Light-trail Test Bed for IP-Centric Applications

Nathan A VanderHorn, Srivatsan Balasubramanian, Mani Mina, Arun K. Somani,
Dependable Computing and Networking Lab
High Speed Systems Engineering Lab
Iowa State University
{nvander, vatsan, mmina, arun}@iastate.edu

***Abstract*—The internet transport infrastructure is evolving towards a model of high speed routers interconnected by intelligent optical networks. In this paper, we review current optical networking architectures and describe a new concept proposed in [1] called light-trails. We develop light-trails as a novel and amenable control and management solution to address IP-centric communication problems at the optical layer. We implemented a test bed to demonstrate light-trail feasibility. We also present three medium access control protocols for light-trails and evaluate their performance. The goal of light-trails and our solution is to combine commercially available components with emerging network technologies to provide a transparent, reliable and highly scalable communication network.**

***Index Terms*—Test bed, Media Access Control, Light-trail, Optical Network**

I. INTRODUCTION

The past decade has seen a networking paradigm shift from connection oriented communication to high bandwidth IP-centric data traffic. The shift is being driven by new applications in home entertainment, business communications and avionics networks. High speed broadband internet access coupled with video on demand and high definition television technologies are enabling digital home networking. Driving business and local area networks is the increasing requirement for networked data storage and access. Avionics networks are also seeing an increase in IP-centric applications including in-flight entertainment, multifaceted sensor arrays and high definition flight displays for both commercial and military aircraft. The availability of such applications depends heavily on the ability to move data in a fast and reliable manner without significantly increasing operation costs.

Advancements in fiber optic communications have increased the capacity of a single fiber to over a few terabits per second utilizing Wavelength Division Multiplexing (WDM). WDM is an approach that exploits the huge bandwidth available in the fiber by splitting it up into multiple non-overlapping channels and allowing users to transmit on each channel at the peak electronic rate. The advancements in fiber optic technology and the introduction of new IP centric applications have motivated researchers to explore new networking designs. A network architecture that enables high speed provisioning, accommodates multi-granularity traffic and supports high data rates will be the key enabling technology for future networks.

In this paper, we give an overview of conventional network paradigms. We then discuss a new network architecture called light-trails [1]. We develop light-trails as a solution to next-generation networking that provides bit rate and protocol transparency, sub wavelength grooming capabilities and high speed provisioning. We present the light-trail test bed and propose three different media access controls to regulate light-trail operation.

The paper is organized as follows. In section II, we briefly introduce some of the widely studied optical network architectures, bring about their main features, merits and demerits and motivate the reasons for showcasing the new light-trail architecture as a possible contender for next generation networks. Section III illustrates the light-trail system architecture, with a description of the light-trail test bed to follow in section IV. In section V, we present three different medium access control mechanisms for light-trails and evaluate their performance. Section VI discusses related and future work. The conclusions follow in section VII.

II. CURRENT NETWORK TECHNOLOGIES

Advancements in fiber optic technology and the proliferation of IP-centric applications in the past decade have mobilized researchers to explore new solutions for next-generation networks; the following sections outline some of the different perspectives. We believe that each of the technologies discussed below have their own respective niche and hence instead of one technology completely replacing the other, it is highly likely that they will co-exist in the evolving internet.

A. Optical Circuit Switching

Among various optical networking paradigms, optical circuit switched (OCS) networks have been the area of intensive research in the past few years. OCS networks operate on the notion of creating a dedicated circuit or lightpath from source to destination. A lightpath consists of an optical channel between a node pair that may be routed transparently through multiple intermediate nodes. A lightpath is realized in the physical fiber optic network using optical line terminals, optical cross-connects and optical add/drop multiplexers. Circuit switched networks are often called wavelength routed networks and are typically characterized by circuit lifetimes of the order of months or years once provisioned.

Although widely used in today's networks, the OCS concept is not consistent with the packet switching philosophy of the

The research reported in this paper is partially supported by National Science Foundation grants 0087746, 0434872 and 0306007

internet. The majority of internet traffic is fractional and if not groomed efficiently may lead to severe underutilization of circuits. Incorporating grooming functionality can alleviate resource requirements but brings along with it concerns related to scalability, complexity, transparency, cost and delay, that are normally associated with electronics. Thus, there is a need for a technology which supports a granularity of transmission and switching finer than that of a full wavelength.

