EFFICIENT NETWORK SWITCHING HIERARCHY

Stephen French*, Dr. Jean-Francois Labourdette, Dr. Krishna Bala and Patricia Miller-Pittman
Tellium Inc., 2 Crescent Place, Oceanport, NJ 07757
*sfrench@tellium.com, Tel: (732) 923-4297, Fax: (732) 483-3007

Abstract

History and economics have set the stage for the move to grooming at 2.5 Gbps granularity in the core of optical mesh networks. The evolution of core grooming speeds to the STS-48 granularity level is a continuation of a cycle that began with the introduction of narrowband cross-connects into the network over 20 years ago. Since then, cross-connect grooming granularity has increased from 64 Kbps, to 1.5 Mbps, 51.8 Mbps, and finally to 2.5 Gbps with the introduction of ultra-broadband cross-connects in 2000. History has shown that switching and grooming speeds increase as transport speeds increase (Figure 1). The primary reasons for scaling grooming speeds with transport speeds are improved network scalability, performance and, most importantly, cost. History has also shown that product designs are specialized - in terms of capacity, scalability, reliability, performance and cost, for particular segments of the network. Network segmentation and specialization is why products designed for the edge do not perform well in the core. Based on the facts stated in this paper, many carriers and industry experts have concluded that an optimal network architecture supports hierarchical transport and grooming speeds, with faster switching (granularity) speeds in the core of the network than at the edge. A hierarchical architecture, with segment specialized products, will persist in the network future just as it has in the past.

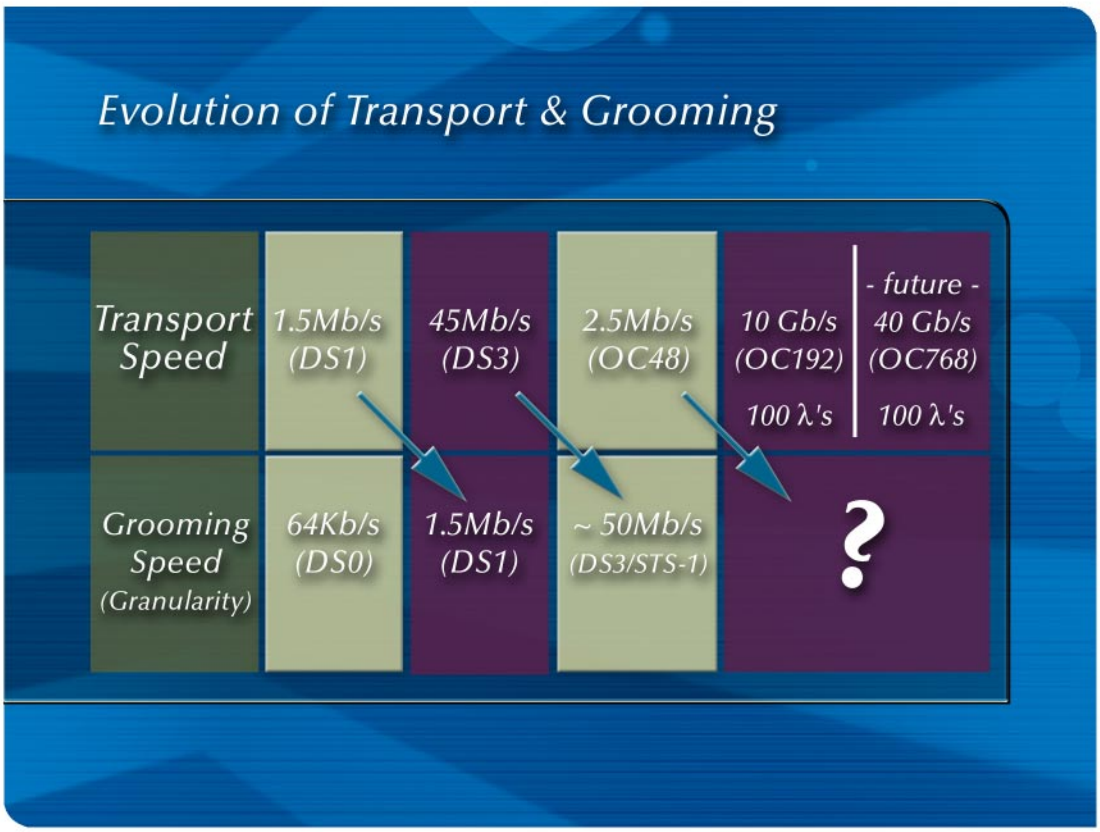

Figure 1: History has shown that switching and grooming speeds increase as transport speeds increase

1. Introduction

Large networks have always been organized in multi-level hierarchies. It has been a service provider's dream to accommodate all services at all rates with a single box that is scalable, manageable and low-cost. However, practical considerations such as hardware and software scalability and manageability have always led to hierarchical network architectures, taking advantage of optimizing each layer independently. "All-purpose boxes" may be well suited for enterprises and some metro applications, but not for core applications that require specialized, carrier-grade products.

