Report on the availability and characteristics of equipment for next-generation networks

IST-2001-34925

Study into European Research and Education Networking As Targeted by *e*Europe

SERENATE

Deliverable no. D9

Delivery type: Report
Security: Public document

Authors: Valentino Cavalli (editor), John Dyer, Michael Enrico, Eoin Kenny, Michal Przybylski, Roberto Sabatino, Stanislav Sima

Date: 22 April 2003

Contents

1. Executive summary

1.1. Introduction

The basic service provided by National Research and Education Networks (NRENs) nationally to their users is a best-efforts IP service. The equivalent service is extended across Europe, and to research networks in other world regions, by the GÉANT[*] network. The characteristic of a best efforts IP service is that it offers ubiquitous connectivity, but that is all. There are no guarantees of performance. In parallel, with best-efforts IP, there have been a number of national and pan-European initiatives to offer guaranteed performance between end-locations in the form of Virtual Private Networks as well as a multicast service which provides network-based broadcasting capabilities. There is a generally increasing demand from users for higher performance and/or more predictable services. This study examines the way telecommunications technology is likely to develop in the next five years, its ability to meet user demands, and the effects this can have on the implementation of research networking in Europe.

1.2. Current networking environment

The best-efforts IP service is normally provided by routers accessed by, and interconnected by, leased circuits. Historically, both nationally and internationally within Europe, leased circuits were provided by monopoly telecommunications operators. As a consequence, there was considerable reluctance, on the part of these operators, to provide access to leading-edge technology. Service provision, as defined by speed of operation of connectivity provided and availability was generally rationed and expensive. The liberalisation of the European telecommunications market place has changed this picture quite dramatically in the last four years. For many locations in Europe, it is now possible to gain access to leased connectivity, which offers the maximum performance technically available today (currently 10 Gb/s). In addition liberalisation has, in some locations, allowed direct access to physical connections, typically fibre-optic cables. This has enabled some NRENs to implement their own transmission technology rather than relying on services provided by telecommunications operators. These factors are changing the technical options available for constructing research networks in Europe.

A further factor of importance is the emergence of groups of users with potentially very large demands for connectivity between a limited number of locations. There have been several reasons for this, notably:

i) The enormous reductions in cost, particularly for international connectivity, have enabled a number of research activities to consider network-based solutions for their connectivity needs where previously these were too expensive. This meant that, in the past, either research co-operation was geographically constrained or that alternative “non-network” based communications such as the physical transport of magnetic tapes were employed

ii) The very large, and increasing, costs of research infrastructure have led to a much more European approach to research and, as a consequence, a significant increase in the demand for pan-European research connectivity. The European Research Area is the political manifestation of this trend

iii) The development of distributed computing power, capable of exploiting high-capacity wide-area connections, and the standardisation of these capabilities in the Grid computing initiatives.

All the above factors mean that the simple model of a basic, “best–efforts,” IP service provided by routers, is no longer sufficient to meet the service requirements of an environment where there may be very large flows of data between a limited set of locations, and which requires predictable and defined performance. In addition, the option of a direct implementation of transmission technology, as an alternative to leasing capacity from telecommunications operators, opens new technical opportunities for the provision of service.

---

[*] http://www.dante.net/geant/

1.3. Basis for the study

The study focuses on four areas of technology namely:

i) Routers. These are currently the basic building blocks of an IP network. They are packet switches that operate and process on a per-packet basis.

ii) Optical switching devices. These are relatively new products designed to switch streams of data on a per-stream basis. In this context a stream is defined as a synchronised stream of bits.

iii) Transmission equipment. This equipment is responsible for transmitting data, in the form of a bit stream, between switches or routers.

iv) Network control techniques. These enable network operators such as NRENs or DANTE to manage the various elements that are used to construct networks. When only providing best-efforts IP services, based on interconnected routers, there is limited network intelligence, and network control techniques are not very important. As equipment becomes more complicated and varied, the ability to manage and control network elements becomes a significant issue. A particular question here is the extent to which such control techniques can operate across networks managed by different network operators.

A questionnaire, covering aspects such as development and deployment plans for higher-capacity transmission (40Gbp/s and higher speeds), optical and photonic switching developments, and network management and control, was completed by 12 equipment vendors and 2 research laboratories. The questionnaire was used mainly as a guideline for subsequent face-to-face discussions. These discussions were targeted at understanding the developments, issues and benefits of the technologies mentioned above. The results were analysed from two perspectives: from an equipment type perspective (routers, switches, transmission equipment) and from the perspective of the way these different elements could be combined to create future networks, paying particular attention to speed of operation (40 Gb/s and higher speeds), network management and the implications these have on future network architectures.

1.4. Summary of findings

1.4.1. Routers

IP routers are mature products and offer a wide range of functionality (multicast, IPv6, differentiated services, MPLS). These features are becoming common on most routers targeted at research backbone networks. In most cases they are, or will become conformant to standards and will be able to operate so that they fully utilise the transmission rate of the circuits to which they are connected. In many cases this is already possible. In terms of developments, for interfaces operating at 40 Gb/s and higher speeds, a number of manufacturers already offer the switching capacity, i.e. the ability within the router itself to support interfaces operating at 40Gb/s, although interface cards offering single-channel 40 Gb/s are not yet available. This is due

i) to low demand for interfaces operating at this speed,

ii) high prices (several vendors stated that a single 40-Gb/s interface was likely to be more expensive than the aggregate cost of four 10-Gb/s interfaces)

iii) the lack of commercial transmission services operating at 40 Gb/s.

For the next two years it is expected that a 40-Gb/s connection to an IP router will be delivered by four 10-Gb/s channels. A common view is that demand for routers in the research community will be the main driver for availability of commercial, single-channel 40-Gb/s systems

1.4.2. Optical switching devices

In the last few years a lot of attention has been given to all-optical switching. Sometimes the term “all-optical” has been confused with devices that offer optical interfaces but operate partially using electrical technology. For example, devices that are able to switch and multiplex Gigabit Ethernet (GE) and SONET frames are optical switches, insofar as they offer optical interfaces, but they use electrical technology to carry out the switching. These devices are also called O-E-O devices (optical-electrical-optical) and have been available for some time, although they are undergoing significant developments in terms of scalability and granularity of the services they can offer. The main issues that need to be resolved for these switches over the next few years relate to standardisation of functionality and interoperability.

The term “all-optical” relates to switching equipment that operates entirely using optical signals, and therefore is also referred to as O-O-O devices. They are independent from (or transparent to) the signal that is being carried over an optical channel (GE, 10 Gb/s, 40 Gb/s,…). These switches use a variety of complex light-switching techniques, such as MEMs (Micro Electromechanical Mirrors), liquid crystals and other “proprietary” methods. While the technology for all-optical switching is improving rapidly, and all-optical switches are available at reasonable prices, there remain some fundamental issues with this technology. These include re-routing of light paths with or without electrical regeneration. The distances between amplifier sites may be different between a main path and a re-routed path. As a result, optimal engineering rules, required by this analogue technology, may not be met in case of re-routing, and the switches themselves also introduce relatively high attenuation of optical signals.

In addition devices that operate using electrical technology remain needed for multiplexing and bandwidth grooming, allowing bits streams with various combinations of speed of connectivity to be multiplexed and switched efficiently, although these functions may be confined more towards the edges of the network. The main advantages of all-optical technology relate to protocol independence, lower unit cost and lower operational costs, particularly in support of 40 Gb/s and higher speeds.

1.4.3. Transmission technology

1.4.3.1. Speed

Transmission technology developments have demonstrated that speeds of 40 Gb/s, and higher, are technically possible. In fact, speeds up to 600 Gb/s transmission have been demonstrated in a laboratory environment. Offering 40 Gb/s, commercially, at attractive prices is a different matter, influenced by numerous factors. These include general market conditions, and mass production, which will drive prices up or down, as well as many technical details. From a technical perspective 40 Gb/s is much more complex to implement than 10 Gb/s. Transmission degradations, such as dispersion, cross talk and attenuation, are significantly more pronounced than is the case for lower speeds. Consequently, techniques to compensate for these effects must be developed, and better amplification methods, compensation for Polarisation Mode Dispersion (PMD) and complex Forward Error Correction (FEC) methods need to be deployed. This will all increase the price for integrated systems operating at 40 Gb/s in a full operational environment. It is thought that, over the next few years, the improvement of Forward Error Correction techniques will be the main enabler for commercial availability of 40-Gb/s transmission systems.

Commercial 40-Gb/s systems will certainly be available soon. It is unclear where we will see their first appearance. Transmission equipment vendors interviewed felt this would be in the market of Ultra Long Haul (up to 4,000 km) transmission systems, while router vendors foresee that this would be in the metropolitan area in 2003.

1.4.3.2. Reach

Many NRENs are currently deploying, or planning to deploy, their own fibre. This is enabled by the increasing developments of transmission technologies, which, even if components are not yet off-the-shelf, allow NRENs to adopt a “do-it-yourself” approach towards the network infrastructure, in contrast with the traditional approach of buying connectivity from carriers. The reach of transmission equipment is of crucial importance for understanding and planning appropriate network architectures that satisfy the needs of NRENs without involving excessive costs. Unfortunately, NRENs still have limited experience with transmission equipment. Reach depends on many factors including fibre type, fibre quality, bit-rate of each wavelength, the number of wavelengths transmitted in parallel, amplification and transmission technology used, FEC and other components.

An issue with transmission is the need to regularly boost the signal between circuit endpoints. Transmission technology is improving significantly in relation to the spans between amplification (boosting the size of the signal to offset attenuation) and regeneration (reconstituting the signal to eliminate signal distortion). Current transmission systems for 2.5 Gb/s or 10 Gb/s require regeneration of the signal after four or five amplification stages. This means: after approximately 400 km. Newer transmission technology will enable 10 Gb/s to be transmitted up to 4,000 km without regeneration. For 40 Gb/s this range is expected to be up to 1,000 km. There is no current knowledge of the relative costs of such equipment. The provision of amplification and regeneration equipment in the network is operationally complicated and expensive, since it can impose the requirement to install and operate equipment in remote locations where NRENs have no reason to be present. Increasing span lengths will enable new network architectures for future research networks, where the ownership and management of long-distance fibre spans, without the need for amplification and regeneration, becomes a possibility worth serious consideration. This is known as “Nothing-in-Line” operation, reflecting the absence of any amplification and regeneration equipment between the terminating points of the fibre. Successful experiences in this direction are taking place in Europe already, as demonstrated by CESNET, the Czech NREN, in reaching 230 km with Nothing-In-Line (NIL).

1.4.4. Network control capabilities

The next-generation research networks are very likely to include a mixture of networking elements (routers, optical and/or photonic switches, multiplexing devices and possibly transmission equipment). In addition the service portfolio that will be offered to users is intended to allow greater user control over network resources and performance. All of these trends imply much greater real-time operational-control network resources. To achieve this, NRENs will have to introduce new element managers in their network management systems. They will have to become accustomed to different protocols, traditionally used by telecommunications operators, which differ considerably from the techniques used in IP. Advances in TMN/Corba (telecommunications operators’ world) and SNP/Corba (IP world) will assist in the integration of different network management systems, although, at present, these standards are immature and require much development.

The organisational structure of research networking is expected to remain unchanged. This will require a much more co-operative approach, among network operators, to resource allocation and control if end-to-end services, crossing the management domains of individual NRENs, are provided to end-users. The development of G-MPLS, and its availability in IP routing and switching equipment, will facilitate developments in this direction. However, despite standardisation efforts, most implementations, especially for switching equipment, are proprietary and non-interoperable at present. A fundamental issue still unsolved is the inter-domain operation of research networks.

1.5. Conclusions

A simple extension of the current service model of offering best efforts IP, at higher speeds of operation, will not meet emerging user requirements.

In the area of transmission technology, it is likely that 40 Gb/s systems will emerge in the next two years. It is perceived that these will not necessarily be cost-effective and the use of parallel slower-speeds wavelengths is more appropriate at present. The potential exploitation of dark-fibre is heavily dependent on the reach and economics of “Nothing-In-Line” systems. This is currently limited to spans of less than 250 km.

There are developments in both routers and optical switches that suggest that a combination of these elements can effectively be used to provide a more flexible and manageable network structure. In the case of switching these will be based on O-E-O devices. Developments of O-O-O technology will require considerable additional effort before they result in useful products

Developments in network control suggest that it will be possible, in the future, to provide management functions that cross domain boundaries. However, this will require the emergence of standardised implementation of network management and control functions, particularly in the area of G-MPLS.

2. Introduction

This report is part of SERENATE, the Study into European Research and Education Networking As Targeted by *e*Europe, contributing to European policies, social objectives and economic development by providing inputs on initiatives that could help to keep European research networking at the forefront of worldwide development. The objective of SERENATE is to provide input to the formulation of policies by the European Commission, but also to national governments and funding bodies, the management of universities and research institutions, and the National Research and Education Networks (NRENs). It focuses on the technology building blocks that are used to construct research networks and considers how developments in these building blocks will change the structure and technical organisation of research networking.

This report presents a study on the characteristics of equipment for next-generation networking, in particular routing, switching and transmission equipment available today, to understand what will be available in 5-years' time. It also investigates emerging technologies, some of which are still in the laboratories and will only become available in a longer timeframe.

The study is based on a series of individual meetings with leading equipment manufacturers and research institutions actively involved in the EU-funded OPTIMIST* project (a Thematic Network focusing on the development of photonic technologies in Europe). Additional desk research was carried out. The following organisations participated in the meetings, which were held between 13 November and 6 December 2002:

- Alcatel
- Calient
- Ciena
- Cisco Systems
- Corvis
- Juniper Networks
- Lucent Technologies
- Marconi
- Nortel Networks
- PhotonEx
- Tellium
- Wavium
- University of Essex
- University of Gent.

Key technical and strategy personnel from the suppliers were interviewed in these meetings by SERENATE partners and consultants from PSNC, CESNET and HEAnet. A set of questions to provide guidance for discussion during the meeting was sent in advance. The questionnaire, which is reproduced in Annex I to this report, was developed in collaboration with the optical networking group of TF-NGN (the task force that looks into tests and experiments for the introduction of next-generation networking technologies on GÉANT) and National Research and Education Networks. The questionnaire was only used as a framework for the interviews as not all questions were relevant to all suppliers.