Another type of circuit switching that is gaining attention as a way to offer quality of service and traffic engineering in next generation IP-based optical networks is Multiprotocol Label Switching (MPLS) [2]. MPLS is designed to provide a way of grouping individual traffic flows together for transport over a core MPLS-enabled network. The basic idea is to add a routing label that MPLS enabled routers can easily understand and route, thus creating a label switched path (circuit) from source to destination in which intermediate routing decisions can be made more quickly based upon the label instead of the IP address. Low utilization can occur in MPLS networks if the ratio of cross-connect configuration time to burst duration is high. In addition, traffic provisioning and MPLS path reservation can become very difficult in large scale networks.

B. Optical Packet Switching

Optical packet switching (OPS) [3] achieves efficient network utilization since it leads to high statistical multiplexing gains and is amenable for traffic engineering. Each packet arriving at an optical packet switch consists of a header and a payload. The switching fabric processes the header and routes the data to the appropriate output port on the appropriate wavelength. Contention in OPS networks is possible both in the control plane and in the data plane. Contention in the data plane may be dealt with in three dimensions - space (deflection routing), time (optical buffering) and wavelength (converters), however, contention in the control plane typically results in lost packets.

Pure OPS technology in which packet header recognition and control are performed entirely in the optical domain is still many years away. Cost effective switches with high port count, low cross talk, and nanosecond switching times are currently only being researched. The requirement for rapid synchronization, scalable packet level parsing mechanisms and the lack of fast and large random access memory units prevent implementation of sophisticated optical router architectures that are possible in electronics.

C. Optical Burst Switching

Optical burst switching (OBS) [4] is a hybrid approach between coarse-grained circuit switching and fine-grained packet switching. It is based on the concept of decoupling the control plane from the data plane. In OBS, data packets are aggregated into large bursts (~ tens of KB) at the network edge before being transported through the burst switched core. A header cell is sent to reserve resources at each switch in the core network. After a predetermined offset time that accounts for header processing and switch configuration at intermediate nodes, the data packet is sent. Contention can occur when two bursts have to be scheduled on the same wavelength channel of an outgoing port. Mechanisms like wavelength conversion, deflection routing, optical buffering and data scheduling are used to minimize and resolve contentions.

The problems concerning OBS are similar to that of MPLS in that path setup is required before transmission which may add to total network delay and additional overhead. In addition, the edge router for burst switched networks needs to implement burst assembly, disassembly and queue fairness algorithms. Thus, the control unit design may become challenging at high data rates.

D. Optical Light-trails

IP traffic is known for its burstiness and high variability; the existing architectures discussed above are not effective in handling such traffic patterns. As a solution to providing high network resource utilization, seamless scalability and network transparency, we discuss light-trail technology [1, 5-8]. The goal of light-trail is to minimize active switching, maximize wavelength utilization, and offer protocol and bit-rate transparency to address the growing demands placed on WDM networks. Light-trail technology is a physical layer architecture that combines commercially available optical components to allow multiple nodes along a lightpath to participate in time multiplexed communication without the need for burst or packet level switch reconfiguration. The following section describes the light-trail architecture in more detail.

III. LIGHT-TRAIL NETWORK

A light-trail is similar to a lightpath in that it requires the establishment of an optical circuit between a chosen source and destination. The key difference is that some intermediate nodes can also receive and transmit data on the same channel. Fig. 1 shows a typical multi-wavelength light-trail node. For each wavelength, a light-trail access unit (LAU) that consists of a splitter, shutter, combiner and power compensator (typically, a semiconductor optical amplifier (SOA)) are

