40G and 100G Overview

This report deals with the Client and Transport i/o including the Optical Interworking area.

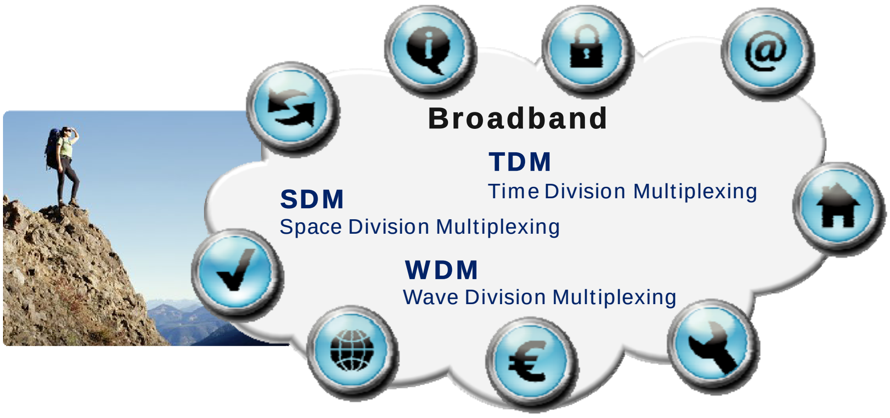
The Author:	Tony Breach, Optical Network Manager
Date:	20-03-2009
Version No.:	1.0

Table of Contents

40G and 100G Overview

1 Introduction

What is the driving force for the global information community with regard to broadband communication advancement – demand for higher bandwidth with a twist of QoS at a lower cost compared bit by bit to yesterdays broadband systems. In general the discussions are focused on the capacity across the board - however this can be achieved in many ways, but focusing on the physical connectivity it is necessary to distinguish between Space Division Multiplexing (SDP), Time Division Multiplexing (TDM) and Wave Division Multiplexing (WDM).



The SDM approach is based on keeping the bit rate the same but use more fibers – durable for short geographical distances and it is the most straightforward upgrade alternative. However it has two main drawbacks. It requires more fibers and for long geographical distance optical amplifiers for each fiber set – driving the cost up.

The TDM approach is based on increasing the transmission bit rate on the same fiber – durable for intermediate distances (up to 80 km). Unfortunately the two major photonic system impairments chromatic dispersion (CD) and polarization mode dispersion PMD are reducing the reach when increasing the bit rate. However several technological achievements like dispersion compensation, RAMAM amplification can reduce the impact of those impairments. The biggest challenge for TDM is how to increase the bit rate.

The WDM approach maintains the bit rate but is adding more wavelengths at the same or higher bit rate. WDM does suffer some disadvantages, like it requires more advanced optical amplifiers with a flat gain in order to support all waves, expensive lasers and receivers, advanced optical add drop multiplexers - although TDM multiplexing components has the same drawback regarding cost.

Clearly, it is possible to deploy a combination of SDM, TDM and WDM as well, as they are complementary.

This paper provides an overview of current 40G and future 100G TDM and WDM solutions where the main emphasis is on 100G.

In addition the political environmental agenda is demanding that everything has carbon dioxide sustainability - this is briefly discussed with regard to 100G systems.

The 100G discussion is complemented with the introduction of Generalized Multi-Protocol Label Switching (GMPLS) for the WDM control plane leading to an automatic adaptive WDM network – potential leading to new architectural design approach for all network layers in relation to automatic restoration.

The closing of this paper is a discussion about where does 100G fit into the Next Generation Network (NGN) pyramid structure supporting the cost reduction and environmental sustainability.

2 Driving the Capacity needs

The major drivers for 10G and now 40G deployments was the need to connect core routers, high capacity demanding users communities like radio astronomers and high-energy physicist; better economics versus yesterday's high end transmission capacity and network capacity exhaust. These have been the dominant cited factors most often given when moving towards higher transmission speeds.

So why not utilize the WDM layer (Extended C-band) which provides a lot of potential with a transfer rate close to 1 Tbps based on commodity 10G wavelength (NRZ Modulation) at a relatively low cost. The straight answer to this is that the transport network optimization is cost reduction driven, as outlined below and illustrated in Figure 2-1.