Non-disclosure agreements were signed with some of the companies and as a consequence, technical details are sometimes described in this report in a non-attributable form.

The report is structured by product classes. Some suppliers have products falling into more than one category. The input to the various sections is based on the following groups of suppliers:

- Routers (packet switches): Alcatel, Cisco, Juniper, Marconi, Nortel

- Switching equipment: Alcatel, Calient, Ciena, Cisco, Corvis, Lucent, Marconi, Nortel, Tellium, Wavium

- Transmission equipment: Alcatel, Ciena, Cisco, Corvis, Lucent, Marconi, Nortel, PhotonEx.

---

* http://www.ist-optimist.org/

Optical networking has several consequences for network management and network architecture, and these are considered in detail in sections 6 and 7 of the report. Section 8 expands on the options for alternative network architecture by discussing the feasibility of procuring dark fibres, whereas section 9 provides an analysis of emerging technologies. The final section adopts a different approach and, based on the information provided in the previous ones, draws general conclusions for the SERENATE study on key issues like network capacity, integration of optical and IP control plane, intelligent optical networking and all-optical networking.

3. Routing equipment

In a study of the expected evolution of packet switching technology (the key component of which is the packet-based "router"[1]) a number of discrete features need to be considered. For this study, the features are as follows:

- scalability (of nodes and links)
- functionality
- interoperability with other networking components
- logical partitioning (virtual routers)
- management.

This section does not explicitly discuss IPv6, which is arguably the next and most far-reaching step technology change that research networking and the global Internet will experience. The reason is that it is assumed that many networks will start deployment of IPv6 in a production environment during the coming years. In readiness for this, most router vendors have already implemented IPv6 in their software releases and it is expected that optimised hardware support for handling IPv6 will become available during the next year or two. Many research networks have well-advanced plans for introducing production IPv6 services and expect to do so during 2003. For example, the GÉANT backbone should offer IPv6 service in a "dual stack" mode (alongside IPv4 services) by mid-2003. Therefore, the expectation is that router vendors in general already have implementations of IPv6 that are approaching production quality.

3.1. Scalability

This section discusses the **scalability** of routers in terms of nodes and links. Node scalability concerns the packet forwarding performance of a router and the number, type and transmission capacity of the line interfaces that can be accommodated by the router chassis. Link scalability is concerned with the implementation of high-capacity links between routers and new developments, with respect to interface types, that may enable closer integration with high-capacity transmission systems.

3.1.1. Node scalability

Firstly, the **packet forwarding performance** of a router (independent of the number, type and transmission capacity of the line interfaces) is addressed. Often this function is performed by a "packet forwarding engine". In modern, high-performance routers the packet forwarding function is performed in hardware by highly specialised network processors. In addition, novel parallel processing architectures and parallel switching fabrics are being implemented to further increase the capability of these packet forwarding engines. In old (now largely obsolete) routers, packet forwarding was often performed by general-purpose processing hardware that made use of a shared-bus architecture to connect the line interfaces. This usually meant that the packet forwarding capability of such a router fell below the net capacity of a full complement of line interfaces, thereby imposing a performance bottleneck. Latest-generation routers often are capable of forwarding all levels of ingress traffic up to the level where there is a maximum complement of the highest-capacity interfaces, all of which are operating at line speed and at full-duplex.

Usually, a router chassis is based on a modular architecture such that the slots that accommodate the line interfaces have a maximum associated transmission capacity and can physically house either a single line interface corresponding to this maximum transmission capacity or multiple interfaces (often of different types) that collectively correspond to the same maximum capacity. Usually, the latter case does not scale down to line interfaces at the lower rates (e.g. STM-1 or 155 Mb/s and below). This is because it is very difficult to fit onto the available area of a printed circuit board the necessary number of discrete electronic components required to drive multiple line interfaces. Equally it is difficult to fit the necessary number of physical connectors on the available area of front plate. Thus a 10-Gb/s-capable slot would probably not be able to accommodate 64 STM-1 interfaces. This does not necessarily represent a problem for highly developed national and regional research networks where upgrades are continually being made. However, it can prove problematic when extending research network coverage to less-developed countries where link speeds remain low (< 155 Mb/s).

---

[1] As with many networking terms, the term "router" can be taken to mean a number of types of networking elements with very different functions - especially when the term is used in product names. For example, Lucent's "LambdaRouter" product is an all-optical (photonic) switch. In this section, the term "router" will be taken to mean an IP packet forwarding device.

When scaling up to larger routing nodes (e.g. those that need more capability than can be offered by today's fully populated, high-performance chassis) it is necessary to implement multi-chassis designs. Doing this with a router that has been primarily designed as a single-chassis-only device can be very difficult, if node scalability and non-blocking throughput is to be maintained. The resulting node designs can be complex, expensive and difficult to manage. However, high-performance routers of a new generation are now emerging that have been designed to be connected together in multi-chassis configurations in which linear scalability of the non-blocking nodal throughput is maintained. Typically, these multi-chassis configurations will consist of the requisite number of dedicated line card chassis connected using proprietary optical connections to one or more specialised switching fabric chassis. Through such configurations, the "Terabit Router" has already been demonstrated and is on the verge of becoming commercially available.

Today's state-of-the-art (single-chassis) high-performance routers already have “40-Gb/s ready” slots and can accommodate up to 32 10-Gb/s line interfaces. The non-blocking packet forwarding capability of these routers matches this interface count (as described above), giving a full-duplex performance of 320 Gb/s. This is expected to increase, over the next one or two years, to the point where a single chassis router has many tens of 40-Gb/s-capable slots with a full-duplex packet forwarding performance to match. This would result in the availability of single-chassis multi-terabit routers which, when used as the building block in a multi-chassis configuration, could yield a router node capable of handling a few tens of Tb/s of non-blocking packet forwarding.

One potential problem with these large, multi-chassis router nodes is that they will be very complex to construct and maintain. It is unlikely that they can be built and upgraded by anybody other than engineers from the router vendor. Although this is often common practice for the established telecommunications operators and operators of large commercial data networks (especially for the more complex telecommunications systems like DWDM transmission systems), it may well be unfamiliar for NRENs which are more used to a “do-it-yourself” approach to installation and operations. Providing an easy-to-use extension strategy is, however, a key requirement for scalable routers. It can be achieved if this requirement has been taken into account from the conception of the multi-chassis router.

Another point about these large multi-chassis routers is that they will probably bring with them a new paradigm in the approach to their management. This is discussed further in the section on router management.

3.1.2. Link scalability

As router nodes become capable of handling larger volumes of traffic, so too must the links grow in capacity in order to avoid the need for many parallel links that will require the routers to load-balance the IP traffic. The highest-capacity, single-channel, router interfaces available today operate at 10 Gb/s. These have been available for several years and are largely of the packet-over-SONET/SDH (POS) type, although 10-Gb/s Ethernet is now emerging following its standardisation by the IEEE. The next step up in link capacity is expected to be to 40 Gb/s, especially because state-of-the-art, and emerging, router chassis already have 40-Gb/s-capable line interface slots. However, currently there are **no** 40-Gb/s interfaces available for these large routers. (Instead, a line interface card that fully utilises a 40-Gb/s slot has four parallel 10-Gb/s interfaces on it.)

There are a number of reasons for this:

- **No significant customer demand**. Until the recent telecommunications downturn, it was thought that there would be by now a significant commercial demand for single-channel 40-Gb/s capability in routers. This has not materialised and the only users tentatively talking about 40-Gb/s capabilities are those from the research and education networking community.

- **Technical difficulties**. These are discussed in greater depth in other sections of this report (see sections 4, 5 and 9), but basically the complexity (and hence cost) of high-speed networking components does not scale linearly with speed. 40-Gb/s systems really stretch the current technology. For example, new ASICs are necessary to do Forward Error Correction (FEC) at 40 Gb/s.

- **Cost**. It is currently still questionable whether single-channel 40 Gb/s is more cost effective than four times 10 Gb/s. From a cost point of view, 40-Gb/s interfaces will become attractive when the price comes down to around 2.5 times that of 10 Gb/s.

- **No commercial 40-Gb/s transmission services**. Even if router vendors were to go ahead and develop single-channel 40-Gb/s interfaces for their high-end routers, none of the telecommunications operators are offering services suitable to connect them. This leads to somewhat of a "chicken and egg" situation - who will take the first step? Interestingly, a number of vendors of transmission equipment (as used by the operators) seem to be of the opinion that it is the router vendors who will provide the main drivers for the first deployments of 40-Gb/s transmission and switching systems.

Vendors of large core routers have plans to introduce a 40-Gb/s single-channel interface when the remaining technological problems have been solved and when it makes sense commercially to do so. This is not expected to be until the end of 2003 at the earliest. When these interfaces become available, it is expected that they will first appear as Very Short Reach (VSR) capabilities and will be intended for high-capacity intra-POP links. In the meantime, where single-channel router-router links operating at speeds greater than 10 Gb/s are required, the only choice will be to use the vendor-proprietary link bonding schemes that the vendors of most large core routers already offer today.

An issue addressed during the SERENATE equipment study was the introduction by router vendors of new interface types, in particular, interfaces that offer the potential for omitting intermediary systems often associated with transmission systems. In most wide-area, high-speed IP backbone networks the routers are connected to each other with high-speed packet-over-SONET/SDH links (e.g. 622 Mb/s, 2.5 Gb/s or 10 Gb/s) or Gigabit Ethernet or 10-GE links. Very often these links are running over a DWDM transmission system. The interfaces between the routers (or other equipment) and the DWDM line termination equipment (e.g. the transponders and multiplexers) are necessarily based on common standards. This means optics operating at wavelengths of 850 nm, 1310 nm or 1550 nm depending on the fibre type (multimode or single mode) and reach (short, intermediate or long). The DWDM-router interfacial optics are often described as being "white" (less often they are also called "black and white" or "grey"). A transponder module shifts the wavelength of the "white" optical signal from the router to the appropriate "coloured" DWDM wavelength (which actually complies with an ITU-T standard) before it can be multiplexed (with other coloured wavelengths) and the multiplex then sent down the transmission fibre. The conversion is optical-electrical-optical (O-E-O) and the electrical stage is expecting to see a signal with a certain framing (e.g. GE or SONET/SDH). Transponders are costly components and, due to their O-E-O nature, are not transparent to signal framing. The question is: can a router be equipped with coloured optical interfaces thereby allowing the elimination of the seemingly unnecessary transponders in the DWDM line termination system?

The answer is that this situation may come into being at some point, but it is highly unlikely that it will ever happen in a mixed-vendor environment. The reason is that DWDM transmission systems are, by their very nature, complex, proprietary, analogue optical systems. The only "standards" that exist are the ITU-T "grids" that specify the exact wavelengths of the optical carriers used in coarse and dense WDM systems. Other aspects of WDM optical transmission such as power levels, modulation schemes, pulse shaping, multiplexer insertion losses, are not standardised. As the physical characteristics of every (dark) fibre span are unique, all but the most simple of optical WDM transmission system must be custom-made. Hence the chances of being able to get vendor-heterogeneous optical transport systems to work are limited (especially where the technology is being pushed to its limits in terms of capacity). A transponder and multiplexer from one vendor will almost certainly not work with optical amplifiers from another. A possible exception to this situation is where the systems are not being pushed to their limits of capacity. For example, a CWDM or metro DWDM system (with a small, well-separated channel count operating at lower bit rates) may allow some vendor interoperability, but there would never be any end-to-end performance guarantees in such a situation.

There is some evidence of limited integration between transmission equipment and client devices (e.g. routers and switches). At least one optical switch vendor has developed its equipment in close co-operation with a vendor of DWDM transmission equipment. As a result, the switch vendor has interfaces with coloured optics that can be connected directly to a DWDM multiplexer of the transmission equipment vendor. For routers, there is evidence that some vendors (more likely those that also have WDM transmission equipment in their portfolio) may start to introduce interfaces with coloured optics in the form of GBIC-style WDM transceiver modules that plug into a common interface card (for a given speed and framing). This approach will enable inventory costs to be reduced.

An alternative approach is the universal interface card (for a given electrical framing) with a tuneable transmitter. Such an interface is still some way off since the key component is a tuneable laser with high stability. Although these are being developed, the technology is far from mature.

3.2. Functionality

This concerns the **functionality** of a router in terms of its network capabilities - often reflecting where the device is intended to be placed in the network (core, edge or somewhere in between). A router must be able to forward packets, based on tables of routing information that are maintained locally on each router. Where these routing tables are very simple it is possible to configure them manually. This is likely to be the case at the very edge of an IP network close to individual end-users. Moving closer to the core, a router's routing tables will become more extensive and potentially more dynamic. Manual configuration is no longer possible, so the router has to be able to interact with other routers in the network in order to automatically update its routing tables. This is done through the operation of routing protocols such as OSPF or IS-IS within a domain (or autonomous system) and BGP between domains. BGP is the standard Internet inter-domain routing protocol which is responsible for propagating all the Internet backbone routing updates. A core router that maintains a full Internet routing table has to maintain a table containing something of the order of 110,000 routes and potentially up to a few tens of peerings with other BGP-enabled routers (especially in a situation where there is a full mesh of internal BGP peerings within a domain). Clearly routing is a very important function in addition to packet forwarding, particularly for core routers. Contemporary router designs have separated the route processing function as much as possible from the packet forwarding function. This makes more resilient router designs possible where a total failure of the route processing function does not necessarily impact the performance of the packet forwarding function. This has come to be called "non-stop forwarding" and is of increasing importance as routers get ever larger and handle higher volumes of traffic. This paradigm will continue as the next generation of large core routers become "carrier-class" having left behind their low resilience, non-redundant "enterprise" beginnings.

Resilient packet forwarding and routing are not the only functions that a router could be required to perform. Other functions are related to the areas of traffic engineering, security and application of value-added IP services.

