lambdaMON – A Passive Monitoring Facility for DWDM Optical Networks

Jörg B. Micheel

NLANR/MNA[1], San Diego Supercomputer Center/UCSD
10100 John Hopkins Dr, 92092-0505 La Jolla, CA, USA

joerg@nlanr.net, http://pma.nlanr.net/~joerg/

Abstract. This paper presents lambdaMON - a novel approach to passive monitoring of very high performance optical networks based on dense wavelength division multiplexing (DWDM). The approach offers very attractive cost/benefit scaling properties, which are further refined by introducing state-of-the-art transparent fiber switching equipment. The rapid pace at which we intend to implement lambdaMONs opens new opportunities to apply passive monitoring facilities for debugging, troubleshooting and performance analysis of novel protocols and applications. To the best of our knowledge, this is the first attempt at designing a passive monitoring facility for optical networks. We report detailed architectural parameters, measurements and experience from laboratory tests and initial field deployment.

1. Introduction

Optical networking, based upon dense wavelength division multiplexing (DWDM), is rapidly becoming the technology of choice for the high-performance networking community, as well as major national and international commercial network providers.

Optical networking uses individual light rays (also referred to as colors, wavelengths, carriers, or channels) to carry very high-performance point-to-point connections over long distances. The capabilities offered by optical networking fundamentally challenge traditional means of time division multiplexed IP services, which have dominated the development of the Internet for the past two decades.

Traditional circuit-based provisioning, such as OC3/OC12 leased lines, ATM, MPLS, or virtual private networks (VPN) are rapidly being replaced by fiber optic communication channels, or lambdas. The opportunities for improving connectivity and performance between regions, countries, or very demanding end user communities, such as the high energy physics research community, are tremendous. At present, it is difficult to imagine the full impact that DWDM technology might have for end-to-end networking, especially when used to support very sensitive and demanding 10-Gigabits/second and quality-of-service aware user applications.

With networks being rolled out at a rapid pace, new protocols being developed and field tested, and new applications emerging, the stakes for success are high and the risks that have to be taken by the research community are substantial. A readily

---

[1] This work was supported by the U.S. National Science Foundation Collaborative Agreement ANI-0129677 (NLANR/MNA, 2002) with subcontract to the University of Waikato in New Zealand

available monitoring facility is highly desirable as a means to locate, debug, troubleshoot and resolve potential problems. In this regard, **passive systems have a unique advantage: they explore the network as is, without interfering with the actual traffic data pattern as generated by end systems or as modified through intermediate devices, such as switches and routers.** In addition, passive monitoring systems present an excellent means to understand and address issues at network layers two to seven.

In this paper we present lambdaMON – a new passive monitoring technology that enables the debugging and troubleshooting of applications and protocols that operate over long-distance DWDM networks. While the lambdaMON is fundamentally a new passive network monitoring technology, it preserves the traditional means to precisely collect and analyze workload and performance characteristics of edge, access, and backbone network links.

The rest of the paper is organized as follows. In the next section we provide a brief overview of traditional passive monitoring systems, their advantages and shortcomings. In section 3 we look at the specifics of DWDM technology as deployed in the field today and determine the point of instrumentation for lambdaMONs. In section 4 we outline the constraints for the lambdaMON architecture. Section 5 looks at implementation challenges. Section 6 summarizes the achievements in architecting and designing lambdaMONs

