# The Lambda Grid – Supporting Mass Storage Systems and Technologies over a Dynamic Optical Network

Abdella Battou
*Lambda OpticalSystems*
*abattou@lopsys.com*

Michael Fox
*Lambda OpticalSystems*
*mfox@lopsys.com*

Leonard Chin
*Lambda OpticalSystems*
*lchin@lopsys.com*

## Abstract

*This paper describes an all-optical network architecture, Lambda Grid, for efficient, scalable, and cost-effective Mass Storage Systems. The architecture is based on wavelength services. A new concept of logical DWDM is introduced and its hardware implementation is described.*

## 1. Introduction

Bandwidth users always manage to find ways to fill what ever is available only to find they need more. Current, emerging, and dreamed applications consume bandwidth in increasing amounts, taking advantage of the unprecedented drop in price for network connections and the near-elimination of any distance tax for services. Rich application suites are emerging with their own QoS, QoR, and security requirements that take advantage of this lower cost bandwidth, customized to increasingly specialized user domains.

Traditionally, as each new class of service has been identified to support applications with common attribute requirements, it has been possible to provide an overlay network to uniquely address the new requirements. With the dramatic pace of application introduction, the increasing divergence of service requirements over all service attributes, and the formation of new user communities of interest requiring high bandwidth connectivity, it is no longer possible to provide a new network overlay for each specific class of service. Under this scenario, the network becomes too complex and network bottlenecks prevent realization of the full value of applications.

What is needed is a vast simplification of the network, a distillation of the network into two layers: a service layer on top of an optical grid layer, as shown in Figure 1.

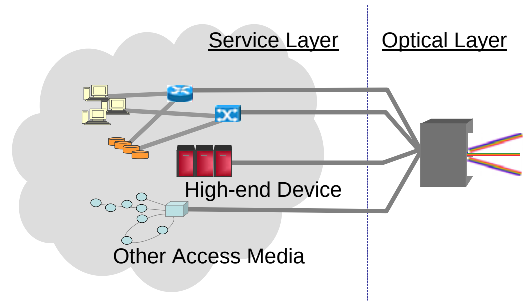



**Figure 1. Network simplification into service and optical layers**

## 2. Our work

In this paper we undertake the design of a complete metro optical network architecture called the Lambda Grid suitable to mass storage systems. We describe the architecture benefits, the services built on top of a Distributed Optical Control Plane based on GMPLS, as well as the hardware node architecture.

### 2.1 Lambda Grid – Bandwidth, Latency, Resiliency and Cost

It is often the case that an individual optical channel will pass through several nodes before reaching its destination. In today's network, these pass-through nodes are sites where grooming is performed and, because the signal is groomed, service level intelligence is also required to insure that QoS, QoR and security requirements of the application are adhered. This requires deployment of electronics at this site, electronics that can introduce packet processing as well as queuing delays, fixed and variable, and adds cost to the overall deployment. In the past, due to the high expense of transport, it made sense to design networks to minimize the number of optical channels by grooming the traffic into the fewest number of wavelengths at very node.

More recently, particularly with the advent of DWDM, the cost of transport as a contributor to total service cost has fallen. In this new environment it makes more sense to eliminate the need for intermediate grooming.

It is, after all, the grooming electronics with the service and network layer intelligence that is most difficult to modify when new services are introduced. Upgrading these elements or, in more extreme cases, introducing new parallel elements is costly and time consuming if it must be completed for each new application class that is introduced.

A radical new approach to networking is needed; one that keeps the service and network level intelligence at the edge, reducing the number of locations where upgrades are required when new classes of applications are introduced to only those endpoints where the application is desired. We call this new, simpler network the Lambda Grid. It must be available to form the foundation of networks where rapid service creation in a variety of rates and formats, with a diversity of QoS, QoR, and security requirements. The Lambda Grid is available today. Necessary innovations in optical switch technologies, DWDM, and GMPLS have matured to the point where it is possible to realize a network that can set up connections between any points, at any data rate, independent of format, temporarily or permanently, with unique QoS/QoR/security requirements.