Traffic engineering is arguably of most importance in the core of a network and is often intimately related to routing. Internal routing protocols (IGPs) have been extended over the years to accommodate intra-domain traffic engineering features and much of the complexity associated with BGP (and its troubleshooting) is associated with inter domain traffic engineering. Beyond the simplistic intra-domain traffic engineering that is often performed through the manipulation of IGP metrics, Multi-Protocol Label Switching (MPLS) is starting to be used to solve traffic engineering problems. However, many operators of large IP networks still regard MPLS as an immature and non-scalable technology. It is envisaged that take-up of MPLS-based traffic engineering will be slow during the coming years and hence the incentive for vendors to strengthen their implementations is somewhat diminished.

Security features are often implemented through the use of "Access control Lists" (ACLs) or firewall filters. These enable highly specific individual traffic flows, or groups of flows, to be identified and various actions to be performed on these identified flows. For example, flows from a particular application can be identified and blocked, or only traffic flows from a particular source can be passed (all others being dropped), and so on. This kind of functionality can be highly processor-intensive, especially where high volumes of traffic are being filtered. In some (mainly older) router architectures this can lead to extensive degradation of overall packet forwarding performance. It is not really a function that is needed in the core of a network (where high-volume packet forwarding is the priority) and it is much more commonly found closer to the edge of a network.

Application of value-added services takes the security features described above to the next level and often entails the addition of service elements like session authentication, personalised stateful firewalls, VPN encryption, quality of service (QoS) features, "captive portal" style session redirection, etc. These elements can be used as the building blocks for value-added subscriber access services (e.g. enhanced broadband access). They are clearly very processor-intensive and probably more of interest to commercial network operators whose customer base consists of large numbers of individuals or small groups of end-users. Traditionally, niche vendors have served this demand with specialised router equipment. The more mainstream router vendors are increasingly making efforts to extend their portfolios into this space, sometimes by acquisition and sometimes through in-house development. Two examples are the acquisition of Shasta Networks (vendors of a broadband service node) by Nortel Networks and the more recent acquisition of Unisphere Networks by Juniper Networks.

In general, NRENs do not have any requirement for services that utilise these value-added features. However, there is an important trend to note here. As IP networks grow in size and volumes of traffic carried, the kinds of routers that were previously only to be found in the core (in terms of forwarding and routing capability) will increasingly be used at the edges. However, it may not be a case of simply physically migrating the older core routers towards the edge because, although they may have the necessary forwarding performance, they may not have the capability to perform the other functions (packet filtering, etc) that are potentially required there. Therefore, when replacing edge routers NRENs may not be able to recycle older core routers and should consider upgrading to newer mid-range routers.

3.3. Interoperability with other networking components

Traditionally, routers have only really interacted with other routers (e.g. passing routing information or signalling for reservation of resources) when it comes to influencing the behaviour of other network elements. Intervening components such as "layer-2" switches (e.g. Ethernet, ATM, frame relay and SONET/SDH switches) and transmission systems (e.g. native SONET/SDH or DWDM) have largely remained transparent to the routers[2] and are configured independently[3].

The new paradigm of the Advanced Switched Transport Network (ASTN) promises to change this situation with the introduction of transport networks that are capable of dynamically establishing "on-demand" end-to-end dedicated bandwidth channels between client devices (e.g. IP routers). The development of the Optical Internetworking Forum (OIF) user network interface (UNI) has given rise to one of the methods by which this can be achieved. A router signals (via the UNI) to the transport network that it wants a channel and the transport network then takes care of all aspects of establishing this channel including its physical routing. This is often referred to as the "overlay" model - the client devices (routers in this case) have no knowledge of the topology of the intervening transport network and cannot exercise any influence over the physical routing of the newly created channel.

As explained in section 6.2, an alternative approach to the overlay model is one in which the routers have a full knowledge of the topology of the intervening transport network and can force a particular path to be taken by the channel when it is established. This is often referred to as the "peer" model because the client devices must now peer with the transport network in order to exchange routing information. The Generalised Multi-Protocol Label Switching (G-MPLS) architecture that is being developed in the Internet Engineering Task Force (IETF) supports both the peer and overlay model.

It is likely that future generations of IP routers will support both the peer and overlay models. Router vendors show signs of having more affinity with the peer model, partly due to the fact that G-MPLS is essentially an extended form of MPLS - a technology with which most router vendors are already very familiar, having supported it for the last few years. The overlay model will clearly be favoured by transport network operators since it allows them to exercise a greater control over the traffic engineering of their network than would be possible with the peer model.

3.4. Virtual routers

Traditionally, a router has been a monolithic device, behaving as a single logical packet forwarding entity, which can route traffic between any of the line interfaces in the physical device. By the same token, the device belongs to a single autonomous system and is exclusively managed by the network management entity responsible within the associated management domain. It is possible, through the use of more advanced router software and secure element management systems, to partition a physical router into a number of logically separated routers that can be independently managed and even be parts of different autonomous systems. The resulting logical router partitions are sometimes called "virtual routers”. Virtual routers are created by a device manager and subsets of the total complement of physical interfaces in a given device (that is capable of being partitioned) are assigned to each virtual router. In addition to the line interfaces, it should also be possible to partition and assign other resources found on a router. These might include: switching capacity, processors or processor time, buffers, shared memory, and so on. Importantly for a router, routing tables within a virtual router can be kept separate from those within other virtual routers residing on the same physical platform. Once a set of resources has been

---

[2] A slight exception to this notion of transparency takes the form of the exchange of certain types of transmission overhead such as parts of the SONET/SDH overhead used for monitoring transmission performance.

[3] A notable exception to this rule is the interaction between a Switched Virtual Circuit (SVC) enabled ATM network and a suitably capable ATM interface in a router. In this case, a router can signal to the network (via the ATM Forum UNI interface) to establish an ATM SVC to another ATM-enabled router or IP host over which it can then pass IP traffic. However, SVC-enabled ATM networks are currently not widely deployed (and never really have been).

assigned to a virtual router, users of other virtual routers (even those residing on the same physical device) cannot make use of them. Similarly, the management of these virtual routers is also kept separated by the router Element Management System (EMS).

This kind of partitioning of a physical device into multiple independent logical devices has been available for a few years now on server hardware and on certain kinds of devices that exhibit router functionality such as edge aggregation (or subscriber management) systems. Notable examples are the Subscriber Management Systems devices from Redback Networks and Nortel Networks' "Shasta" broadband service node. These devices are very much intended to go at the edge of networks and have niche applications. Typically, they are used to aggregate heterogeneous broadband access services from thousands or tens of thousands of subscribers, and allow the provision of highly tailored IP services on a per-subscriber basis.

Until now, resource partitioning has not been available on large core routers. This will change with the emergence of the next generation of very large core routers. Thus, it will be possible for an operator of a large IP network to provision networks of virtual routers and allow its users to manage them autonomously. For research and education networks, this approach may prove to be a useful way to support Grid projects as described in sections 6 and 7 below.

3.5. Router management

Most routers have in the past been managed through a Command Line interface (CLI) and their configurations have been stored as flat text files, in which each line corresponds to a particular aspect of the overall configuration of the device. This is still largely the case today, even though some router vendors have implemented management systems in a separate box, with embedded database functionality. The general trend over the last few years has seen the development of Graphical User Interface (GUI) based "point-and-click" management systems that have not proved popular with IP network management staff.

Contrast this approach with the situation for transmission and switching systems, which have usually been managed using closed, proprietary, GUI-based element and network management systems. It is envisaged that large multi-chassis routers, due to their complexity and the large size of any flat text-based configuration files, will also be managed using a GUI-based EMS, and CLIs may no longer be available, or be limited to low-level debugging and unusual configuration actions. This situation may not only arise for very large routers with large port counts but may also be the case for smaller routers with higher levels of functionality. This has already been seen to be the case for complex edge aggregation devices such as those described above where a client-server EMS (using Java-based management clients) often fronts a complex CLI to a system whose flat text configuration is lengthy, complex and not very readable.

Element and network management systems are discussed further in section 6 of this report.

3.6. Conclusions

In summary, the interviews showed that routers today widely support functionality like differentiated classes of service, multicast, IPv6, MPLS-based VPNs and G-MPLS. All router vendors are following standards, but there is a need to improve interoperability. Inter-domain functionality is still questionable. Routers are already scalable to terabits capacity, in multi-chassis platforms. 40-Gb/s back-plane support and slot capability exists today. In terms of expected developments, 40-Gb/s interface capability is planned, but not yet available.

4. Switching equipment

Some of the most important elements of modern broadband core networks are cross connects, sometimes referred to as switches. The main function of cross connects is to provide easy and flexible bandwidth and connectivity management for network providers.

Generally cross connects can be divided into three categories:

1. Digital/Electrical Cross Connects (DXC/EXC) are devices that receive, process and send data in the electrical domain; these are mostly old SONET/SDH cross connects, with copper interfaces.
2. Optical Cross Connects (OXC) are digital cross connects equipped with optical interfaces; they are sometimes referred to as O-E-O (Optical-Electrical-Optical) switches and are widely available from many vendors.
3. Photonic Cross Connects (PXC) are devices that receive, process and send data in the optical domain and are sometimes referred to as O-O-O (Optical-Optical-Optical) switches; currently only few vendors produce these devices.

4.1. Optical cross connects

Most of current OXCs are sophisticated devices that are able to work as SONET/SDH cross-connects as well as ATM and Ethernet / Gigabit Ethernet switches with VLANS, CoS and other protocol specific functionality. The name "optical cross connect" indicates that the broadband interfaces are optical, i.e. connected via optical fibre to other devices, but the internal processing is done in the electrical domain. In order to extend their functionality, OXCs can also have several electrical (copper) interfaces, including Fast Ethernet and Gigabit Ethernet over copper.

Optical cross connects use well-developed ASICs, allowing for fast processing of high-bit-rate signals reaching the switching speed of 640 Gb/s and higher for the largest devices. With such a high switching capacity the switching granularity remains very fine, mostly being SDH VC-3 (or SONET STS-1 or 49 Mb/s) or SDH VC-4 (or SONET STS-3 or 150 Mb/s) (RFC 3255). These devices also allow for additional functionality, including link bundling, bandwidth grooming and protection, and Generic Framing Procedure (GFP ITU-T G.7041). GFP is designed to map different services like Ethernet, IP/PPP, Fibre Channel, Ficon, Escon and others into SONET/SDH and OTN networks, as well as LCAS (Link Capacity Adjustment Scheme), which provides more efficient bandwidth usage in combination with virtual concatenation.

	Total system switching capacity	**interface speed[4]**	**granularity**
Alcatel, 1674 Lambda Gate	approx. 5 Tb/s	150 Mb/s – 10 Gb/s, 10, 7 Gb/s	VC-4 (150 Mb/s)
Ciena, Core Director	640 Gb/s (38 Tb/s in multi rack)	155 Mb/s – 40 Gb/s	STS-1 / AU-3
Cisco, ONS 15454	not available	155 Mb/s – 10 Gb/s	VC-4 , switch
Corvis, OCS	240 Gb/s / 11.5 Tb/s	155 Mb/s – 10 Gb/s	STS-1, VC-4
Lucent, Lambda Unite MSS	640 Gb/s	155 Mb/s – 40 Gb/s, GE, 10-GE	VC-3 (49 Mb/s), STS-1, VC-4
Marconi, MSH2k	320 Gb/s, 960Gb/s (2.88Tb/s in multi-racks)	155 Mb/s – 10 Gb/s	VC-4
Nortel, OPTera HDX	640 Gb/s (3.85Tb/s in multi rack)	155 Mb/s – 40 Gb/s	STS-1, VC-4
Tellium, Aurora	1.28 Gb/s (20 Tb/s in multi rack)	2.5 – 10 Gb/s	STM-16 (2.5 Gb/s)

Table 1. Switch survey

---

[4] optical interfaces only

Table 1 shows that although there is a marked difference between the single-rack chassis, which typically have 640 Gb/s capacity, and the multi-rack systems which typically have Tb/s capability, offerings within each class are remarkably consistent.

Protection mechanisms are important parameters of OXCs. As these devices evolved from standard SDH cross connects, their protection options are similar to SDH protection. Modern OXC devices use the following protection mechanisms:

- MS-Spring/BLSR
- SNCP/UPSR
- 1+1, 1:N APS/MSP
- Meshed restoration (with G-MPLS)

The main drawback in the current OXC world is the lack of standardisation, which means that running and maintaining a network with devices of different vendors may be extremely difficult. The following are the most important incompatibilities:

- **Bandwidth grooming** is performed in a proprietary fashion. It will therefore not work between different machines.

- **G-MPLS,** where only a small subset of identified G-MPLS protocols has been standardised, and vendor implementations are still proprietary. Although G-MPLS interoperation has been demonstrated at SUPERCOMM conferences, it is not yet fully operational in off-the-shelf products. It should be noted that G-MPLS protocol development and standardisation in the IETF could (and probably will) be significantly different from the protocols developed and standardised under the ASTN umbrella in the ITU-T. This is primarily due to the various standards bodies addressing different customers or needs.

- **Colour DWDM interfaces** are being introduced to reduce the cost of connecting OXCs to DWDM (so that no DWDM transponders are necessary anymore). This solution is only valid for equipment of a specific vendor.

Even though vendors announce high G-MPLS compliance, most of them require centralised management systems to set up and take down connections. There is limited support for NNI and UNI protocols in routers interfaces (here considered as path terminators).

Interoperability between OXCs from different vendors can only be achieved at the level of well-standardised interfaces like SDH, GE and ATM. Problems still exist with advanced MPLS features.

4.2. Photonic cross connects

The concept of “all-optical networks” emerged in the late 1990s when the telecommunications sector was very buoyant. The expectation, at the time, was that development of research and prototypes into commercial products would take place very rapidly. The global downturn in telecommunications investment resulted in much of the expected progress being delayed, and most of it still appears to be very much on-hold. The standardisation bodies and academic research groups have continued their work on the specification of the requirements for optical networking.

This situation has led to start-ups and older companies filling a small but emerging niche market providing new-generation optical network equipment, including 40-Gb/s transmission equipment and photonic cross connects. New products included photonic cross connects from Calient (Diamond Wave), Corvis (Optical Switch) and Lucent (Lambda Router), which are key elements for the all-optical network concept. These products have been available already for a few years, but due to adverse market conditions a few of them have been discontinued.