- Reducing the number of wavelength (managed entities) leading to reduced network complexity
- Better fiber and wavelength utilization
- Reducing network cost by increasing statistical multiplexing efficiency
- Future proof systems, scalable to manage the expected bandwidth demand

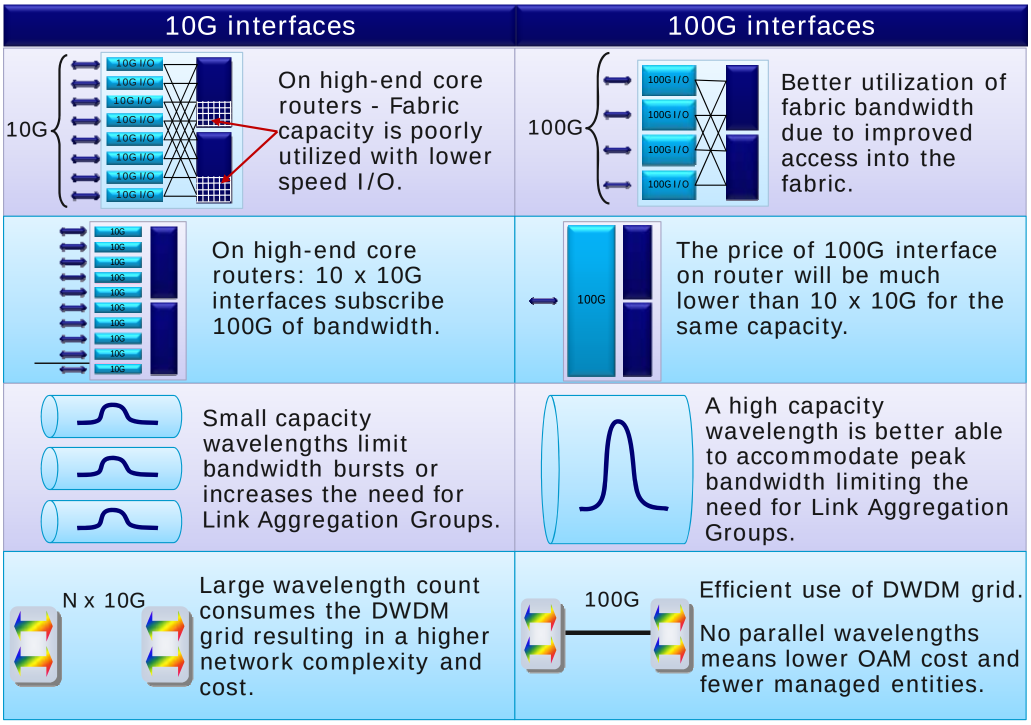


Figure 2-1: 10G vs. 100G pros. and cons.

Although there is always a trade off from all these statements – should the network operator go for more wavelengths or just increase the wavelength capacity, which narrows down to CAPEX and OPEX leverage of pros and cons seen from the individual operator's perspective.

Hence the major drivers for 100G transmission will be similar to the drivers for moving from 10G to 40G, but the economic factors seems to be one of the main focus areas.

The community in general has recognized this difference early, knowing a large investment would be required to bring 100G into reality in a cost-effective manner.

If the cost drive is right, once 100G is commercially viable and standardized, network operators will quickly capitalize 40G investments and adopt 100G transmission for their future deployments. The prediction is that after 2010, the move from 40G to 100G Ethernet will start to take off rapidly pulling along the 100G transport layer.

3 40G

The market acceptance of 40G was delayed by the cost and by the slow development process during the period around the burst of the IT bubble. Many suppliers chose different implementations, creating 40G autonomous solutions resulting in a fragmentation of the market for components and resulting in overall higher costs.

So far the industry has been unable to achieve the rule-of-tomb in terms of a price/bandwidth ratio of 2½ times the price for a fourfold increase in bandwidth.

Nevertheless, 40G has an excellent near-term window of opportunity between now and 2011. Several commercial operators worldwide have started deploying 40G production backbones from several different suppliers.

In continuance of that - NORDUnet are in the process of upgrading NORDUnets IP Core infrastructure based on Juniper T1600. The router interconnection wavelengths is carried by NORDUnets Alcatel-Lucent 1626 Light Manager running a mix of ULH 10G and 40G wavelength, as illustrated in Figure 3-1.

Consequently NORDUnet is deploying both client and transport 40G interfaces – expected completion Q2 2009.

The statement is therefore that 40G is seen as general available for production in all areas and is the next natural step towards 100G. However, new 40G transmission features is doubly to be seen from the suppliers, due to the fact that it is expected that they will capitalize their investments. Instead it is anticipated that extensive resources will be put into 100G and 1T development.

Figure 3-1: NORDUnets 40G IP Core.

4 100G

Component and equipment suppliers, network operators and standards bodies like IEEE, the International Telecommunications Union (ITU) and Optical Internetworking Forum (OIF) have come together in order to fostering a healthy 100G ecosystem. The intention of this broad inclusion should result in an accelerated introduction of 100G ULH solutions that should meet industry performance, size, cost, and power requirements.