Major attributes of this solution include

a) Dynamic connectivity with high bandwidth and low latency using GMPLS to enable fast service discovery and velocity.

b) Ability to redistribute bandwidth statically or dynamically as new storage comes on-line or more bandwidth is required for an immediate large transfer.

c) Low latency. There is no queuing in the path and minimum latency is guaranteed across the Lambda Grid.

d) No congestion. The Lambda Grid uses either static dedicated wavelengths between servers and disk arrays or on-demand wavelengths to satisfy irregular large transfers. Wavelength services isolate traffic and provide immunity against congestions. Large clusters for example can grab wavelengths on-demand to satisfy huge transfers without affecting regular daily jobs.

e) Last but not least, consolidation of all services over a single elegant, cost-effective, and scalable optical infrastructure.

## 2.2 Node Architecture

Any flat switching architecture ultimately faces the scalability challenge, and wavelength switching is no exception. While today's 10 Gbps or 40 Gbps wavelength circuits may seem more than sufficient for voice and commercial video applications, e-science experiments together with high-fidelity applications in intelligence and defense can easily require *several* wavelength circuits at a time. For such demanding scenarios, even transparent wavelength switching is challenged. Switching between $N$ fiber links with $W$ wavelength channels in each fiber requires an $NW \times NW$ wavelength cross-connect fabric: With wavelength densities approaching hundreds, and as network topologies migrate to dense mesh structures with high nodal degrees for ultimate survivability ($N$=5 or 6 being typical numbers), a flat wavelength cross-connect can easily require *a few thousand* optical ports, clearly posing a scalability problem.

*Combined waveband and wavelength switching* is the next-generation switching architecture that comfortably scales with increasing wavelength demand. A waveband refers to a group of contiguous or separate wavelengths in a fiber, and waveband switching refers to switching a group of wavelength circuits with overlapping routes as a group, using only two optical ports, in contrast with wavelength switching, where each wavelength circuit is switched individually. Figure 2 shows the hardware architecture of the LambdaNode 2000 designed to support wavebanding.

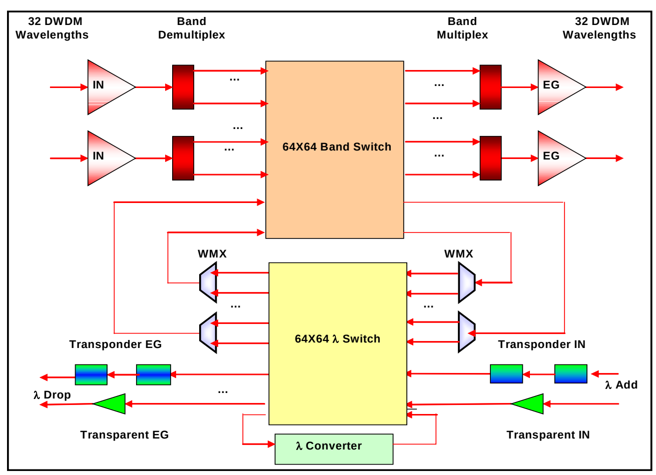



**Figure 2. LambdaNode 2000 hardware architecture**

Another way to view the banding as a service is that of logical DWDM. This architecture permits the deployment of logical DWDM enabling the capability to dynamically reconfigure the network in order to redirect bandwidth where it is needed. The flexibility and savings of wavebanding architecture are covered in [1].

