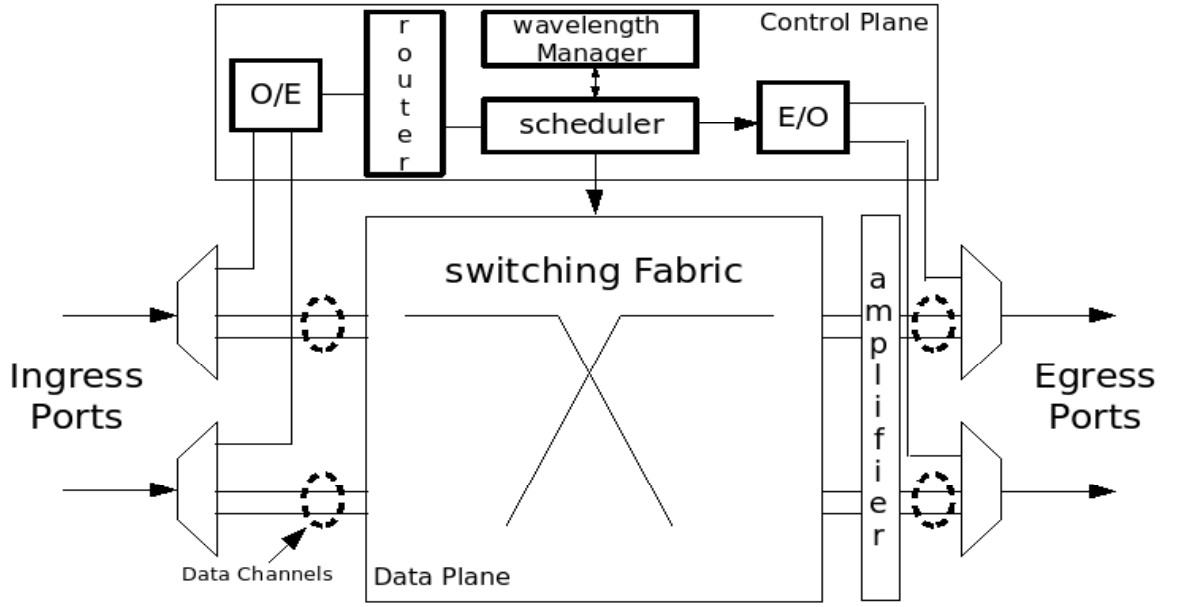
Figure 2.5. Architecture of core switch node [1]

The data burst remains in the optical domain and is switched to the output ports by the switching fabric using the data channel scheduler. If fiber delay lines (FDL) are used they can assist in resolving contention without which there may be burst loss, since the burst may be sent only after the control packet is processed.

2.2.2.1 Signaling

The OBS network signaling diagram is shown below in Figure 2.6. The diagram shows the source (ingress) edge node and the receiving (egress) edge node as **S** and

15

D, with the intermediate relaying core nodes as **N1** and **N2** switching the traffic. Whenever the source node **S** is ready to transmit a burst, it sends a control packet which forms the wavelength reservation setup. On receiving this signal from the ingress edge, the core node processes the BHP and sets up a wavelength to send the signal to the next hop node for its setup. This process is continued until the destination is reached. The source transmits the data burst with the initial offset time equal to $T_{\text{offset}}^{(\text{min})}$. This offset is changed on each relaying node depending upon the network conditions. The data bursts are delayed using fiber delay lines (FDL) until the control BHP packet is processed and a channel assignment is made that is governed by the T_{proc} BHP processing time. Once the burst passes through the output port of the node, the channel is released.
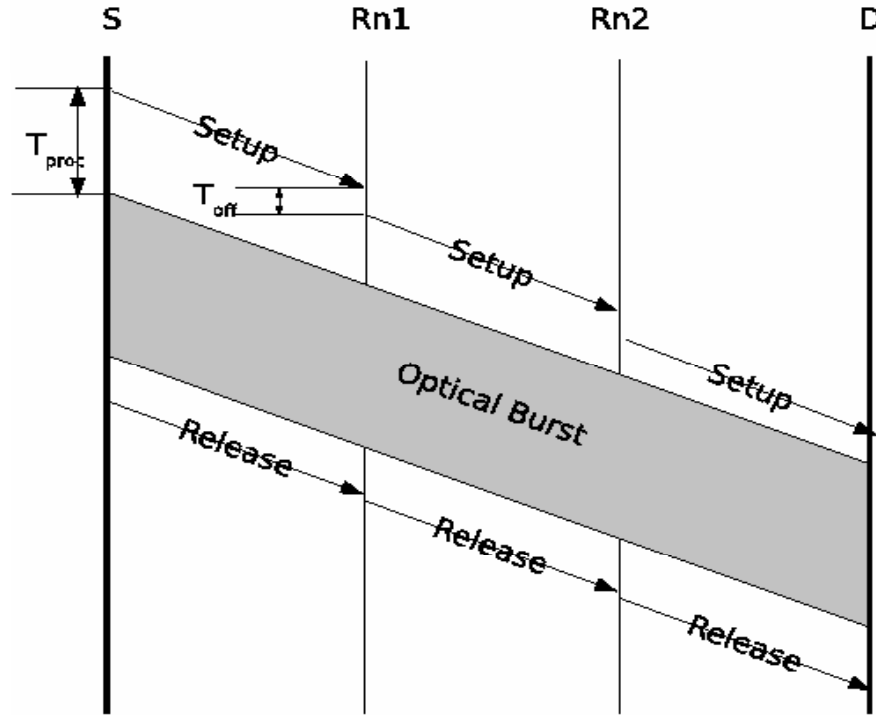

Figure 2.6. Signaling in OBS networks

From Figure 2.6 we see that the signaling mechanism is governed by the timers that assure proper transport of the burst. These three important parameters are the minimum offset time T_{offset}, BHP processing time T_{proc}, and optical switching time T_{OXC}. The configuration of initial offset time should be made such that the burst reaches the receiver with the suitable offset time after the BHP packet is processed, using the Equation 2.1:

$$T_{\text{offset}}^{(\text{min})} = nT_{\text{proc}} + T_{\text{OXC}} \tag{2.1}$$

where $T_{\text{offset}}^{(\text{min})}$ is the minimum offset time between the BHP control packet and the first bit of the burst to reach a receiving node, n is the number of nodes in the path including the ingress and the egress nodes, and T_{proc} is the BHP control packet processing time taken by a node. T_{OXC} is the delay for a node to configure its switching fabric, once the command from the scheduler is received, to transmit the first bit of the burst from the input ports to the output ports.

2.3 Wireless Access Networks

The wireless local area network (WLAN) has many advantages over conventional wired access, the most important of which is to enable untethered operation and mobility. This comes at the cost of bandwidth and reliability, due to the properties of the wireless channel that include noise, interference, distortion, and fading. Wireless access networks are primarily deployed as WLANs and WMAN (wireless metropolitan area network). 802.11 is the standard for WLANs and is therefore used as the basis for the simulations in this thesis. The 802.11 standard specifies the physical and MAC layers.

2.3.1 802.11 Access Networks

802.11 is the standard for local-area wireless access. It comprises of two layers: the physical layer and medium acces control (MAC) layer which also provides link-layer framing. The physical layer uses a number of coding techniques and the direct sequence spread spectrum (DSSS) modulation techniques depending on the variant in use (e.g. 802.11g or 802.11n).

The 802.11 medium access control (MAC) layer provides reliable transfer of data by implementing link-layer ARQ (automatic repeat request) acknowledgements and retransmissions. Thus, 802.11 marks the effects of moderately unreliable channels. Contention is resolved using carrier sense multiple access/collision avoidance (CSMA/CA).

In the 802.11 distributed coordination function (DCF), the source senses the medium, and if the medium remains idle for a DCF interframe sequence (DIFS) interval, the source accesses the medium. If the medium is busy to access, the source backs off randomly to a value within the contention window. To improve fairness and starvation, the back-off timer values are set low for nodes that waited for long times.

802.11 wireless networks operate in two modes: infrastructure and ad hoc mode. In infrastructure mode, all communication is a single hop from the wireless node to an access point connected to the Internet. In ad hoc mode, multiple hops are needed, requiring a routing protocol.

2.3.2 Transport Layer Protocols and Issues

The *transport layer* is responsible for end-to-end transmission of messages. It is also responsible for flow control and error control of the end-to-end data transmission. The transport layer can operate in two modes: connectionless mode and connection-oriented. Connectionless mode is useful when reliable transmission is not required and it eliminates the overhead of connection management. User datagram protocol (UDP) [19] provides this transport layer service in the Internet. Connection-oriented mode provides a reliable service end-to-end ensuring proper transmission of messages. Transmission control protocol (TCP) [2] is the reliable transport protocol for the Internet. A breif overview of these two protocols is presented below.

2.3.2.1 User Datagram Protocol

UDP is a connectionless transport layer protocol offering unreliable end-to-end transport service with minimal control overhead. The characteristic features of UDP protocol include:

- no connection establishment

- higher goodput due to less overhead

- no congestion control

- unordered delivery of packets to the application layer

- no handshaking involved between source and destination

Figure 2.3.2.1 shows the UDP segment format with which the protocol transmits the data. The header consists of five fields: the source port that specifies the data originating port, the destination port specifying the destination to which the segment is sent, the length of the segment and the checksum value that permits the receiver to check the integrity of the data.
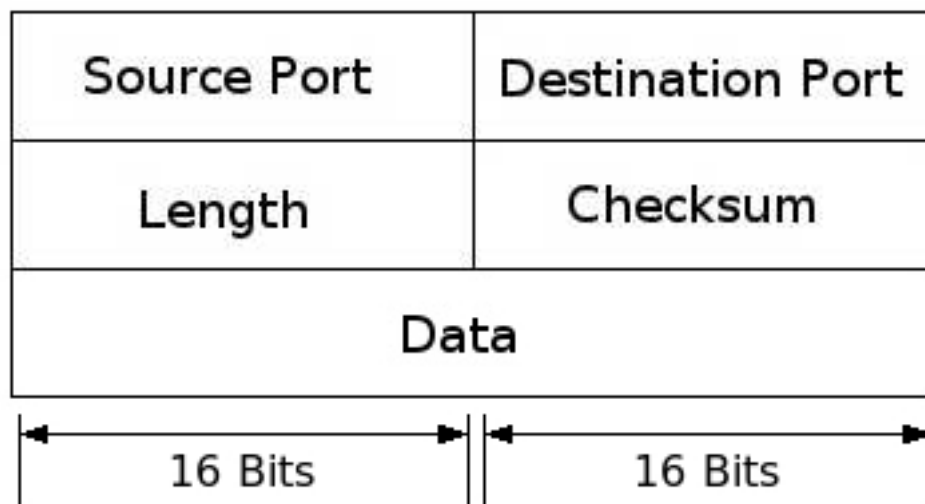
Figure 2.7.   UDP segment

2.3.2.2 Transmission Control Protocol

TCP provides a connection-oriented reliable service. TCP uses a three way handshake between the source and the receiver by which a connection is established. As shown in the Figure 2.8, the source sends a synchronization (SYN) segment initiating a connection to the destination. On obtaining the SYN, the receiver replies with an acknowledgement (ACK) to the source. The source then acknowledges the destination and data segments can be transmitted.

Figure 2.8. Three way TCP handshaking [2]

Some of the significant features of TCP include:

- connection oriented protocol

- point-to-point transmission of data

- end-to-end reliability for data delivery

- explicit connection start and teardown (SYN/FIN)

- congestion control and flow control mechanisms add more reliability

2.4 Flow Control and Congestion Control in TCP

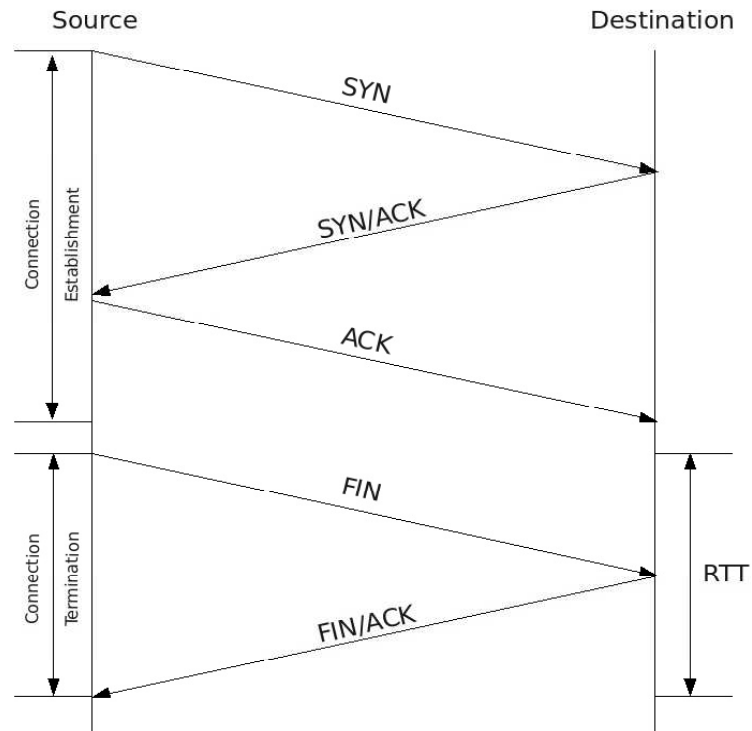Two important mechanisms which affect heterogeneous networks while using TCP protocols are *flow control* and *congestion control* [20–22]. When segments are assumed lost, TCP waits for one round trip time (RTT) before sending again. Estimation of RTT is very important and affects directly the performance of TCP. The RTT is computed using the formula given by Equation 2.2. Error Control in TCP is based on byte sequence number, retransmission time out (RTO) and acknowledgements. TCP uses the RTT to set the retransmission timers before retransmitting the segment. The timer expiration is caused if the ACK is not received, which is considered to be packet loss by the sender.

$$RTT_{new} = \alpha \times RTT_{old} + (1 - \alpha) \times New\_Sample \qquad (2.2)$$

Congestion control in TCP is an additive-increase and multiplicative-decrease (AIMD) process by which the window size is adjusted to affect the throughput based on network conditions. The window size used for flow control is a minimum value between congestion window and receiver advertised window size. The window size starts from one data segment or a particular initial value and is increased exponentially upon the receipt of every successful ACK until the transmission pipe is full. This increase in window size is known as slow start and the estimation on capacity of the pipe is called slow-start threshold. Upon reaching the pipe bandwidth, the sender goes in to a congestion avoidance phase where the window size is increased by one segment upon every round trip time.