In a high level perspective this process have been split in three generalized categories covering the client, transport and the interworking area controlled by IEEE, ITU and OIF as visualized in Figure 4-1.

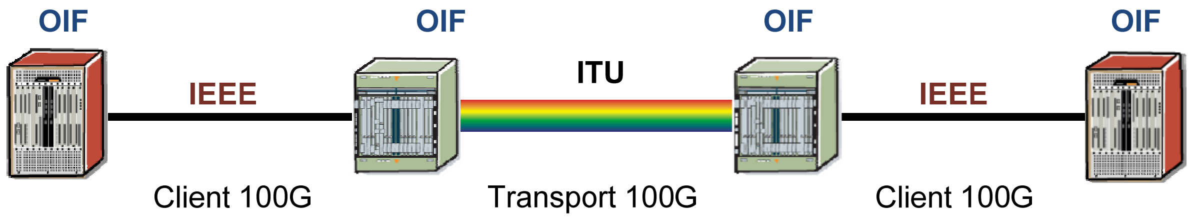


Figure 4-1: OIF, IEEE and ITUs responsibility areas.

From a standard perspective, IEEE, ITU-T and OIFs work will not be complete before 2010 at the earliest. The work until now is outlined in the following sections.

4.1 100G Client Interface

In 2006, the IEEE 802.3 working group formed the High Speed Study Group (HSSG) and found that the Ethernet ecosystem needed something faster than 10 Gigabit Ethernet (GbE).

For the first time in the history of Ethernet, a HSSG determined that two new rates were needed:

- 40 Gbps for server, computing and switch interconnection
- 100 Gbps for LAN/WAN aggregation and interconnection.

The IEEE P802.3ba 40 Gbps and 100 Gbps Task Force was formed in January 2008 to develop a 40 GbE and 100 GbE draft standard.

Included in this effort would be the development of the physical layer specifications for communication across backplanes, copper cabling, multi-mode fiber, and single-mode fiber. In February, 2009, the Task Force generated Draft 1.2 of the amendment to the IEEE 802.3 specification. Draft 1.n is currently under review.

The IEEE P802.3ba approved objective list covers the following areas:

- Support full-duplex operation only

- Preserve the 802.3 Ethernet frame format utilizing the 802.3 MAC
- Preserve maximum and minimum frame size of current 802.3 standard
- Support a Bit Error Rate (BER) better than or equal to 10^{-12} at the MAC/Physical Layer service interface
- Provide appropriate support for Optical Transport Network (OTN)
- Support MAC data rates of 40 Gbps and 100 Gbps
- Provide Physical Layer specifications as outlined in Table 4-1

Distance and Medium	40 GbE	100 GbE
1 m Backplane	40GBASE-KR4	
10 m Copper	40GBASE-CR4	100GBASE-CR10
100 m MMF	40GBASE-SR4	100GBASE-SR10
10 km SMF	40GBASE-LR4	100GBASE-LR10
40 km SMF		100GBASE-ER10

Table 4-1: Physical interface distances and medium.

The Standard Approval timeline for IEEE P802.3ba is depicted in Figure 4-2.

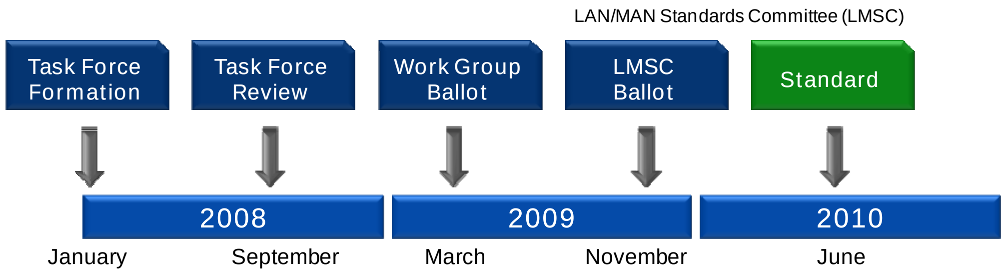


Figure 4-2: Approval timeline for IEEE P802.3ba.

4.2 100G Transport Interface

ITU Study Group 15 (SG 15) is addressing the extension of the G.709 OTN standard beyond the current 43 Gbps. So far the proposal for OTU-4 is 3 x ODU-3 at approximately 130 Gbps, as well as a rate optimized for transporting 100 GbE - approximately 112 Gbps. However, given that the IEEE has yet to define the exact frame format and encoded date rate for 100 GbE no decisions could ultimately be made.