PXCs are quite different from OXCs. The most important distinction between an OXC and a PXC is that the former processes data in the electrical domain and converts it to the optical domain at the output interface, while in PXCs the signal remains in the optical domain. A few technologies are used to build PXCs: the most well-known include MEMs (Micro Electromechanical Mirrors) and liquid crystals, while others still remain proprietary technology. Due to their operation principles, PXCs are only able to switch data from one port to another (i.e., by reflecting the light beam with micro-mirrors) or (for some products) multiple other ports in a multi- or broadcast fashion. They are not able to perform grooming or fine-granularity switching as there is no data processing inside the switch.

There are multiple approaches to PXCs: most types of devices are not integrated with transmission equipment and therefore require a DWDM system terminal at the end of the fibre, bringing the different wavelengths back to the electrical domain by means of transponders. In other PXCs, any given port on the switch takes a fibre pair with all the wavelengths in it, optically unravels and switches them to the desired output port, where they are regrouped optically, and sent onto the output fibre associated with that port; consequently there are no DWDM terminals and no electronics in the path: the signals remain in the optical domain all the time.

There are some difficulties associated with the use of PXCs in multi-lambda optical networks. In this context multi-lambda means that the network is built with DWDM equipment so that multiple parallel optical channels are available on each link. In some points of the network, PXCs may be used for wavelength routing in order to provide better network flexibility. As PXC devices operate in the optical domain, they do not perform signal regeneration or wavelength (lambda) conversion. As a result, the effects listed below may impair the network performance:

- **Wavelength blocking** – this situation happens if a number of sources, wishing to use the same wavelengths, are trying to use one link. Due to its nature, all wavelengths in a DWDM system have to have different colours. In this case only one wavelength can be served. The technique called “wavelength conversion” may provide a solution by shifting the remaining wavelengths so that they can be transmitted in single link. Unfortunately, for fully optical networks “wavelength conversion” has to be done in the optical domain and the necessary components are not yet available on the market. Currently conversion is done either using transponders terminating DWDM links close to the PXC, so the switching is performed in “white light” (conversion is therefore unnecessary) or can be done by tuneable lasers, which in fact also perform O-E-O conversion. Such methods however do not provide the benefits of “all-optical switching”. Another way is to integrate PXC and DWDM into one chassis, which additionally reduces the cost. According to some vendors, careful engineering of fully optical networks causes wavelength blocking to appear only if the network is loaded more than 80-85%, depending on topology, number of wavelengths etc.

- **Indefinite wavelength route length –** depending on the current network routing configuration, a given wavelength may use different links to propagate through the network. Therefore, the physical distance between transponders may vary. Currently DWDM devices have to be engineered and tuned for specific links. Usually there is only a small margin for adjustments. In case of all-optical networks, two wavelengths routed from the same source to the same destination may traverse completely different physical links, made of different fibres, with different line equipment, number of wavelengths etc. In such networks – especially with increasing per-wavelength bit rates - it is extremely hard to provide suitable amplification and hence to control polarisation and chromatic dispersion as well as signal level. All these issues are technological challenges that will probably not be solved soon by all vendors. However, the emergence of Ultra Long Haul DWDM systems, with embedded advanced self-adjustment functionality, may solve that problem. The problem of indefinite wavelength route length will of course not exist if the transmission characteristics of maximum length links are better than those required by the DWDM transmission system.

- **High loss** – the signal attenuation of PXCs is around 7-8 dB. This corresponds to approx. 30-45 km of fibre distance. Such a high loss will probably limit the use of PXC to metropolitan areas initially. The study shows that some vendors are building PXCs with technology allowing for almost no-loss operations. This technology however remains confidential and proprietary.

As PXCs do not process data electrically[5], their main advantages are:

a) bit rate independence (the limit is currently 40 Gb/s, but the development of new PMD control methods will support higher bit rates as well);

b) protocol independence (they can switch any known and future protocol without exceeding the maximum bit rate);

---

[5] In some solutions an optical splitter and a special electronic monitoring card can be used to check the parameters of the signal, such as BER etc.

c) wavelength independence (they can switch any wavelength from a broad spectrum, including O, E, S, C and L band);

d) lower cost than OXCs (the price depends on the number of ports; it can be estimated as approximately 3,000-7,000 euro per port for stand-alone PXCs);

e) low footprint and power consumption (it is usually much lower than for an OXC, and does not change if higher signal bit rates are switched).

The above factors make PXCs interesting devices for future broadband core, even if today their use is limited to “patch-panel with management functionality”.

	Switching capacity	**Interface speed**	**Range of port numbers supported**	**Signal. support**
Diamond Wave	40 Tb/s	≤ 40Gb/s, tested 80 Gb/s	8-4096	G-MPLS
Optical Switch	11.5 Tb/s	≤ 40 Gb/s	1-6[6]	G-MPLS, UNI

Table 2. Main parameters of PXC devices

This table is very short, featuring only two products. Even though many companies started development of PXCs, few continue production and many other projects have been stopped. On the other hand, this short summary shows the potential of new technology – including extremely high switching capacity, high number of ports and high interface speed. An additional feature is that the footprint of PXCs is much smaller than OXCs with much higher capacity, lower cost and power consumption.

4.3. All-optical networks

The previous sections described the main features of OXCs and PXCs. Generally manufacturers of OXC devices tend to point out that PXCs are not yet mature. They lack necessary functionality and their use leads to much complexity in network design. On the other hand, PXC manufacturers agree that the best use of their products is in conjunction with advanced Ultra Long Haul (ULH) DWDM systems and TDM equipment allowing for bandwidth grooming.

The following list describes some of the most important limitations for building all-optical networks that have been identified:

- Several vendors work actively in this direction, even though in some cases uncertain and depressed market conditions have been causing a reduction of research funding, forcing companies to put their research on hold.
- Lack of, or very costly technology for, wavelength conversion. Tuneable lasers are not easy to build. They are sensitive to changes in their operating environment and the wavelength shifting range of such devices is limited. Inability to convert between wavelengths leads to non-optimal network usage, mainly due to wavelength blocking.
- Re-routing of wavelengths may lead to the situation where different optical channels (wavelengths) in single fibre have different route lengths, which may make amplification and dispersion control extremely complicated. There may be many constraints for routing protocols – such as maximum route length and necessity to amplify and perform dispersion control at the optical channel level, rather than at the fibre level.
- The quality of the signal is very difficult to control as there is no electrical processing in the network.
- External TDM devices are necessary for bandwidth grooming.
- DWDM systems are analogue. Interoperability between different DWDM terminals of different vendors is not supported yet because of the lack of standardisation in areas such as wavelength power, modulation, FEC mechanisms and wavelength grids. In the case of early-production optical devices, this lack of

---

[6] One such port can support up to 140 wavelengths (C-band only) or even 284 (C- and L-band). Matrix scalable to switch 1136 optical channels.

standardisation leads to incompatibilities between the equipment of different vendors and the need for lengthy manual fine-tuning of each optical link.

4.4. The future of switching equipment

Several innovations are being introduced in the area of circuit-switched optical networks, with various degrees of development. The main developments that are expected to become available in the next five years regard G.709 interfaces on routers and switches (standardised FEC), wide support for Generic Framing Procedure (GFP), 40-Gb/s upgrade for all OXCs, G-MPLS standardisation and interoperability, integration of PXC/OXC and DWDM into single chassis, fast optical circuit switching, optical burst switching and virtual switches. Details of these developments are provided in the following.

- **G.709 interfaces on routers and switches (standardised FEC).** Digital Wrapper (G.709) is a new standard that defines networking for the new Optical Transport Network (OTN) which includes management of single and multiple wavelengths. The purpose of the G.709 protocol is to carry a client signal inside a DWDM link. This functionality is achieved by encapsulation of the client channel (i.e. STM-64 – 10Gb/s) into the slightly faster (approx. 10.7 Gb/s) DWDM channel. Thus the customer signal remains undisturbed and an extra space is available for standardised control bits and Forward Error Correction (FEC). FEC is fully standardised in G.709, but G.709 gives vendors options to develop proprietary enhancements (Ultra FEC). It is envisaged that, in the very near future, this will be overcome by the availability of standardised FEC chipsets (see section 9 on future and emerging technologies).
- **GFP support.** Even though it seemed likely that Digital Wrapper (G.709) would become the dominant framing mechanism for transporting any type of signal over a DWDM network, some vendors' view on this has shifted because it became apparent that most carriers will not write off their current investments in SDH/SONET-based systems and will not replace this familiar framing technique for something new. In order to transport non-SDH based signals, a different approach seems to prevail now. This technique consists of using Generic Framing Procedure (GFP, G.7041) to map any signal into an SDH frame (or several ones through virtual concatenation) and transport it over the existing SDH/DWDM networks.
- **40-Gb/s upgrade of OXCs.** Currently most OXCs are using 10-Gb/s interfaces, since 40-Gb/s interfaces are not in commercial production. The consequence of this is that the cost of 40-Gb/s chipsets, and therefore interfaces, is extremely high and likely to deter all but the most committed early adopters.
- **GMPLS standardisation and interoperability.** Full standardisation of the G-MPLS protocol set is something to be expected in a few years' time. It is very likely that the current set of protocols will be replaced by more advanced ones. It is possible that G-MPLS concepts will be adopted in SDH networks.
- **Integration of PXC/OXC and DWDM into single chassis.** An important element of the price of optical networks is related to laser interfaces. Integration of PXC/OXC and DWDM into a single chassis could significantly lower the price of the optical network, by reducing the number of interfaces.
- **Fast optical circuit switching and optical burst switching.** Fast optical circuit switching and optical burst switching technologies are based on the assumption that the optical network connections can be set up for a specific user and for a specific (usually very short) time. This is a similar approach to the ASON standard, but stresses more the switching speed (therefore, provisioning time). In the ASON standard the provisioning time may reach minutes and the connection duration may last from hours to years. Burst packet switching aims at providing very short-lived connections in a very short time – just long enough to deliver one optical packet or a packet burst. IP data is assembled into optical packets at the edge device, then the optical path is established between source and destination device. A packet (or burst of packets) is sent through the link in very short time, and the path is taken down. The same process is repeated when the next data arrives. In this way burst packet switching can add more dynamism to optical networks as multiple optical packets from multiple sources would be delivered to multiple destinations in a single time unit. Appropriate classification according to QoS, routing and switching and optical label swapping is performed at optical network and Core Optical Packet Switches. This development requires specialised interface cards, able to work in a bursty manner. The most likely beneficiaries of the development of such equipment are Grids, which seem to need such a solution.
- **Virtual switches.** This is an extension to switched equipment of the same concept described in section 3.4 above. A virtual switch is a mechanism allowing the partitioning of a real physical switch into a number of logical, independent devices, each having dedicated resources.

The long-term future of switching equipment (beyond five years) will see the emergence of **fully optical control planes**. This improvement is necessary for development of all-optical networks. Unfortunately, this is not going to happen in the near future – full optical control planes require optical computers, which are not yet available. It is estimated that this development will happen not earlier than 20-30 years from now.

4.5. Conclusions

Optical Cross Connects (OXC) scale to hundreds of Gb/s and higher switching capacity, by using advanced ASICs. Bandwidth grooming is performed with proprietary techniques, thereby affecting interoperability. Interfaces support GFP and colour DWDM interfaces, but there are only some proprietary examples of the latter and they will only work with same vendor's transmission equipment. Support for G-MPLS is provided but implementations still have proprietary features, although some interoperability of OXCs between vendor equipment has been demonstrated.

Currently, very few Photonic Cross Connects (PXC) are available and little investment is put in their development. The earliest envisaged use of smallest products is as "remotely manageable optical patch panel". PXCs enable a significant saving on O-E-O conversions, hence they offer smaller footprint, power consumption and cost. Other advantages are bit rate, protocol and wavelength independence as well as the ability to scale to tens of Tb/s switching capacity. The main difficulties of PXCs are re-routing of wavelengths leading to optical channels having different route lengths, control of amplification and dispersion and QoS control, need for external TDM devices for bandwidth grooming. Finally, interoperability with other vendor equipment is still an issue.

5. Fibres and transmission equipment

5.1. Fibre types and capacity

In an optical transmission system, information is transmitted over the fibre in the following way: electrical signals are converted into light signals that travel down the fibre until they reach a “detector”, which then changes the light signals back into electrical signals.

The light source emits light pulses at particular wavelengths. A wavelength is also referred to as a lambda or channel. The terms lambda, wavelength and channel are often used interchangeably. Though the wavelengths are invisible to the eye, they are nonetheless often referred to as “colours”.

Fibre capacity is very high. State-of-the-art transmission equipment supports data rates of 10 Gb/s in one lambda and 128 lambdas, which corresponds to 1.28 Tb/s total capacity over a single fibre. A fibre route could have eight ducts with each optical cable containing 192 fibre pairs, so the capacity could carry over more than a Petabit per second (1 Pb/s equals 10^{15} b/s). In practice much less capacity is used, and we see cases of cables with six fibres utilised for single-colour transmission rate at 10 Mb/s.

The current trend in carrier systems for long-haul optical transmission is to provide higher bit rates per lambda as well as to increase the number of lambdas supported in a single optical fibre. A capacity of 10.2 Tb/s over a single fibre has been demonstrated in the laboratory. However, many non-linearity problems occur when systems are operated at such limits (see the discussion in section 9.4), requiring the latest or next generation of fibres with inherent high-quality properties.

The properties of the first generation of Single Mode Fibres (SFM) are specified by the ITU G.652 recommendation (so-called standard single-mode fibres). Most of single-mode fibres used today in Europe probably conform to the G.652 recommendation, and the same is true for submarine optical cables. New fibre routes are usually equipped with G.655 fibres, which are more suitable for 10 Gb/s and higher transmission rates.