2.5 TCP Performance on OBS networks

The performance of transport protocols in OBS (optical burst switching) networks has been researched. There are a number of research papers investigating shortcomings of TCP over OBS. TCP performance is affected mainly by the three reasons discussed in [23]: packet drops, end-to-end delay changes, and throughput changes. Some of the major parameters affecting the performance of TCP over OBS networks [15,23–26] are burst size, burst timeout, and fiber delay lines, burst switching affects the transport layer performance due to the burst assembly delay at the edge node [23]. If λ is the arrival rate and B is the maximum burst size at the edge node, then the average bursting time is B/λ. For low arrival rates the bursts are generated using timeouts. Therefore, assembling the packets is an essential aspect when using TCP. A number of algorithms have been proposed to determine how packets are assembled in OBS network using TCP [27]. Two significant algorithms have been proposed to define the assembly of packets in OBS networks: fixed assembly period (FAP) [28] and adaptive assembly period (AAP) [28]:

- FAP: A simple general method by which a fixed assembly period AP is defined at the edge node, and works well under uniform traffic conditions.

- AAP: If the traffic is bursty as in the case of TCP, the FAP algorithm does not take into account the traffic to form the burst which leads the way to AAP based assembly. The AAP assembles the burst depending on the traffic conditions by varying the burst size adaptively with respect to the traffic.

It has been shown that the performance of FAP is poor due to the large assembly period and delay for data segments and ACKs resulting in a poor per-

formance [24]. As the number of connections into the ingress edge increases the performance of the network also is better. Better performance is reported with the AAP mechanism with higher goodput given a sufficient number of TCP connections. However, both mechanisms perform poorly when the arrival rate is low and with a low number of TCP connections. This shows the inability of TCP to drive itself when there aren't enough TCP inputs to the edge nodes to assist in rapid burst formation. The main aim of this thesis topic is to eliminate this shortcoming by employing a new bursting methodology.

The TCP throughput is modeled as [25]

$$\text{Throughput}_{\text{TCP}} = \frac{\text{MSS}}{(\text{RTT}\sqrt{\frac{2bp}{3}} + T_o\min(1, 3\sqrt{\frac{3bp}{8}})p(1 + 32p^2))} \tag{2.3}$$

Where MSS is the maximum segment size in bits, RTT is the round trip time, and b is the number of packets acknowledged. The RTT is defined as

$$\text{RTT} = 2 \times (T_{\text{assembly}} + N_{\text{hops}} \times T_{\text{hop}} + T_{\text{disassembly}}) \tag{2.4}$$

where T_{assembly} and $T_{\text{disassembly}}$ are the assembly and the disassembly times at the OBS edge node, T_{hop} is the propagation delay on each link, and N_{hops} is the number of links.

The performance of TCP also depends on the burst loss in the network, which throttles the sender rate by the congestion control mechanism. Increasing the burst size might solve this problem, as fewer bursts are generated to be dropped [23].

In summary, the performance reported in prior work consists of two main effects: delay penalties and correlation benefit. The delay penalties reduce the transmission rate of a TCP connection proportional to the ratio between the burstification period and RTT. High correlation benefit is obtained using high-rate sources since burst loss and gain occur in a single congestion window.

2.6 TCP Issues over Heterogeneous Networks

TCP increasingly must cope with different transmission technologies and mediums due to the heterogeneity of today's Internet, which uses optical, satellite, wired, and wireless links for transmission of data. These technologies contain inherent differences such as bandwidth, delay and loss causing a variety of issues to which TCP must react. Most of the literature about TCP considers single subnetwork technologies rather than providing a realistic scenario of the Internet containing dissimilar subnetwork types [3, 29, 30]. Unlike the fiber optical backbone, wireless links frequently have high bit error rate (BER) due to channel characteristics and impairments. The connection between two wireless nodes may be weak and episodic, which can stall TCP and cause connections to time out. Link asymmetry is another factor that causes packet losses in TCP due to the ACK compression and selfclocking nature of TCP.

The future scenario of TCP over OBS and wireless hybrid networks has been discussed in [31]. The RTT of the network depends upon the burst manager configuration consisting of burst timeout (BTO) and burst size. The BTO is a dominant parameter when the traffic load is light, since the bursts are generated only when it expires. On the other hand, the burst size is a dominant parameter

when the load on the ingress edge is high. The performance of TCP depends upon burst loss and the link loss due to the optical and wireless links. The modeling of the network should be done with attenuation to the wireless links since most errors are caused due to the wireless links.

2.6.1 Parameters governing TCP Performance over OBS

OBS network operations depend on different parameters which affects the transport layer performance. The parameters are classified into two categories: end-to-end and hop-by-hop.

2.6.1.1 End-to-End Network Parameters

The end-to-end network parameters include maximum burst size, burst time-out, offset time and burst header packet (BHP) processing time. While maximum burst size provides the amount of data required to form a single burst at the ingress edge, the burst timeout is the minimum amount of time a node waits for the burst to form. The offset time is the time between the BHP packet arriving at the core node and the first bit of the data burst being received. The BHP processing time is the time taken for a node to process the BHP and reserve a wavelength for subsequently arriving data burst. These parameters govern the performance across the network.

2.6.1.2 Hop-by-hop Network Parameters

The hop-by-hop network parameters consist of the bandwidth of the data and control channels and number of data and control channels for each link. Thus, the composition of all of the link parameters affects the end-to-end TCP performance.

Chapter 3

End-to-End Bursting Model

This chapter explains in detail the proposed solution to the problem of TCP performance in heterogeneous networks with low traffic arrival rates. Section 3.1 provides an introduction and Section 3.2 the end-to-end approach for enhancing TCP performance. In Section 3.3 the architecture of the proposed solution is described. Section 3.4 describes the network and node model that will be used to simulate the proposed solution.

3.1 Introduction and Problem Statement

In this thesis we use the concept of *realms* [7] that are introduced by the postmodern internet (PoMo) architecture. Realms consist of nodes that share common network layer properties, as well as trust and policy. Communications between realms are made through the inter-realm gateway. In this thesis, wireless access and OBS (optical burst switching) core realm models are developed for next generation Internet architecture. An important aspect of this approach is

to develop an unified end-to-end communication mechanism that can match the requirements of the application layer as in applcation layer framing (ALF) [32]. Using this approach, the nodes in the wireless realms will be able to construct bursts based on the application layer requirement that are transmitted end-to-end through the OBS core.

This section provides the functional overview of the end-to-end communication paradigm that addresses the disadvantages of current implementations discussed in the previous sections. Detail on its components and architectural design are provided, followed by a discussion on advantages and issues with this approach.

With the current implementations, when dissimilar realms are interconnected, such as wireless access and OBS core networks, the wireless nodes will send IP packets to the wireless gateway connected to the edge router of the backbone. The ingress gateway queue the packets and forward them to the edge node for burst and assembly. At the egress edge router, the burst is debursted and the IP packets are sent to the egress gateway.

Such implementations suffer from improper synchronization of TCP and OBS parameters as discussed in Chapter 2. We propose a better structure for integrating such networks, which requires modifications to the architecture of both access and core networks to provide better end-to-end transport.

3.2 End-to-End Inter-Realm Bursting

The end-to-end inter-realm bursting mechanism is developed to unify the transport layer communications across realms. Some of the advantages of this

model are:

- elimination of burst aggregation and disaggregation at optical edge node

- better control over TCP congestion control and CWND sizes

- elimination of ACK self clocking problems by forming the burst at the edge
  nodes

The burst assembly process includes transmitting burst from the source wireless node with a maximum transmission unit (MTU) based on the application layer requirement. By doing so, the burst aggregation and desegregation is eliminated at the optical realm edge nodes.

3.3 End-to-End Bursting Architecture

Figure 3.1 depicts the bursting mechanism performed by the edge nodes in OBS networks. TCP at the source frames the segments and sends them to the network layer to route them to the next hop neighbor. The network layer performs this operation using the underlying link layer. The link layer establishes a logical connection with the neighbor and a physical connection using the physical layer. The packet gets relayed until it reaches the edge router of the OBS backbone.

The ingress edge node, frames the bursts and switches them through the core switch fabric to the egress edge node. The egress edge debursts into the original packets, which are then transmitted to the receiving wireless node through the access technology. On the reverse link, the ACKs are transmitted back from the
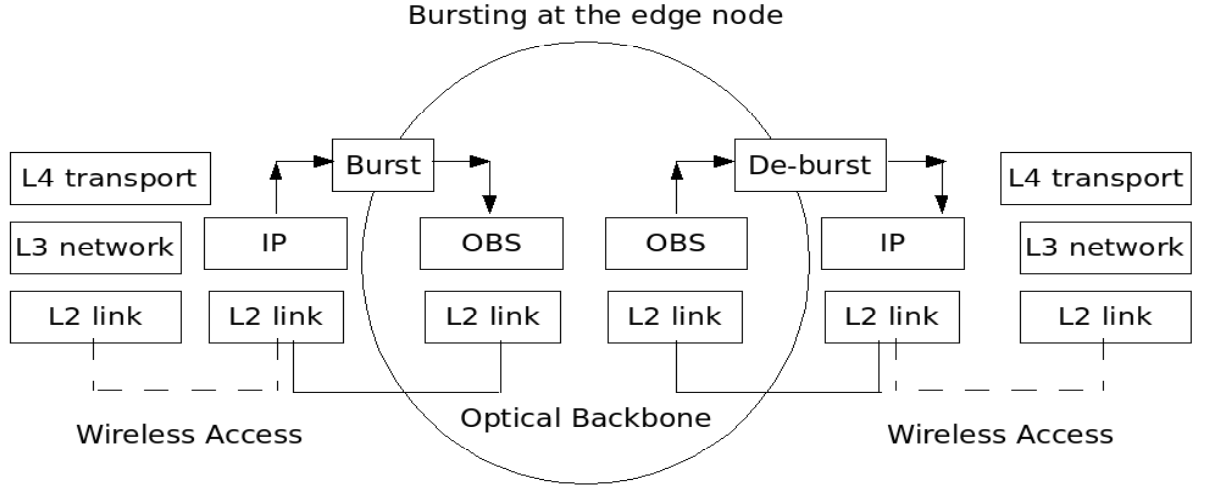
Figure 3.1. Edge bursting model

receiver to the source node, which enables TCP to maintain steady transmission rate. The burst switching at the OBS backbone is governed by the burst size and BTO (burst timeout) configured at the ingress edge router and the offset time between control and data packet configured by the edge router and transmitted appropriately through the core switch fabric, as described in Chapter 2.

In this model, the optical backbone network performs the burst generation at the edge router. Neither the source nor the destination node has any control over the rate of transmission of data in the core and is entirely dependent on the backbone network configurations. To obtain more control over the data being sent,we propose that burst assembly should be done at the end system rather than at the edge router as shown in Figure 3.2. In this case, the transport layer frames jumbogram segments that are transmitted end-to-end from the access network to the backbone and vice versa.

This, the process of bursting is moved to the end system, and the access-

network jumbogram segment is directly transferred into the optical backbone as an equivalent burst. Therefore the burst ans deburst blocks are not present at the OBS edge nodes. Note that the other functionality such as BHP and offset processing still must occur.
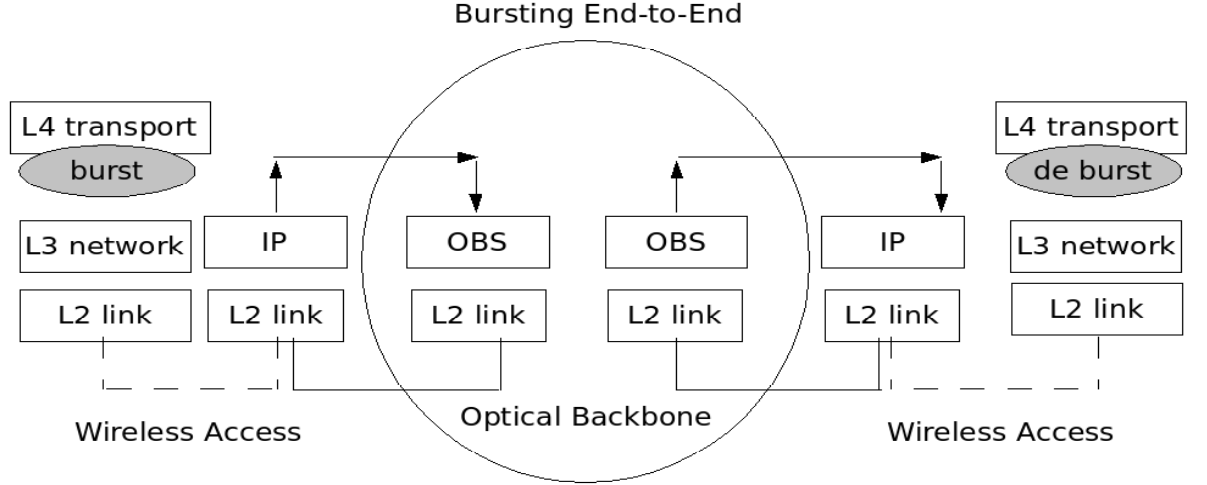

Figure 3.2. End-to-end bursting model

3.4 End-to-End Unified Bursting Model

In this section, we describe the end-to-end bursting model that will be used to simulate the proposed architecture.

