A novel Burst assembly architecture for QoS provisioning in Optical Burst Switched multicast communications

Terrance Frederick Fernandez and T. Megala

Department of Computer Science and Engineering
Pondicherry Engineering College, Puducherry – 605 014, India

Abstract- In TCP over OBS communications, the existing architecture namely "Destination based Burst Aggregation" cannot guarantee the delivery of Real Time traffic as the packets belonging to different priorities are assembled into a single Data Burst. In this paper, a new architecture is proposed at the ingress of the OBS network called "Packet Priority Classifier and Burst Transcription based Burst Aggregation", in which packets with Real Time Constraints are aggregated into a Data Burst labeled as Delay Sensitive Bursts (DSBs) and the others assembled into Best Effort Bursts (BEBs).

Keywords: Optical Burst Switching (OBS), Multicast Routing, Burst Aggregation, TCP over OBS, Real Time traffic, guaranteed delivery.

1 INTRODUCTION

Internet is indispensible to our lives these days. Daybreak in most houses checking emails, downloading and listening to online radios. Driven by these human needs, the Internet grows exponentially and doubles its traffic volume every nine months before 1997 and every six months after that [1]. This is traffic doubling. The year 1990, hypertext and its supporting protocols made the World Wide Web, accessible to the public users. Introduction of optical fiber cables decreased the cost of transmission bandwidth and thus made Internet affordable to laymen. Optical fiber allows multiplexing many signals at different centre wavelengths onto a single fiber. This is called "Wavelength Division multiplexing (WDM)". Fine wavelength spacing of these signals is known as "Dense Wavelength Division Multiplexing (DWDM)". WDM features 8, 16 or 32 wavelengths on a single fiber. There are First Generation Optical Networks and Second Generation Optical Networks. In optical networks, at each switching point, the optical signal is converted, processed and forwarded as electrical signal, then modulated and converted as optical signal. This is known as an "Optical-Electronic (O/E/O) conversion". This is a feature of First Generation Optical Networks. Second Generation Optical Networks does not require the need of O/E/O conversion. It provides switching and routing services at the optical layer. These networks are known as "All optical networks (AON)". An all-optical path is called a "Light Path (LP)" [2]. The main attraction of AON is routing of optical data signals without the need for conversion to electrical signals and, so it is independent of data rate and data protocol [3].

AONs are classified along into three types to carry traffic across optical core:-

- Broadcast and select architecture (Single hop networks) – Access Networks
- The wavelength routing architecture (Optical Circuit Switching)
- The Photonic architecture (Optical Packet Switching) [4].

Optical networks which performs optical routing and traversal of multiple switches for typical connections are called "wavelength routing networks (WRN)". WRN is a form of OCS, where an end-to-end connection is established between source and destination hosts before data transfer [5]. In OCS, an LP is established that forms a logical path between edge nodes. In OPS, data is broken into smaller chunks called packets and are switched in all-optical form. Packets in these networks must carry fixed duration and network must be slotted [5].

1.1 Optical Burst Switching:

OBS is a promising candidate to efficiently utilize WDM and to fulfill demands in today's Internet traffic [1]. It is a hybrid switching that resembles wavelength routing and packet switching. An OBS network consists of Electronic Edge Nodes (EN) and Optical Core Nodes (CN). The input EN is called as "Ingress Node" and the output EN is called as "Egress Node". Data is broken into variable sized entities called bursts thus overcoming the demerit of the OPS architecture. A burst is defined as the Digitized talk spurt or the data message [2] that has intermediate granularity between a packet and amount of data in a circuit. A Burst consists of the payload called as Data Burst (DB) and a Burst Header Packet (BHP). DB is switched all-optically, while BHP undergoes O/E/O conversion at each hop. BHP precedes DB and is sent on a different wavelength with an "Offset Time (OT)". This OT allows the processing of the control packet and configuring switches until the arrival of data [2].