G.655 is a standard for Non-Zero Dispersion Shifted Fibre (NZDSF). G.655 optical fibre is designed for use with multi-channel Dense Wavelength Division Multiplexing (DWDM) systems. G.655 specifies a non-zero dispersion shifted fibre, which represents the optimum compromise between chromatic dispersion, which degrades pulse shape at high data rates and results in Inter-Symbol-Interference (ISI), and Four Wave Mixing (FWM), an undesirable interaction between high-power optical signals at different wavelengths in the same fibre.

The ITU G.692 recommendation describes optical channel plan and spacing (e.g. 50 GHz, 100 GHz, 200 GHz) for wavelength division multiplexing in fibres.

High-capacity, long-distance links require multi-parameter optimisation including the number of lambdas and their frequency spacing, the modulation format, the distance between optical amplifiers, the power level at the fibre input, non-linear effects such as self-phase modulation (SPM), cross-phase modulation (XPM), four-wave mixing (FWM), the choice of the type of fibres and dispersion compensating fibres and their arrangement.

Chromatic dispersion of standard single-mode fibres, which constitute a large fraction of the installed fibres is about 17 ps/(nm.km) at a wavelength of 1550 nm. The effect of chromatic dispersion is known to scale with the square of the bit rate and should, therefore, be carefully compensated in high-speed systems (at 10Gb/s and higher transmission rates). The most mature and promising technique is the use of dispersion compensating fibres (DCF). DCFs are customarily used as discrete modules, which can be inserted at amplifier stations and therefore allow for an easy upgrade of installed systems towards higher bit rates. The general rule for efficient management of dispersion and fibre non-linearity in DWDM systems is to keep the local dispersion sufficiently high, to reduce non-linear effects (XPM, FWM), while ensuring that the total accumulated dispersion is close to zero for each channel over the link to suppress Inter-Symbol-Interference (ISI).

The choice of fibre types for long distance transmission is usually quite limited by existing fibre trunks. Some fibres fabricated before PMD control was introduced (roughly before 1992) are used in Europe, particularly in western and northern regions. Such fibres are used up to transmission rate (basic rate) of 2.5 Gb/s and require highly expensive compensation equipment to be able to support higher bit rates.

5.2. Transmission components

The main components of a fibre optic transmission system are fibre and line terminators, i.e. transmitter and receiver. Transmitters accept coded electronic signals, convert them into light signals (modulation), and send them down the fibre. Semiconductor laser diodes (LDs) can be used for light generation and are suitable for long distance transmission. The light source emits light pulses at particular wavelengths. LDs can transmit them in the 1550-nm range and have modulation capabilities up to 10 GHz. External modulation of light is used at bit rates up to, and higher than, 10Gb/s. Expensive LDs are not needed for short-distance transmission.

Receivers are placed at the other end of the fibre. Receivers use a photo-detector to convert the incoming light signal back into an electrical signal. The wavelength designation of the receiver must match that of the transmitter. Important characteristics are Saturation, Sensitivity and Bit Error Rate (BER) for the digital level and Signal-to-Noise Ratio (SNR) for the analogue level. The Bit Error Rate is the number of errors that occur between the transmitter and the receiver. The Saturation defines the maximum received power that can be accepted. If too much power is received, the result is a distortion of the signal, causing poor performance. Sensitivity is the minimum power that must be received on an incoming signal. A weak signal can cause misread bits or low SNR.

Wavelength division multiplexing (WDM) is a method enabling a single optical fibre to carry multiple wavelengths (channels).

Optical Add and Drop Multiplexer (OADM) is equipment that adds or removes traffic (i.e. some number of channels) from an optical circuit without requiring conversion to electrical signals. As the traffic to a network node grows, OADM can be used for optical bypass of transit traffic. Advanced OADM equipment eliminates the need for complex engineering through its add/drop capability for any or all optical channels carried by the fibre pair without affecting the rest of the transit traffic. They are scalable to supporting up to 284 optical channels, 100 % dynamic add/drop capability and built-in dynamic gain flattening (DGF). OADM equipment can be upgraded to optical switches (see section 4) to support multiple fibre routes, and switch optical channels without performing O-E-O conversion.

Introduced in the late 1980s, Optical Amplification (OA) redefined the economics of optical networks by extending the distance between costly Optical-Electrical-Optical (O-E-O) regenerators. A typical O-E-O network might use optical amplifiers every 40-100 km, depending on the type of equipment, and have O-E-O regeneration every 400 to 500 km. O-E-O regeneration requires separate conversion and amplification for each channel. Erbium (EDFA) and Raman amplifiers are frequently used for pure optical Long Haul (LH) transmission (e.g. 2,000 km) and Ultra Long Haul (ULH) transmission (e.g. 4,000 km), i.e. without O-E-O conversion.

The purpose of O-E-O conversion is restoration, reshaping and retiming (3R) of the optical signal. This is not completely possible with optical devices available on the market today. However, ongoing research activities are promising and we may expect pure optical regenerators in the future.

Several vendors provide complete systems for metro, LH (up to 2,000 km) and ULH (up to 4,000 km) fibre transmission. The transmission rate per wavelength is usually up to 10 Gb/s today, but 40-Gb/s transmission systems are being prepared for overland transmission. Depending on capacity or interconnection demands, up to 128 wavelengths may be transmitted on a single fibre. Vendors of routers and switches usually incorporate some transmission components into systems designed for metropolitan networks, in a few cases also for LH networks. From the economic point of view, the most important factors are price, footprint for small configuration and scalability. Due to their analogue nature LH and ULH pure-optical multi-wavelength transmission systems are rather vendor specific and deployment requires some engineering work, depending on fibre parameters. If multi-vendor interoperability would become available, price advantages could potentially be achieved by combining equipment from different vendors.

5.3. Reach

Fibre, splices, connectors and all other equipment between the transmitter and the receiver have physical characteristics (attenuation, dispersion etc.), which distort the optical signal. Excessive distortion means unsuccessful transmission. For this reason, devices for optical amplification, dispersion compensation and regeneration are used for longer distances.

5.3.1. Reach of current interface cards in routers and switches

In the metropolitan area network, interface cards delivered with available routers or O-E-O switches (with transmitters and receivers) can be used with G.652 dark fibres for a maximum distance of 125 km at 2.5-Gb/s transmission rate per lambda (the maximum reach of former-generation equipment was 80 km) and for a maximum of 80 km at 10 Gb/s per lambda.

The transmission capacity is higher with the use of multiplexers and demultiplexers for DWDM systems. Cheaper CWDM systems with lower density and wider band may be used, if needed, for further amplification (e.g. “upgrade” of distance). Transmitters and receivers are available for bi-directional transfer on single fibre with reach up to about 125 km and transmission rate up to 1 Gb/s. Transmitters and receivers for transmission rates at 40 Gb/s or higher speeds are technically available, but not yet commercially deployed. In such transmission systems there are no intermediate devices between the transmitter and the receiver (no in-line devices).

With G.655 fibre and narrow band, higher reach can be achieved at 10-Gb/s (and higher) transmission rate per lambda. For example, transceivers for Ethernet data can drive fibre up to 120 km. The following are typical distances and prices for Ethernet transceivers: Fast Ethernet (100 Mb/s) transceivers for driving fibre up to 80 km cost about 700 euro per end. Gigabit Ethernet transceivers for driving fibre up to 60 km cost about 2,000 euro per end. But the prices of transceivers are dropping dramatically now. As an example, 10 Gigabit Ethernet transceivers chip sets capable of driving 40 km of fibre are being sampled at less than 100 euro. These transceivers can usually be controlled and managed by standard LAN management systems. There are numerous companies that sell fibre transceivers. Most Gigabit Ethernet equipment manufacturers include long-haul lasers that can also be directly attached to the dark fibre.

5.3.2. Reach with Nothing-In-Line (NIL) method

Reach can be further extended without placing transmission equipment in any points (premises, huts) in line. This means that additional optical amplifiers are only placed at both ends of the dark fibre and connected directly to the termination equipment. As an example, CESNET is using NIL connection with G.652 fibre with a length of 189 km at 1 Gb/s transmission rate with standard Cisco GSR 12016 GE card and Keopsys 24 dBm EDFA optical amplifier as booster (post-amplifier) on the transmitter side. The list price of this amplifier is around 18,000 euro. The line has been carrying production traffic between Prague and Pardubice since 17 May 2002 without problem. The same device was tested with Cisco OC-48 card (2.5 Gb/s) and with narrow-band DWDM equipment. In this test the reach was extended up to 230 km by adding another optical amplifier (pre-amplifier) at the receiver side, and up to 280 km by further adding a Raman amplifier on the receiver side. Additional tests are currently being planned for 10-Gb/s transmission rate using both G.652 and G.655 fibre. Equipment from one of the vendors interviewed has reached up to 350 km (depending on fibre and capacity) without in-line devices while supporting up to 80 wavelengths. By completely eliminating in-line amplification in point-to-point links, this equipment supports fibre network architectures that span aquatic, terrestrial or mixed geographical terrain where the deployment of in-line amplifiers is costly or not possible. An example is provided by France Télécom, who upgraded one of their network links to 4 x 2.5Gb/s. This link, which now connects the French mainland to the island of Corsica, will provide France Télécom with four times more capacity, resulting in a lower cost per bit of information.

5.3.3. Long Haul and Ultra Long Haul transmission systems

Repeated in-line amplification is necessary in pure optical Long Haul (LH, up to 2,000 km) and Ultra Long Haul (ULH, up to 4,000 km) transmission systems. The distance between amplifiers is called a span. Usually the span length is 40-100 km, depending on the type of equipment. In some situations a greater span (200-250 km) is needed, for example in the submarine sections (in order to avoid underwater amplifiers). For a limited number of lambdas it is possible to build transmission lines with greater span, e.g. 200 km instead of 100 km at 10-Gb/s bit rate per lambda. A trade-off to achieve such a span for a high number of lambdas (e.g. 80) is using Raman amplification and Forward Error Correction (FEC). Maintaining the quality of multi-lambda signals through amplification steps is crucial. Signal distortion will accumulate over repeated spans. In general, cheaper equipment per amplification step is needed in LH when compared with ULH systems.

Concerning long haul transmission, with 10 Gb/s, one can find equipment in production networks today that is capable (after extension) of carrying over 280 channels of 10G each over distances of more than 3,000 km. With the present technology, 40-Gb/s systems could reach 1,100 km (with system margins) by using the CS-RZ format with 100-GHz spacing and 80 km long fibre spans. These ULH systems would need dynamic dispersion compensation at the receiver, active gain equalisers, Raman amplification and Polarisation Mode Dispersion (PMD) mitigators (depending on the line fibre).

5.4. Economics and future expectations

Currently it is possible to build a pure-optical fibre network with point-to-point fibre lines 4,000 km long. Pure-optical means that O-E-O conversion takes place at the user interface only. Pure-optical fibre networks should be considerably cheaper than legacy networks for given long distances and transmission capacities. A typical transmission system architecture is illustrated in figure 1 below.

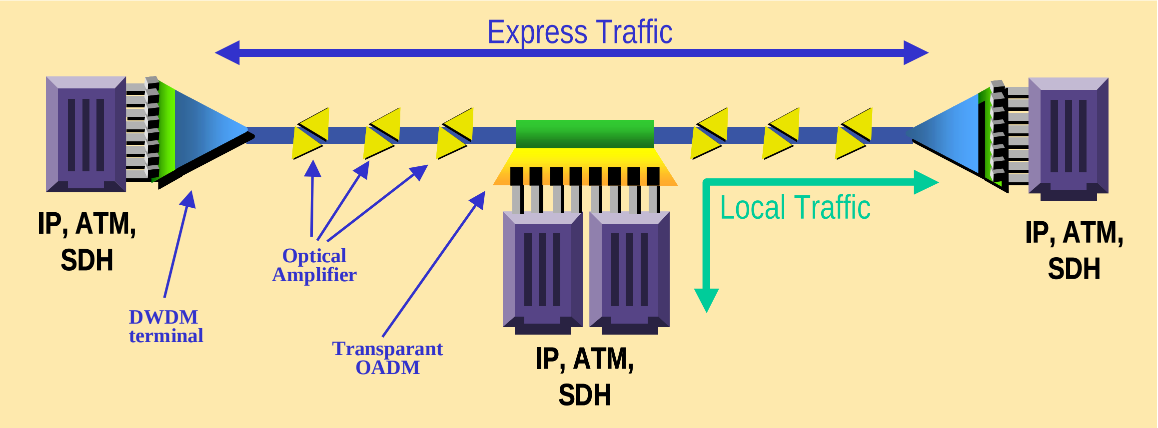

Figure 1. Optical transmission system architecture

The benefits of such a network include the removal of the transponders and electrical interfaces from intermediate sites (such as the OADM site above) and regeneration sites. In a traditional architecture, where circuits may pass through several nodes before reaching their destination, transponders and electrical interfaces are needed at every link of the path. In a true all-optical network only one set of transponders is required.

Long Haul and Ultra Long Haul system components are expensive, but seem to be more cost effective than legacy transmission systems, by removing the need for intermediate O-E-O conversion. In fact, in a traditional architecture, all signals must be regenerated in every node as well as in regeneration sites, because there is no way to only terminate one or a few wavelengths and let the others continue. As a consequence, covering a long distance, with shorter-reach optical systems, could be more expensive than using an optical system with long or ultra long reach, because in the latter ones there is no need for the same amount of regeneration.

Prices of equipment are relatively important for the economics of transmission systems, but the overall costs heavily depend on the network architecture and topology. The basic approach is to connect geographically neighbouring universities by fibres. NRENs should carefully analyse the pros and cons of any solution going beyond this, i.e. using an overlay fibre structure, because such a solution is usually more expensive. It would be feasible and cheap, albeit not simple, to independently cross national borders to build regional fibre networks. This situation requires a new and careful analysis of the architecture and topology of the transmission layer (typically a transmission structure overlaying the NREN’s topology) at the European level.

O-E-O conversion in transmission systems is quite cheap for single-colour systems but very expensive for transmission systems with a high number of wavelengths, because they need separation (demultiplexing) of colours, followed by conversion of each colour and multiplexing. Legacy LH and ULH multi-wavelengths transmission systems are very expensive because of frequent O-E-O conversion. But, as seen above, with the current availability of pure-optical transmission systems for distances up to 4,000 km (and more in an experimental phase) the situation is gradually changing. Moreover, we can expect further price advantages if production quantity increases.