In the past few years we have witnessed tremendous growth in traffic volume, driven primarily by the exponential growth in data traffic. In order to accommodate the growth in traffic volume and shift in traffic mix, service providers are upgrading their core transport infrastructure from traditional SONET/SDH rings to optical mesh networks. As carriers begin to build their next-generation transport infrastructure, they need to ask themselves – what is the right granularity of grooming for the core of my optical transport network? More specifically, should carriers continue to groom traffic at STS-1 (51.8 Mbps) granularity in the core, as they have since 1988 when core transport speeds were at DS3/STS-1 (45/51.8 Mbps), or is it time to move to core level grooming at STS-48 (2.5 Gbps) granularity now that core transport speeds are at 10 Gbps? (*Throughout this paper we use the term STS-48 [or OC-48 when referring to the corresponding optical transport speed] to indicate either STS-48 or STM-16 electrical signals at 2.5 Gbps).*

In this paper, we argue that the right granularity for grooming in the next-generation core optical network is STS-48 (2.5 Gbps), while STS-1 (51.8 Mbps) grooming is better suited for the edge of the network and legacy applications. In today's TDM world, both wideband (VT1.5 - 1.5 Mbps) and broadband (STS-1 – 51.8 Mbps) grooming are required in particular segments of the network to manage different traffic levels – serving distinct but complementary roles in the network. Similarly, as transport speeds move primarily to 10 Gbps and eventually scale to 40 Gbps, and DWDM channel counts continue to increase, we expect STS-1 and STS-48 grooming switches to also perform distinct but complementary roles at the network edge and core, respectively (see Figure 2).

This paper is organized as follows: In Section 2, we present a brief history on hierarchical network grooming/switching. Next, in Section 3, we discuss what types of services and applications are driving core traffic growth. Section 4 covers the evolution of transmission speed and Section 5 discusses core scalability and manageability. Section 6 analyzes the impact of different grooming granularities on provisioning, restoration and management performance. Section 7 is the conclusion, and reviews the economic, performance and scalability issues supporting the migration to 2.5 Gbps grooming/switching in the optical network core.

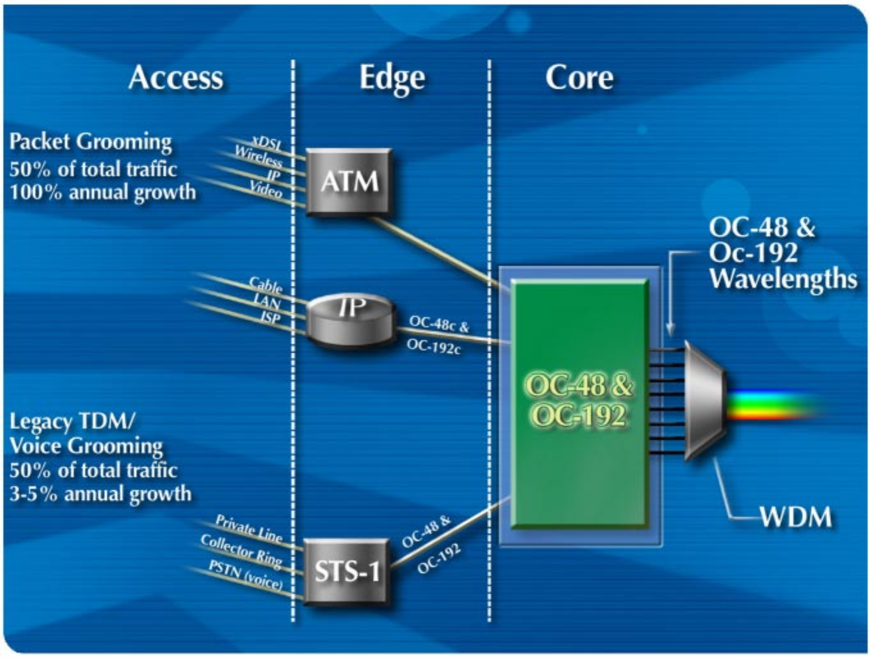

Figure 2: Optimal grooming in optical mesh networks: Edge AND Core

2. History of Hierarchical Network Grooming

Although 'grooming' has received significant marketing coverage in 2001, one needs to remember that grooming devices have been utilized in networks for over 20 years. Grooming has and will continue to be a vital requirement in carrier networks. Grooming has evolved over the past two decades, starting with narrowband (64 Kbps) cross-connects in 1981, then wideband (1.5 Mbps) and broadband (45 Mbps) cross-connects in the late 1980s, and finally ultra-broadband (2.5 Gbps) cross-connects, referred to as optical core switches, which were introduced in 2000 (see Figure 3).