Buffers are absent at core nodes, headers can be processed at slower speeds, synchronization requirements are relaxed in OBS [2]. OCS is suitable for constant rate traffic (voice traffic) but unsuitable for dynamic traffic [2]. Faster header processing and strict synchronization are required in OPS due to lack of optical buffers. Only company that offers commercial OBS products is “Matisse networks” as the technology is still immature [6] and thus good for research. OBS can cut-through optical switches i.e.., the burst may be transmitted before it is completely assembled or dissembled at intermediate nodes. The switching speeds of electronics cannot keep up with the transmission capacity offered by optics [6]. Also optical switches handles large number of switching ports when compared with electronic switches.

2 BACKGROUND

2.1 Burst Contention and Dropping

Burst Contention is a phenomenon that occurs in OBS when two or more incoming bursts contend for the same output wavelength at the same link and at the same time [6]. When contentions occur between bursts, one of them successfully reserves the channel and the remaining drops. This phenomenon is called as Burst Dropping. This occurs due to the absence of Optical buffers at the OBS Network as the DBs traverse an all-optical path.

2.2 Burst Assembly:

Data packets from various traffic sources are assembled into a variable length entities called Bursts. The DB is assembled at the Ingress node before they are sent to the network. This assembly is made based on some assembly algorithms. They are Timer based Burst Assembly Mechanism, Burst length (threshold) based Burst Assembly Mechanism, Hybrid/Mixed Burst Assembly Mechanism and Optimized/Adaptive Burst Assembly Mechanism.

A burst is created and sent into ON at periodic time intervals and Bursts are of variable lengths for Timer based Burst Assembly Mechanism [2]. The burst lengths are fixed and are generated at non-periodic time intervals for Burst length based Burst Assembly Mechanism [2]. Combine packets of different classes to same bursts. Packets are placed from head of the burst to the tail of the bursts in order of decreasing classes. This is called as composite burst assembly algorithm [2]. In mixed burst assembly mechanism, the bursts are assembled and sent either if the timer expires or the burst length is reached [7]. Dynamic adaptive threshold on burst length is set in order to optimize the overall performance in OBS for QoS sensitive traffic [7].

2.3 Traffic distribution algorithms:

Traffic distribution in OBS can be done at the Burst level [3] by making a path selection for each newly incoming burst at ingress node according to a calculated distribution probability within a time period. It can also be made at the Flow level [5] by making a path selection for each newly incoming flow at ingress node according to a calculated distribution probability within a time period and all bursts within a particular flow will follow the same path. Balance the traffic load by shifting incoming bursts along the primary and secondary paths. Probe the sent packets periodically and sent through the least congested path [5]. Always aim to minimize Burst loss and average transfer delay [5].

- Equal Proportion Multipath Routing (EPMR).
- Hop Length Multipath Routing (HLMR).
- Adaptive Alternative Routing (AAR).
- Gradient Projection Multipath Routing (GPMR).

2.4 Scheduling for OBS:

The various scheduling algorithms in OBS include:

- Horizon Scheduling.
- Latest available unused channel with Void filling (LAUC-VF).
- Minimum Starting Void (MSV).
- Constant Time Burst Resequencing (CTBR)

The latest time at which a channel is currently scheduled to use is called as a “horizon”. With the above the horizon scheduler selects the channel with latest horizon from the set of channels whose horizons are less than burst arrival times. The time complexity here is O (log h) [3].LAUC-VF keeps track of all voids in a channel and schedule bursts in one of the voids. If more than one void can fit a burst then the one with latest beginning time is assigned. The link utilization is higher than horizon scheduling but gets slower with huge number of voids. The time complexity here is O (log m), where ‘m’ is no. of voids [5]. MSV uses a geometric approach (binary search tree) that minimizes the distance between the starting time of the void and starting time of the burst. The time complexity is O (log m), where ‘m’ is no. of voids [5]. CTBR has the time complexity of O (1) called “Optimum Wavelength Scheduler”. Instead to process burst as soon as the BHP arrive, we delay scheduling of the bursts and process in order of expected Burst arrival time. BHP is processed not in its arrival times but at the arrival times of Data bursts [5].