2. Traditional Passive Monitoring Systems

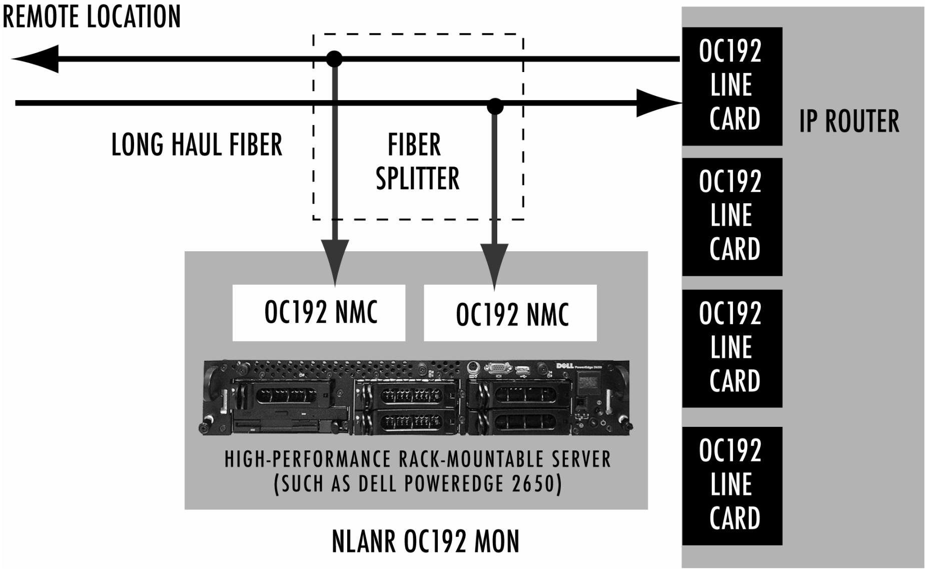

Figure 1. Classic OCxMON monitoring setup.

Passive network monitors interoperate with the live network at the link level (see Figure 1). They are considered a vendor independent means of gathering data as they do not depend on features that would otherwise have to be provided by active network equipment operating as routers, switches, hubs or end systems [1]. Operating at the medium dependent physical layer, these OCxMON systems are equipped with link layer specific network measurement cards (NMCs), which reimplement all layer one and two functions, such as deserialization, packetization and various packet encapsulations, and then pass the data on to analysis-specific functions, such as arrival time stamping, selective filtering and payload discard or flow state analysis. Such functions are executed by reconstructing and accessing information that is specific to network layers three to seven. If real-time analysis applications are used, the monitor may deliver a complete solution by means of a graphical user interface.

The biggest strength of passive monitoring technology (i.e., link layer dependency) is at the same time, also one of its biggest weaknesses. Every emerging technology advance demands that a new set of PC cards (NMCs) be developed in order to support a compatible interface to the network link. Passive network monitors are, by design, lower-cost devices, and their implementation is hence based on off-the-shelf components. Therefore NMCs are typically designed and implemented after a given link layer technology has been rolled out in the field, and as a result they are generally available only for the second half of the life cycle of any high-performance (backbone) link layer technology. This means that the use of passive monitors, so far, has been restricted to traffic and workload characterizations of mature networks.

While passive monitors are extremely powerful once deployed in the field, their widespread use faces some steep resistance for both cost and technology reasons. Being a per-link facility limits their deployment to dedicated research environments for cost reasons, thus preventing them from becoming a more general-purpose operational network facility. Examples of larger scale deployment include the infrastructures operated by NLANR/MNA and CAIDA. Perhaps the largest known infrastructure in a commercial setting is operated by Sprint ATL's IPMON research group. In addition to the financial obstacles, the process for installing fiber optic splitters still presents a technical hurdle for most users new to passive monitoring, and misconfigurations are frequent. At best, the monitor will be unable to collect data on one or both of its interfaces. At worst, the network link itself will be unable to operate, or remain intermittent and unreliable – unacceptable for any type of network operations.

As a result of these obstacles, passive monitors have never been used for locating, debugging, troubleshooting and eliminating end-user application, protocol, and performance problems. For a random end-user or problem the chances of one (or more) passive monitors being present, available, and accessible along an end-to-end networking path are very small.

3. DWDM Optical Networks

DWDM harnesses a spectrum of lambdas within the third optical window (1520 to 1620 nanometers) for long-distance high-performance data transmission. This spectrum has been chosen for its low attenuation and for its ability to amplify signals at the optical level without the need for regeneration, which would otherwise force a technically expensive optical-electrical-optical conversion.

Major amplification technologies in use today are erbium doped fiber amplifiers (EDFA) and Raman pump lasers. Unlike EDFAs, which can be purchased and integrated as modules into the fiber path, Raman amplifiers use the entire long-distance fiber span as a medium, with the pump laser located at the receiver of the optical transmission line. Due to the way these amplifiers work, the third optical window has been subdivided into the C (1520 to 1560 nanometers) and L (1565 to 1620 nanometers) bands. With the use of EDFAs, fibers spans of up to 600 kilometers (375 miles) can be achieved. EDFAs combined with Raman amplifiers will reach up to 2000 kilometers (1250 miles) for production use.