Electronic *narrowband cross-connects* (64 Kbps granularity) were first introduced in 1981. By the mid 1980s, as the core network bandwidth grew, it became clear that narrowband (1/0) cross-connects were not suitable for backbone traffic engineering, since the backbone was running at DS3 (45 Mbps) rates.

Electronic *wideband cross-connects* (1.5 Mbps granularity) were introduced in the late 1980s to replace the functionality of DS1/DS3 multiplexers, whose connections were hard-wired to both the narrowband DS0 and broadband DS3 core networks.

Electronic *broadband cross-connects* (45 Mbps granularity) were also introduced in the late 1980s to replace manual patch panels as the vehicle for routing office circuits onto long haul transmission equipment and for routing pass-through circuits off of one long haul system and onto the next span.

As history has shown, grooming granularity has increased to support traffic engineering of backbone networks and to offer higher service rates in the network. When DS3 and lower rates dominated the network, carriers originally questioned the economics of broadband (45 Mbps) cross-connects. The broadband cross-connects' popularity increased with carriers as they accounted for the operational penalties of sending unchannelized DS3 services through sub-rate grooming devices and as their customers began to experience the superior performance of network-based restoration services.

Broadband cross-connects grooming at the DS3 (45 Mbps) level, or equivalent SONET rate – STS-1 (51.8 Mbps), were introduced in the late 1980s when core transport speeds were at a similar rate. With core transport speeds now at 10 Gbps, nearly 200 times faster, continuing to groom the core at the same level as in the late 1980s seems quite inadequate.

History has shown that grooming speeds increase as transport speeds, average circuit size and overall network traffic increases. Carriers will need to evolve their core grooming speeds to 2.5 Gbps for the same reasons as they moved from narrowband to wideband and broadband crossconnects:

- Emerging application drivers – what feeds the core
- Evolution of transport speeds
- Scalability and manageability
- Performance – provisioning, restoration and management
- Total network cost

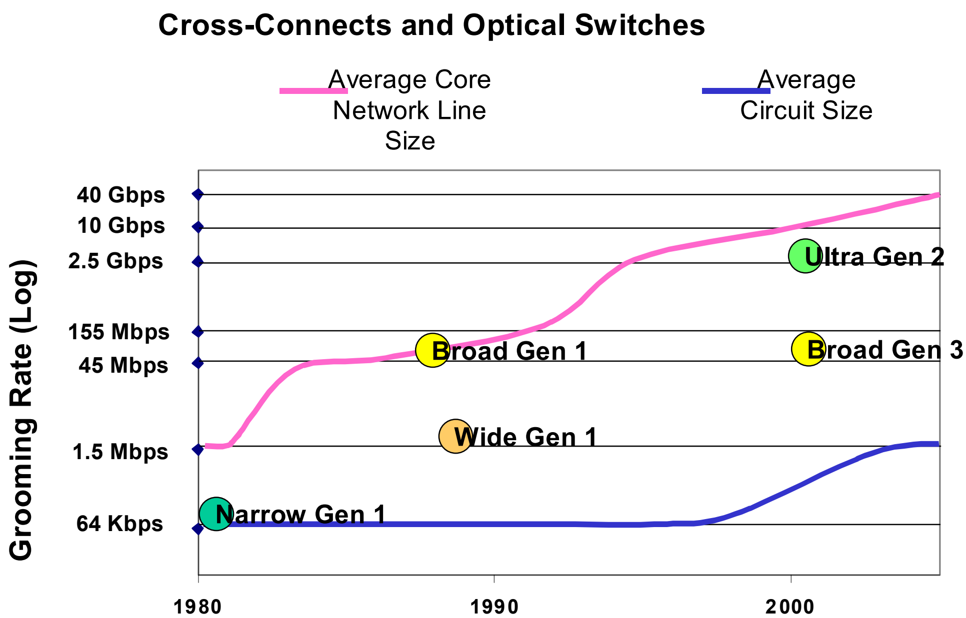

Figure 3: The evolution of transport and grooming speeds

3. Driving Core Traffic Growth

The dominant traffic carried in today's network is evolving from legacy voice and leased line services to data services, pre-dominantly IP. According to a recent Dell'Oro report, OC-48 and higher-speed router port shipments are expected to more than quadruple between 2001 and 2004 (see Figure 4). TDM aggregation switches, optimized for legacy voice and leased line services, and acting as edge devices, groom signals at lower bit-rates, VT1.5 (1.5 Mbps) and STS-1 (51.8 Mbps), and typically feed into the core at rates of OC-48 and above. Furthermore, high-speed trunking between edge service platforms (IP routers, ATM switches, DCSs, SONET/SDH ADMs, MSPPs, & GigE platforms) for sub-OC-48 services is quickly moving to OC-48 and OC-192 speeds. Additionally, the wavelength services market (metro & longhaul) is forecasted to increase sharply over the next five years, reaching almost $8 billion by 2005. [1] Consequently, the natural rate of grooming granularity in the core is STS-48 (2.5 Gbps).