2.5 Multicast Communication Paradigm

Routing refers to the process of selecting communication paths in a network to send network traffic. The routing consists of various delivery semantics that govern the delivery of data. They are:

- Unicast
- Broadcast
- Multicast
- Anycast
- Geocast

In unicasting, messages from source are forwarded to a single specific destination. Broadcast delivers messages to all nodes in the network. Anycast delivers messages to any one out of a group of nodes, typically the one nearest to the source. Geocast delivers messages to a geographical area [8]. In multicasting, messages from source are forwarded to a group of destinations called multicast group. The cardinality of destinations in a multicast group is referred to as the group size. Multicast tree corresponds to a communication session established between source and fixed number of destinations which are already given [9].

2.6 Network Traffic

Network Traffic refers to the data in the network. Some on them have specific time constraints while others don't. Based on this they can be classified as:

- Sensitive Traffic
- Best-Effort traffic

Sensitive traffic is traffic that is to be delivered on time. This includes VoIP, online gaming, video conferencing, and web browsing. Shaping schemes are generally tailored in such a way that the quality of service of these selected uses is guaranteed, or at least prioritized over other classes of traffic. This can be accomplished by the absence of shaping for this traffic class, or by prioritizing sensitive traffic above other classes. Best effort traffic is all other kinds of Non-Real Time traffic. This is traffic that isn't sensitive to Quality of Service metrics. A typical example would be peer-to-peer and email applications. Traffic management schemes are generally tailored so best-effort traffic gets what is left after sensitive traffic.

3 NETWORK SCENARIO

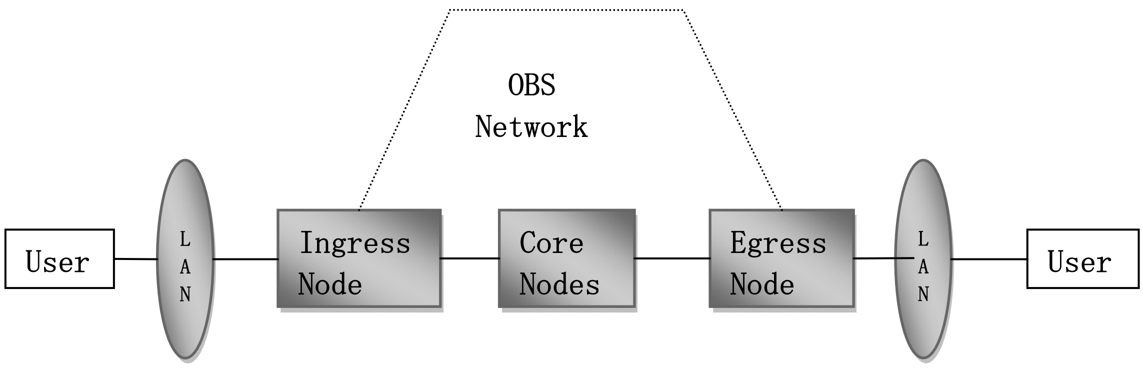

Fig 1: TCP/OBS Network

Figure 1 show OBS network architecture. OBS networks consist of edge nodes and core nodes connected by WDM links. The input traffic from multiple senders is assembled at the source edge node and is transmitted as bursts through high-capacity WDM links over the optical core. A destination edge node, upon receiving a burst, disassembles and delivers the data packets to the corresponding receivers. Source edge nodes are called as ingress nodes and destination edge nodes are called as egress nodes.

In OBS networks, data is transmitted as Data Bursts. Each burst consists of multiple packets (IP packets for an IP over WDM network or ATM packet), and a Burst Header Packet (BHP) is generated for each payload Burst. BHP contains the routing information to be used by core routers, the burst offset time, the data burst length, the data wavelength carrying the burst, and the QoS information. A burst and its BHP are transmitted separately on different channels. In a burst transmission scenario, a BHP is sent out ahead of the burst by an offset time in order to reserve resources at each intermediate core node along the path of the burst. While the BHP is processed at each intermediate node, the burst cuts through the pre-configured nodes all-optically. The offset time must ensure that, at each intermediate node, the BHP is processed prior to the burst arrival. Hence, the delay burst at each intermediate node is not required.