ITU SG15 agreed that it would continue to track progress within the IEEE HSSG, and the resulting Task Force via the mutual collaborating, and define the new OTU4 rate accordingly. Still under the assumption that the IEEE decide to stick with 64B66B as the line code for 100GbE (resulting in an encoded line rate of ~ 103 Gb/s).

Specifically in relation to IEEE HSSG, the ITU-T address transport of 100GbE & 40GbE WAN over the OTN, collaborating with IEEE to ensure consistent transport of future standardized 100G interfaces as illustrated in Figure 4-3.

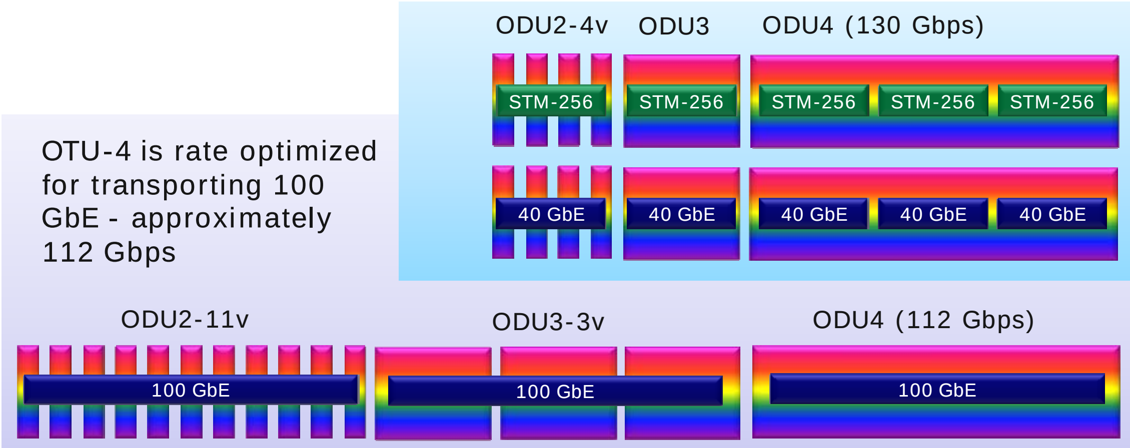


Figure 4-3: ITU SG15 Extension of the G.709 OTN Standard with OTU-4

The client interface developments points towards aggregation of 10 times 10GbE or 3 times ODU-3 and on top of that supplier specific split into n times STM-x/OC-x and xGbE. Where all interfaces will be based on provisionable single CFP or XFP units.

Expected performances for the 100G Transport interface:

- Transparent reach greater than 1.000 km
- Channel spacing equal 50 GHz
- Channel count equal to the extended C-band
- Compatible with existing installed DWDM base
  - Smooth integration into 50 GHz systems
- Compatible with existing 10G & 40G channels

The Standard Approval timeline for the extension of OTU-4 will follow IEEE P802.3ba time line – where the standard is expected to be released summer 2010.

4.3 Optical Internetworking for 100G

Consequently, a critical mass of component and equipment manufacturers came together through the OIF in May 2008 to develop a consensus on a technical approach for 100G DWDM transmission in ULH applications. Dual-Polarization - Quadrature Phase Shift Keying (DP-QPSK) modulation with a coherent receiver was the chosen approach. In July 2008, the OIF created two new projects - Forward Error Correction and Backplane as additional specification aspects of the DP-QPSK implementation direction.

4.3.1 100G LH DWDM Integrated Photonics Components

With a number of possible modulation techniques available to use at 100G, the identification and consensus with regard to a specific one, would gives the industry an important starting point for hardware design and reduce development risk.

DP-QPSK modulation with a coherent receiver was the chosen 100G modulation technique for integrated photonic components.

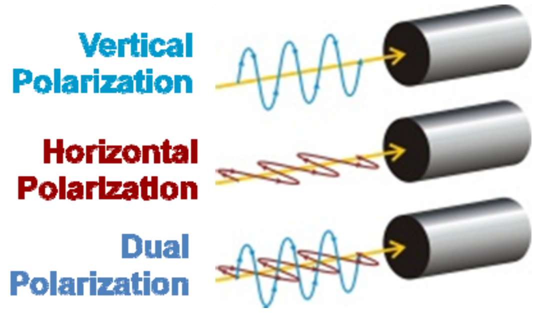


Figure 4-4: DP-QPSK modulation schema.

Dual polarization enables the transmission of more channels on a given fiber by supporting a signal in two polarizations, as shown in Figure 4-5. One signal is sent in the horizontal polarization and the other in the vertical polarization. However both signals are in the same frequency, and they are polarized 90° from each other so they will not interact.