3.4.1 Optical Backbone Network

The OBS network model [17] is modified to simulate the proposed solution such that the optical ingress edge does not do burst aggregation, but still converts the packets into the optical bursts generating the BHP (burst header packet)

necessary for transmission in the OBS core network. The burst assembly and disassembly delays are removed from the edge nodes and are moved to the source and destination TCP processing. The modified model is shown in Figure 3.3.
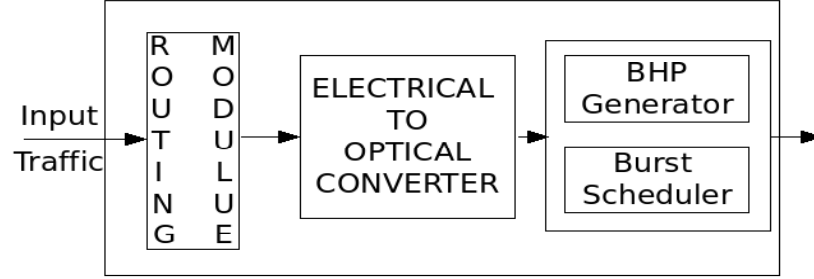

Figure 3.3. Edge node architecture

3.4.2 Wireless Access Network

The 802.11 wireless access network consists of nodes with the ability to burst end-to-end from source to destination access networks. The nodes operate in either infrastructure or ad hoc connectivity mode. The wireless node transport layer architecture is modified to obtain a bigger burst packet generated at the source [33]. Once the transport layer obtains the information, packets are burst accordingly. Figure 3.4 shows the wireless node model with the maximum segment size (MSS) adaptation implemented at the transport layer. The components and their functionalities are now explained. *Wireless Channel*: In untethered wireless networks, the nodes communicate among themselves using the wireless channels. The channel conditions depend upon various factors such as attenuation, interference, fading and topology characteristics, which determine the communication
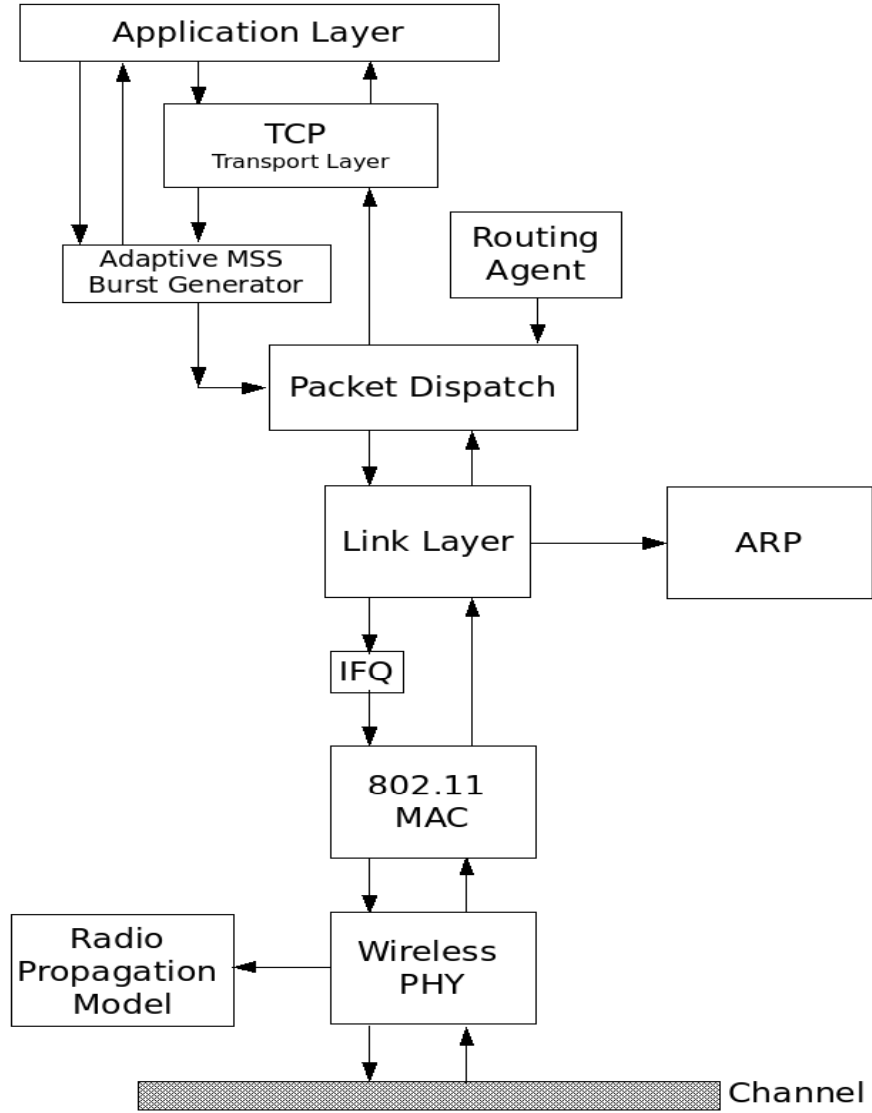
Figure 3.4. Wireless node architecture

process. The channel component is a shared medium, requiring medium access control. Depending upon the error model on the channel, the number of packets received is varied to simulate bit error rate (BER) characteristics of the channel.

Wireless PHY: The wireless physical layer (PHY) model provides the ability to decode a packet with respect to the radio propagation model and the channel

strength.

Radio Propagation Model: Alternatives for the radio propagation of the channel include free-space, two-ray ground, and shadow for multipath fading effects. The free space model can be used during the absence of any reflections, governed by the formula:

$$P_r = \frac{P_t \; G_t \; G_r \; \lambda^2}{(4 \; \pi \; d)^2 \; \times L} \tag{3.1}$$

where P_r is the received power at the receiver, P_t is the transmitted power at the source, G_t and G_r are antenna gains of the transmitter and the receiver, λ is the wavelength, d is the distance at which the received power is computed, and L is the system loss. The two ray ground model considers two paths: direct line of sight (LOS) path and one reflected, given by the formula:

$$P_r = \frac{P_t \; G_t \; G_r \; h_t^2 \; h_r^2}{(4 \; \pi \; d)^2 \; \times L} \tag{3.2}$$

where h_t and h_r are the heights of the transmit and receive antennas respectively.

802.11 MAC: The MAC (medium access control) layer of the wireless node takes the burst generated at the transport layer encapsulated in a link layer frame and transmits it over the wireless channel to the neighboring node. In the reverse direction, it receives the packet and sends the payload to the link layer for further processing. The MAC layer also notifies the routing protocol about the link strength and failures.

Link Layer and ARP: When the link layer receives a packet from the network layer, it sends an address resolution protocol (ARP) request to obtain the MAC

address of the next hop node to which the packet is sent. After ARP resolution, the network layer packet is encapsulated into a frame and put on the interface queue (IFQ). If a frame is received from the MAC layer, it checks for the packet type. If it is an IP packet, it is sent to the dispatcher, otherwise it is sent to ARP.

The ARP block maintains an ARP table that contains the required mapping between IP address and the corresponding MAC address necessary for 802.11.

Routing Agent: The routing agent is a part of the network layer that makes the decisions for the packet to reach the destination node. In this model we use DSDV routing [34], as will be described in Chapter 4.

Adaptive MSS Generator: The adaptive maximum segment size (MSS) determines the size of the burst to be generated. The jumbogram bursts are generated at the end systems such that they are equal to the burst size configured at the OBS edge node, so that the edge node does not wait for BTO (burst timeout) before transmitting.

TCP Transport Layer: The transport protocol of the wireless node is either UDP or a standard TCP Reno [35], depending upon the scenario.

Chapter 4

Simulation Implementation

This chapter describes the simulation implementations of the conventional edge bursting and the proposed end-to-end bursting model. The scenarios considered for different simulation runs are explained in detail. A detailed description of the simulation environment is given in Section 4.1. The simulation setup for the edge bursting model and end-to-end bursting model is presented in Section 4.2 and Section 4.3 respectively.

4.1 Introduction to Simulation Environment

All simulation scenarios are based on the wireless-optical model, where the access for end users is wireless and the backbone is optical. The objective is to analyze how these fundamentally different technologies perform when they interact with each other. Hence, the simulation scenarios are targeted towards studying the differences between the two realms. The overall objective is to select an efficient bursting mechanism for the wireless-optical network.

The simulation models are developed using ns-2 [36] simulator, version 2.28. The stock version is modified by adding OBS (optical burst switching) module [31,37] and patches to increase the TCP maximum segment size (MSS) [33,38]. The topology consists of a 1500 m × 1500 m area with two identical wireless realms, each consisting of 10 wireless nodes interconnected by an optical realm, as shown in Figure 4.1. Designated base stations act as gateways to access wireless realms from the optical backbone and vice versa. The simulated optical realm uses conventional circuit switching and has a ring topology consisting of edge routers and core switches. The edge routers are the nodes that are connected to the core switch fabric, which assemble and schedule data packets across the core fabric to the receiving edge node. The core fabric consists of high-speed switches that switch traffic across the edges. The designated gateway, connected directly to the edge routers of the wired backbone, translates data packets between the realms. The simulation scenarios include varying propagation delays in the optical network representing different backbone networks (from local area network (LAN) to intercontinental wide area network (IWAN) type optical links). The propagation delay thus, varies from a microsecond to a second.
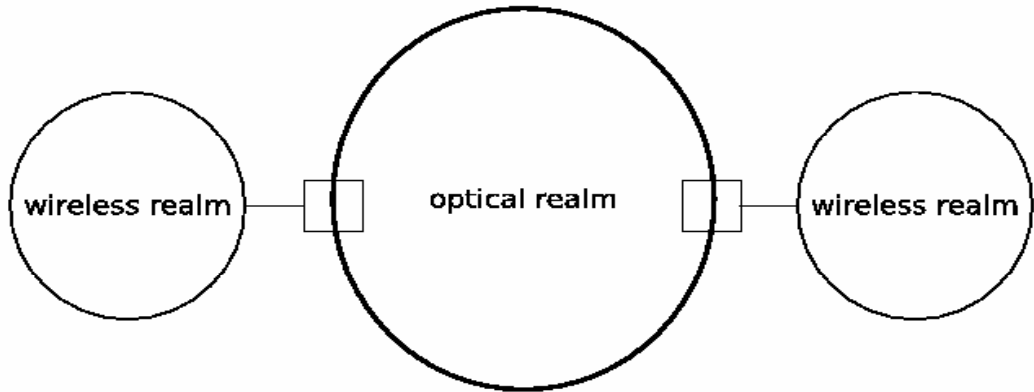

Figure 4.1. Heterogeneous realm model

Optical wavelength division multiplexing (WDM) network simulator (OWNS) [37] modeled optical link is used for optical simulations. The wireless realm is modeled using traditional 802.11b MAC and physical interface with hierarchichal addressing and droptail interface. Both ad hoc and infrastructure mode of operations in the wireless access network is simulated.

The performance of this heterogeneous network is evaluated for varying, transport protocols (UDP and TCP variants) and applications (CBR (constant bit rate), FTP (file transfer protocol)) [39]. The analysis is based on performance measures, such as goodput,and end-to-end delay.The simulations are repeated for both edge and end-to-end bursting.

The tool command language (TCL) scripts for the ns-2 simulator are generated with matrix laboratory (MATLAB) script generator using design parameters as shown in the Figure 4.2. Each simulation scenario is repeated 10 times to obtain the average value. The simulated output trace files are post processed using principal extraction language (PERL) scripts to obtain throughput, delay and goodput data that is finally plotted using MATLAB.

4.2 Edge Node Bursting Model

For edge bursting model, he OBS realm gateways are assumed to have large memory buffers and high processing power to process network traffic at a rate of several gigabits/second. The wireless gateway is connected to the subnet of the optical edge node. This gateway then relays the traffic to the destination node in the wireless realm over a single hop or multi hops depending upon the location of the destination node. The routing protocol in the wireless realm is
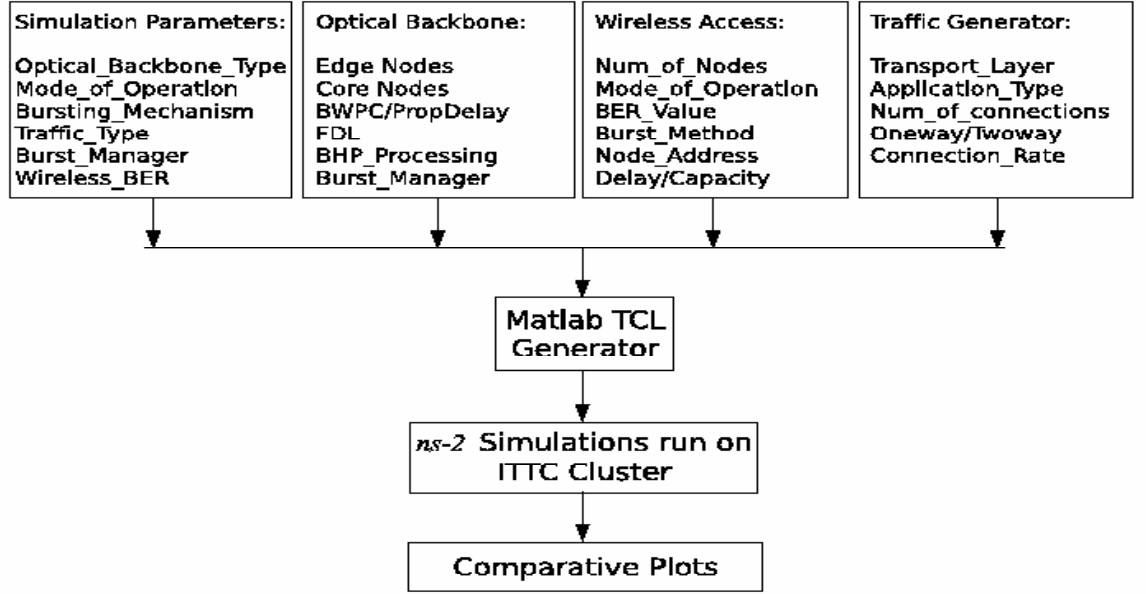
Figure 4.2. Implementation process diagram

destination sequenced distance vector (DSDV). Hierarchical addressing/routing schemes are used to attach the wireless nodes to their corresponding gateways. The addressing scheme, as shown in Figure 4.4 similar to IP address has a 3 level heirarchy: **Domain.Cluster.NodeID**, where the domain defines wireless or optical domain, the cluster represents a group of nodes within the given domain and the node id is the address of an individual node.