4 EXISTING BURST AGGREGATION

The existing aggregation mechanism is Destination based in which data from different TCP sources and belonging to a same destination are aggregated and are assembled into a single Burst by one of the assembly mechanisms, i.e.., timer based or threshold based or both. The assembled bursts are fed into the Scheduler. This Scheduler selects the channel on which the DBs are to be sent. The Control Header for a DB is created in the BHP module and they are sent. The DBs are sent along the Data Channel and the Control Header is sent along the Control Channel. In this existing Burst Aggregation mechanism, data packets with Real-time QoS constraints could be aggregated with non-critical ones and that's is a major reason to go for a new Burst Aggregation approach.

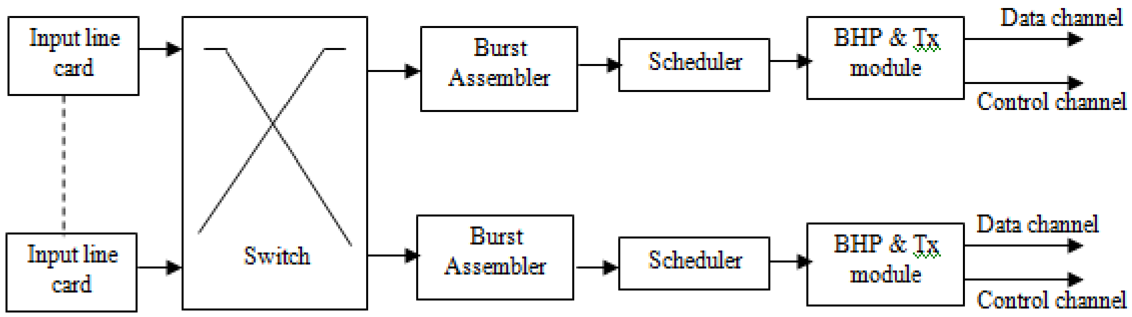

Fig2. Destination Based Burst Assembly Architecture

5 PROPOSED BURST AGGREGATION

In multicasting communications, contentions between Bursts are more frequent when compared to the unicast communications. These contentions results in dropping of the Bursts and hence the throughput of the entire network is affected. These dropping of Bursts at the core nodes and a further retransmission could affect the delivery of data at the receiver, if the Burst contains packet belonging to time critical applications.

To overcome the above demerit in the existing OBS architecture, a novel Burst aggregation approach is implemented at each ingress node and is done in two stages namely:

- Packet Priority Classification
- Burst Transcription

A Packet Priority Classifier (PPC) is a classifier that can receive incoming packets from different traffic sources such as IP, ATM or SONET networks and it could prioritize then as either High Priority Packets (HPP) or Low Priority Packets (LPP). Such a priority classifier is implemented at the first stage of the proposed burst aggregation algorithm. This is done in order to preempt high priority packets to the low priority ones as the HPP are usually the ones that are belonging to the delay sensitive applications. The low priority packets are those that do not have real time constraints and can be preempted by the HPP. The output of the priority classifier is fed to the assembler. Unlike the traditional approach, here in the proposed burst aggregation mechanism two assemblers are implemented and they are High Priority Assembler (HPA) and Low Priority Assembler (LPA). The typical Hybrid / Mixed Burst Assembly Mechanism is implemented at this ingress node for both assemblers. The HPPs are assembled at the HPA and they are labeled as Delay Sensitive Burst (DSB). The LPPs are assembled at the LPA and they are labeled as Best Effort Burst (BEB).

Second stage in proposed burst aggregation is called as Burst Transcription. In an OBS network Burst Contention is a common phenomenon thus leading to Burst Losses. So the ultimate aim must be to avoid contention among high priority burst or in other words to provide efficient reactive measure to tackle loss of Delay Sensitive Burst (DSB). So the output from the HPA is fed to the Burst Transcription stage where these burst are transcript or duplicated and they are called Duplicated Burst. No Optical Splitter is necessary for duplicating an optical burst [10]. Now the original burst is sent to the priority burst scheduler to be scheduled before sending it to the network. The Best Effort Burst bypasses this stage and directly fed to the priority burst scheduler.