The use of lambdas within the C and L bands has been standardized by the International Telecommunication Union – Telecommunication Standardization Sector (ITU-T). ITU-T Recommendation G.694 defines a grid of wavelengths rooted at 1552.52 nanometers (193.1 THz). The spacing between carriers is implementation dependent and includes 200 GHz, 100 GHz, 50 GHz, and 25 GHz options. The choice of grid spacing controls the number of channels that can be supported by any single channel. For instance, with a 50 GHz grid, up to 80 channels can be supported in the C band. Due to the modulation noise band and necessary isolation between channels, a 50 GHz grid limits the digital carrier signaling frequency to about 20 GHz. For technical reasons it is unlikely (but not impossible – see [6]) that carriers beyond OC192/10-Gigabit-Ethernet will be deployed on a 50 GHz grid any time soon.

DWDM networks are presently built in a static setup as single-vendor single-product implementations. The simplest configuration will involve a pair of DWDM terminals operating at either end of a bidirectional long distance fiber (see Figure 2). A DWDM terminal supports access to the individual carrier wavelengths by means of a transponder, which supports the connection of traditional carrier class equipment (SONET OC12, OC48, OC192) or local area networking gear (such as 1-Gigabit and 10-Gigabit Ethernet devices).

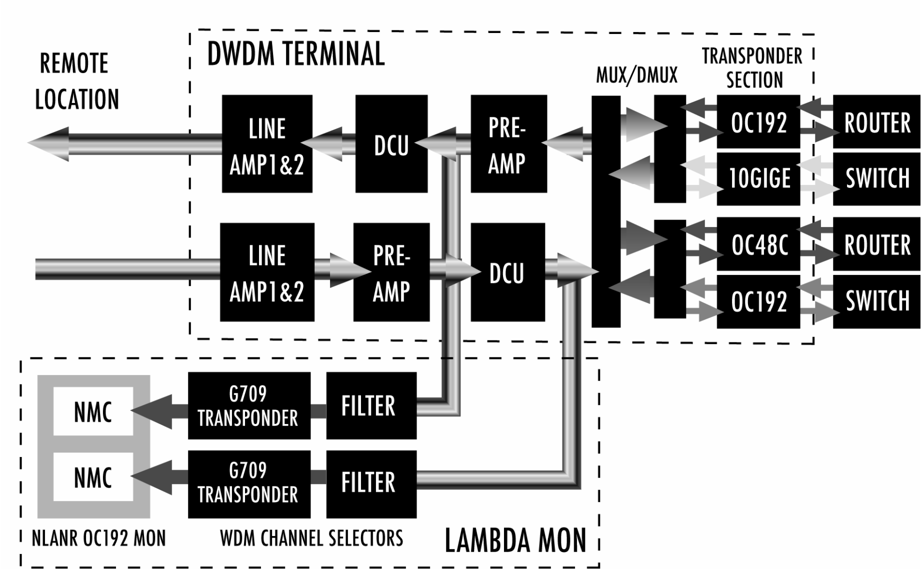

Figure 2. DWDM terminal with lambdaMON setup.

The transponder converts a traditional SONET/SDH or Ethernet LAN PHY signal via a G.709/G.975 encoder by employing a Reed-Solomon (RS[239,255]) forward error correction (FEC) code to lower the expected bit error rate (BER) on the transmission link. The use of FEC supports the operation of longer fiber spans without the need for regeneration, which in turn makes the entire system significantly more cost effective. The G.709 encoded signal operates at an approximately 7 percent higher rate relative to the original 10 Gigabit carrier and is launched via a laser at a specified wavelength into the first stage multiplexer (MUX). It is this signal that the lambdaMON will pick up once the signal has passed through additional, but optional, stages of multiplexers and amplifiers (AMP), but before entering the dispersion compensation unit (DCU).

For the inbound direction, all lambdas will pass through one or more stages of the demultiplexer (DEMUX), with a single channel (still G.709 encoded) eventually reaching the decoding section of the transponder, where it will be converted back into the original SONET/SDH or Ethernet LAN PHY signal.

All of the passive modules forming a DWDM terminal, such as MUX, DEMUX and DCU, introduce a device dependent attenuation, typically between 5 dB and 10 dB. However, the operating range of transmitters, receivers and other parts of the system is typically limited to between +5 dBm and -20 dBm. Therefore, signals will have to be strengthened at least once within the terminal, which is the role of preamplifiers (PRE AMP).

The point of instrumentation at a DWDM terminal requires a fully multiplexed, non-dispersed signal that is strong enough to tolerate the additional attenuation introduced by the fiber optic splitter and lambdaMON components, such as the tunable channel filter (TCL). Therefore, the exact location to instrument is the right-hand side of any DCU and past any preamplifiers in the direction of light travel (see Figure 2).