In this simulation model the conventional burst assembly and disassembly is done at the optical edge nodes. The wireless access network transmits its data as 1 KB packets which are aggregated at the edge nodes according to the burst manager configuration. The burst manager sets the burst size and burst timeout value at the edge node, that determines the nature of the burst generated. In the simulation model, the burst size is configured at 64 KB with burst timeout

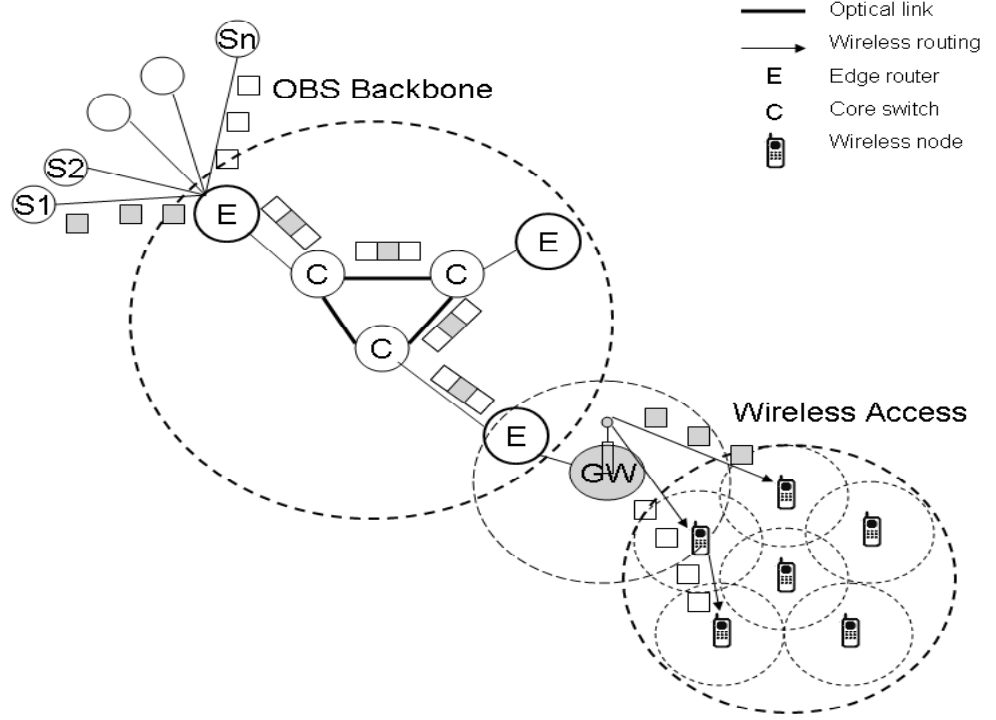varying from 1 ms to 1 s. This simulation scenario is shown in Figure 4.3.

Figure 4.3. Heterogeneous network with rate adaptation at the gateway

The simulation scenario is run for 50 times to obtain the statistical confidence intervals. All data packets destined towards a given edge are combined in a single burst. We simulate two types of application for the access network nodes: CBR and FTP with UDP and TCP Reno as the underlying transport mechanism. The ingress node aggregates these packets and bursts them to the egress edge through the core fabric. If the generated traffic from the access network is not sufficient to fill the burst size, the edge node pads the frame to create the burst and transmits it upon burst timeout expiration.

Table 4.1. Parameters governing edge node assembly

| *Characteristics*	*Methods*
burst Mechanism	edge node bursting
access network mode	infrastructure vs. ad hoc mode
application / transport layer	CBR / UDP and FTP / TCP
optical network delays	1 μs to 100 ms
wireless access topology	1500 $\times$ 1500 meters
access realm	2 networks with 10 nodes
simulation time	1000 seconds
application run time	500 seconds
queuing	droptail
burst size	64 KB
TCP MTU size	1 KB
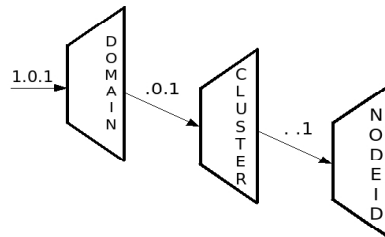
Figure 4.4. Hierarchical addressing and routing

For the simulation scenario shown in Figure 4.3, the **NAM** representation is shown in Figure 4.5. Table 4.3 lists the different parameters that are used for the simulation. The performance of this model for varying MTU sizes is evaluated through goodput and end-to-end delay plots.

Disadvantages

- burst aggregation can lead to increased delay when the burst timeout or burst size does not match the generated traffic rate

- TCP performance may be affected due to the queuing of SYN-FIN (synchronize-finish) packets leading to delay in connection establishment

- high core capacity and traffic load demands, higher processing power and memory at the gateways to assemble and disassemble bursts.

4.3 End-to-End Bursting Model

Modern Internet supports a MTU size of 1500 bytes, which is the traditional Ethernet frame size. Previous research suggests that high-speed networks require higher MTU's to effectively utilize link capacities [33]. The proposed end-to-end bursting is simulated to compare it against the edge bursting model. Hence, the simulationparameters, in this case are identical to that of the edge bursting model, so that meaningful comparisions can be drawn. However, the transport layer features of the wireless nodes are modified to support end-to-end bursting.

The modifications are made to the MTU size of the TCP segment which was set to 1 KB in the edge bursting model. The module is tweaked to accommo-
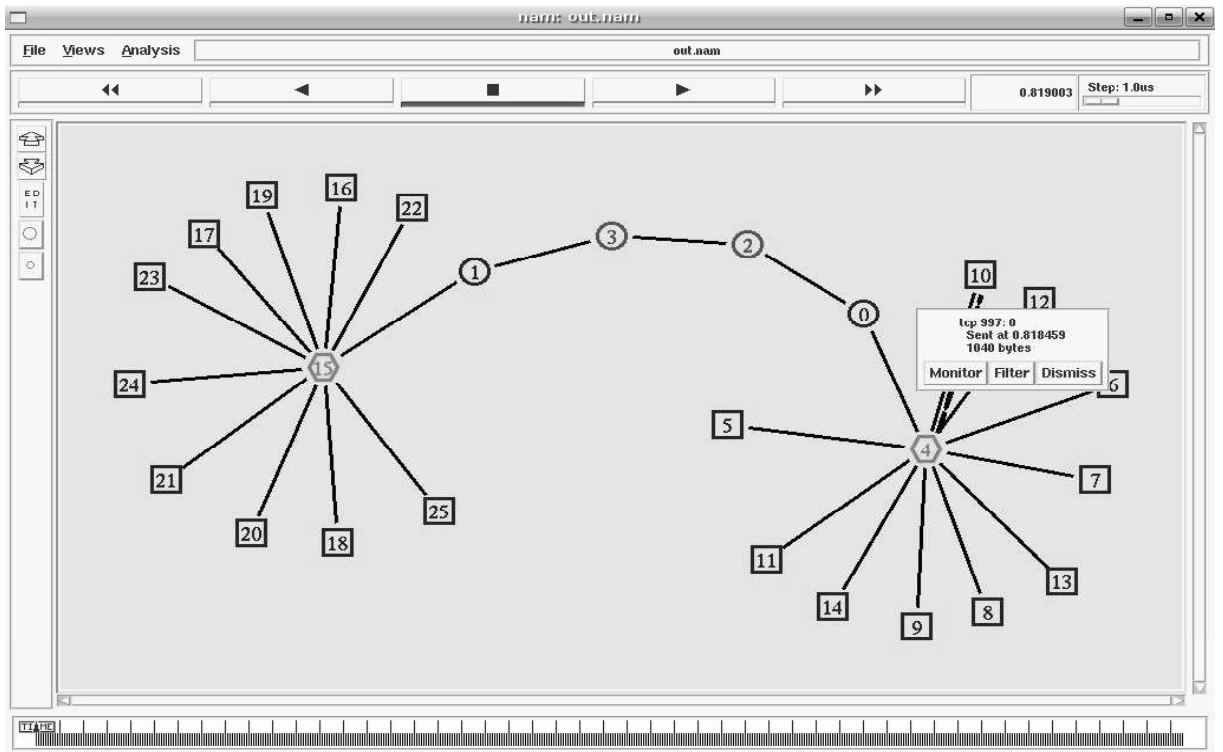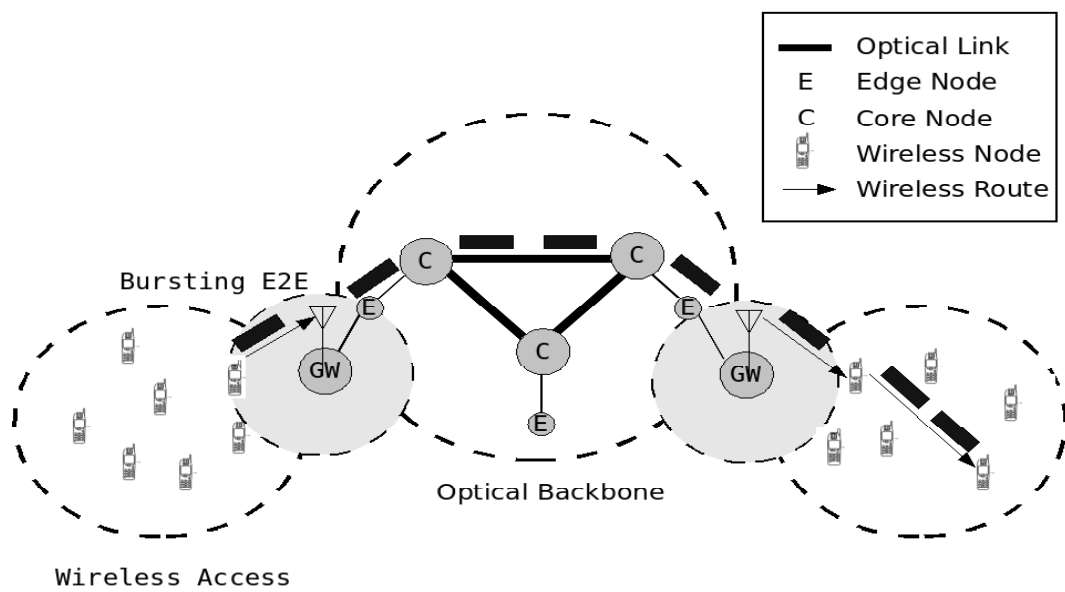
Figure 4.5. OBS burst simulation screenshot

Figure 4.6. End-to-end bursting model

date bigger MTU (maximum transmission unit) sizes, so that the bursts can be transmitted directly from the source nodes. The burst size at the optical edge node is set to equal the MTU size, so that there is no burst aggregation at the the edge nodes. Multiple runs were made by simulation paramenter such as, the burst timeout, MTU size, mode of wireless access, optical network propagation delay.