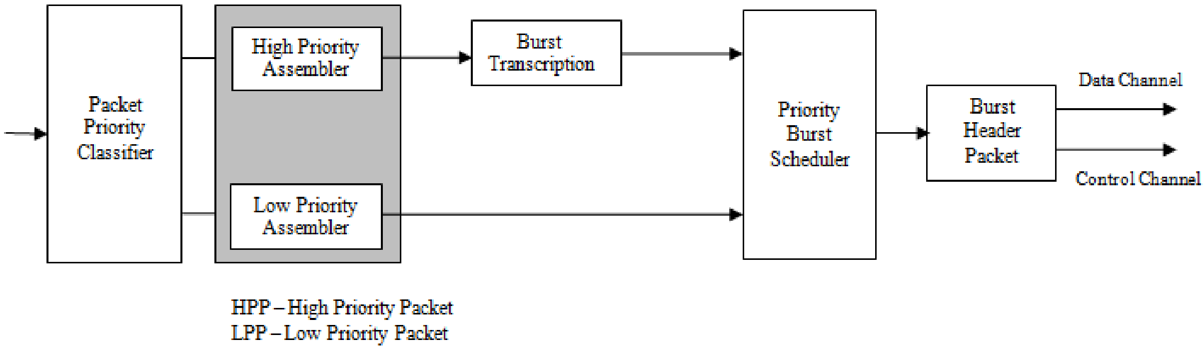

Fig 3: QoS provisioned Burst Assembly Architecture

In priority burst scheduler an adaptive scheduling algorithm is used that considers voids along the scheduled communication channels with minimal voids at the end. High preference is given to unscheduled channels. After the scheduling the bursts it is sent to the Burst Header Packet (BHP) module, where a burst header is created for each incoming burst. The burst header is sent along the control channel and Just Enough Time (JET) signaling mechanism is used to reserve the channel for the corresponding data burst that is sent along the data channel to the network.

At the core nodes, in the event of burst loss due to contention a Negative ACKnowledgements (NACK) is usually sent to the source. But here the NACK is sent to the ingress node. Here the burst is checked for its priority. If the Burst size is too small then it is automatically regarded as low priority, because fragmented bursts are deflected multiple times and they waste network Bandwidth. If the burst is a DSB then a duplicated copy of the burst is present at the Burst Transcription stage. This is transmitted to that core node where it is dropped. This could reduce the Transmission Delay of the burst and more importantly retransmissions from the source can be reduced. So that End-to-End delay is reduced.

6 SIMULATION RESULTS

The work is simulated in ns2 using nOBS [11] patch. Simulations are carried out on a 14 node NSFNET topology with 17 optical links. Source and the destination set are selected randomly and they follow uniform distribution. Real-Time Packet Loss probability is the ratio between the Real Time packets lost at the OBS cores to the Real Time packets transmitted into an OBS Network. Figure 4, shows that the average value of Real-Time Packet Loss probability for the proposal is 0.1 approx, while the existing mechanism carries a 0.5 probability for RT Packet losses. Real-Time Packet throughput is the amount of RT Packets transmitted per unit time. Figure 5, shows that the proposal yields more than twice (almost thrice) the throughput achieved through existing aggregation mechanisms. Latency is the total delay incurred by the burst during Burst transfer and will be the sum of transmission delay, propagation delay and queuing delay. Figure 6; show that the latency is minimized when the proposed burst aggregation is used.