ULH technologies are eliminating O-E-O conversion for network traffic except at source and destination points. Reducing the requirements for regenerators, electrical switches, and their associated costly elements can potentially lead to saving 20 to 40 percent on capital costs and 50 to 60 percent on operational costs of long haul platforms, compared to costs of legacy transmission systems. Moreover, all-optical networks enable carriers to deliver wavelength services on demand. Some vendors' product suites support on-demand provisioning and reconfiguration. This simplifies network management and speeds up configuration tasks.

Expected developments in the next 2-5 years go in the following directions:

- further increasing reach of fibre transmission system;
- decreasing price of fibre transmission equipment as a result of technology advances and of increased production quantity;
- a move to Ethernet framing and wide deployment of 10-GE in Storage Area Networks (SAN), local, metro, LH and ULH networks;
- increasing proportion of customer-empowered fibre networks and decreasing proportion of carriers networks, increasing number of homes and institutions connected by fibre;
- vendors' awareness that NRENs are an emerging market segment, important for new equipment development, testing and evaluation;
- availability of equipment for multiple 10-GE transmission on dark fibre;
- 100-GE standard and first implementation;
- further integration of supercomputers, long-reach transmission equipments and wide-area networks (illustrated for example by the project to connect the four biggest supercomputers in the United States to create by 2004 one biggest supercomputer);
- improved collaboration of researchers from network equipment producers, supercomputer producers and NRENs;
- wider support of fibre infrastructure (leading to deployment of advanced fibre types) by municipalities in a way similar to other infrastructure (water or gas plumbing, heat conduction, electric power transmission etc.).

5.5. Conclusions

The study shows that capabilities of current state-of-the-art DWDM transmission equipment far exceeds the bandwidth needs expected in the next few years, possibly with the significant exception of R&D networks. The limitations of such systems are mostly due to their analogue nature. Standardisation developments are progressing, but the only "standards" are ITU grid wavelength specifications. As a consequence there is still little vendor interoperability amongst transmission components.

The analysis of equipment made clear that the reach of transmission systems results from a complex equation depending on fibre type (G.652, G655….), capacity of each wavelength and number of wavelengths, amplification and transmission technology used, FEC and lots of other physical parameters. Reach with Nothing In Line (NIL) is possible as showed by the experience of CESNET, with pre- and post-amplification (using RAMAN amplifiers) being able to reach up to 280 km at 2.5 Gb/s. Longer reach (350 km) with NIL was demonstrated too. Long Haul (up to 2,000 km) and Ultra Long Haul systems (up to 4,000 km) require amplification at each span (from 40 to 200 km depending on the number of wavelengths) but can reduce costs of O-E-O regeneration. 40-Gb/s transmission systems can reach more than 1,000 km without regeneration with 80 km spans.

6. Network management of networks at 40 Gb/s and above

National Research and Education Networks (NRENs) have been pushing the boundaries of new optical technologies to the limit in order to provide bandwidth for new applications such as Grids. As networking speeds have increased, so too have the requirements on optical networking and the requirements to manage, configure, collect alarms and operate these networks.

This report looks at the technical problems associated with developing networks at speeds of 40 Gb/s and higher. In order to build networks with 40+Gb/s links new network designs and topologies and network management paradigms have to be developed.

At 40+ Gb/s a whole new family of optical networking equipment, like DWDM, optical switches, add-drop multiplexers, amplifiers etc., is required to build such networks. It is possible to manage a 40-Gb/s based network today by using traditional SDH architectures. However, the introduction of more complex equipment into the network poses new problems of how to manage, configure and operate it. There are two main groups developing standards for optical networking, the IETF's IP-over-optical working group and the ITU's Study group 15, (Optical and transport networks).

As explained in section 3.3, the IETF is developing G-MPLS, which details both an optical networking architecture and the signalling protocols required to manage the optical and IP management layers. The ITU is developing an Automatically Switched Transport Network (ASTN) G.8070 and the protocols required for this network. A lot of work is still required to complete both the ITU and IETF standards and as there is overlap between the two bodies, it is unclear as to which standards will be used and developed further. The following picture shows the main standards organisations working on optical networking and the interactions between these organisations.

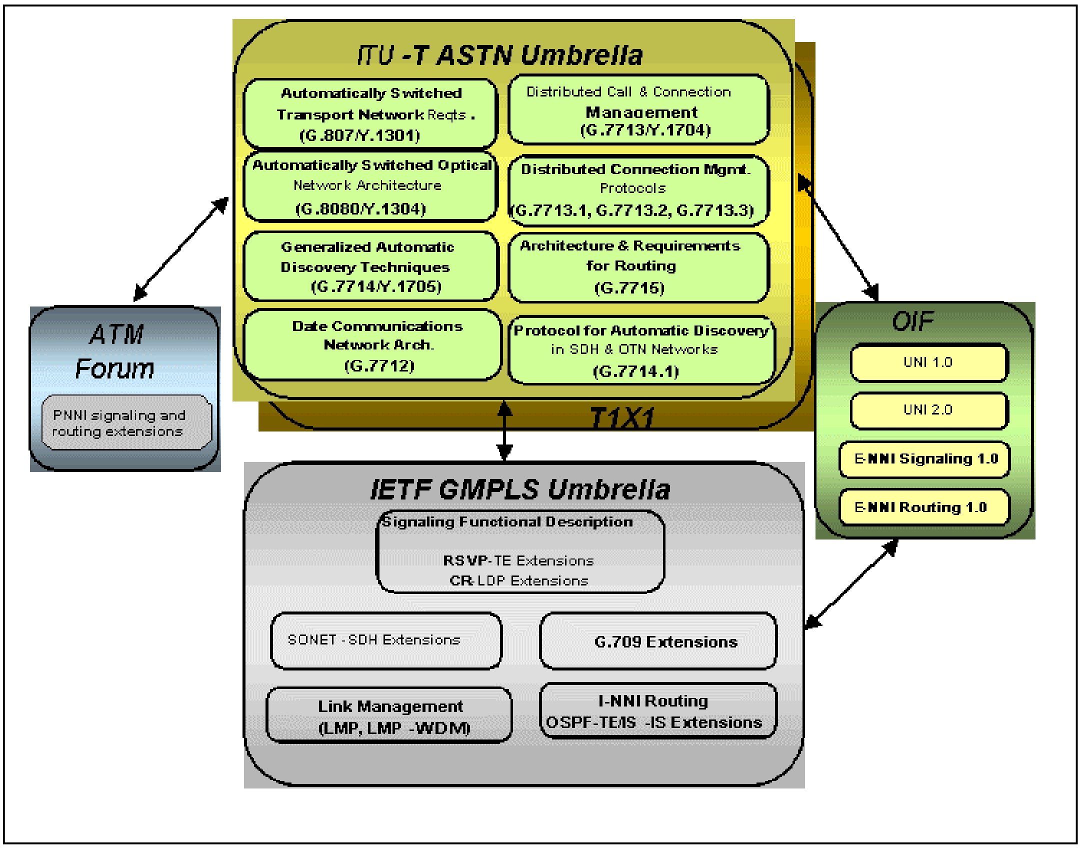

Figure 2. Optical networking standards

In line with the ITU TMN (Telecommunications Management Network) standards we will not consider the control plane to be part of the network management system and therefore we will examine the issue of managing the network separately from the way in which it is controlled.

6.1. Network Management Systems (NMS)

A Network Management System (NMS) contains collective information about all the network elements within the system. The NMS would usually consist of several distributed applications (alarm aggregators, CLIs, polling applications etc). It is the collection of all these that makes a NMS. The Telecommunications Management Network (TMN) architecture is traditionally sub-divided into five conceptual layers. These range from the bottom Network Element Layer (NEL) to the top Business Management Layer (BML). These layers are only conceptual, and tools and processes can cover more than one layer. The TMN model has traditionally been used by telecommunications operators and companies that have to manage very large complex multi–vendor equipment networks.

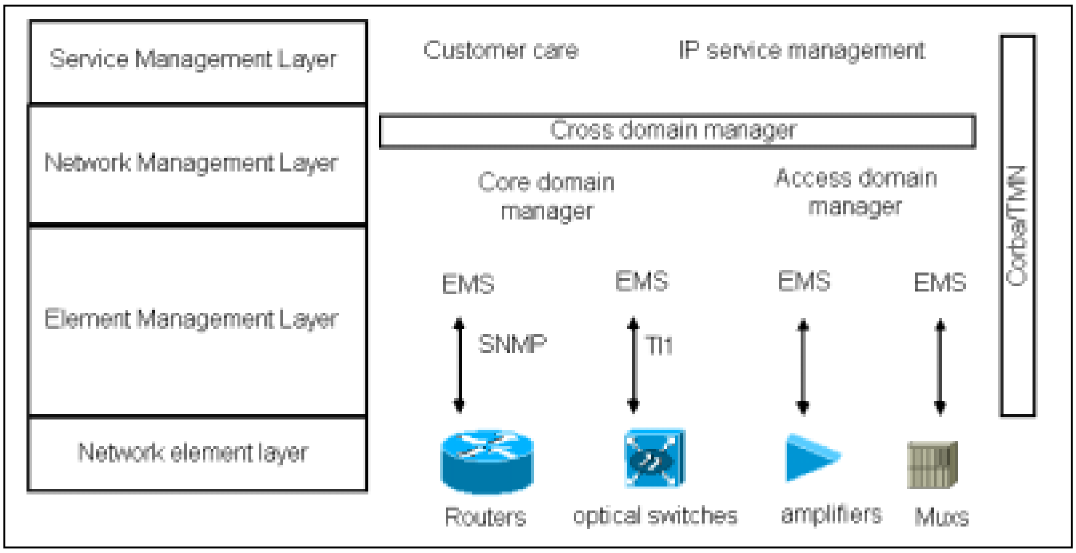

Figure 3. A typical TMN architecture

The IETF has developed its own approach to network management and does not use the same 5-layer model as the ITU. The framework developed by the IETF is called the Simple Network Management Protocol (SNMP).

Traditionally each networking layer and its resources where managed by dedicated management systems, each with a view of only their own resources. In order to manage all these different managers, many telecommunications companies and NRENs have developed dedicated proprietary management systems to integrate new networking equipment into their existing management systems. These systems are extremely expensive to develop and maintain, as most of the APIs are proprietary. However, this is changing with the introduction of new middleware software like the Common Object Request Broker Architecture (CORBA). CORBA has well-understood APIs, which can be used to combine several separate network management systems under one network management system.

NRENs only have to contend with managing single-platform networks of, mainly layer-3, IP equipment. NRENs then use their NMS to configure or provision services using SNMP to distribute information to the correct network elements. This approach works well for NRENs but it requires the support of SNMP and managed information bases to be developed for network elements. Unfortunately a lot of optical equipment manufacturers are tailoring their equipment for the telecom market, which does not traditionally use SNMP for network management. Hence NRENs will have to incorporate new protocols such as CMIP, Q3[7] and CORBA into their network management systems to manage optical equipment, such as optical switches etc.

Indeed, incorporating all these new devices under a single network management layer will require much system integration between existing network management systems and new management systems. A lot of work is ongoing at the moment to make open APIs to enable system integrators to integrate new element managers into their management systems. There exists a telecommunications domain task force within the Object Management

---

[7] Q3 is a management interface standardised by ETSI and ITU-T, the US equivalent, being standardised by ANSI is TL1.

Group (OMG) whose mission is to generate CORBA-based standard services that are relevant to the telecommunications industry. The CORBA/TMN inter-working specification details how CORBA-based management systems may interact with systems based on GDMO/CMIP and SMI/SNMP. It also includes a specification to give CORBA-based management systems the powerful management capabilities of CMIP. CORBA has gained widespread acceptance within the telecommunications equipment industry as a solution for solving interoperability issues in OSS.

6.2. Control planes

Much work is still required to complete the ITU-T ASTN umbrella of standards. The Optical Internet Forum (OIF) has developed UNI and is currently standardising the E-NNI interface to the optical network, which are complementary to the ITU-T optical transport network. The optical UNI includes signalling for connection establishment, neighbour discovery, fault detection and others. The ITU has made no decision on G-MPLS as the signalling protocol for the control plane of the optical transport network. Indeed the ITU-T are extending and enhancing a number of protocols to serve Service Providers including ASTN RSVP-TE, CR-LDP and PNNI. The status of these protocols is the following:

- OIF UNI G-MPLS addresses the client/user signalling – i.e., represents the call management portion. OIF used the base G-MPLS signalling protocol and has extended/modified it to support UNI 1.0. It supports both RSVP-TE and CR-LDP based signalling protocol options. Expected enhancements should support further functions in UNI 2.0 (e.g., bandwidth modification, support for Ethernet signal types).
- OIF E-NNI G-MPLS. Work is starting in specifying an implementation agreement for E-NNI signalling specifications (a close linkage between ITU-T Rec. G.7713.x series is expected). In March 2003, the OIF arranged the first transport network solution integrating UNI1.0 and NNI, with 12 vendors.
- OIF R2.0 is also aligning itself on Link Management and Discovery (G.7714).
- IETF G-MPLS. G-MPLS is continuing to evolve as new requirements impact its development. A “Toolkit” approach with various options has been adopted; this protocol set is not tailored according to interface type. It provides RSVP-TE and CR-LDP based signalling protocols. The IETF is continuing to discuss technology specific extensions (e.g., SONET/SDH, G.709).
- ITU-T ASTN. The work is moving quickly on G.7713.1, G.7713.2, G.7713.3 addressing PNNI, ASTN RSVP-TE-based and G-MPLS CR-LDP-based signalling, respectively. The ITU-T is also actively working the specifications G.7712 (Data Communications Network), G.7714 (Link Management and Discovery) and G.7715 (E-NNI Routing).

The equipment vendors interviewed had various views on G-MPLS and the overlay and peer model (see also section 3.3). Most equipment manufacturers support both the overlay and the peer model. While most manufacturers have G-MPLS implementations or are developing them, the more traditional manufacturers believe the overlay model is the most useful, with an UNI interface controlling signalling across the network.