Characteristics	Methods
burst mechanism	end-to-end bursting
access network mode	infrastructure vs. ad hoc mode
application / transport layer	CBR / UDP and FTP / TCP
optical network delays	1 μs to 100 ms
wireless access topology	1500 $\times$ 1500 meters
access realm	2 networks with 10 nodes
simulation time	1000 seconds
application run time	500 seconds
queuing	droptail
TCP MTU size	64 KB

The NAM representation of the simulation scenario in Figure 4.6is shown in Figure 4.7. It also shows a sample TCP burst generated at the end system with 64 KB packet size.
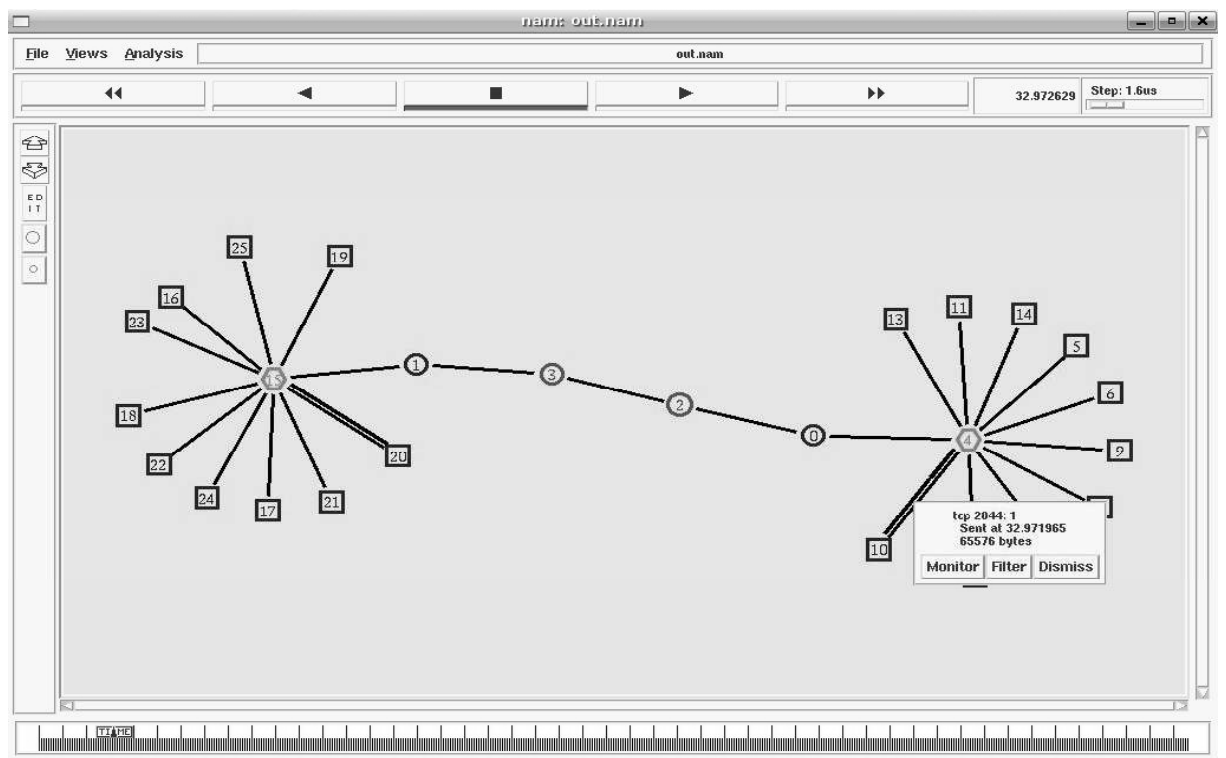
Figure 4.7. End-to-end bursting simulation model screenshot

Chapter 5

Results and Analysis

This chapter presents the performance analysis comparing the proposed end-to-end bursting solution with the edge bursting mechanisms. The chapter starts with describing the performance metrics used to analyze the simulation scenarios in Section 5.1. Plots are provided which compares the goodput, delay and TCP performances in subsequent sections from Section 5.2 to Section 5.4. The remaining sections analyze various performance aspects: throughput in Section 5.2, delay in Section 5.3, end-to-end TCP performance in Section 5.4, and analysis with wireless channel errors in Section 5.5. Finally, Section 5.6 summarizes the impact of end-to-end bursting on TCP performance.

5.1 Metrics and Plots

The simulation results are analyzed for corresponding metrics to understand the effects of TCP performance using the edge bursting and end-to-end bursting. Following are the metrics that are used to validate the simulation results obtained from post processing the traces. Recall that the simulation scenarios are run for

10 times and averaged to obtain 95% confidence.

Goodput is the rate at which packets successfully received by the application layer, after re-transmissions, if any have occured [40]. Goodput is calculated using the following formula:

$$\text{goodput} = (P_{\text{RX}}) \times P_{\text{size}}/I_{\text{measurement}} \tag{5.1}$$

Where P_{RX} is the number of packets received at the receiver per measurement interval $I_{\text{measurement}}$. P_{size} is the packet size delivered to the application layer at the receiver.

End-to-End delay is the time taken for a data segment to reach the receiver from the time instance it was generated at the sender.

TCP sequence plots are used to understand the performance of TCP over the network, in which the slope of the sequence number plot gives the network goodput.

5.2 End-to-End Throughput and Goodput

The first set of simulations evaluate the throughput of the proposed end-to-end bursting. The 802.11 access networks are configured in infrastructure mode and ad hoc mode to understand the effects of multihop routing on throughput. The burst size is configured to be 64 KB at the edge nodes and the TCP maximum transmission unit (MTU) is configured to be 1 KB for edge bursting and equalized to the 64 KB burst size for end-to-end bursting, to avoid aggregation delay at the ingress OBS edge. To capture the effect of the optical core delay, propagation delays ranging from 1 μs to 100 ms are simulated.

Table 5.1. CBR over UDP for edge and end-to-end bursting

Parameters	Configurations
burst mechanism	edge vs. end-to-end bursting
access network mode	infrastructure vs. ad hoc mode
burst size	64 KB
TCP MTU size	1 KB (edge) or 64 KB (end-to-end)
burst timeout	100 ms
application / transport layer	CBR over UDP
optical network delay	1 μs to 100 ms

5.2.1 UDP Throughput

This section explores the throughput of CBR (constant bit rate) traffic over UDP. Recall that UDP is connection less with no error, flow and congestion control. Table 5.1 shows the simulations carried out with UDP transport. Throughput is almost identical in all cases, as shown in Figure 5.1. CBR over UDP protocol performance performs well since the delay due to burst timers at the OBS edge do not affect the transmission rate of the source. Thus, the source fills the pipe at the desired rate at all times resulting in a straight line performance of end-to-end throughput. The end-to-end bursting protocol also performs equally well, since the only difference is where the bursts are formed.

5.2.2 TCP Goodput

On the other hand the performance of TCP is governed and scheduled by timers and the indows used for error, flow, and congestion control. Table 5.2 sows parameters for both end-to-end and edge bursting for FTP over TCP. Small burst timeout of 1 ms essentially does not aggregate packets at the edge is used
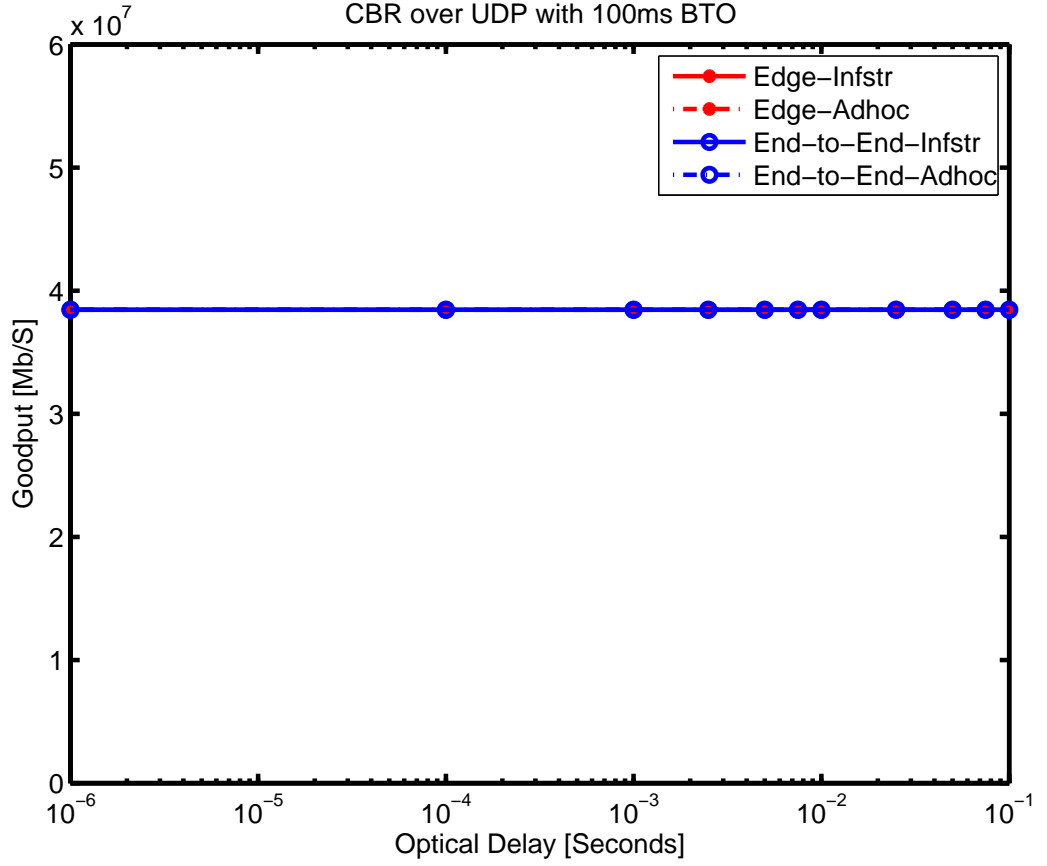
Figure 5.1. End-to-end throughput plot for UDP

to simulate end-to-end bursting. in conjunction with a 64 KB MTU jumbogram.
Figure 5.2 is plotted for one way propagation delay on the x-axis and application
level goodput on the y-axis. We see that the performance of the edge bursting ap-
proximately matches the performance of end-to-end bursting using infrastructure
mode on the 802.11 wireless access network. But in ad hoc mode, the performance
of edge bursting is lower due to the buffering of 64 KB jumbograms at multiple
hops. The confidence interval for 95% confidence with a propagation delay of 10
μs using end-to-end bursting is ± 2.48 Mb/s.

The goodput for end-to-end bursting stays near the 802.11g maximum of 54

Table 5.2. FTP over TCP with 1 ms BTO

Parameters	Values
burst mechanism	edge vs. end-to-end bursting
access network mode	infrastructure vs. ad hoc mode
burst size	64 KB
TCP MTU size	1 KB (edge) or 64 KB (end-to-end)
burst timeout	1 ms
application / transport layer	FTP over TCP
optical network delay	1 μs to 100 ms

Mb/s for propagation delays from 1 μs to 1ms, beyond which it drops to reaching 34 Mb/s for a delay of 10ms. This shows that the only high propagation delays affect the goodput, and is not a concern for LANs and MANs. In the case of the edge bursting, the goodput is high when the delay is small and falls off as we approach MAN latencies due to the delay incurred in queuing and burst assembly at the ingress edge.

Table 5.3. FTP over TCP with 10 ms BTO

Parameters	Configurations
burst mechanism	edge vs. end-to-end bursting
access network mode	infrastructure vs. ad hoc mode
burst size	64 KB
TCP MTU size	1 KB (edge) or 64 KB (end-to-end)
burst timeout	10 ms
application / transport layer	FTP over TCP
optical network delay	1 μs to 100 ms

When the values of BTO (burst timeout) is increased to 10 ms and 100 ms as shown in Tables 5.3 and 5.4. We see in Figures 5.3 and 5.4 that TCP cannot
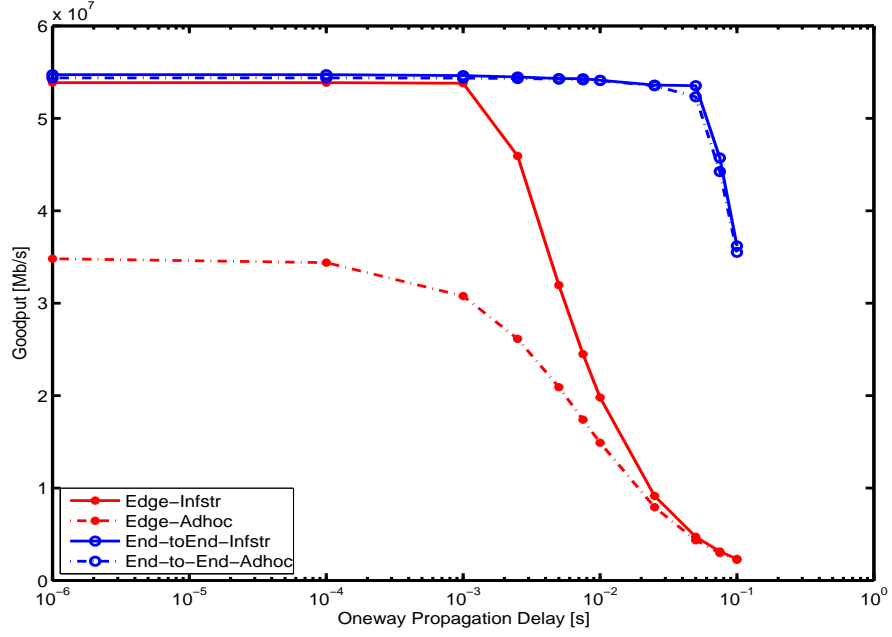
51