Data traffic is statistically groomed and packed by IP routing devices into concatenated OC-48 and OC-192 trunks - without the need for STS-1 grooming. There is no benefit gained from sending a concatenated 2.5 Gbps or 10 Gbps circuit through an STS-1 grooming broadband switch, and in fact, it may harm the circuit quality and drive up the operational costs. This argument follows through with emerging Gigabit and 10-Gigabit Ethernet services. With the majority of network traffic being data, and with Internet traffic still growing at over 100% per year (RHK, May 2002), it has become increasingly apparent that STS-1 grooming broadband switches do not belong in the core.

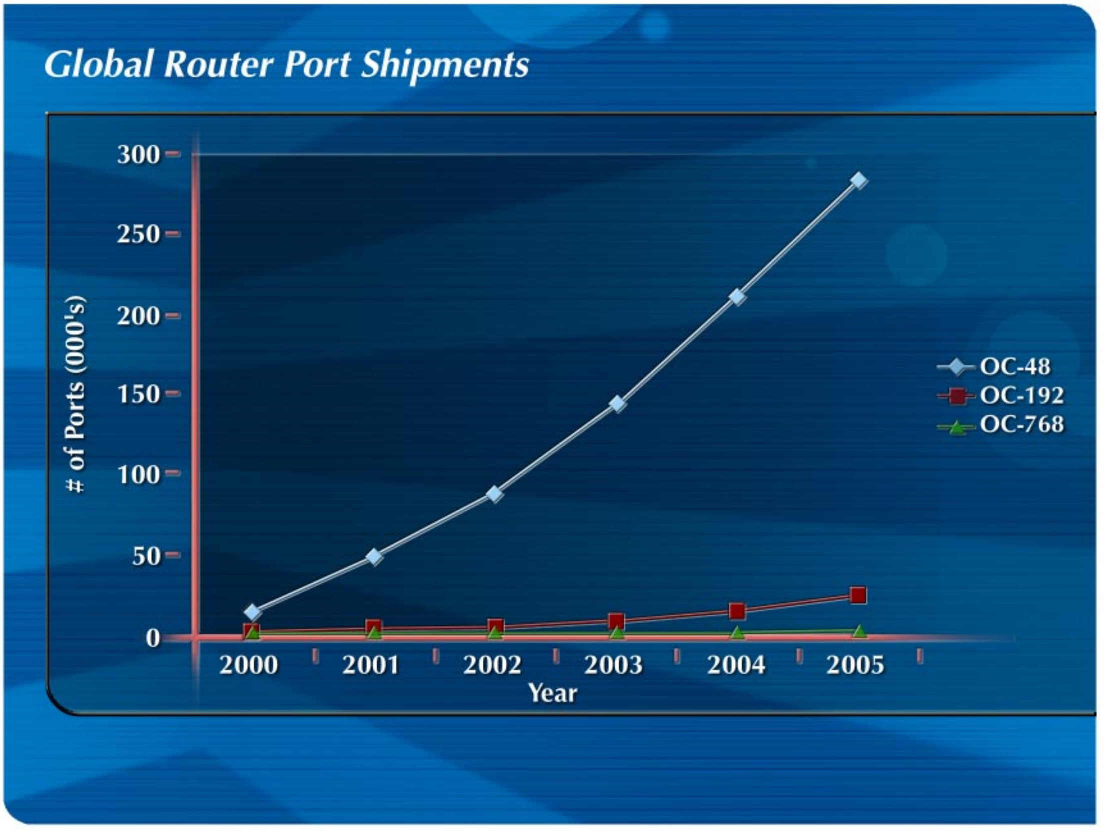

Figure 4: Router port shipment forecast (Source: Dell'Oro)

4. Transmission Speed Evolution

Transmission equipment at the core of the network has experienced a dramatic increase in both channel counts per fiber and transmission speed per channel over the last few years. The core of the transmission network is evolving from carrying 10's of OC-48 wavelengths per fiber to carrying 100+ OC-192 channels per fiber using long haul DWDM systems. Furthermore, the TDM hierarchy is not at its end. It is now clear that the network will migrate to 40 Gbps per wavelength. At STS-48 granularity, a 40 Gbps signal is comprised of 16 blocks, as opposed to 768 blocks per 40 Gbps signal with STS-1 granularity. This difference illustrates the complexity of grooming at the STS-1 layer - approximately 50 times that of managing at the STS-48 layer. A KMI research report (see Figure 5) reinforces the fact that the network core operates at 2.5 Gbps and higher speeds, showing that sales of OC-48 and higher-speed DWDM systems will represent over 95% of shipments between 2000 and 2005. With the evolution in transport speeds to OC-48 and higher, core grooming speeds are naturally driven to the STS-48 level.