Unlike traditional modulation, where the signal is either 1 or 0 one bit per timeslot, QPSK enables the modulation of two bits per time slot because the signal can be in one of four different phases. The signal could be a (0, 0); (1, 1); (0, 1); or (1, 0).

Theoretically, DP-QPSK provides two times QPSK, which equals four bits pr time slot seen by the receiver resulting in four times the bandwidth.

In essence it is the combination of dual polarization, which provides twice as much information pr time slot, and then combined with QPSK, providing twice as many bits pr time slot.

Effectively, the electronics are dealing with a signal that is a factor of four slower, and that information can be processed by general available digital signal processing (DSP) technology and thereby reducing the requirements to the optical interworking area only to transfer speeds at 25-28 Gbps for the Laser modulator.

4.3.1.1 DP–QPSK Paired with a Coherent Receiver

In the OIF implementation DP–QPSK is paired with a coherent receiver, which is the most challenging aspect of the system.

In the past, optical receivers received signals by detecting the intensity of the light, which was turned on and off by means of using a more basic modulation format known as on–off keying. The typical receiver operates much like an AM radio making the coherent receiver more like an AM/FM radio, complete with local oscillator. The local oscillator is tuned to the frequency of the incoming signal and only extracts that information. In other words, the coherent receiver locks onto both the frequency and phase of the DP–QPSK signal and is able to accurately recover the incoming bits.

In addition the coherent receiver is also used to compensate for optical impairments like chromatic dispersion and polarization mode dispersion and thereby eliminating the need for expensive optical compensation techniques. The coherent receiver, therefore, enables a better Optical Signal to Noise Ratio (OSNR), which, in turn, increases the distance the signal can travel without regeneration.

4.3.2 SERDES-Framer Interfaces for 100G

OIF has launched a Serializer/Deserializer (SERDES) Framer Interface (SFI) project. SFI connects Physical Layer devices to other Physical Layer devices like in a back plane or on-board connectivity.

The SFI project will define a scalable interface from an optical module to a framer device for line rates from 40 Gbps to 100 Gbps and beyond. The SFI project forms an important contribution to the requirements for transport of high-speed data via the OIF's Common Electrical Interface (CEI). The current, ongoing CEI-25 project establishes electrical signaling rates for high-speed backplanes.

In order to explain the interrelation between the different areas the OIF system reference model is used based on a 100G scenario as illustrated in Figure 4-5.

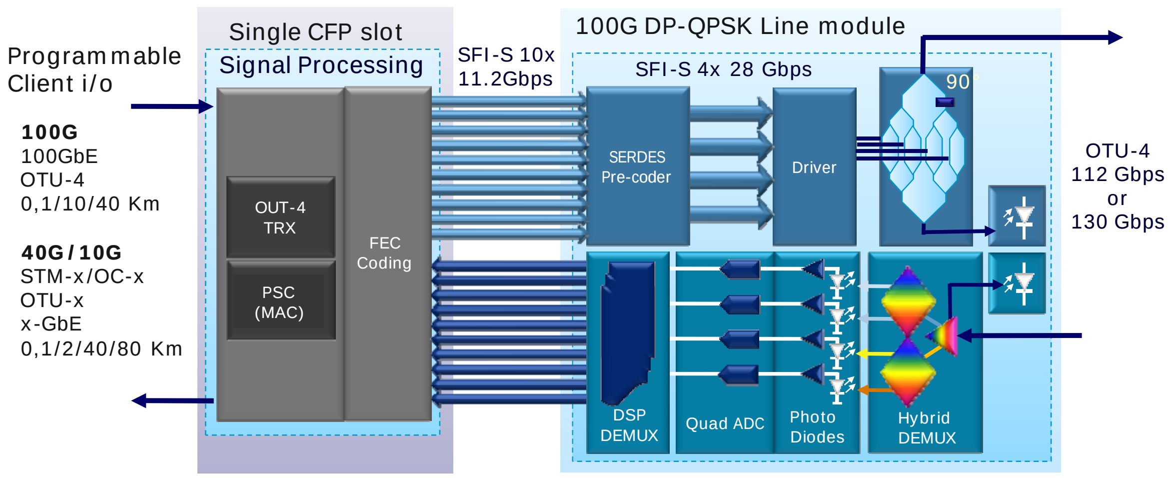


Figure 4-5: OIF system reference model based on a 100G scenario.

The signal processing consists of a physical layer Framer module having dual conversion functionality either as OTU-4 or MAC layer. The framer module can be integrated with a FEC module or this can be an adjacent module interconnected via SFI into 10 x 11,2 Gbps.