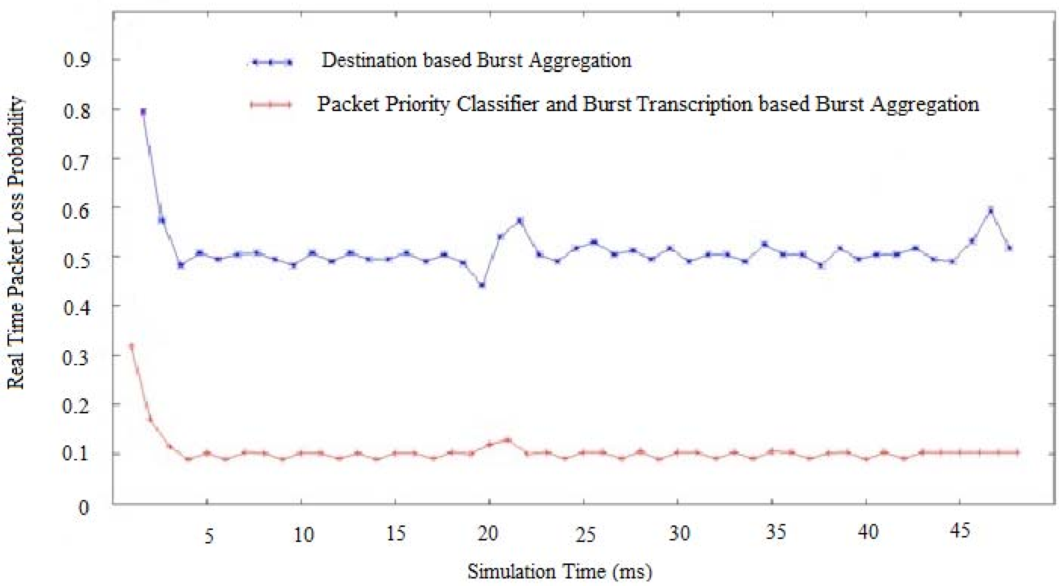

Fig 4: Real Time Packet Loss Probability vs. Simulation Time

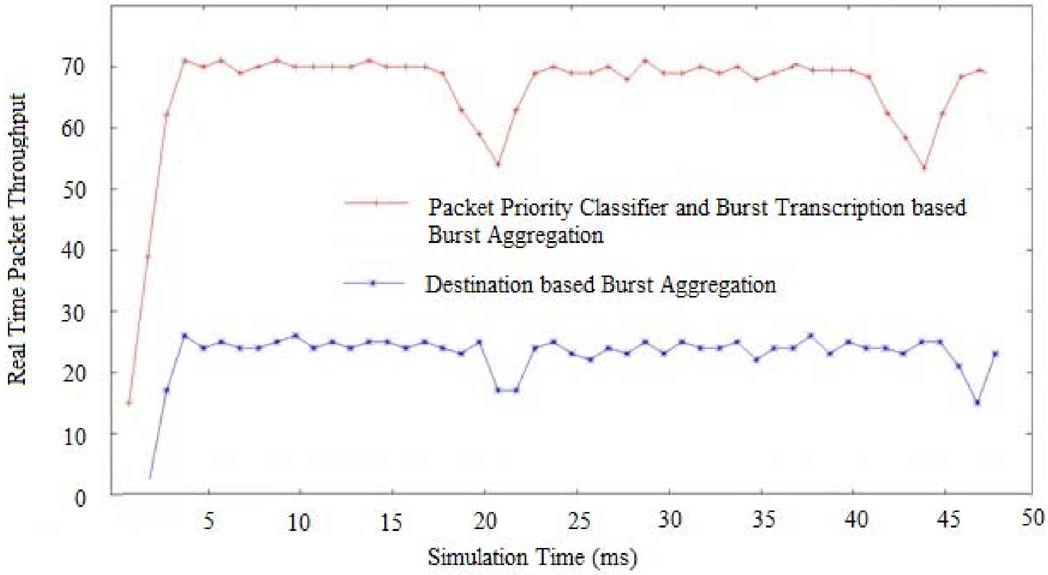

Fig 5: Real Time Packet Throughput vs. Simulation Time

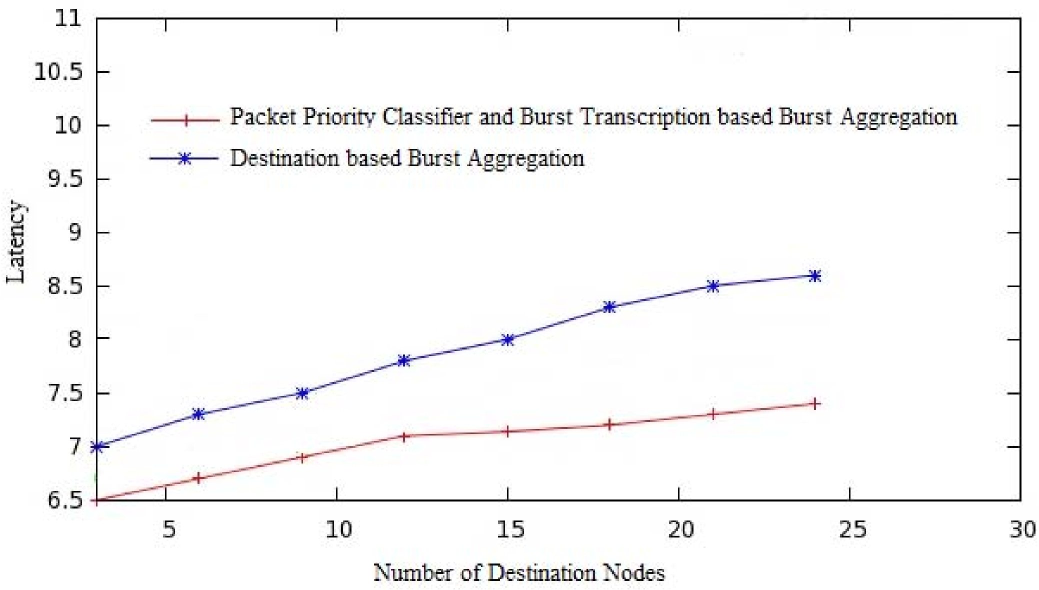

Fig 6: Latency vs. Number of Destination Nodes

7. CONCLUSION

Figure 4, 5 and 6 shows the simulation results obtained from experiments for the proposed work that confirms that resending cloned DSB offers faster delivery than the data retransmission requested at the Source. Thus, overall throughput of the network is increased by proposed architecture with minimal latency and losses.

REFERENCES

[1] Avranil Das, Shamik Chakraborty, and Alok K Das, Senior Member,IEEE, "QoS Performance of OBS in WDM/TDM Networks", *2009 IEEE*.

[2] Farid Farahmand1, Jason Jue , Vinod Vokkarane,"A Layered Architecture for Supporting Optical Burst Switching", *Proceedings of the Advanced Industrial Conference on Telecommunications, 2005 IEEE.*

[3] Georgios I. Papadimitriou, *Senior Member, IEEE,* Chrisoula Papazoglou, and Andreas S. Pomportsis, "Optical Switching: Switch Fabrics, Techniques, and Architectures", *Journal of lightwave technology, Vol. 21, No. 2, February 2003.*

[4] Tetsuya Miki (Member IEEE), "Optical Transport Networks", *Proceedings of the IEEE, Vol. 81, No. I I , November 1993.*