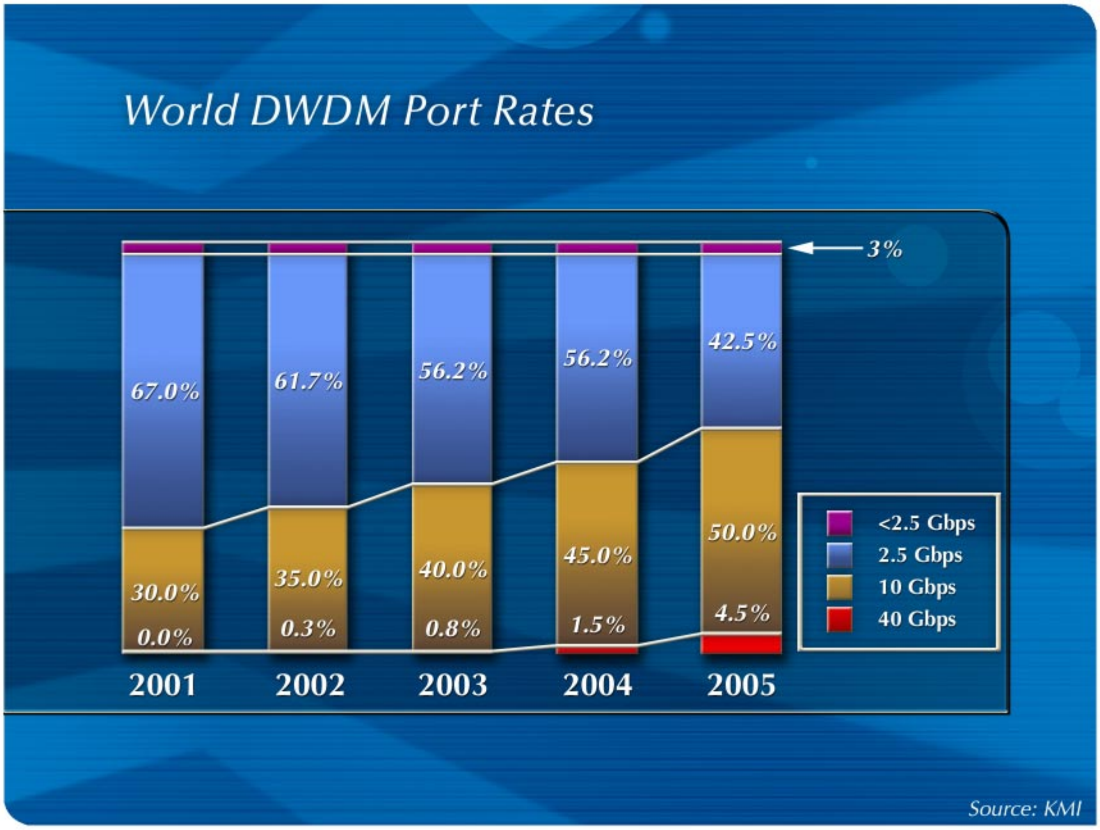

Figure 5: Worldwide DWDM transceiver forecast - metro and longhaul (Source: KMI)

5. Core Scalability and Manageability

History has repeatedly demonstrated that core grooming speeds increase in-step with increases in core transports speeds for reasons of scalability and manageability. As traffic volume grows and the size of the network increases, core grooming granularity naturally increases to keep network complexity under control. The complexity of managing the core of a large optical mesh network at edge grooming rates of STS-1 is at least 48 times more complex compared to an optical core that supports STS-48 grooming (see Figure 6). Grooming at STS-48 granularity allows operators to scale their network, taking advantage of higher port count STS-48 switches, and dealing with the right-size "block" as total traffic and network size increase.

Scaling limitations of STS-1 switches are mostly due to the tremendous increase in software complexity involved with handling STS-1 signals. This includes path computation across the fabric and subsequent management of those paths, which becomes at least 48 times more complex for an OC-48 equivalent connection through an STS-1 fabric compared to an STS-48 fabric. Maintaining the synchronization among all the STS-1 signals traversing the fabric also becomes extremely difficult as the size of the switch increases. The challenge for OC-192 (10 Gbps) is at least four times higher (192 STS-1s), which is why STS-1 switch vendors have had great difficulty in developing and deploying OC-192 interfaces that support STS-1 grooming granularity. OC-192 interfaces for STS-48 based switches have been shipping to customers since Q1 2001.