DWDM transmission lines can be built from just one, or two independent, fiber runs. In the event of a fiber cut, the DWDM system will, within a defined period of time, automatically switch from the primary fiber to the secondary. Since protection switching is handled at the line amplifier section of a DWDM terminal or device, the lambdaMON architecture is not affected by whether the network owner chooses to operate in protected mode or not.

Since modern DWDM sites are built in a modular fashion, similar components as are presently found in DWDM terminals will also be present at optical line amplifier (OLA), regenerator (REGEN) or optical add/drop multiplexer (OADM) sites. The lambdaMON architecture fundamentally permits deploying monitors at any of those DWDM sites, however, since passive monitors are best operated remotely with a legacy network connection, there is not much point in placing them at OLA or REGEN sites. OADM sites are very similar in nature to DWDM terminals and will support the proposed architecture.

4. lambdaMONs

A lambdaMON is a bidirectional 10 Gigabits/second device capable of collecting and analyzing packet header data at line rate from any one (at a time) of the active wavelength carriers at a single DWDM link.

A multi-feed lambdaMON, or lambdaMON node, is an advanced configuration permitting a lambdaMON to monitor any one (at a time) of a number of DWDM links at a given location.

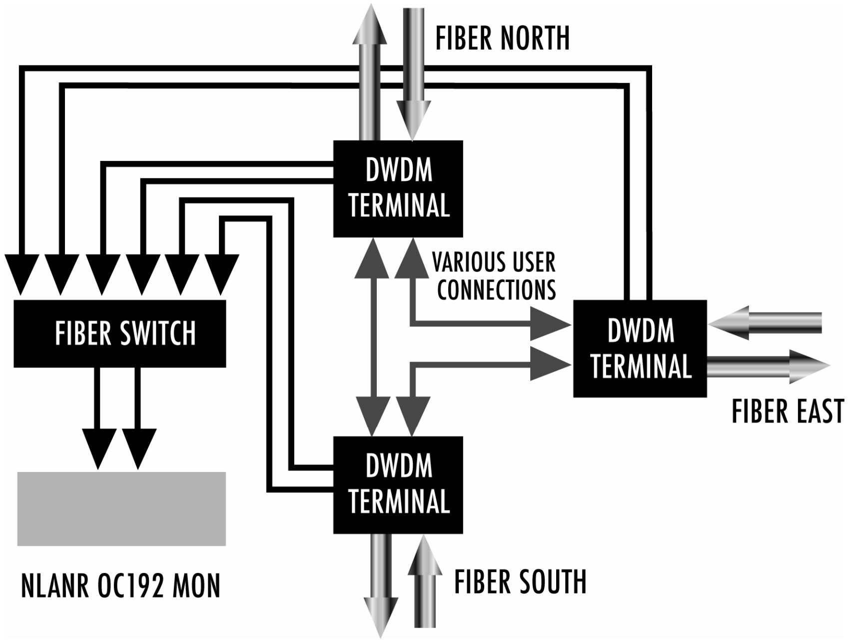

Figure 3. Multi-feed lambdaMON or lambdaMON node.

Our design for lambdaMONs and lambdaMON nodes is driven by the following considerations:

1. We are keen to break the technology dependency cycle that lets passive monitors miss the first 18 months of the life cycle of any high performance network link layer technology.
2. We want to demonstrate the utility of passive monitoring systems for troubleshooting and debugging of networks, applications and protocols. Such a facility is critical within the first year or two after network rollout, much less once the network is mature.
3. We want to retain the ability to collect and analyze workload profiles of mature networks.
4. We believe that there is no strict need for a static association between network links and monitors. It appears that the community would rather have a passive monitoring facility available on demand, and possibly on short notice, if and when such a requirement arises.
5. We believe that the success of passive monitoring technologies has been hampered by cost considerations as well as technical obstacles. We can break those constraints by designing a readily available facility central to the network.

Based on those constraints, we derive our lambdaMON architecture as follows:

1. Make best use of existing OC192MON technology, which permits loss free data capture and real-time analysis at 10 Gigabit LAN/WAN PHY and OC192c Packet-over-SONET network links.
2. Address the per-link constraint (and associated costs) of traditional OC192MON deployment by designing a system that will permit dynamically tuning into any one (at a time) of the active lambdas at a given DWDM link.
3. Increase the coverage on a given optical infrastructure by introducing a transparent fiber switch to tap into multiple DWDM links at a site (see Figure 3).
4. Retain the option to increase analysis power by introducing additional lambdaMONs to the fiber switch in the future. As opposed to a static per-DWDM-link setup, this arrangement will permit the analysis of multiple lambdas at a single DWDM link, if required.