The Signal processing module is connected to the 100G DP-QPSK Line Module via SFI-S. The Tx signal is pre-coded via SFI into 4x 28 Gbps electrical NRZ data streams used by the Laser DP-QPSK Modulator. The Tx line signal is then generated as an optical wavelength carrying an OTU-4 of 112 Gbps

The Rx line signal is received by an advanced Optical hybrid DEMUX directing the four phases of the wavelength to one of four specific photo diodes and thereby providing full information on the real and imaginary parts of the signal along the polarization axes.

Each photo diode is connected to and advanced Analogue Digital Converter (ADC). The latest ADC technology is a single-chip DP-QPSK coherent receiver for 100G optical networks; with four-channels of 56 Giga Samples/sec 8-bit ADCs integrated with logic and memory to perform the complete receive physical function when connected directly to the photo diodes.

In addition a single-chip approach avoids the need to transfer terabits-per-second of data between ADC and DSP whereby reducing power consumption, silicon area, and the number of I/O pins. The DSP conduct advanced processing including post compensating of optical impairments and delivers the Rx data stream via the SFI to the signal processing module. The processing module corrects and frames the Rx data stream to the appropriate link layer either OTU-4 or MAC.

Physical Layer devices are generally located on the same card and therefore the electrical requirements for SFI support relatively short chip-to-chip or chip-to-module connections. On the other hand, Link Layer devices are frequently not on the same card as the Physical Layer devices, and therefore System Packet Interfaces must traverse a backplane. CEI defines electrical requirements for both “Short Reach” and “Long Reach” interfaces, supporting both SFI chip-to-chip and chip-to-module applications, and SPI backplane applications.

4.3.3 Forward Error Correction for 100G DP-QPSK LH

The OIF members have defined a common basis for a Forward Error Correction (FEC) Encoder, where the project aims to develop a code with improved performance over current industry practice in order to meet the strong performance goals of the 100G transport application.

While agreements on key technologies for 100G transmission like DP-QPSK ultimately can enable interoperable solutions, OIF has defined that full DWDM interoperability is outside their current work scope.

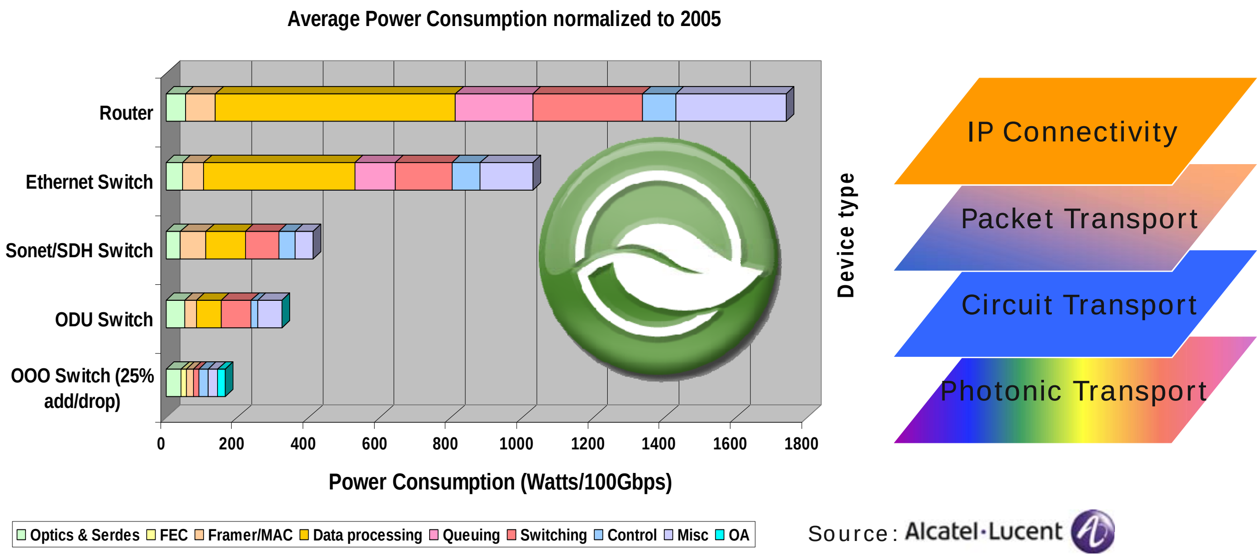
Instead, the OIF is focusing on creating a larger market for component suppliers benefiting from the common implementation approach, and interoperability of the component building blocks within a 100G DWDM transponder, and thereby providing opportunities for product differentiation.

This differentiation is expected to result in autonomous suppliers applications with regards to the DSP modules of the coherent receiver. The main parts of this will be software implementations – the community should therefore push for having the potential to disable FEC or even return to a simple standardised format. That approach would remove another stone from the path towards Supplier vs. Supplier interoperability at the transport layer for 100G solutions.