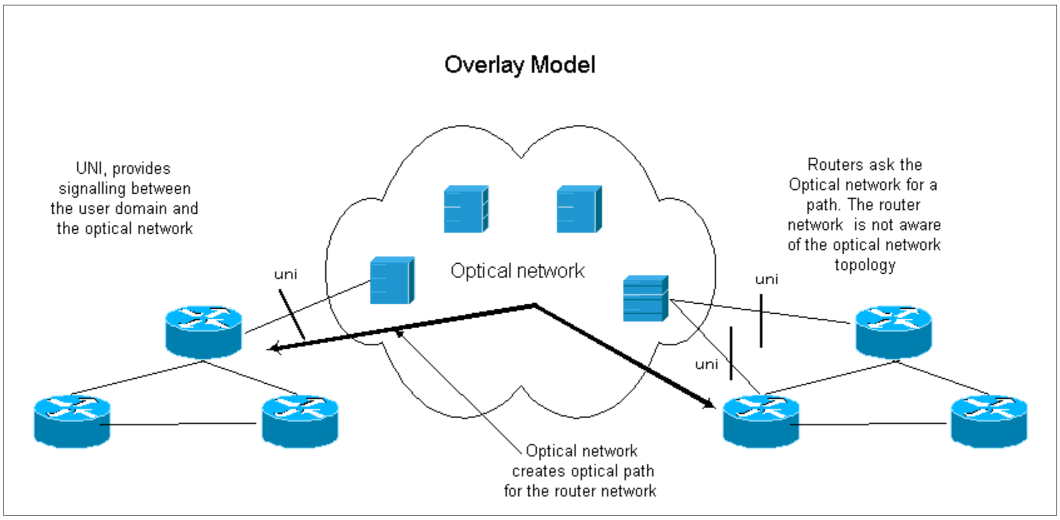

Figure 4. Overlay model

Most of the traditional telecommunications equipment providers favour the overlay model as it simplifies the network management. The overlay model has it limitations. However, it is a step in the right direction. By having two control planes, the data layer and the transport layer, it simplifies management and the complexity of the network.

In the peer model, optical switches and routers act as peers, using a uniform and unified control plane to establish label-switched paths across them with complete knowledge of the network resources. The peer model allows complex layered networks to scale by building a forwarding hierarchy of interfaces, from fibres all the way up to routers. Label switched paths (LSPs) can be established within each layer and " nested " within others so that an LSP beginning and ending on optical switch interfaces may contain many LSPs within it that begin and end on routers.

The peer model has several disadvantages, related to management. Firstly it is one large control plane, with many network elements. Hence each network element will have to maintain information relating to the entire topology of the network. This also adds to the complexity of the network and increases the difficulty of troubleshooting.

Router vendors are in general pushing the peer model and G-MPLS, as they naturally favour routers over optical equipment when it comes to control mechanisms. The peer model does have a scalability problem because of the amount of topology information that has to be maintained by each network element. The peer model is a lot more complex than the overlay model but it does have significant utilisation and optimisation advantages.

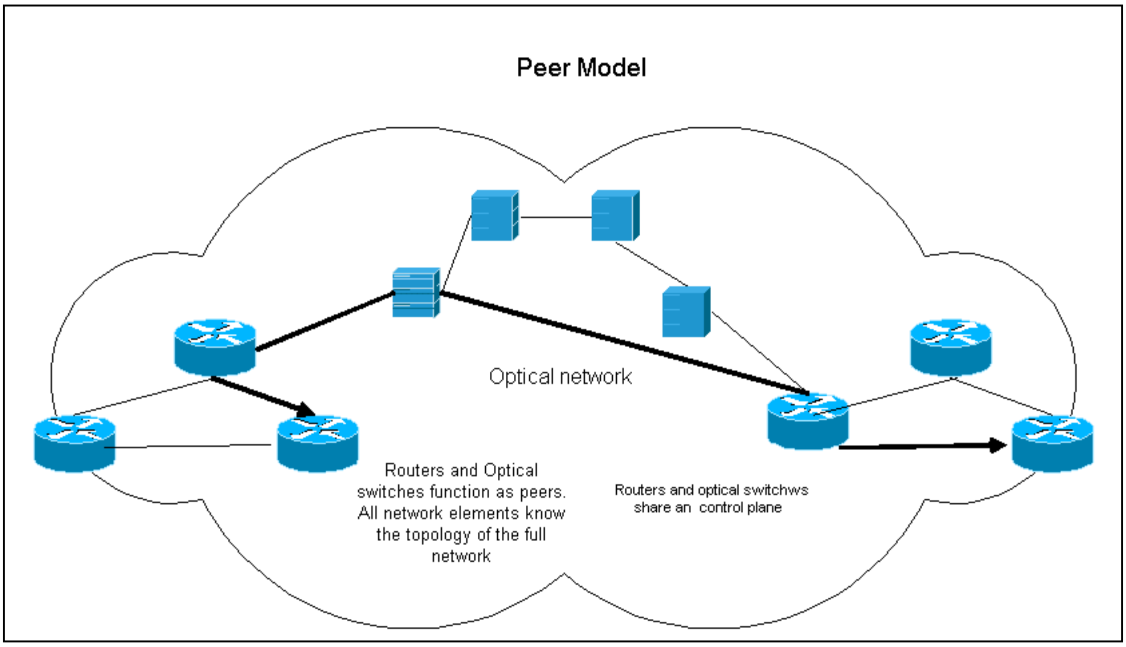

Figure 5. Peer model

The hybrid model represents a middle ground between overlay and peer, making use of the most suitable methods from both approaches. From the overlay model, the hybrid takes the support for multiple administrative domains. From the peer model, the hybrid takes support for multiple technologies within a single domain. Hence, topology discovery, route computation and light-path provisioning are all best performed in a centralised manner, whereas neighbour discovery and network protection are best implemented in a distributed fashion. In this model the IP and optical networks retain their clear demarcation points and exchange reachability information.

Mature signalling protocols give users the ability to signal and switch/route wavelength "on-demand" and, therefore, to dynamically provision bandwidth needed by specific applications. Many solutions for "intelligent optical networking" are available but, very often, they are put together in a vendor-proprietary fashion and there is still a lot of centralised network management needed. Only one of the vendors interviewed has claimed that they have a truly "intelligent optical networking" product available now, utilising a distributed PNNI-based routing and signalling scheme to dynamically allocate network resources. Multi-vendor intelligent optical networks have not yet been deployed, but various laboratory trials of systems interconnected via O-UNI show that such systems might be commercially available within the next eighteen months.

6.3. Management tools

As in traditional networks, standard tools such as ping and traceroute are needed for debugging and performance monitoring of G-MPLS networks.

Fault correlation between multiple layers
In the peer model many layers may be involved in the control and the transmission of G-MPLS data and control information. It is therefore required that a fault in one layer is passed on to the adjacent higher or lower layers in an effort to notify them of a possible fault. However, due to the nature of these many layers, the number of alarms may grow by an uncontrollable amount, where it damages other network elements by the sheer volume of alarms. Alarms will have to be managed so that the NMS only receives the minimum number of alarms and only the most significant ones. Another problem is that, in case of a failure, several layers may attempt the restoration of the link at the same time, so that different protection schemes may interact with each other. This effect should be carefully examined as it may lead to service degradation.

Provisioning systems
CLIs and SNMP have been the provisioning protocols of choice for most NRENs. As new optical network elements are introduced to the NREN networks it is essential that services can still be brought up in an orderly and timely fashion. In order to do this, new provisioning tools will have to be customised to the NRENs' requirements. A new generation of GUI applications already allows for Label Switched Path (LSP) and backup LSP creation, including LSPs traversing light-paths (lambdas) in a full optical cloud. Such LSP/Lambda set can be route-optimised for the whole network, thus allowing for better network utilisation.

6.4. Conclusion

NRENs will have to introduce new element managers into their network management systems in order to manage future data and transmission networks. The management of this new equipment will require protocols traditionally used by telecommunications companies rather than NRENs.

Several advances in TMN/CORBA and SNMP/CORBA integration will help to integrate new and existing network management systems but these standards are immature at present and a lot of work has still to be completed.

Both the ITU and the IETF are working on new network architectures for optical networks. It is still very early to predict which architecture will become the de-facto standard. It is still uncertain if we will have one or two control planes. What is certain, is that existing control planes are not suitable for the optical networks of the future. Creating a standardised control plane for optical networks will enable NRENs to introduce new features like end-to-end light-path creation. What this control plane will be, is still uncertain, but initially an overlay model with an optical UNI is an excellent starting point.

7. Network architectures

In this section we describe possible network architectures for future research and education networks taking into account technological developments outlined in the previous sections and the needs of researchers mentioned below. We will only provide this overview for the backbone and will not deal with access techniques, as these have been only marginally addressed in this study.

7.1. Requirements for future research networks

There will be the continued need to support traditional applications, i.e. those applications that have limited bandwidth requirements (for example up to 50 Mb/s) and for relatively short time scales. These applications may require some kind of differential treatment, which can be fulfilled by Premium IP or Less-than-Best Efforts. For these kinds of applications, there is no new requirement imposed upon today's networks. The techniques used to deliver service will be improved (for example advanced management for QoS) but the basic techniques are the same as today, with the addition of support for IPv6.

The current need to support Virtual Private Networks to facilitate research on networking technology will remain in the future, but the bandwidth requirements are likely to be higher. Currently VPNs are provided to projects such as ATRIUM[*] and DATATAG[**], which have bandwidth requirements of 622 Mb/s. We can expect these to become requirements for Gigabit Ethernet, 10-GE, STM-16 or even STM-64. This is a significant new requirement for the research networks of the future.

Above all, there is the requirement to support at least 40 Gb/s, on aggregate, on core links. This requirement is driven by the combination of the following three factors:

- organic growth of user traffic;
- VPNs with higher capacity requirements;
- support for Grid-like applications or projects.

The support for Grid-like applications is the main contributor to the requirement for 40 Gb/s on core network links. These applications exhibit a completely new traffic profile when compared with traditional applications: they require high-volume data transfers at high data rates. The Grid-like applications are not yet ready to exploit the availability of high-capacity research networks (national and pan-European), apart from some proof-of-concept examples, which are themselves already capable of generating single flows of up to 500 Mb/s. This situation is changing rapidly and we may expect a significant number of these applications in the next few years, which will lead to a dramatic increase of demand and utilisation of bandwidth – hence the requirement for 40 Gb/s aggregate capacity in core networks.

Another important requirement for networks of the future is to have a network architecture that enables a different network-funding model. Today, at least in the pan-European case, users require and benefit from any-to-any connectivity all over Europe and therefore users contribute to the cost of the network accordingly. In the case of GÉANT each network that connects to it contributes to its cost partly on their access capacity and partly on their geographic location. This translates to some form of flat fee for access to GÉANT, regardless of the actual usage.

In future we expect that some Grid applications or projects with very large bandwidth requirements will only require connectivity between certain sites in Europe and will require contributing to the cost of the network according to the coverage they obtain and use from it. We expect this phenomenon to be true both at the pan-European and the national network levels.

The scenario is that few projects will use the vast majority of network resources but only in a limited number of locations. Therefore, the next-generation network architecture must take this into account and facilitate different funding models in order to enable and encourage usage of shared networks by projects.

The requirements outlined above lead to the consequence that appropriate resource allocation techniques must be deployed. For example, if a Grid project requires 1GE of connectivity between two locations, mechanisms need

---

[*] http://world.alcatel.be/atrium/

[**] http://datatag.web.cern/ch/datatag/

to be in place that enable the allocation of that level of resource and that ensure that the resource is used by who requested it.

The concept of customer-empowered networks is gaining some momentum within the research community, and there are already some examples, such as CA*net 4 in Canada that are moving in that direction. These are networks where end-users can control and allocate resources between end-points, with little or no intervention from backbone operators. More details on customer empowered networks and a discussion about pros and cons for NRENs to acquire dark fibres are provided in section 8 of this report.

7.2. Network architecture options

In this section we will explore three different network architecture options. For each network architecture there is already evidence that it is capable of supporting the requirements set out in the previous section. We will examine a network architecture that is almost identical in nature to the majority of current research networks, apart from the capacities and router technology used. These are the shared IP-only networks. Then we consider networks that are able to offer switched point-to-point connections as well as an IP service, or hybrid networks. In both cases the raw material, connectivity, is provided via the traditional lease agreements currently used with telecommunications operators. Finally, we explore the case where NRENs have a more direct role in the provisioning of the basic connectivity, i.e. where they take (almost) full control of the fibre infrastructure.

7.2.1. Shared IP-only network

In this context, shared IP-only refers to a network in which all connectivity is handled by IP routers. This is the current set up for GÉANT and most NRENs. All user traffic transits IP routers in the core backbone. QoS and VPN features are offered using techniques available on IP routers (DiffServ for QoS, MPLS for VPNs). MPLS techniques are certainly able to support VPNs with bandwidth requirements of 2.5 Gb/s. Provided that the IP routers are able to support 40 Gb/s on aggregate on a point-to-point link, between two locations (this is possible as outlined in the section on routers), the emerging Grid applications can also be supported on this kind of network architecture.

Figure 6 below outlines the example of a pan-European IP-only network providing connectivity for traditional users, Grid users and VPNs.

The core of the pan-European network is composed of 4x10-Gb/s links to provide an aggregate of 40 Gb/s. This may evolve to a situation where there is one (or more) 40-Gb/s circuit(s) between locations when the router interface cards are available and when operators offer a single channel 40-Gb/s service. On the interface to the NRENs, n x 10 Gb/s are available. The NREN networks are built in a similar fashion, and they offer services to a large number of traditional users. In some cases, the NREN networks act as interconnects for regional networks. In the example depicted there is a Grid application, running between Grid users in NREN-1 and NREN-2. Similarly there is a VPN running between users in NREN-1 and NREN-2.