Figure 5.2. TCP goodput with BTO = 1 ms

drive the network to perform better with edge bursting, since more segments must be aggregated before the timer timesout. This is due to the delay caused by the ACKs in the reverse path. In the case of end-to-end bursting the performance is uniformly higher due to the ability of TCP to fill the pipe as discussed in Section 5.4. The impact of BTO is largely seen on the goodput performance until the delay of 10^{-2}. Above which the high optical delay at the OBS core reduces goodput. In end-to-end bursting, goodput reaches maximum, since the TCP is not dependent on the BTO at the OBS edge node.
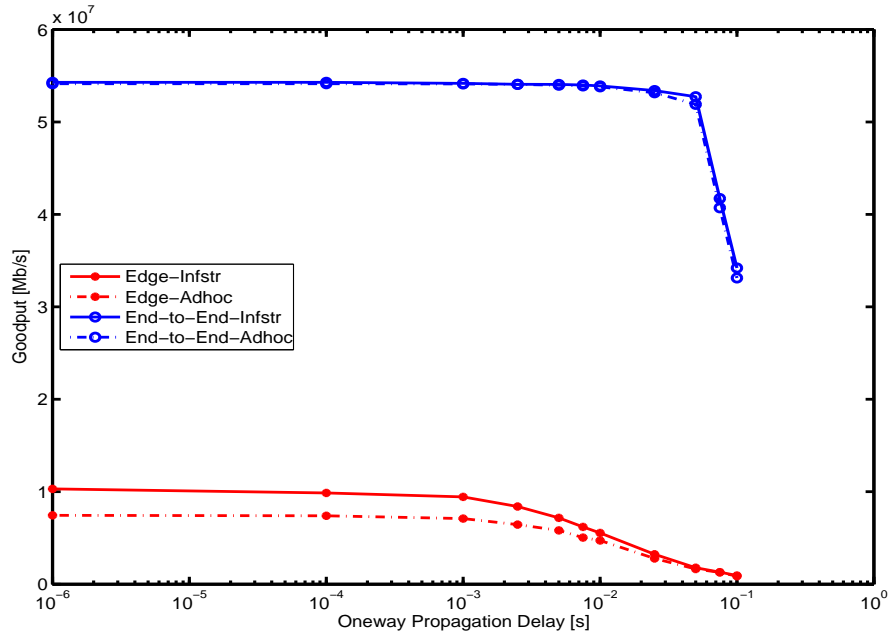
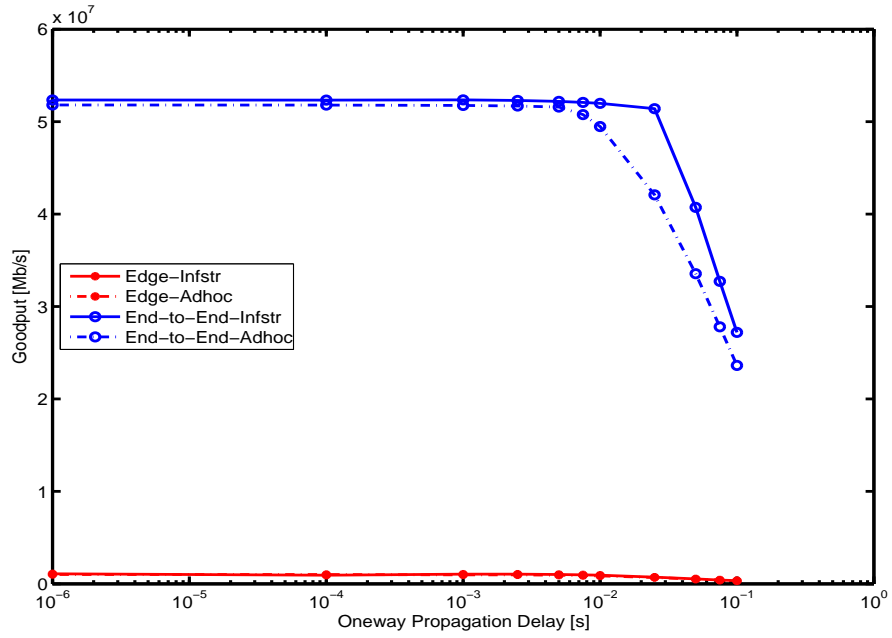**Figure 5.3.** TCP goodput with BTO = 10 ms

Figure 5.4. TCP goodput with BTO = 100 ms

Table 5.4. FTP over TCP with 100ms BTO

Parameters	Configurations
burst mechanism	edge vs. end-to-end bursting
access network mode	infrastructure vs. ad hoc mode
burst size	64 KB
TCP MTU size	1 KB (edge) or 64 KB (end-to-end)
burst timeout	100 ms
application / transport layer	FTP over TCP
optical network delay	1 μs to 100 ms

To understand the relation beteen burst size and timeout with the edge bursting mechanism, the burst size is varied keeping the BTO a constant and the performance of a single TCP connection is analyzed as shown in the Table 5.5.

Table 5.5. Goodput for different burst sizes over varying BTO

Parameters	Configurations
burst mechanism	edge bursting
access network mode	infrastructure mode
burst size	2 to 64 KB
burst timeout	1 ms, 10 ms
application / transport layer	FTP over TCP
optical network delay	1 μs to 100 ms

In Figure 5.5, the goodput is measured with 1 ms to 100 ms BTO and plotted over varying burst sizes from 1 KB to 64 KB. Goodput of 54 Mb/s is obtained when minimal or no aggregation occurs at the edge node. When the BTO is increased to 100 ms, traffic at the ingress edge is insufficient to generate bursts within one BTO resulting in poor goodput. To analyze better the tradeoff, intermediate BTO

54

values of 5 ms and 10 ms are plotted, these curves show an unexpected dip at burst sizes from 8 to 24 KB. For low burst sizes, goodput depends on burst size. When the burst size is high, the bursts are triggered by BTO expiration.
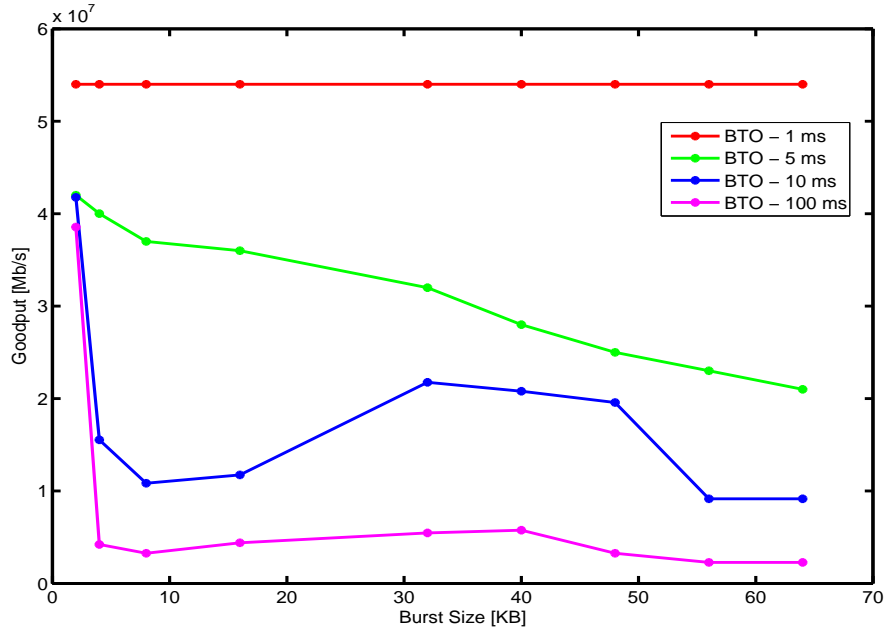

Figure 5.5.   Burst timer and burst size tradeoff
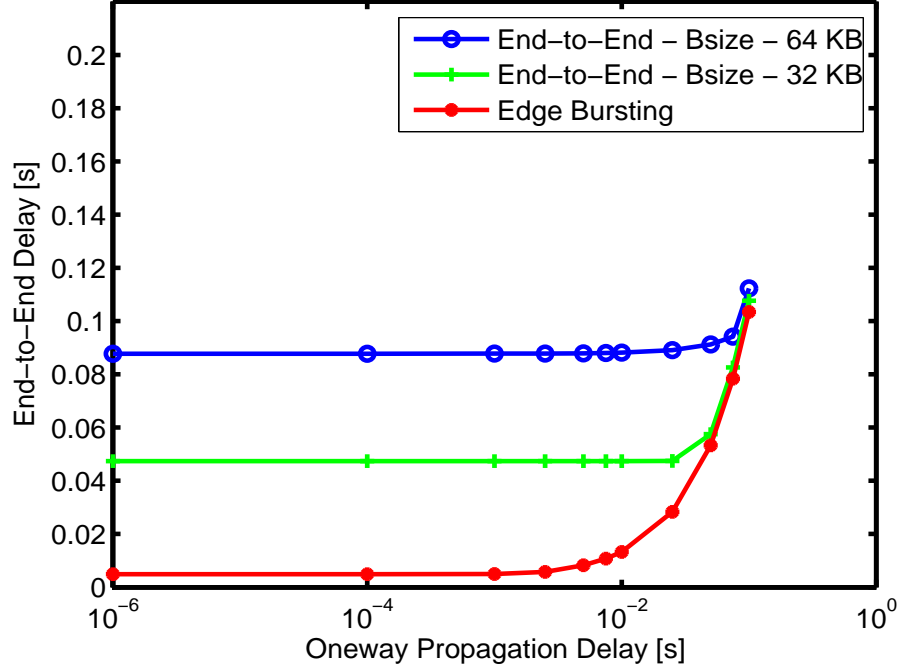
5.3 Delay Analysis

Figure 5.6.   End-to-end delay plot BTO = 1 ms

To further understand the performance of different bursting mechanism, end-to-end delay is plotted for the simulation scenarios, averaged over all packets for every simulation. Figure 5.6 shows the delay plot for both the bursting mechanisms. The plot shows that the average delay of edge bursting is less than the delay for end-to-end bursting, which is due to the fact that the plots are not normalized for packet sizes. The edge bursts shows delay for a TCP-MTU packet size of 1 KB whereas the end-to-end bursting delay is for TCP-MTU of 32 to 64 KB. The delay increases with the core network propagation delay, across the plot.

From the plots of end-to-end delay the performance of end-to-end bursting is nearly similar for all values of BTO. Even though we transmit larger bursts
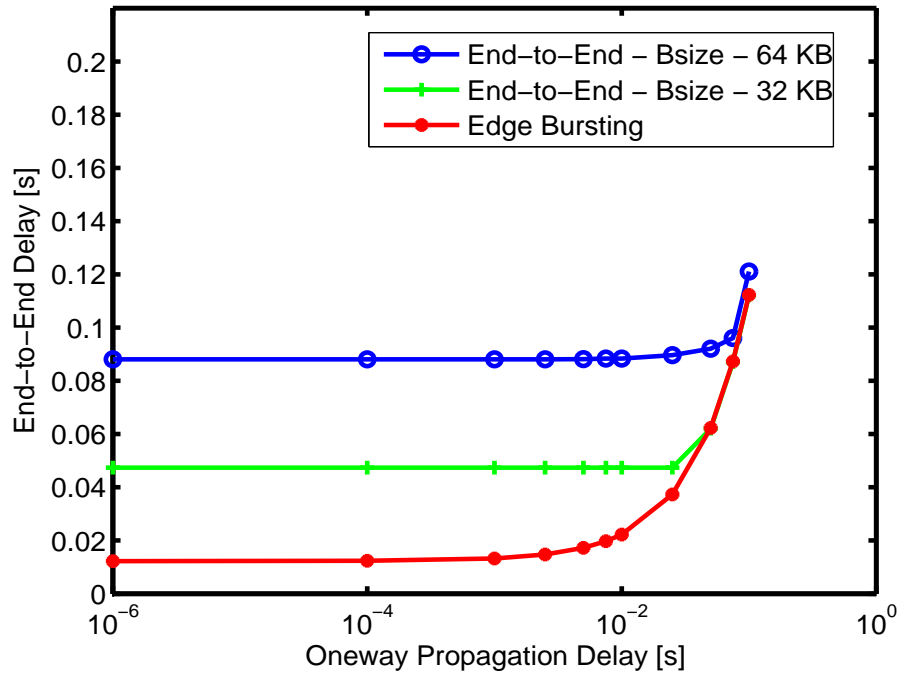
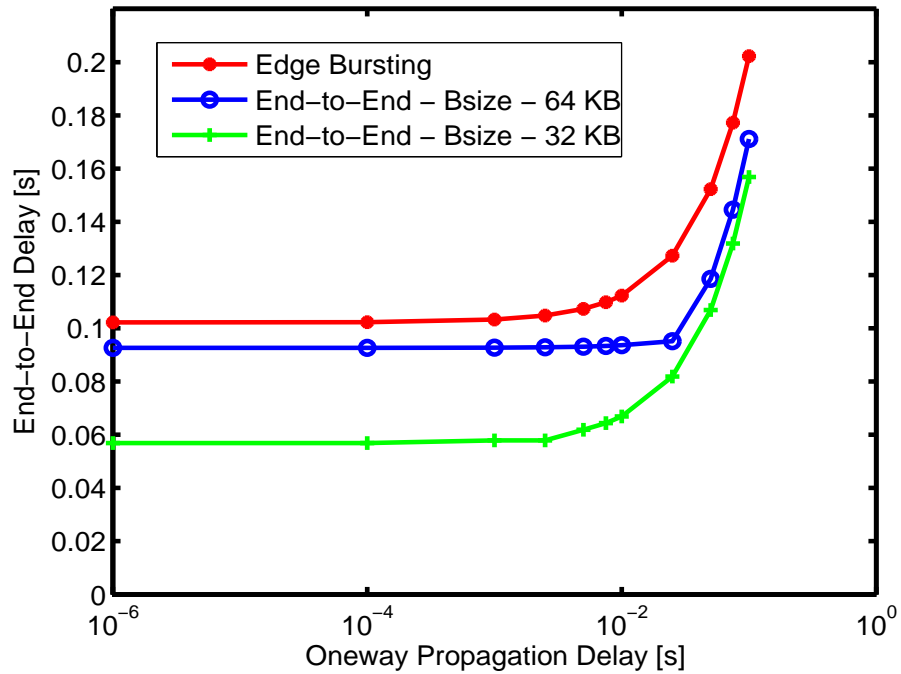**Figure 5.7.** End-to-end delay plot BTO = 10 ms

Figure 5.8. End-to-end delay plot BTO = 100 ms

57

end-to-end, the delay is higher when compared to transmitting packets with edge bursting. Therefore, a trade off can be made in choosing the burst size. The plots show that the delay for a 32 KB MTU is half the delay for 64 KB for 1 ms LANs and 10 ms MANs.
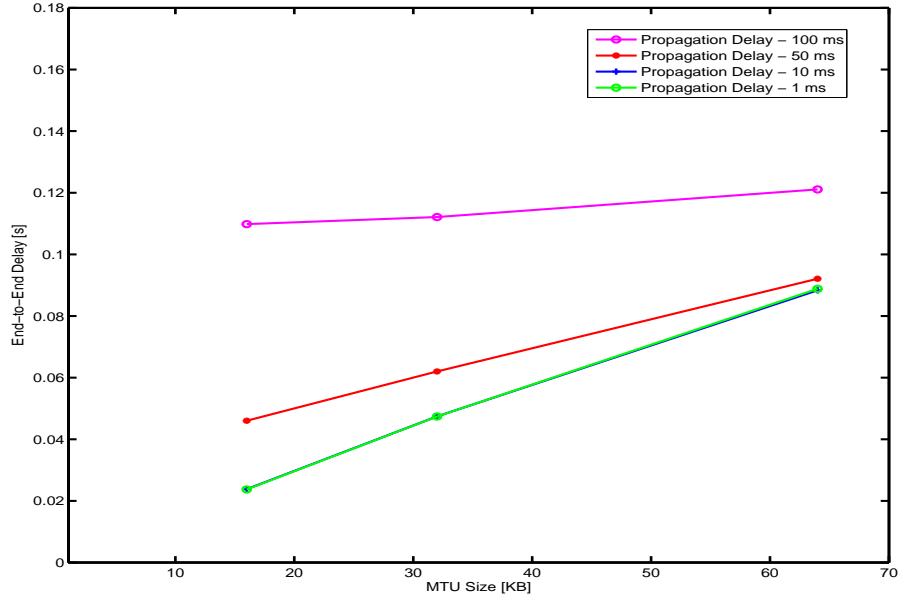

Figure 5.9. Delay over packet size variation with BTO = 1 ms

As the BTO is increased to 10 ms the end-to-end delay for edge bursting increases, shown in Figure 5.7, reducing the network goodput seen in the last section. But the reduction in goodput and the increase in the delay are not proportional, showing the poor performance of the edge bursting causing TCP to reduce its rate due to the delay incurred by the burst assembly process at the ingress edge. Figure 5.8 shows the end-to-end delay with a BTO of 100 ms. The higher BTO is reflected in higher delay curves, as expected.