## 2.3 Services

Generalized Multi-Protocol Label Switching (GMPLS) provides a unified framework to control packet, time slot, wavelength, waveband and fiber switching across multiple switching technologies (layers). Being a superset of the widely-deployed and proven MPLS technology, GMPLS is equally applicable to both packet-switched and circuit-switched networks. It is an IP-based control plane standard that enables all participating network nodes to discover the network topology and dynamically set up connections. The basic GMPLS functionalities can be summarized as

- Neighbor discovery and link management (Link Management Protocol - LMP)
- Routing with traffic engineering extensions (OSPF-TE, ISIS-TE)
- Signaling (RSVP-TE with GMPLS extensions)

Using the above three core functionalities, GMPLS offers several other applications that bring even more flexibility into the network;

- Recovery: A large array of service recovery options are available in GMPLS, ranging from *proactive* (protection) methods that designate the backup resources in advance, to the more dynamic and *reactive* restoration methods, which start the recovery process after the failure.
- Make-before-break: The ability to reroute a live circuit with minimal impact on the end user has been present in MPLS and by extension in GMPLS

GMPLS has been developed by the Internet Engineering Task Force (IETF), and is a mature standard now (RFC 3945, RFC 3471 and RFC 3473 are the central documents). Various extensions have been developed to extend its reach to different switching layers including SONET/SDH, wavelength, waveband, and most recently Ethernet as a ubiquitous layer-2 transport, with the last one still under study.

GMPLS is the driving force behind dynamic provisioning. A recent study boasts a 50% reduction in OPEX through GMPLS [2].

| **Service Levels** | **Quality** | **Resiliency** |
|---|---|---|
| Basic unprotected service | 10-12 BER | Service restored after repair |
| 1+1 protected service either revertive or non-revertive | 10-12 BER | 50ms switchover, failed path restored after failure |
| Auto-restored service | 10-12 BER | Service restored within 5 seconds after failures |

These services are accessible through the network management system and part of a software suite called LambdaCreate.

## 2.4 Lambda Grid Virtual SAN Application

Today, one the biggest challenges faced by most enterprises is how to deal with an ever-growing demand for storage capacity (60% annually) in a cost-effective way. The majority have started deploying high-speed dedicated storage networks, SANs, to centralize all storage within the data center into a virtual pool that offers high performance and high availability.

Furthermore, regulations such as the Sarbanes-Oxley Act are pushing enterprises to look into remote data vaulting to replace their physical offsite data storage. Larger enterprises with multiple sites across a wide-area-network (WAN) are consolidating all their services over a single WAN. The above requirements translate into the challenge of extending SANs over WAN seamlessly, reliably, and cost-effectively. These new requirements have to be implemented with in mind that data is mission critical and any downtime can have serious business consequences. Any delay in the retrieval of stored data can cause downtime and consequently impact business.

In summary, enterprises are looking for scalable, low delay, high bandwidth, and resilient connections from server to storage either from SAN to SAN or from server to remote data vaulting. These connections need to be dynamic to mitigate any failures in the WAN or remote sites.

### 2.1.1. VSAN Architecture

We propose the Lambda Grid as the solution because

(a) The Lambda Grid is scalable. The Lambda Grid offers great performance scalability in 3 dimensions – distance, rate, and number of wavelengths. Today, Dedicated or on-demand wavelengths can be deployed between any server and any storage device within 1000 Km. This distance can be easily increased by introducing newer amplifiers such as Raman Amplification or a hybrid EDFA-Raman amplification. Furthermore, the LambdaNode 2000 is a fully transparent all-optical switching platform capable of 10Gbps and 40Gbps wavelengths. With 32 wavelengths per fiber, 7 fibers per bay, and 32 local Add/Drops wavelengths, it has a capacity of 2.4Tbps.

(b) The Lambda Grid enables VSANs that are independent of storage size, number of servers, distance between servers and storage.

(c) It is cost-effectively managed with a single centralized point of management using LambdaCreate Service Delivery System (SDS).

(d) Is it highly available with different quality of resilience. 1+1 provides built-in redundancy and dynamic access to it. Auto-restoration provides dynamic cost-effective access to replicated services.

(e) It is protocol and bit rate independent and as such will support heterogeneous servers and data storage, as well as any protocol whether it is Fibre Channel or the new incoming Infiniband.

(f) The transparency of the Lambda Grid built around LambdaNode 2000 is one of the key attributes to this solution as it will accommodate seamlessly future storage technologies.