[5] Yong Liu, Gurusamy Mohan, Senior Member, IEEE, Kee Chaing Chua, and Jia Lu, " Multipath Traffic Engineering in WDM Optical Burst Switching Networks" , *IEEE Transactions on Communications, Vol. 57, No. 4, April 2009.*

[6] Onur Ozturk, Ezhan Karasan, Member, IEEE, and Nail Akar, Member, IEEE, "Performance Evaluation of Slotted Optical Burst Switching Systems with Quality of Service Differentiation, *Journal of lightwave technology, vol. 27, no. 14, July 15, 2009.*

[7] Vinod M. Vokkarane, Jason P. Jue, and Sriranjani Sitaraman, “Burst Segmentation: An Approach For Reducing Packet Loss In Optical Burst Switched Networks”, *2002 IEEE.*

[8] Tomasz,"Geocast—Geographic Addressing And Routing", Proceedings Of The 3rd Annual ACM/IEEE International Conference On Mobile Computing And Networking,pp 66-76, 1997.

[9] Balagangadhar G Bathula And Vinod M Vokkarane, "Qos-Based Manycasting Over Optical Burst-Switched (Obs) Networks", IEEE/ACM Transactions On networking, . Vol.18, No.1, pp - 271-283, 2010.

[10] Vinod M. Vokkarane, Jason P. Jue, and Sriranjani Sitaraman, “Burst Segmentation: An Approach For Reducing Packet Loss In Optical Burst Switched Networks”, *2002 IEEE.*

[11] Guray Gurel , Onur Alparslan and Ezhan Karasan, “nOBS: an ns2 based simulation tool for performance evaluation of TCP traffic in OBS networks”, 2007