5 Reduction of the Carbon Dioxide Footprint

Bandwidth pr Subscriber increases faster than Power pr Bandwidth decreases. This is not environmentally and economically sustainable. Technology improvements and optimized system design alone will not solve the problem but a part of the solution could be using the lowest possible network layer. Nevertheless, the power consumption will rise significantly in electronic based network components.

Looking more specific at the network components, routers have the far worst power profile and will dominate the overall power profile, while photonic components will help limiting the power profile as outlined in Figure 5-1.



2009 figures are roughly 1KW/100Gbps for Routers & 500W/100Gbps for Switches

Figure 5-1: Average Power Consumption for typical optical and electrical Telco equipment

All though that the photonic layer has the lowest carbon footprint it must be subjected to developments and enhancement in order to reduce the power consumption cost resulting in a smaller interface carbon footprint.

One of OIFs sub project are related to a photonic integration project specifying both transmitter and receiver photonic modules for incorporation into 100G DWDM transponders. Integrating many photonic components into a small number of modules will be critical to reduce the cost, size, and power dissipation of a 100G transponder.

6 Emerging Design Paradigm

The uptake of 100G and the introduction of GMPLS for the WDM control plane seem to be general available at the same time in 2010 – potential leading to new architectural design approach primarily for 100G deployments. In addition theses discussion will also treat the topic with regard to where does 100G fit into the NGN pyramid structure supporting the cost reduction and environmental sustainability.

6.1 Re-emerged Ground-Breaking Protection Schema

GMPLS has become one of the key Control Plan Management Protocols, which have found its way to the photonic area and is expected to take off in the beginning of 2010. This will lead to emerging photonic functionality like:

- Network and Resource Discovery simplifying the installation and commissioning process
- Dynamic Provisioning allowing the offering of bandwidth on demand and end-to-end connection service, where the provisioning has been triggered through the NMS or by the client NE
- Distributed Automatic Restoration enabling sharing protection resources in the network. It differs from traditional protection schemes, such as O-SNCP or O-MSP, where protection resources are dedicated and already assigned, so that they cannot be used to protect other circuits or links.

In addition dynamic and static photonic impairments are taken into consideration by incorporating these figures into the GMPLS calculation process and thereby securing the optimal photonic performance at all times.

Trials have shown that the photonic protection and restoration mechanism can recover a transport link within a few msec. The recovery time for multiple failures is anticipated to be in range of 25 to 50 msec., including photonic adjustment in order to gain optimal photonic performance.

This provides the potential for a ground-breaking way of designing the protection and restoration mechanism for all networks layers, especially the costly use of redundant IP interface pack as illustrated in Figure 6-1.

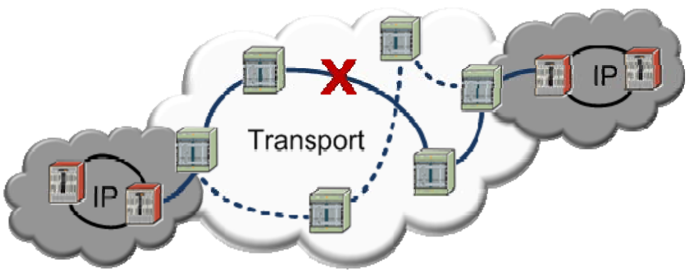


Figure 6-1: Dual IP domains interconnect via GMPLS enabled Transport layer providing protection.

Currently the IP interface packs are serving the objective to protect against external failures.

The protection for external failures can consequently be moved down to the less costly photonic layer, and thereby securing the IP core against external failures.

However the IP interface pack will still pose as single point of failure and the user community should therefore put even stronger requirements towards the suppliers in order to have an even lower MTBF for the IP interface packs than today. In other words, the IP interface pack suppliers should focus their resources towards enrichment and robustness of the IP interface packs, instead of spending development resources on IP interface packs with coherent long reach photonics.

6.2 100G Fitting into NGN Pyramid Structure

Archiving a sustainable hybrid NGN infrastructure in terms of service offering, cost and environmental impact – focus should be put towards the network pyramid structure outlined in Figure 6-2.