Attach the various data centers to the Lambda Grid built around the LambdaNode 2000 and allocate the appropriate number of wavelengths to each data center to realize a virtual SAN (VSAN) as shown in the figure 4 below. The Lambda Grid becomes the Virtual SAN that provides the scalability, performance, high bandwidth and very low latency to satisfy not only SAN requirements but also the requirements of consolidating the voice, data, and video enterprise networks.

Perhaps the most significant advantage of the LambdaNode 2000 is the truly all-optical nature of the switch. Traditionally, each node in a given network represents a mix of disparate network elements, some optical and some electronic, that combine to create a system that is not satisfactorily scalable, not satisfactorily manageable, and which makes timely provisioning virtually impossible.

## 3. Conclusion

Optical networks provide a scalable and cost – effective technology. A dynamic control plane such as GMPLS enables rapid redeployment of bandwidth and prompt recovery of services in case of failures. In the past data centers have implemented static disaster recovery procedures using DWDM. We describe a dynamic and scalable switched architecture using Logical DWDM and dynamic services such as 1+1 service or an auto-restoration service to enable storage virtualization.

## 5. References

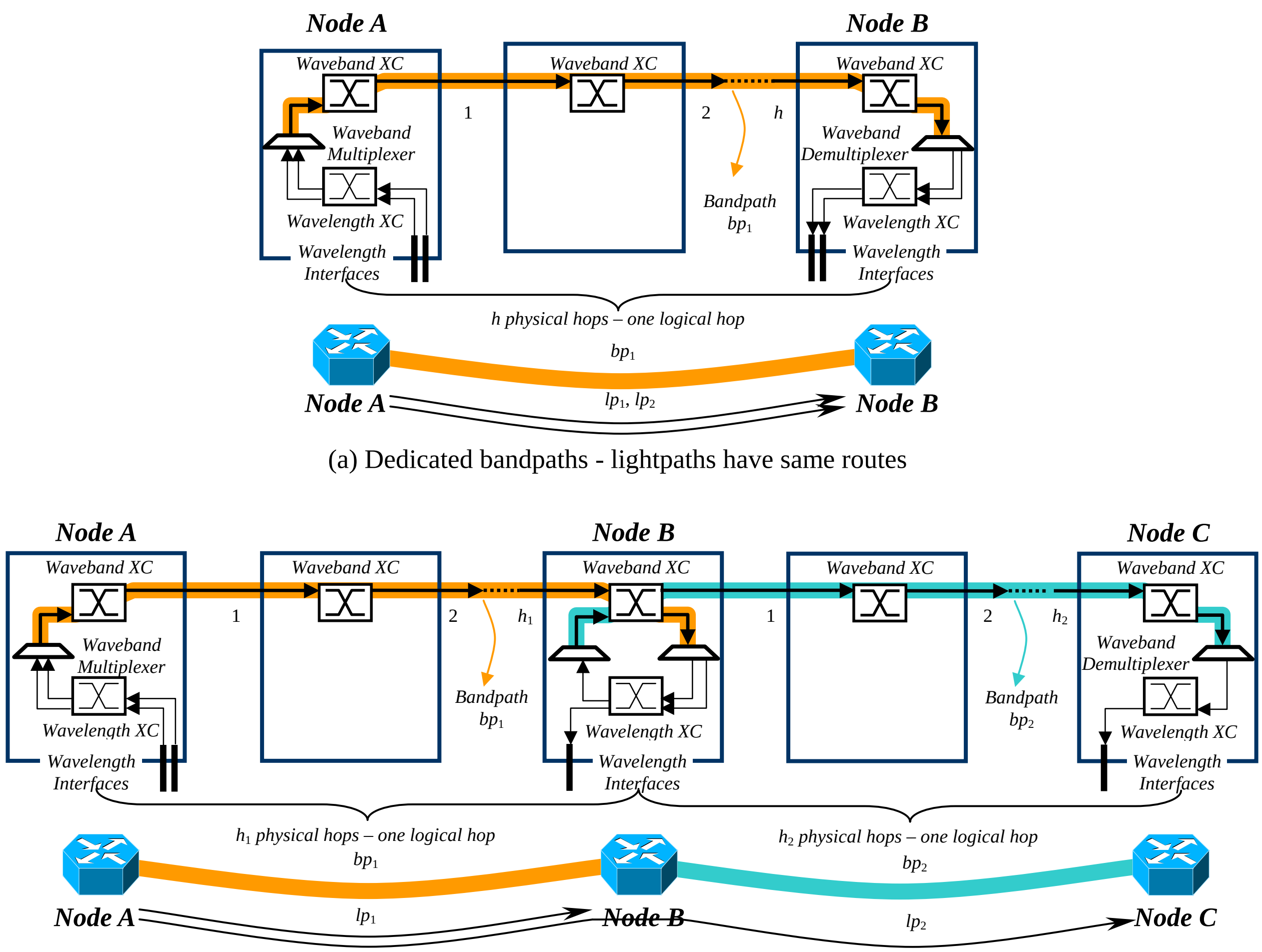

[1] Payam Torab, Virginia Hutcheon, David Walters, and Abdella Battou, *Waveband Switching Efficiency in WDM Networks: Analysis and Case Study*, OFC 2006.

[2] Sandrine Pasquali, Andreas Kirstadter, Andreas Iselt, and R. Chahine, (Siemens AG); Sofie Verbrugge, Didier Colle, Mario Pickavet, and Piet Demeester (Ghent University), *Influence of GMPLS on Network Provider's Operational Expenditures: A Quantitative Study*, IEEE Communications Magazine, July 2005 Vol. 43, N0.7.

[3] IETF RFC 3945, *Generalized Multi-Protocol Label Switching (GMPLS) Architecture*.

[4] IETF RFC 3471, GMPLS *Signaling Functional Description*.

[5] IETF RFC 3473, Signaling *Resource Reservation Protocol – Traffic Engineering (RSVP-TE).*




(a) Dedicated bandpaths - lightpaths have same routes

(b) Shared bandpaths - lightpaths have different endpoints and partially overlapping routes

**Figure 3. Logical DWDM through waveband switching**

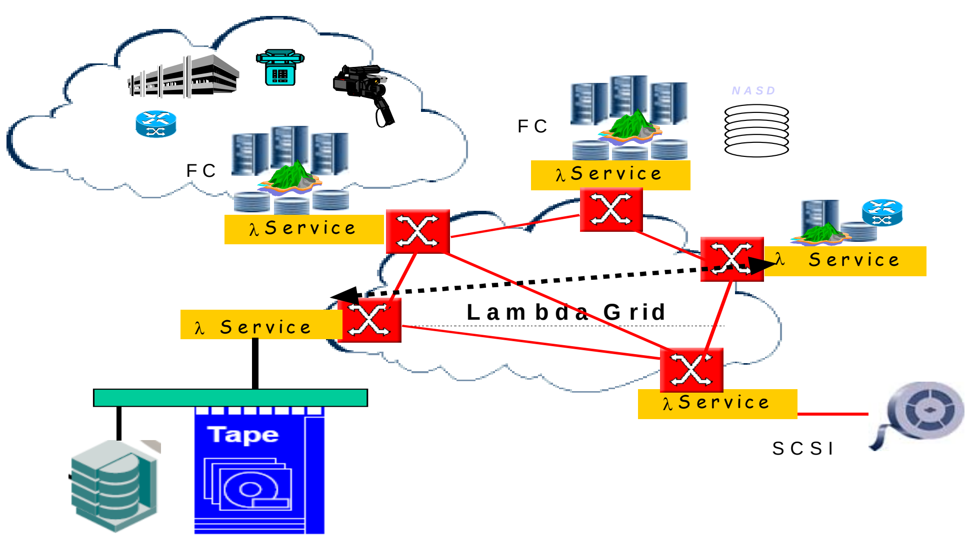



**Figure 4. Virtual SAN**