Today, there are core STS-48 grooming products that are shipping with 512 OC-48 ports, twice the size of any STS-1 switch currently on the market. Furthermore, these core STS-48 grooming switches can scale to 8,192 ports of equivalent OC-48 capacity. This is a factor of two to three times larger than the scalability of STS-1 switches. Therefore, as traffic scales beyond STS-1 switch capabilities, carriers will be forced to deploy several STS-1 granularity switches per office in a one-tier STS-1 architecture, wasting large portions of effective capacity and losing potential revenue streams due to inter-machine ties *(carriers are familiar with this problem from their deployment of wideband cross-connects)*. As overall traffic grows and the traffic mix becomes more heavily weighted towards OC-48 and above data rates, a 2-tier architecture made of STS-1 based switches at the edge and STS-48 based switches at the core, will be more efficient and cost-effective than a single-tier STS-1 based architecture. Based on both internal and independent third-party network studies, overall network capital savings of a two-tier STS-1/STS-48 architecture over a single-tier STS-1 architecture can start from 24% to 36% under reasonable network assumptions of size and traffic. [2] & [3]

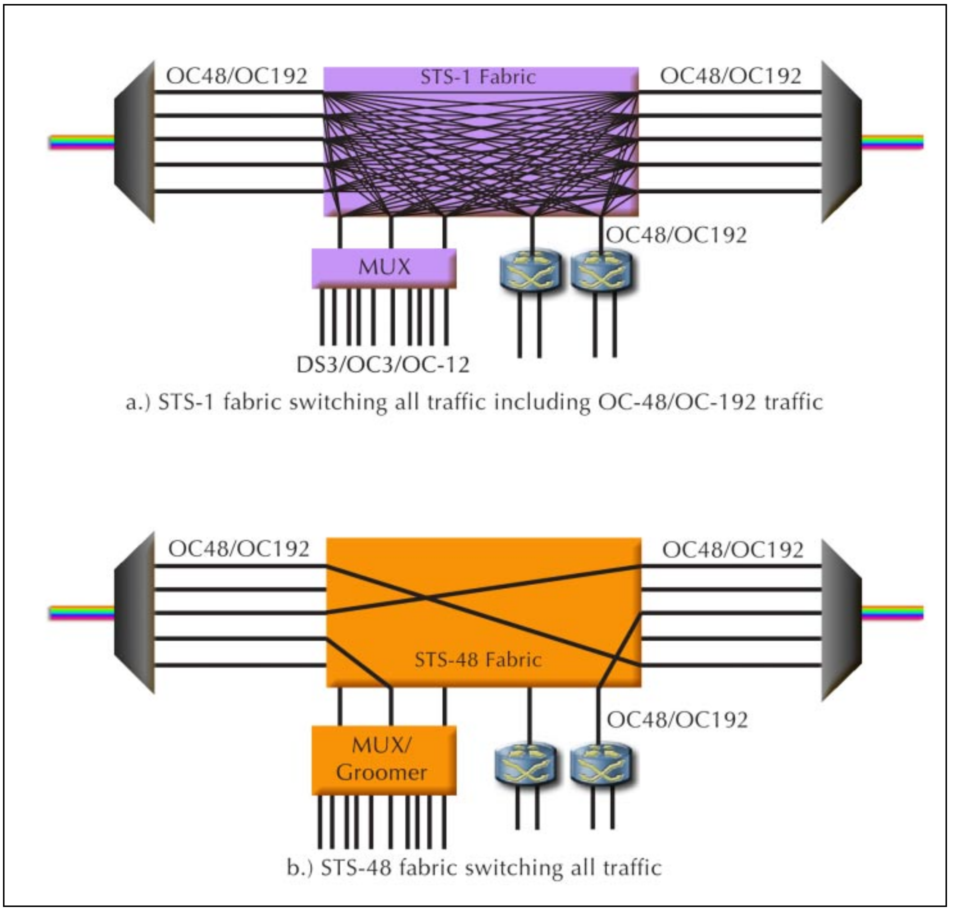

Figure 6: The inherent complexity of an STS-1 based switch fabric

6. Provisioning, Restoration and Management Performance

Critical real-time network functions such as mesh (1) provisioning (2) restoration and (3) network management can be performed much more efficiently and with much better performance at STS-48 granularity. Core grooming at the STS-48 level allows operators to support fast and capacity-efficient shared-mesh provisioning, restoration and management.