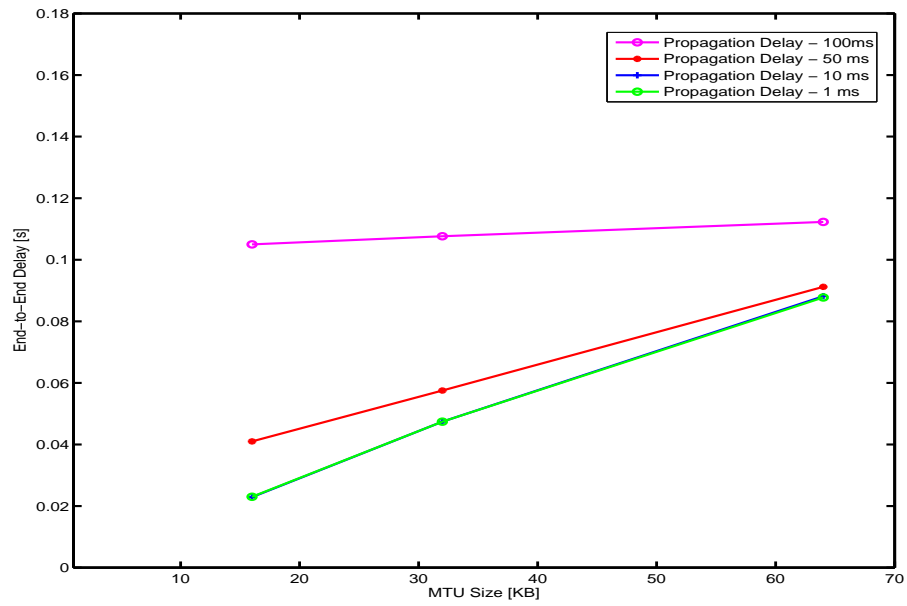Figure 5.9 shows the variation of end-to-end delay with change in burst size.

Figure 5.10. Delay over packet size variation with BTO = 10 ms
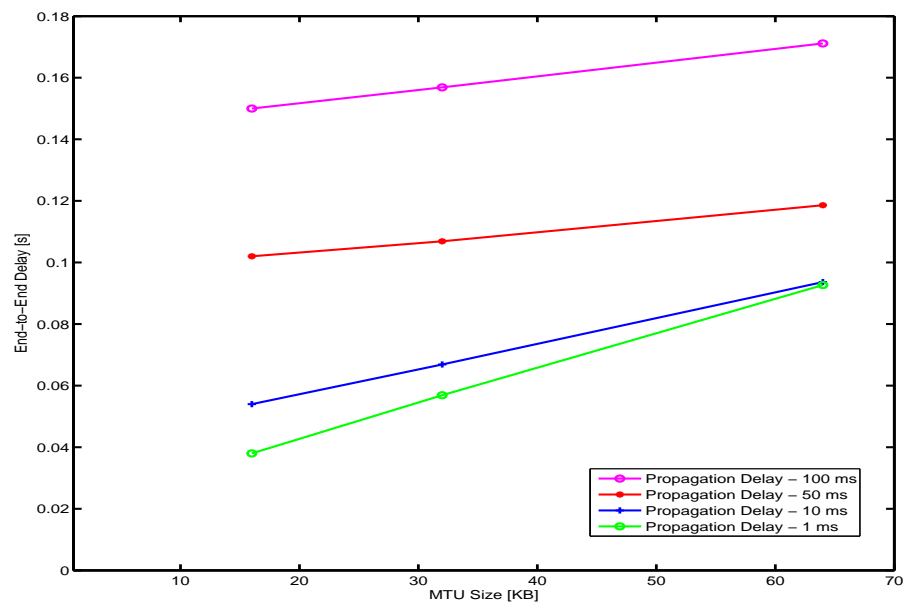

Figure 5.11. Delay over packet size variation with BTO = 100 ms

The BTO is kept constant at 1 ms and the variation of the end-to-end delay is plotted with increasing burst (MTU) size by varying the propagation delay at the optical core network. It is seen that the burst size significantly affects end-to-end delay, when the propagation delay is low, and has less impact as the propagation delay increases. The plot further shows a linear increase in end-to-end delay with burst size in the range plotted.

Similar relationships can be seen in Figures 5.10 and 5.11, where the BTO is increased by one order of magnitude to 10 ms and subsequently to 100 ms.

5.4 TCP Performance Analysis

To better understand the performance of TCP over heterogeneous networks a simulation run was made with two nodes connected directly with a 1 Gb/s wired link. Figure 5.12 shows the effect on goodput with increasing propagation delay. The plot shows that TCP provides better performance when the delay in the network is low, and as the delay increases from 10^{-4} the curve falls down rapidly affecting the overall goodput as expected. This helps explain the reasons for the poor performance of edge bursting scenario when the end-to-end delay is high.

The sequence numbers are plotted for both bursting mechanisms to visualize the same behaviour seen in the two node TCP model. Figure 5.13 shows the sequence number over time for edge bursting for the three BTO values. This confirms that the effect of BTO on end-to-end performance is similar to the effect of propagation delay. The TCP sequence plot for the end-to-end bursting for the three BTO values is shown in Figure 5.14. The plot shows that the model can achieve better goodput even when the delay is high than edge bursting.
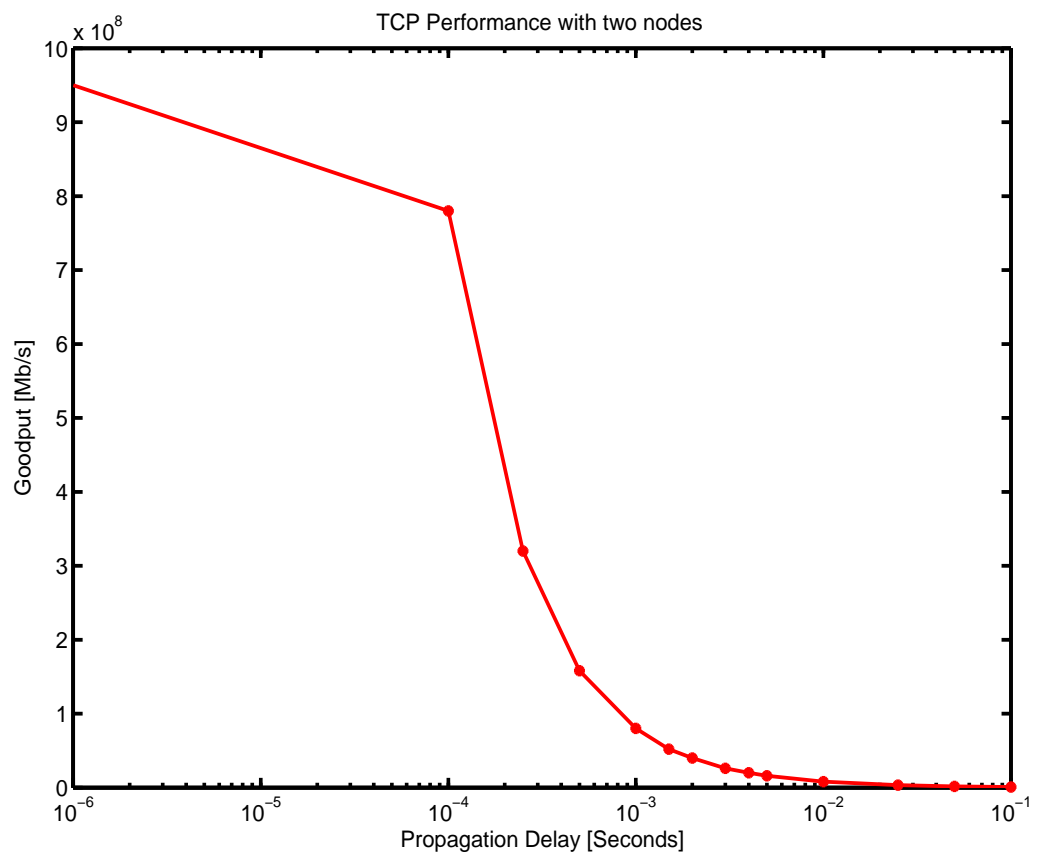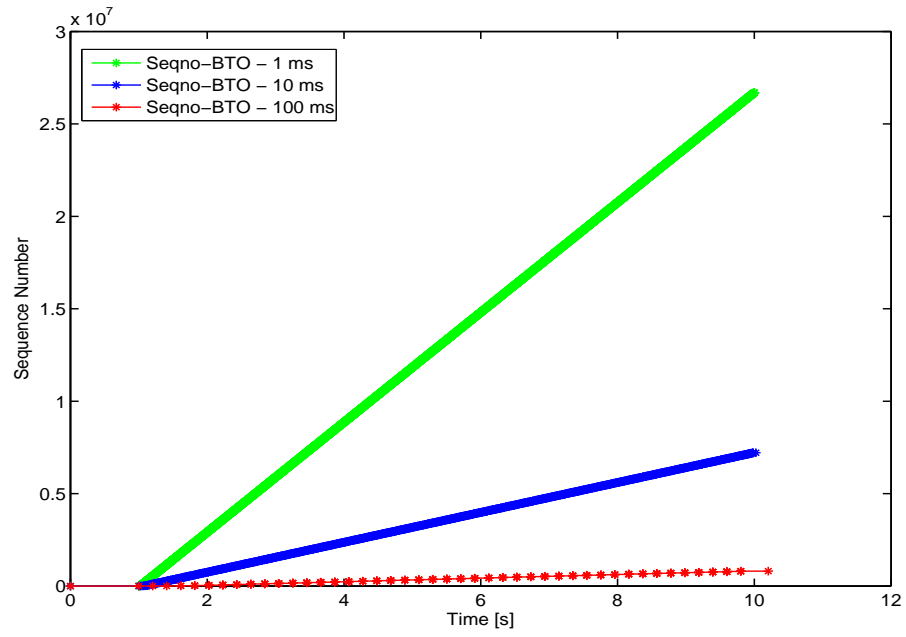
Figure 5.12. TCP performance over 1 Gb/s link

Figure 5.13. Sequence plot for TCP with OBS edge bursting
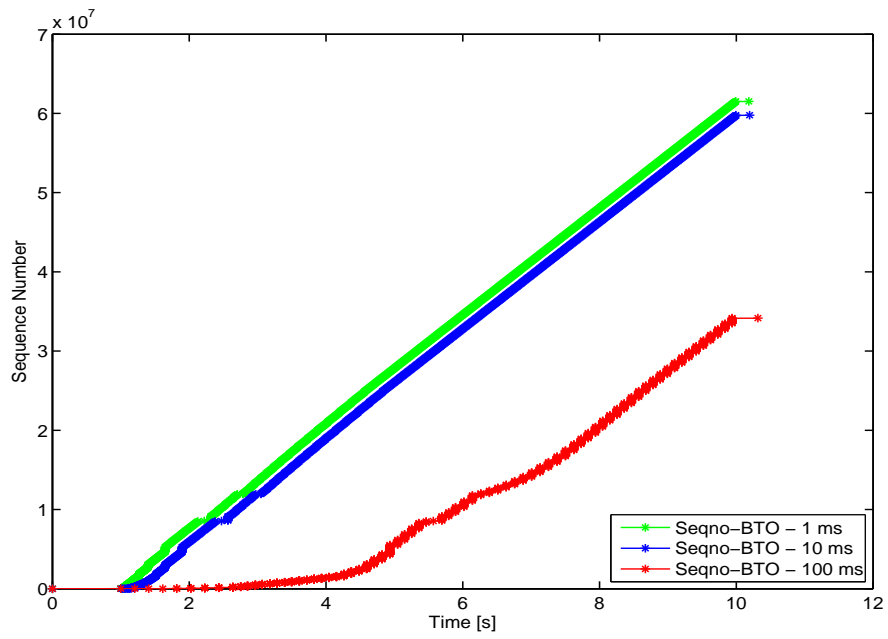

Figure 5.14. Sequence plot for TCP with end-to-end bursting

5.5 Goodput with Wireless Access BER Variation

To obtain the effects of wireless access channel impairments on end-to-end bursting performance, plots are generated for goodput by varying wireless link BER as shown in Table 5.6. The network performance for end-to-end bursting is obtained with different TCP-MTU sizes from 1 KB to 64 KB shown in Figure 5.15. From the plot it can be understood that the goodput of 64 KB burst is good