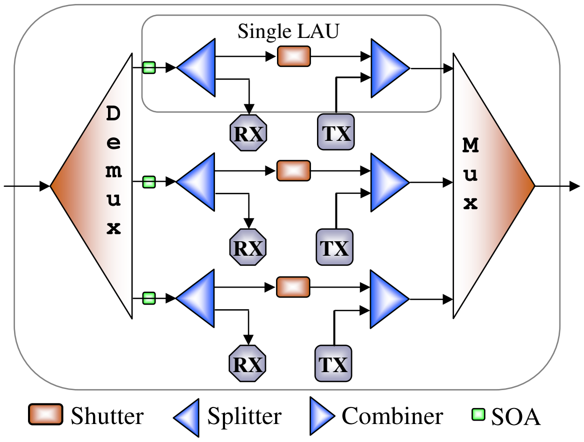

Fig. 1. Multiple wavelength node featuring multiple Light-Trail Access Units (LAU).

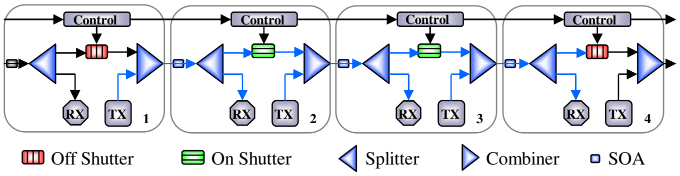

Fig. 2. Four node light-trail – optical connectivity path is displayed in blue.

provisioned. Although WDM can be supported, the remainder of this paper only considers the single wavelength node.

During operation, when data is sourced by a node, based on the light-trail medium access strategy, the optical signal traverses all nodes en route to the destination. At each LAU splitter, a sufficient amount of optical power is tapped from the incoming signal for local processing. The local node may choose to ignore the packet or forward it to higher layer if the packet is destined for it. After the signal passes through the splitter, the remaining signal is sent through the optical shutter: a simple optical attenuator configured to either block a selected wavelength or let the signal pass through. If the node is the first or last node on the trail, the shutter is configured in the off/blocking position and for all intermediate nodes, the shutter is in the on/pass-through position. This configuration isolates the particular wavelength from the rest of the network and enables spatial reuse of wavelengths. Lastly, if the signal is not blocked by the shutter, the transit signal is coupled at the combiner. The coupler is used by the local transmitter to introduce its signal when it accesses the channel.

A four node light-trail is shown in Fig. 2. A light-trail (LT) is defined to be a simple path in a network graph that can support multiple requests subject to the following constraints

- **Containment Constraint:** The light-trail can support any request (i,j) if $i,j \in$ LT and j is downstream of i on LT.
- **Capacity Constraint:** The aggregate traffic supported by all connections in LT is at most the capacity of a single wavelength.

In some sense, the light-trail concept may appear to be similar to the Distributed Queue Dual Bus (DQDB) architecture specified in IEEE 802.6 standard. However, it is important to the note the two key differences between light-trails and DQDB architectures. DQDB is bi-directional, whereas light-trail is not. The light-trail's unidirectional nature is best suited to meet the prevalent asymmetric traffic patterns of the internet and to give the designer the flexibility to establish only those trails that optimally meet the traffic requirements. The second key difference is that the DQDB is a physical topology whereas the collection of light-trails defines a virtual topology embedded over a mesh network. Thus, light-trails lend themselves naturally to be a more general framework to cater to the needs of IP-centric applications.

A. *L-Bone Architecture*

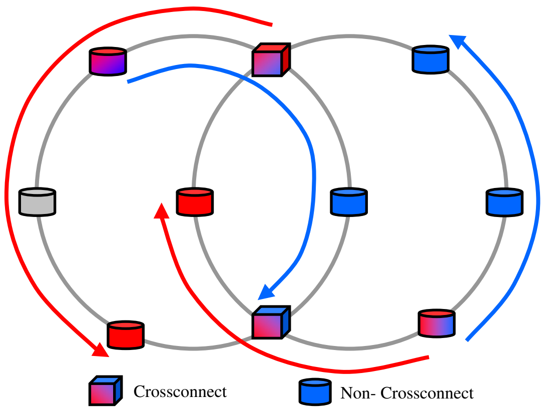

Fig. 3. The L-Bone network is an example of how light-trails can be configured as a complete network solution. The node color indicates which light-trail a node is active on. A grey shade indicates that a node is inactive.