Provisioning - Mesh routing of lightpaths requires diversity of primary and restoration paths as well as appropriate sharing restrictions on restoration capacity to offer restoration guarantees against link and node failures. Optimal mesh routing of thousands of OC-48 lightpaths can be effectively handled in a shared-mesh optical STS-48 granularity core network. On the other hand, the complexity of routing hundreds of thousands of STS-1 equivalent signals in a single-tier STS-1 network architecture, while trying to satisfy diversity constraints and optimize sharing, is a much larger task. Time and memory requirements are 48 times greater per connection, leading to minutes (instead of seconds) and requiring Giga-bytes of memory to compute the route (primary and restoration paths) of a lightpath. Having to sequentially route all STS-1 connections that constitute an OC-48 could potentially increase provisioning times by a factor of 48.

Restoration - Fast, capacity-efficient shared-mesh restoration is a key benefit of an optical core mesh network. [4] The performance of mesh restoration is greatly impacted by the number of signals that have to be restored during a link or node failure event. Where an STS-48 core network would restore all services by handling the 10's of failed OC-48 connections on a link or node (100's of connections for large points-of-presence), an STS-1 granularity switch would have to restore 48 times more connections. Such an increase in the number of signals to be restored would drive the restoration time for shared-mesh restored services from less than 200 msec in an STS-48 network to several seconds in an STS-1 network as shown by our simulation studies. [5]

Using an analytical approximation, the effects of lower switching granularity on shared-mesh restoration performance can be estimated. [5] An analytical approximation, based on simulation results, was developed for a 50 node network switching at STS-48 granularity. In Figure 7, analytical approximation results for the same 50 node network are shown, but with the assumption that switching is performed at the lower STS-1 granularity. This means that, for every OC-48 lightpath to be switched, 48 switching operations have to be performed. Figure 7 shows the approximation results for STS-48 switching compared to STS-1 switcing with the aforementioned bundling factors. As can be seen from Figure 7, restoration using STS-48 granularity can be significantly faster than restoration using STS-1 granularity.

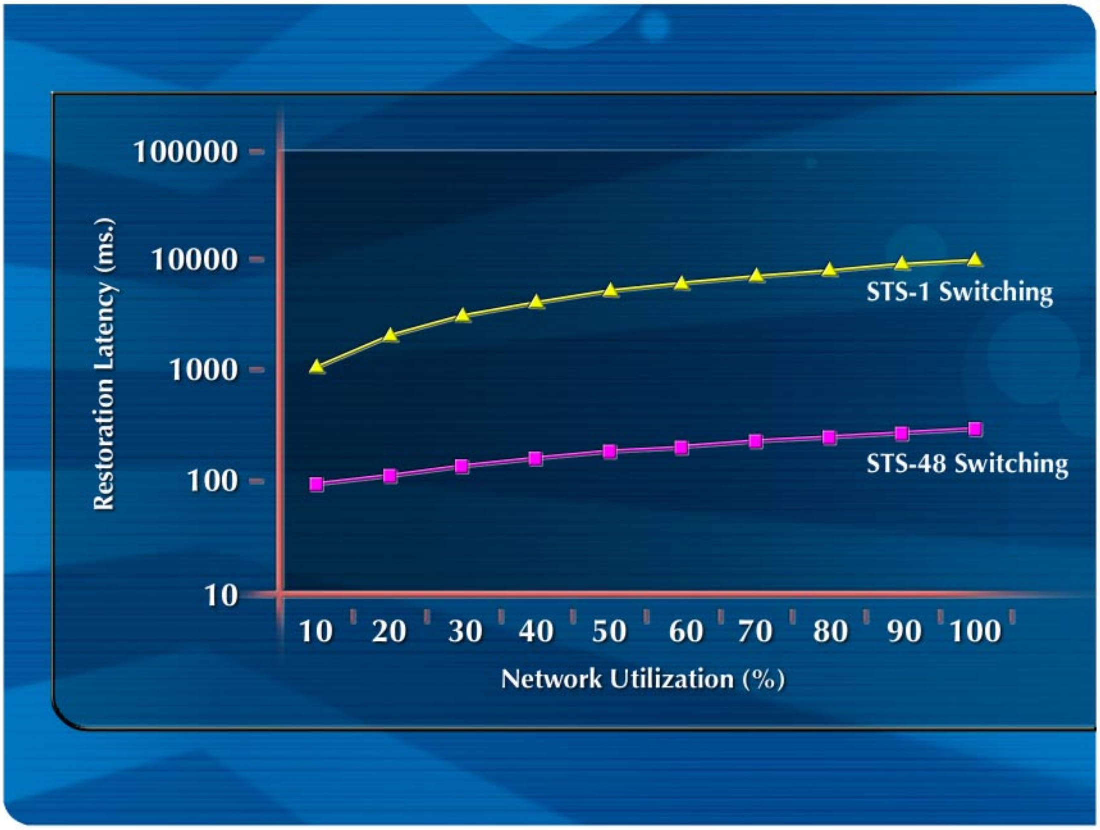

Figure 7: Analytical results comparing STS-48 & STS-1 (switching/grooming) restoration performance for a hypothetical 50 node network

The difficulty of doing shared-mesh restoration in an STS-1 based network could force connections to be protected in a dedicated mesh (1+1) configuration. This would drive the required network capacity at least 50% higher than would be required when shared-mesh restoration is used. [6] With an STS-48 granularity core network, carriers can provide better performance using less dedicated protection than with an STS-1 core, resulting in significant cost savings (see Table).