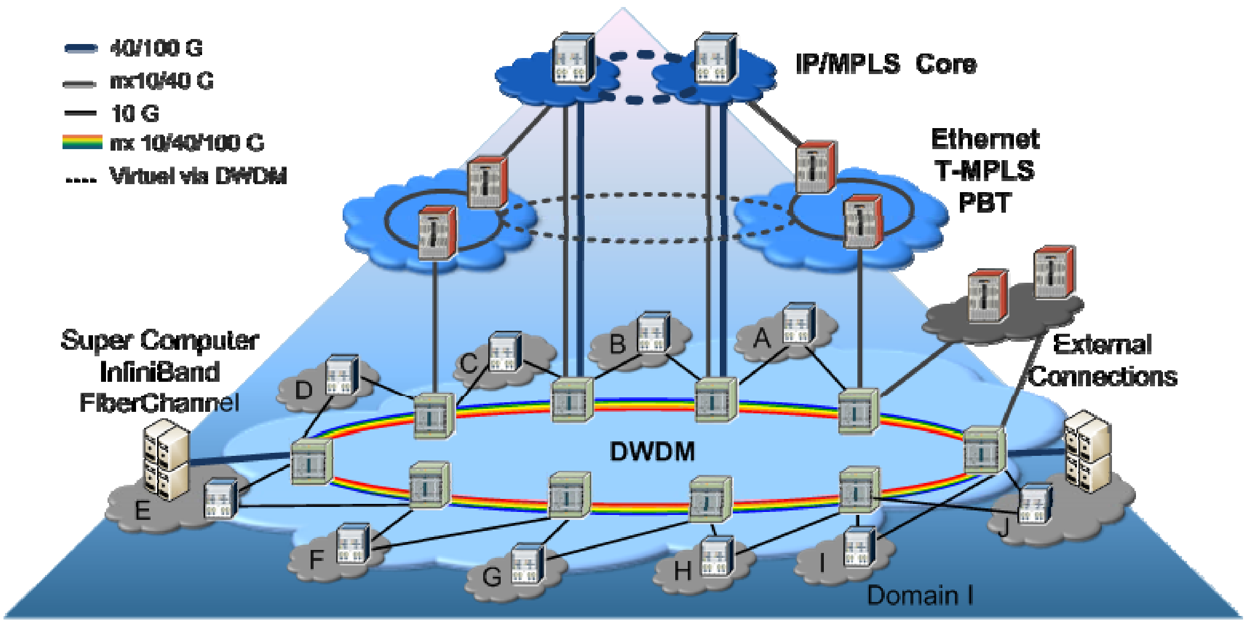


Figure 6-2: NGN Pyramid Structure

Starting from the bottom of the pyramid, it consists of a photonic base layer covering access, metro and core. Moving on to the next level imposes a connection and connection-less transport overlay. Finally, the top level is a centralised IP layer.

The client or campus premises demands are typical mass amount of IP streams, request for dedicated Point to Point connections and VPN solutions.

One specialized flavour of Point to Point is InfiniBand. Over the past few years, it has made a remarkable comeback and has been backed by large companies, using it to develop strategic products like GRID solutions, blade servers, storage, cluster interconnects, etc. In order to significantly decrease the latency and bit by bit utilisation of the transmitted data streams compared to Ethernet.

Many smaller InfiniBand companies are being bought by industry giants, who want to enter the market quickly. For long geographical distance InfiniBand 120 Gbps solutions will have to use OTU-4 (130 Gbps).

6.2.1 100G Initial Deployment

100G deployment will start from the top of the pyramid, where it will replace core 40G interfaces. These 40G interfaces will be pushed down as replacement upgrade in the metro areas and again pushing those interfaces towards the access and thereby reusing the network components leading to a simple cost optimization.

The service offering is based on a wide pallet due to the nature of the hybrid network layer which makes it easy to accommodate additional requirements for intensified network exhaust.

In addition an efficient a sustainable OAM&P environment supported by the improved control plane management in the photonic base layer.

The environmental impact will be more sustainable due to less power intensive electronic devices, due to a reduced transport layer and centralized IP layer. In total supporting the rule of thumb switch were you can route were you must.

7 Summary

The slow to market process of 40G have enabled many different 40G solutions for the transport layer - resulting in fragmentation of the market for components and resulting in overall higher costs.

Nevertheless, 40G is commercial available - where the expected service window will be until 2011.

The industry have come together in order to create a healthy 100G ecosystem, which will be beneficial for the entire community.

This broad inclusion will result in a fast introduction of 100G solutions that will meet industry performance, size, cost, and power requirements.

Initial deployment of 100G is expected to take place in 2010 followed by a quick adaption.

If the cost drive is right, once 100G is standardized and commercially available, network operators will quickly capitalize 40G investments and adopt 100G transmission for their future deployments.

It is expected that after 2010, the move from 40G to 100G Ethernet will start to take off pulling along the 100G transport layer.

It is estimated that 100G will have a service window from 2011 until 2016 where the introduction of 1T solutions is expected to take place.