The L-Bone network is an example network that illustrates how light-trails can be used as a complete mesh network solution.

Light-trails in the L-bone network are established and torn down during the light-trail design phase and are not configured on a packet-by-packet basis. This is done through the use of an out of band control channel which is dropped and processed at each node. The signaling channel carries information pertaining to the set up, tear down and dimensioning of light-trails, and is responsible for provisioning "optical connections," ranging in duration from IP burst to virtual circuits.

Fig. 3 shows an L-bone network. Note that the signal may not be tapped on every intermediate node of the trail. If the node is actively communicating on a trail, the incoming signal is demultiplexed, switched to the LAU, and is reinserted back into an OXC to be multiplexed onto an output fiber. However, if the node is not involved with communication on this trail, then, the OXC lets the signal bypass the LAU and switches it directly onto the output fiber.

The L-Bone network offers full optical connectivity that can share the wavelength among all nodes in time domain leading to dynamic sub-wavelength allocation and multicasting. It is important to note that despite the absence of dynamic switching within light-trails, the granularity obtained is sufficient to provide IP-centric communication bursts in addition to full lightpath connections.

IV. Light-Trail Test Bed

In order to investigate the characteristics, capabilities and limitations of light-trail technology, a single wavelength prototype is developed. Fig. 4 illustrates a block diagram of the four node test bed configuration.

The primary component of the test bed is the Xilinx Virtex 2 Pro FPGA development board. The FPGA device contains high-speed RocketIO SERial DESerial (SERDES) Multi-Gigabit Transceivers (MGT) which provides the serial data stream to the laser modules. Each of the four stations of the test bed assembly maintains block RAM (BRAM) modules for storing transmit and receive data. The BRAM modules of all

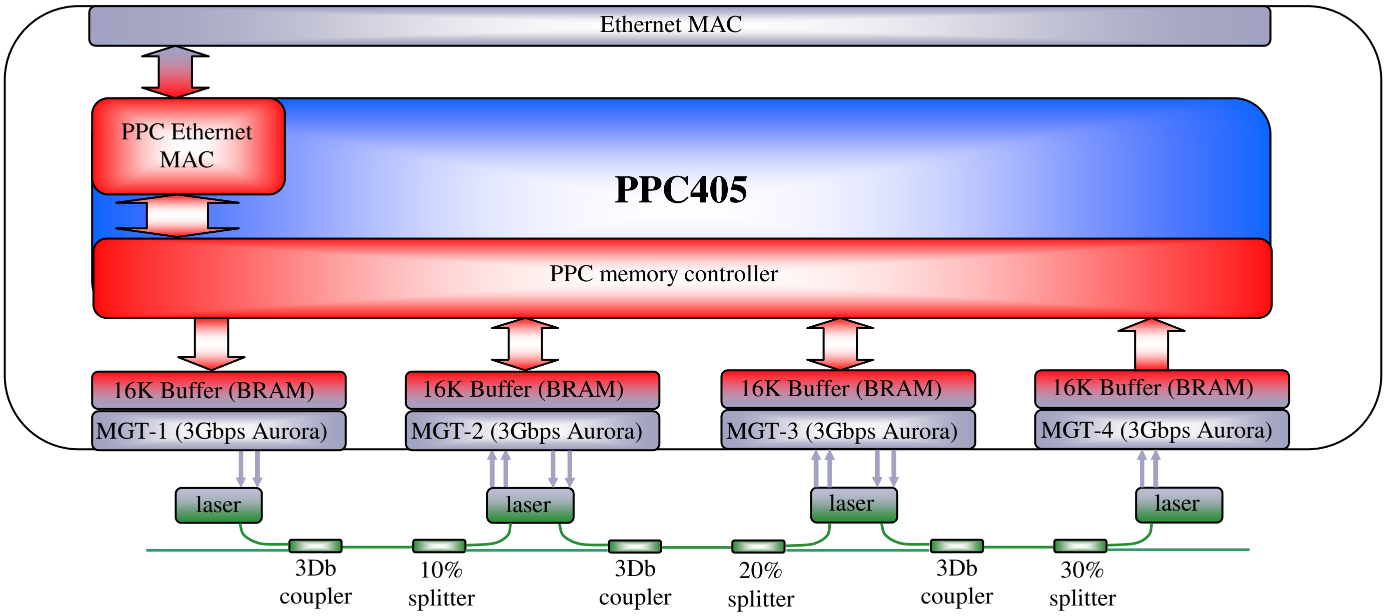

Fig. 4. Four node light-trail test bed configuration as implemented on a Xilinx FPGA development board.