Table 5.6. Goodput with access network BER variation

Parameter	Configuration
burst mechanism	end-to-end bursting
access network mode	infrastructure
TCP MTU size	1 to 64 KB
burst timeout	1 ms
application / transport layer	FTP over TCP
optical network delay	1 μs to 100 ms
bit error rate	10^{-8} to 10^{-1}

until the BER is 10^{-7} above which it falls down rapidly to zero at 10^{-4}. When the channel conditions are lossy. At BER value of 10^{-6} both 32 KB and 16 KB provide an approximately equal goodput, indicating that some efficiency is achievable.
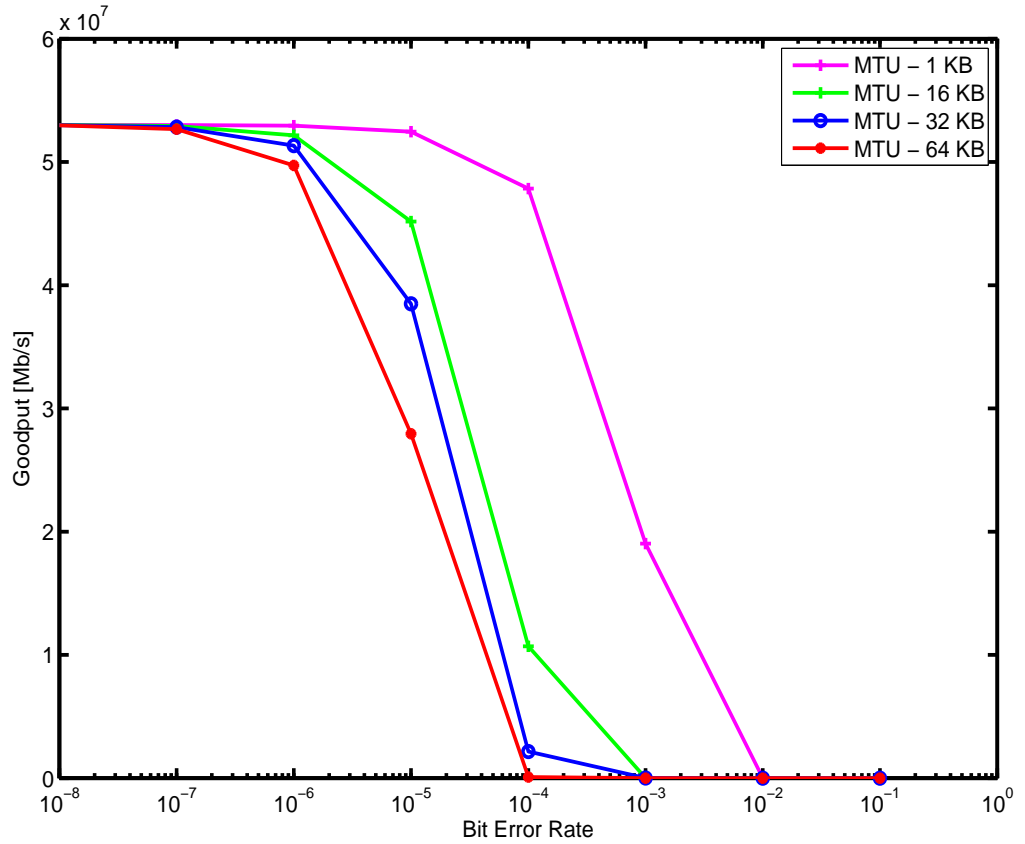
Figure 5.15. TCP performance over access network BER variation

5.6 Summary of TCP performance

Several important observations can be made from the results: TCP performance depends mainly on the timers that are synchronized for smooth transmissions and any delay like ACK (acknowledgement) getting queued on edge node with a high burst timeout value may result in poor TCP performance. The burst size and timeouts should be configured depending on the traffic in the network otherwise it might impact the entire design of the network and its performance. The TCP traffic sent from a wireless node gets delayed more than the RTT (round

trip time), the CWND drops down to a minimum value resulting in poor transmission rates. Bursting at the edge node for heterogeneous scenarios should be done cautiously, taking into account the traffic at the ingress nodes at all ends of the network. Bursting end-to-end provides a higher goodput, but issues like fairness and contention at the node, may reduce the effective goodput.

Chapter 6

Conclusions and Future Work

This chapter discusses the conclusions obtained from the simulation models in Section 6.1. Section 6.2 summarizes the contributions of this thesis. Section 6.3 provides possible future directions.

6.1 Conclusions

Th aim of this thesis was to investigate and to comprehend the effects of TCP over heterogeneous networks with optical burst switched backbone and wireless access networks. It presents a unified end-to-end communication paradigm that effectively utilizes the network and frames data packets according to the application layer requirement. The main contributions of the thesis include design and development of heterogeneous realm model with end-to-end burst mechanism, and to validate their performance using ns-2 simulation results. TCP performance is improved with this method and is better utilized under different scenarios as discussed in the previous chapter. The proposed end-to-end solution would enable higher goodput across the network at low traffic arrival rates and subsequently

better performance at higher rates.

6.2 Research Contributions

A unified end-to-end communication model is implemented with transport layer at the end systems determine the end-to-end data transmission characteristics over wireless and optical realms. The end-to-end bursting model to match application layer requirements is simulated using, a multi-realm with a unified end-to-end communication paradigm. TCP performance characteristics are studied over heterogeneous wireless and optical realms under different bursting schemes.

6.3 Future Work

The thesis will act as a starting point of architecture design for future work of the Resilinets group at The University of Kansas, which includes developing PoMo Architecture for the NSF-Find project. It involves development of optical wireless realm models, incorporating *knobs* and *dials* for providing application driven data framing at the ends with cross-layer optimizations for forward error correction (FEC), and implementing PoMo transport protocol for inter-realm modeling. The results provided in this thesis can further be extended to analyze contention at the relaying nodes, when the packet size is big and, also issues related to 802.11 coordination functions, that could be altered to transmit bigger packets at the wireless end system.

Bibliography

[1] Georgios I. Papadimitriou, Chrisoula Papazoglou, and Andreas S. Pomportsis. *"Optical Circuit and Burst Switching"*. Wiley Series in Microwave and Optical Engineering, Hoboken, New Jersey, 2007.

[2] J. Postel. "Transmission Control Protocol (TCP)-RFC 793". 1981.

[3] C. Barakat, E. Altman, and W. Dabbous. "On TCP Performance in a Heterogeneous Network: a Survey". *IEEE Communications Magazine*, 38(1):40–46, 2000.

[4] W.Y. Ma, I. Bedner, G. Chang, A. Kuchinsky, and H.J. Zhang. "Framework for Adaptive Content Delivery in Heterogeneous Network Environments". *PROC SPIE INT SOC OPT ENG*, 3969:86–100, 2000.

[5] E.A. Brewer, R.H. Katz, Y. Chawathe, S.D. Gribble, T. Hodes, G. Nguyen, M. Stemm, T. Henderson, E. Amir, H. Balakrishnan, et al. "A Network Architecture for Heterogeneous Mobile Computing". *IEEE Personal Communications*, 5(5):8–24, 1998.

[6] A.A. Khokhar, V.K. Prasanna, M.E. Shaaban, and C.L. Wang. "Heterogeneous Computing: Challenges and Opportunities". *Computer*, 26(6):18–27, 1993.

[7] "Bobby Bhattacharjee and Ken Calvert and Jim Griffioen and Neil Spring and James P.G. Sterbenz". Postmodern internetwork architecture. Technical Report ITTC-FY2006-TR-45030-01, Information Telecommunication and Technology Center, University of Kansas, Lawrence, KS, 2006.

[8] B.P. Crow, I. Widjaja, L.G. Kim, and P.T. Sakai. "IEEE 802.11 Wireless Local Area Networks". *Communications Magazine, IEEE*, 35(9):116–126, 1997.

[9] T.I.N.B. Access. "IEEE Standard 802.16: A Technical Overview of the WirelessMAN Air Interface for Broadband Wireless Access". *IEEE Communications Magazine*, page 98, 2002.

[10] J. Postel. "RFC 791: Internet Protocol", 1981.

[11] J.B. Andersen, T.S. Rappaport, and S. Yoshida. "Propagation Measurements and Models for Wireless Communication Channels". *Communications Magazine, IEEE*, 33(1):42–49, 1995.

[12] K. Chen, Y. Xue, S.H. Shah, and K. Nahrstedt. "Understanding Bandwidth-Delay Product in Mobile Ad hoc Networks". *Computer Communications*, 27(10):923–934, 2004.

[13] Tzvetelina (Lina) Battestilli. "Optical Burst Switching: A Survey".

[14] S. Oh, Y. Kim, M. Yoo, and H.H. Hong. "Survivability in the Optical Internet Using the Optical Burst Switch". *ETRI Journal*, 24(2), 2002.

[15] X. Yu, C. Qiao, and Y. Liu. "TCP Implementations and False Time Out Detection in OBS Networks". *INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies*, 2, 2004.

[16] Will E. Leland, Murad S. Taqq, Walter Willinger, and Daniel V. Wilson. "On the Self-Similar Nature of Ethernet Traffic". In Deepinder P. Sidhu, editor, *ACM SIGCOMM*, pages 183–193, San Francisco, California, 1993.

[17] C. Qiao and M. Yoo. "Choices, Features and Issues in Optical Burst Switching". *Optical Networks Magazine*, 1(2):36–44, 2000.

[18] L. Xu, H.G. Perros, and G. Rouskas. "Techniques for Optical Packet Switching and Optical Burst Switching". *Communications Magazine, IEEE*, 39(1):136–142, 2001.

[19] J. Postel. "User Datagram Protocol (UDP), IETF Standard, RFC 768". August, 1980.

[20] S. Floyd and K. Fall. "Promoting the use of End-to-End Congestion Control in the Internet". *IEEE/ACM Transactions on Networking (TON)*, 7(4):458–472, 1999.

[21] M. Allman and W. Stevens. "TCP Congestion Control". 1999.

[22] M. Allman, V. Paxson, and W. Stevens. "RFC2581: TCP Congestion Control". *Internet RFCs*, 1999.

[23] S. Gowda, RK Shenai, KM Sivalingam, and HC Cankaya. "Performance Evaluation of TCP Over Optical Burst Switched (OBS) WDM Networks". *Communications, 2003. ICC'03. IEEE International Conference on*, 2, 2003.

[24] O. González, I. de Miguel, N. Merayo, P. Fernández, R.M. Lorenzo, and E.J. Abril. "The Impact of Delayed ACK in TCP Flows in OBS Networks". *Proceedings of NOC 2005*, pages 367–374, 2005.

[25] M. Casoni, E. Luppi, and M. Merani. "Impact of Assembly Algorithms on End-to-End Performance in Optical Burst Switched Networks with Different QoS Classes". *Proc. IEEE/SPIE 3rd Workshop Optical Burst Switching.*

[26] Optical Burst Switched. "Analysis of a Dynamically Wavelength-Routed".

[27] C. Qiao and M. Yoo. "Optical burst switching (OBS) - a new paradigm for an Optical Internet", 1999.

[28] X. Cao, J. Li, Y. Chen, and C. Qiao. "Assembling TCP/IP Packets in Optical Burst Switched Networks". *Global Telecommunications Conference, 2002. GLOBECOM'02. IEEE*, 3, 2002.

[29] E.H.K. Wu and M.Z. Chen. "JTCP: Jitter-based TCP for Heterogeneous Wireless Networks". *Selected Areas in Communications, IEEE Journal on*, 22(4):757–766, 2004.

[30] Gavin Holland and Nitin H. Vaidya. "Analysis of TCP Performance over Mobile Ad Hoc Networks". In *Mobile Computing and Networking*, pages 219–230, 1999.

[31] I. Martinez-Yelmo, I. Soto, D. Larrabeiti, and C. Guerrero. "A Simulation-Based Study of TCP Performance over an Optical Burst Switched Backbone with 802.11 Access". *Lecture Notes in Computer Science*, 4606:120, 2007.

[32] D.D. Clark and D.L. Tennenhouse. "Architectural Considerations for a New Generation of Protocols".

[33] M. Mathis and J. Heffner. "Packetization Layer Path MTU Discovery". *Work in Progress, June*, 2004.

[34] C.E. Perkins and P. Bhagwat. "Highly Dynamic Destination-Sequenced Distance-Vector Routing (DSDV) for Mobile Computers". *Proceedings of the conference on Communications architectures, protocols and applications*, pages 234–244, 1994.

[35] J. Padhye, V. Firoiu, DF Towsley, and JF Kurose. "Modeling TCP Reno Performance: A Simple Model and its Empirical Validation". *Networking, IEEE/ACM Transactions on*, 8(2):133–145, 2000.

[36] DARPA-VINT UC Berkeley LBL USC/ISI and Xerox PARC. "Network Simulator Version-2 (ns-2)", 1995.

[37] B. Wen, N.M. Bhide, R.K. Shenai, and K.M. Sivalingam. "Optical Wavelength Division Multiplexing (WDM) Network Simulator (OWns): Architecture and Performance Studies". *SPIE Optical Networks Magazine*, pages 16–26, 2001.

[38] B.S. Bakshi, P. Krishna, N.H. Vaidya, and D.K. Pradhan. "Improving Performance of TCP Over Wireless Networks". *17th International Conference on Distributed Computing Systems (ICDCS), May*, 1997.

[39] S. Avallone, A. Pescape, and G. Ventre. "Analysis and Experimentation of Internet Traffic Generator". *New2an*, 4:70–75.

[40] A. Romanov and S. Floyd. "The Dynamics of TCP Traffic over ATM Networks". *SIGCOMM94, ACM Computer Communication Review*.