Optical Switch Fabric Granularity	*Shared-Mesh Restoration Time*	*Protection Type Required for Msec Restoration*	*Percent Idle Versus Working Bandwidth*
STS -1	Seconds	Dedicated	110%-130%
STS-48	Milliseconds	Shared	30% to 70%*

* Lower numbers achieved using pre-emptable traffic on the shared-mesh backup path

Table 1: Protection capacity requirements for STS-1 and STS-48 based core mesh networks

Network Management - Network management scaling limitations and performance issues arise when managing a large number of STS-1 connections in a mesh architecture. An STS-48 core network provides a more scalable network management solution than that of an STS-1 network. For topology and configuration management, an Element Management System (EMS) is typically required to maintain a port-state database, which stores the status and availability information for all of the Physical Termination Points (PTPs) and Connection Termination Points (CTPs) within the core network. If the core is operated at the STS-1 level, there are 48 times more CTPs compared to an equivalent STS-48 network, thus resulting in 48 times more storage overhead. Similarly, excessive communication overhead would also be incurred for service management (i.e., lightpath and mesh sharing databases) and performance monitoring (PM) information, resulting in an extraordinary amount of storage capacity and potentially impacting performance. Our analysis indicates that 250 Giga-bytes of PM data would be sent to and stored in the EMS every day in a 100 node STS-1 network. This compares to 5 Giga-bytes for a core network that uses STS-48 grooming switches.

7. Conclusion

The evolution of core grooming speeds to the STS-48 granularity level is a continuation of a cycle that began with the introduction of narrowband cross-connects into the network over 20 years ago. Since then, cross-connect grooming granularity has increased from 64 Kbps, to 1.5 Mbps, 51.8 Mbps, and finally to 2.5 Gbps with the introduction of ultra-broadband cross-connects in 2000. The following key factors are driving the need to migrate the core grooming speed to STS-48 granularity:

- Bandwidth requirements continue to grow, with Internet traffic growth rates exceeding 100% per year.
- Access bandwidth grooming at STS-1 granularity is best suited for the edge of the network and legacy TDM applications.
- Emerging applications and services, such as IP, are feeding the core at concatenated OC-48 and OC-192 rates and do not require STS-1 grooming.
- Core transport speeds are at OC-48 and OC-192 today and will be at OC-768 in a few years.
- Using STS-48 granularity switches reduces equipment and network complexity, increases reliability, and provides the necessary equipment and network scalability and manageability.
- Reduced complexity improves equipment and network (provisioning, restoration and management) performance.
- Overall network capital savings of a two-tier network (STS-1 and STS-48) over a one-tier network (STS-1 only) can start from 24% to 36% under reasonable network assumptions of size and traffic. [2] & [3]

History has shown that core grooming granularity has increased as application and transport speeds, and network capacity have. The reasons for this are clear and simple – reliability, scalability, performance and most importantly, cost. For carriers planning to build out their optical networks, STS-48 is the right-size “block” for grooming in the core.

8. References

[1] D. Saha & S. French, “Mining Services From the Optical Layer” paper, to be presented at NFOEC 2002, Dallas, TX, September 2002.
[2] TeleChoice, “The Business Case For Optically Switched Networks” web seminar presentation, http://www.tellium.com/webseminars/, March 2002.

[3] Chris Olszewski, “Two Tier Economics” paper, to be presented at NFOEC 2002, Dallas, TX, September 2002.
[4] Ahmet A. Akyamac, “Ring-Speed Restoration and Optical Core Mesh Networks”, to be presented at Networks and Optical Communications 2002, Darmstadt, Germany, June 2002.
[5] Ahmet A. Akyamac, “Optical Mesh Network Modeling: Simulation and Analysis of Restoration Performance” paper, to be presented at NFOEC 2002, Dallas, TX, September 2002.
[6] S. Samieian, Z. Bogdanowicz, R. Ramamurthy, S. Chaudhuri, K. Bala, “Capacity Requirements of Ring, 1+1 and Mesh Configurations”, *Lightwave Magazine*, August 2001.