stations are also connected to the embedded PPC 405 microcontroller. The connection to the microcontroller gives us the ability to load and display the contents of any of the BRAMs through the use of an Ethernet controller connected to a host computer.

The four-node implementation utilizes six fiber optic couplers to complete the light-trail. The fiber optic coupler configuration is shown in Fig. 4. Here, 3dB couplers are used at all nodes to couple the local signal with the upstream transit signal. The test bed also uses optical splitters to tap the signal at the local receiver, these splitting ratios are 90/10, 80/20 and 70/30 for nodes 1, 2, and 3 respectively.

To demonstrate light-trail functionality, packets for each of the first three stations are generated with pseudorandom payloads and loaded into their respective memories (BRAMs). The source-destination pair is randomly generated by the PPC using a permutation matrix of possible connections. The selected source node’s laser is activated and the packet is transmitted. In our experiments, the destination node verifies the contents of the packet and checks for correctness. Multicast and broadcast capabilities are also verified in a similar fashion.

Because light-trails operate on a shared medium, it is important that the design of a complete light-trail solution must include a media access control to govern bandwidth arbitration. The next section discusses three light-trail MAC protocols designed to prevent collisions and to provide bandwidth allocation policies.

V. LIGHT TRIAL MEDIUM ACCESS CONTROL

A number of priority based access protocols have been proposed for unidirectional bus topologies [9, 10]. The following sections describe three medium access control methods for light-trails.

A. Light-trail MAC

A simple MAC protocol based on carrier sensing is proposed for light-trails in [1]. When a node, say X, wants to transmit data, it senses the carrier for upstream activity. If the channel is free, X sends a beacon signal downstream indicating that it has a packet to transmit. If the channel is busy, X waits until the channel is free to send the beacon signal. After some predetermined offset time, called the guard band, X transmits its data. Node X continues sensing the channel while its packet is being transmitted. During transmission, X may hear a beacon signal from an upstream node, say Y, which has data to transmit. Upon hearing this, X terminates its transmission and lets Y's packet pass through. X’s truncated packet is discarded by the receivers since it fails the link level error checks. Thus, by sending a beacon signal before transmitting a packet and by always giving priority to the upstream nodes, the protocol successfully resolves medium access contentions.

The key design parameter that affects protocol performance is the guard band gap. When the decision to stop is made, the feedback control happens through a microcontroller and hence the delay is large (typically at least 50 ns) due to the electronic processing overhead. The guard band (set to 75 ns) should be set large enough to pull out of transmission after sensing the beacon signal and small enough to avoid significant overhead.

B. Light-bus MAC

A new architecture called the light-bus is proposed in [5]. The key difference is that the light-bus switch architecture includes a fiber delay loop at every node. The delay is statically set to the time needed to transmit a maximum size packet. When a node has data to transmit, it first checks the delay line for activity. If the delay line is free, the node transmits its packet. If the delay line is busy, the node waits until the delay line becomes free and then starts transmission. In the middle of a transmission, if an upstream packet reaches