5. Technical Challenges

DWDM is still an emerging technology and some of the equipment in the field today may not easily support passive fiber splitters, and the resulting engineering effort may turn out to be prohibitive to consider the installation and operation of lambdaMONs. As an alternative, existing one percent monitoring ports, which appear to be a standard feature with most equipment, can be used in replacement for the fiber tap. The advantage is that equipment in the field will not experience any service disruption when installing lambdaMONs; however, on the downside, the signal levels are too low for the operation of fiber optic transceivers and other passive optical equipment. Therefore, the signal levels will need to be amplified by an EDFA first, which causes an overall increase in costs. First field tests have shown that signal levels at one percent monitoring ports vary, and a single stage amplifier may or may not be sufficient to increase the signal power level in such a way that a commercially available G.709 transponder can be operated reliably at the expected BER.

Tunable channel filters (TCL) are just becoming commercially available for field deployment, and are restricted to operating in either the C or the L band. In addition, the selection of filters that support stable operation on a 50 GHz grid is presently very small. Field tests have shown that TCLs come at an affordable price and work reliably to the full satisfaction of the project.

It appears to be attractive, from a cost point of view, to integrate G.709 transponders into OC192MONs. Even though there is some loss of flexibility, the cost reductions are significant, specifically in a large scale rollout, like in a distributed infrastructure.

6. Conclusion

We have successfully carried out laboratory tests using a C band setup simulating a total of 600km of long distance fiber terminated by a pair of CISCO 15808 terminals. We amplified signals from the one percent monitoring ports in both the transmit and receive directions via a CISCO 15501 EDFA and successfully transferred 1.4 Terabytes of data over the course of several hours with varying packet sizes and data loads without incurring any bit errors or packet drops.

Our next step, at the time of publication, is to stage a phase 2 prototype of the lambdaMON, which will involve a full setup of the lambdaMON node in Los Angeles, with the active support of CENIC. Such a node may include OC48MON monitoring equipment right from the start, as there are a number of links that are of particular interest to us.

We are also looking at expanding the passive measurement node concept to traditional SONET-based networking. It is expected that we will continue to publish our progress via the project's Web site [7].

Acknowledgement

The author expresses his thanks for the tremendous support received by key staff at the Corporation for Network Research Initiatives in California (CENIC), National LambdaRail (NLR), the Internet2 HOPI team, the Indiana GlobalNOC and TransPAC. This project would have been impossible without the help of CISCO Systems. We also appreciate the support received from Iolon Corporation. Jim Hale of the NLANR/MNA team at SDSC contributed significantly in preparation of the field trials and assisted in the performance tests. Thank you, Jim!

References

1. OC3MON: Flexible, Affordable, High-Performance Statistics Collection. Joel Apisdorf, kc claffy, Kevin Thompson and Rick Wilder. Proceedings of INET'97, 24-27 June 1997, Kuala Lumpur, Malaysia. http://www.isoc.org/isoc/whatis/conferences/inet/97/proceedings/F1/F1_2.HTM
2. Precision Timestamping of Network Packets. Joerg Micheel, Stephen Donnelly, Ian Graham. ACM SIGCOMM Internet Measurement Workshop, San Francisco, CA, USA. November 1-2, 2001. http://www.icir.org/vern/imw-2001/imw2001-papers/23.pdf
3. ITU-T Recommendation G.694.1: Spectral grids for WDM applications: DWDM frequency grid. February 27th, 2003. http://www.itu.int/itudoc/itut/aap/sg15aap/history/g.694.1/g6941.html
4. ITU-T Recommendation G.709/Y.1331: Interfaces for the optical transport network. February 11th, 2003. http://www.itu.int/itudoc/itut/aap/sg15aap/history/g709/g709.html
5. Understanding Fiber Optics: Jeff Hecht. 4th Edition. Prentice Hall. ISBN 0-13-027828-9
6. An ISP's view of an LSR run. Peter Lothberg, Svensk TeleUtveckling & ProduktInnovation AB**,** ESCC/Internet2 Joint Techs Workshop, July 18th – 21st, Columbus, OH. http://events.internet2.edu/2004/JointTechs/Columbus/sessionDetails.cfm?session=1505&event=218
7. lambdaMON prototyping Web page http://pma.nlanr.net/lambdamon.html