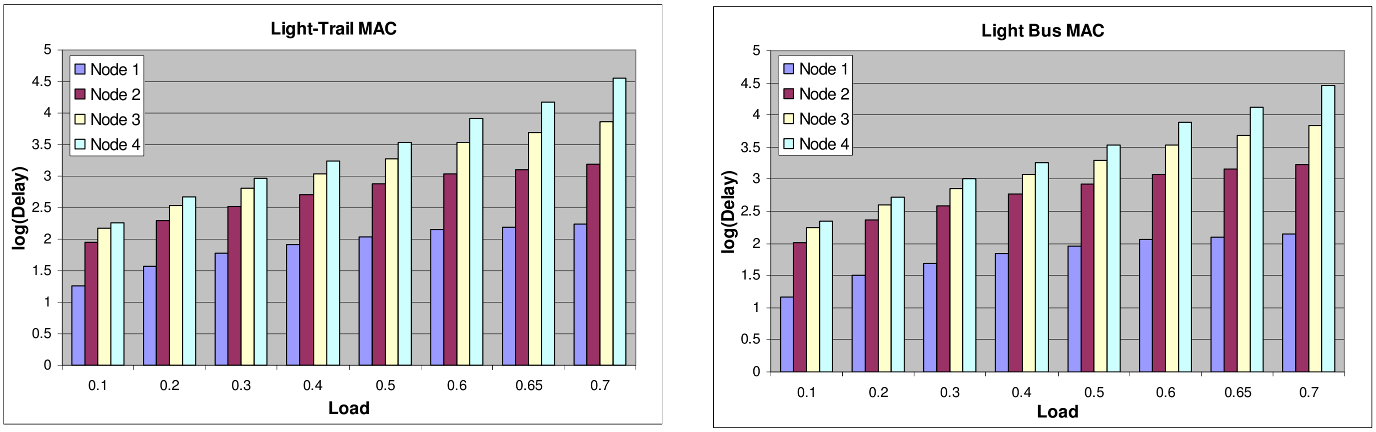

Fig. 5. Queuing Delay Vs Load for a 5 node a) light-trail and b) light-bus with packet sizes uniformly distributed between 500-1500B bytes.

this node, it is buffered in the delay line and the current transmission is completed before the upstream packet exits the delay line.

The key idea behind this protocol is to design an efficient control structure on both the sender and the receiver nodes. The transmitter does not have to worry about aborting and retransmitting packets and the receiver does not have to deal with fragmented packets. The light-bus approach does not lead to wasted transmissions. In the light-trail protocol, a guard band whose duration depends on the rise and fall times of the transponders and the speed of the control electronics is required. This overhead incurred for every packet is avoided in the light-bus protocol. However, the fiber loops may impose extra queuing and propagation delays. The central idea behind the light-bus model is to trade-off performance for simplicity and efficient control.

C. Light-trail and Light-bus Performance Evaluation

Performance analysis of the protocols for different traffic loads is done using discrete event simulation techniques. Poisson arrivals and uniform service times are assumed for the simulations done on 10 Gbps systems. We considered five-node light-trails and light-buses based on the expected length suggested in [1]. In the graphs, the x-axis plots the total offered load expressed as a fraction of the capacity of the wavelength and The y axis plots the log of the delay normalized to the maximum sized burst (and scaled by a factor of 1000). For the five node system, traffic density for nodes 1 through 4 are split in the ratio 4:3:2:1, respectively.

Our observation is that both the light-trail and the light-bus protocols perform well until about loads of 0.7 C. The delay encountered by a packet on a node is more than the delay experienced by a packet on a node upstream to it. The queuing delay is highest for the 4th node while the delays are negligible for all the other nodes. For results shown in Fig. 5, the packet sizes are uniformly distributed between 500 and 1500 bytes. When the packet sizes are small, the guard band required by the light-trail protocol is of the order of the packet transmission time and the light-bus protocol is able to outperform the light-trail protocol.

D. Light-trail Fair Access MAC

Although the light-trail and light-bus MAC are effective in preventing collisions, these schemes do not provide fairness among nodes and do not guarantee a bound on access delay. The reason for this undesirable behavior is because stations at the head of the trail have highest priority, and it is possible that head end stations can monopolize the channel for unrestricted amounts of time. To address this problem, we present the Light-trail Fair Access (LT-FA) MAC. The goal of the fair access MAC is to design a bandwidth allocation scheme that takes fairness and maximum access delay into consideration while providing collision free service.

The LT-FA mechanism uses a round robin scheduling technique with predetermined round intervals to bound access delay. Fairness is offered in terms of an acceptable use policy by negotiating an acceptable allocation estimate for each node; the aggregate of this allocation must be less than light-trail capacity. As the light-trail master, the head node sends a beacon signal to interrupt current transmission and designate the start of a new round. During operation a nodes access is surrendered for a round when either the transmission buffer is empty or the fair access limit (plus an extra amount explained later) is reached. Because internet traffic is bursty in nature, often times, the actual round usage required will not correspond to the negotiated fair estimate, thus, when trail access is relinquished the difference between a nodes actual usage and its fair access portion is passed on to the next downstream node. This additional unused capacity can be usurped by downstream nodes to provide burst mode support at a per round granularity. Further details of the LT-FA can be found in [6].

Fig. 6 illustrates maximum access delay and average queuing delay using the LT-FA mechanism with a round time of 100 slots. The values in Fig. 6 illustrate the effect of increasing Poisson distributed traffic from 12 to 56 percent (12 to 56 slots per round). Deterministic traffic and burst mode traffic, in which the burst arrival time is governed by a Poisson process and the burst size is exponentially distributed with a mean value of B_{size}, remain constant at a rate of 12 percent (An

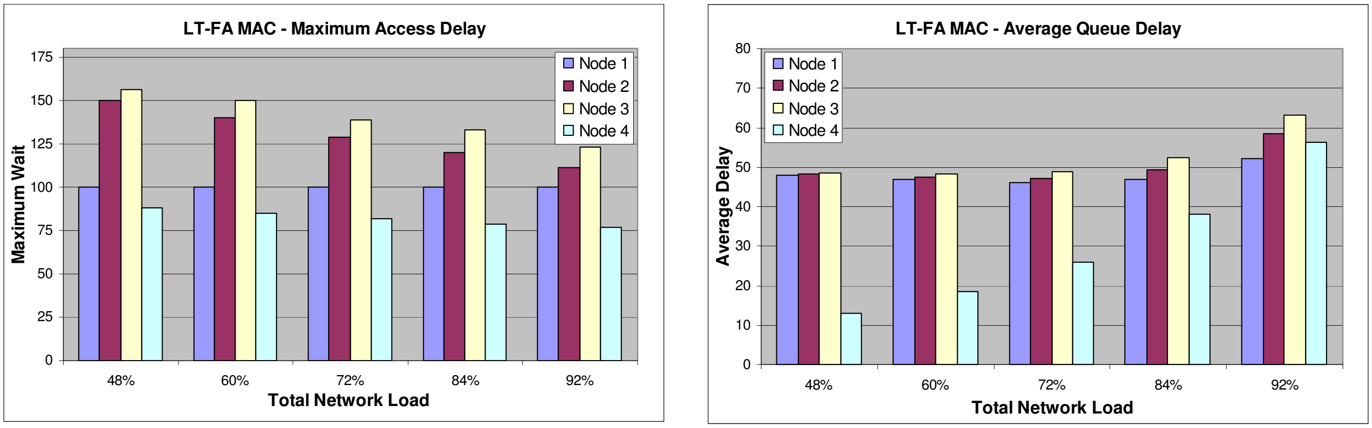

Fig. 6. a) Maximum access delay and b) average queuing delay for a 5 node light-trail configuration with varying Poisson traffic loads.

additional 12 percent of deterministic traffic is allocated to account for laser rise time). All traffic is uniformly distributed among all transmitting stations. The results shown are for an average B_{size} of 5 which gives a burst size range between 1 and 75 slots.

As mentioned earlier, in unidirectional bus networks, nodes near the end of the bus experience higher queuing delays because head stations have the first opportunity to utilize available resources. When compared to p_i-persistent protocol [9] the LT-FA was shown to perform better with respect to average delay experienced by stations near the end of the trail.

VI. FUTURE WORK

As a newly proposed concept, light-trail architecture brings up various issues in designing optical networks for transporting IP-centric traffic. The static light-trail design problem is defined and an ILP formulation is proposed and solved in [7]. However, dynamic routing in light-trails is another topic that merits further research. The resource requirements of light-trail and groomed lightpath networks for both static and dynamic cases need to be compared to identify the types of topology and traffic that could possibly favor one architecture over the other.

Fault management becomes an important issue since the failure of a single light-trail can lead to failures of multiple connections traversing the trail. The survivable light-trail design problem, which takes into account link failure during the design and operation phase, is formulated in [8] and an ILP is presented for a connection based protection. Efficient heuristics for shared and dedicated path protection in large scale optical networks need to be developed.

At the time of print, the light-trail test bed has been enhanced to support streaming media applications between multiple source destination pairs. The proliferation of the test bed to include multiple wavelength support, robust media access control and L-bone crossconnect functionality requires much deeper research into design challenges at the device level as well as at the system level.

Early experiments suggest that we must consider many practical implementation issues such as, measurement and characterization of devices to guarantee system level compatibility, careful planning of amplifier placement to meet power budget constraints yet avoid receiver saturation and design of burst mode lasers to increase network throughput and response time. In addition various properties and design specifications of devices such as, couplers, connectors, shutters, laser on/off switches and fiber types must be considered.

The practical implementation dimension makes the light trail solution a more interesting, challenging, and rich problem. While the implementation aspects introduce limitations and added complexity, such practical problems are worth thinking, planning, and including in the comprehensive light-trail design.

VII. CONCLUSION

This paper is focused on discussing some of the emerging next generation optical network technologies for the internet. We have introduced the merits and demerits of these technologies and presented light-trail technology as a contender to these methods.

The light-trail network and test bed is presented as an all optical access network that combines commercially available components with emerging network technologies to provide a transparent, reliable and highly scalable communication network. In addition, performance evaluation for three light-trial media access controls is discussed. We believe light-trail technology provides attractive properties for next generation all optical networks.

REFERENCES

[1] A. Gumaste, I. Chlamtac, “Light-trails: “A novel conceptual framework for conducting optical communications,” Workshop on High Performance Switching and Routing, Pages 251 – 256, June 2003.

[2] A. Viswanathan, N. Feldman, Z. Wang, R. Callon, “Evolution of Multiprotocol Label Switching,” IEEE Communications Magazine, vol 36, pp.165-173, May 1998.

[3] S. Yao, B. Mukherjee, and S. Dixit, “Advances in photonic packet switching: an overview,” IEEE Communications Magazine, vol.38, pp.84-90, Feb. 2000.

[4] Y. Chen, C. Qaio and X. Yu, “Optical burst switching: a new area in optical networking research,” IEEE Network, vol.18, pp.16-23, May 2004.

[5] S. Balasubramanian, Ahmed Kamal, A. K. Somani, "Medium Access Control Protocols For Light-trail and Light-bus Networks," Proceeding of 8th IFIP Working Conference on Optical Network Design and Modeling, Feb 2-4, 2004.

[6] N. A. VanderHorn, M. Mina and A. K. Somani “Light-trails: A Passive Optical Networking Solution for Wavelength Sharing in the Metro,” Workshop on High Capacity Optical Networks and Enabling Technologies, Dec 2004.

[7] Jing Fang, Wensheng He and Arun K. Somani, "Optimal Light-trail Design in WDM Optical Networks," in the proceedings of ICC 2004, vol.3, pp.1699-1703, Paris, 2004.

[8] W. He, J. Fang, A. K. Somani, “On Survivable Design in Light-trail Optical Networks,” Proceeding of 8th IFIP Working Conference on Optical Network Design and Modeling, Feb 2-4, 2004.

[9] B. Mukherjee, J.S. Meditch, “The pi-persistent protocol for unidirectional broadcast bus networks,” IEEE Transactions on Communications, Vol. 36, Issue 12, Pages: 1277 – 1286, Dec. 1988.

[10] B. Mukherjee and A. E. Kamal, “Scheduling variable-length messages on slotted high-speed fiber optic LANs/MANs using the continuation-bit approach,” Computer Networks & ISDN Systems, special issue on Media Access Techniques for High-Speed LANs and MANs, vol. 26, no. 6-8, pp. 721-744, March 1994.