Users at large are served best by the shared IP network. Using MPLS techniques, it is also possible to deliver service to VPN users. The shared IP network can also serve Grid users. In fact one may argue: what is so special about Grid traffic that it needs something different than shared IP? Technically, there is no strong evidence that this is the case but some verification is still needed. What is important to note here is that to support Grid users on a shared IP network signifies adding several high-speed interfaces on IP routers along the path between the users. The current list price for 10-Gb/s interfaces is in the order of 250,000 euro. This is partly because of the transmission components needed, but mainly because of the complexity of ASIC and software needed to operate all the advanced IP services at high data rates. On the other hand Grid applications seem to require only raw bandwidth at the core network level, with the addition of middleware functions, which are best performed on other equipment, such as Unix workstations, rather than IP routers.

Providing services for a limited number (compared to traditional users) of Grid applications in a limited number of locations can be more expensive than simply providing dedicated point-to-point circuits for the single Grid applications.

Furthermore, the shared IP network does not facilitate different funding models for projects or applications such as Grids as the network itself does not explicitly separate the resources. Different funding models can be achieved, but these would remain unrelated to the network architecture.

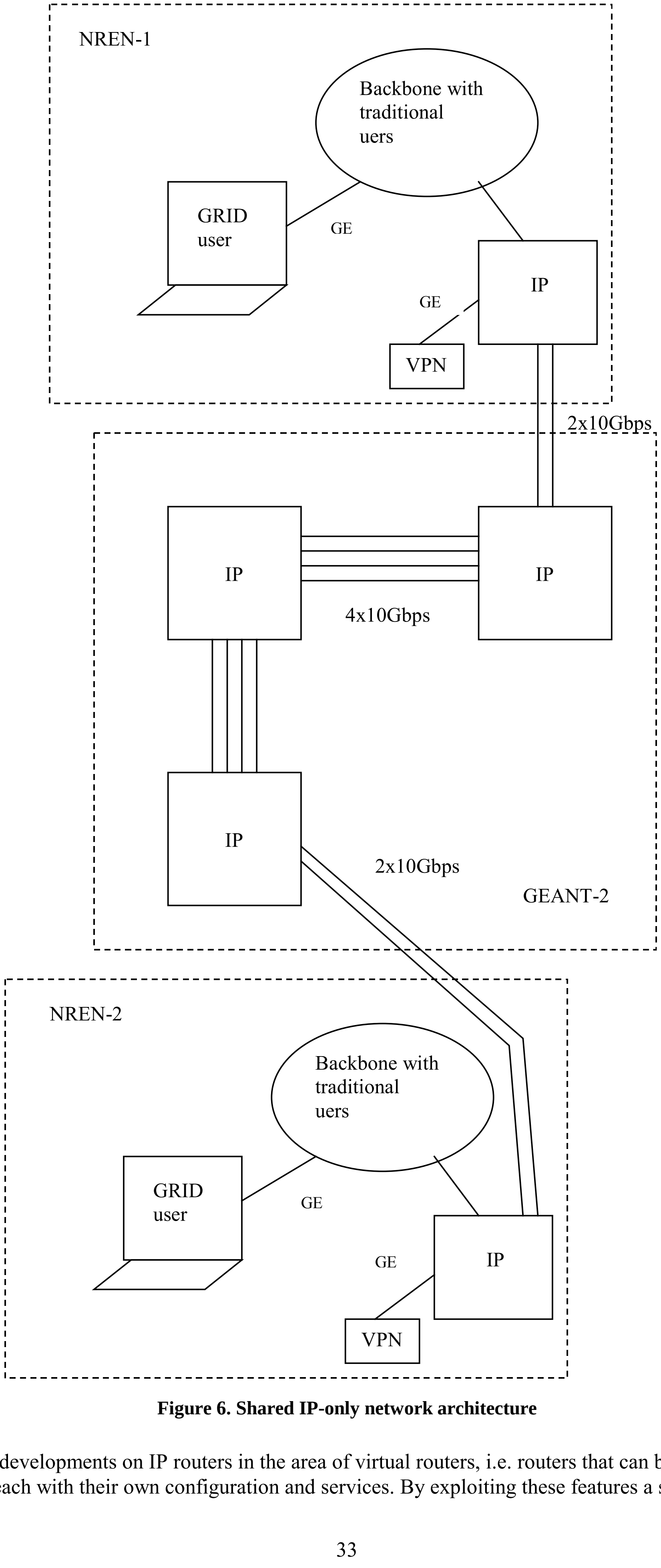

Figure 6. Shared IP-only network architecture

There are some developments on IP routers in the area of virtual routers, i.e. routers that can be divided into logical entities each with their own configuration and services. By exploiting these features a shared IP network can also be used to build a customer-empowered network (see section 3.4 above). Users would have access to virtual routers with which they could manage their connectivity requirements.

7.2.2. A hybrid (IP + switched) network

We call a network that is made of an all-IP component and complemented by the possibility to offer switched connections, a “hybrid” network. In this scenario, illustrated in figure 7 below, the general user still uses a shared IP infrastructure. The shared IP infrastructure can also be used to support users with requirements for QoS. To support Grid or VPN applications, a different approach can be taken. As seen in section 4 of this report, switching equipment that provides Gigabit Ethernet services and SDH grooming is already available (and is being developed). This means that an STM-64 circuit can be split up into several GE and STM-16 (or less) circuits and these can be presented on a dedicated interface to users.

This is a considerably cheaper solution than the all-IP solution presented previously because of the lower cost of interfaces on this type of equipment. In fact, IP routers and switching equipment will have similar transmission and switching components, but the latter do not have nor need the ASICs and software required for advanced IP services.

In terms of network management and allocation of resources, this architecture has its pros and cons. For resource allocation, most switching equipment vendors are in the process of supporting G-MPLS, albeit implementations are not always conformant to the standards and therefore not vendor-interoperable.

This means that connectivity resources distributed between IP routers and switching equipment have the potential of being allocated and managed seamlessly. A drawback is represented by the fact that NRENs are accustomed to use SNMP based tools for network monitoring, whilst SNMP is not supported on most of the switching equipment examined.

Such network architecture facilitates different charging models for usages such as Grids. In fact, especially in the case where a whole 2.5 Gb/s is required by an application, it is quite simple to clearly and unambiguously identify the additional hardware and connectivity resources needed to fulfil that requirement. The differential charging can be derived from the additional resources required.

Two options exist for the operational management of such a network architecture:

- The NRENs manage the IP and switching equipment.
- The NRENs manage the IP equipment and procure a managed switching service from traditional operators. As in all cases of connectivity services, it is most likely wiser to select more than one fibre provider.

In either case, there is a clear demarcation between connectivity providers and switching/routing elements. Similar to the shared-IP only network, the connectivity provider offers a multiple wavelength service. From the NREN point of view, there is a clear distinction, in that single wavelengths can be offered separately to user projects. This is especially the case where the switching elements are optical. Therefore, in this case the network architecture can also be defined as a **multiple wavelength network**.

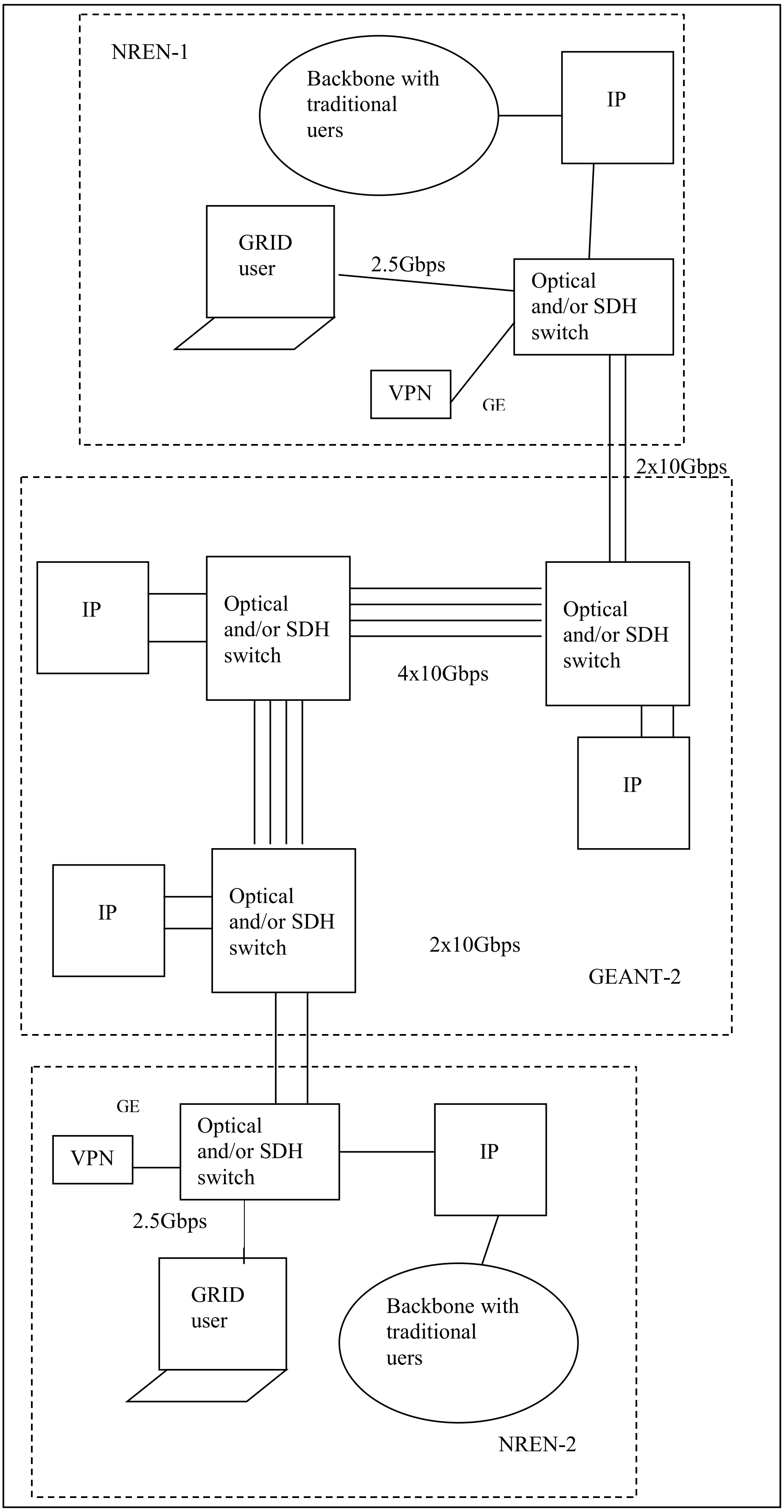

Figure 7. Hybrid network architecture

7.2.3. Fibre network

In this architecture an NREN has control over a fibre network infrastructure. The NREN can provision as many wavelengths (within the limits of the technology used) as required. The NREN can then build its shared IP-only or hybrid network on top of the fibre network and offer services to its users.

This network architecture is very similar to the hybrid network architecture, the key difference being that the NREN also controls the fibre infrastructure. A fundamental consideration to make is: how is the network operated? There are two options:

- The fibre network is operated by a traditional carrier on behalf of the NREN. This represents a new model of working between NRENs and carriers. Traditionally carriers will buy (or sell) fibre pairs, and the owner takes full control of all the sites where in-line equipment is needed. The in-line and line-termination equipment is selected and operated by the owner. In this case third-parties would perform the fibre-based functions on behalf of the NREN.

- The fibre network is operated in full by the NREN. Technical advances in the transmission equipment make this a feasible option even if on an international scale this is still difficult because of the reasons outlined below.

The following diagram outlines a possible pan-European set up.

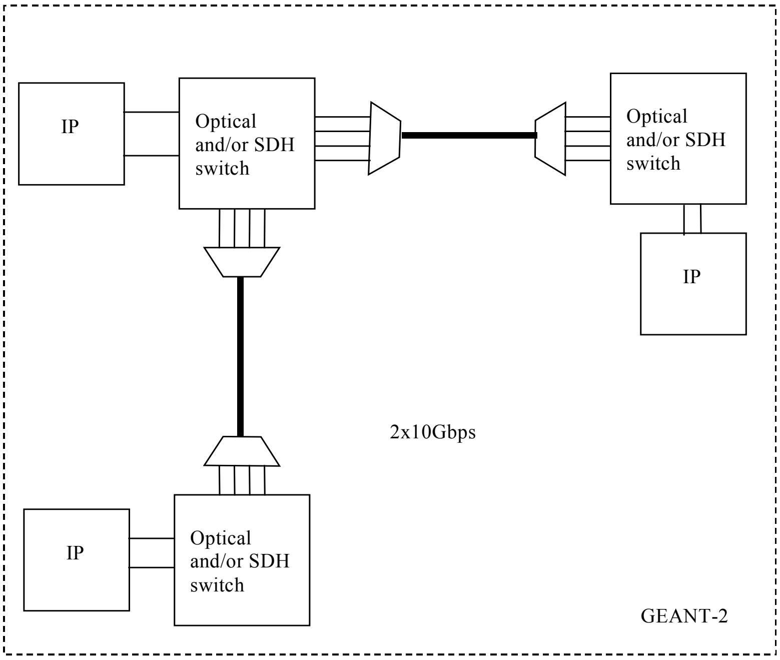

Figure 8. Fibre network architecture

There are already cases in parts of Europe where NRENs operate, in part, a fibre network together with traditional leased circuits. This is true in Poland, the Czech Republic, Slovakia, Ireland, the Netherlands and Switzerland. These NRENs have leased or deployed their own fibre, which is lit by equipment they own. Various reasons made it possible for these NRENs to do so:

- the distances involved are such that only in few cases regeneration is required;
- a cost analysis showed that these cases owning fibre infrastructure was more cost-effective than leasing bandwidth;
- technology developments are such that the management of wavelengths on transmission equipment has been significantly simplified compared to a few years ago.

On an international scale there are still some very important factors to take into account:

- Between many locations the costs of leasing single wavelengths is believed to be lower than fibre leasing and equipment costs.

- Advances in transmission technologies, as well as wide availability of fibre (especially new-generation fibre) enable increasing the reach without regeneration; in wide scale networks it is possible, with the current technology, to provide up to 2.5 Gb/s on 3,000-km links without regeneration, but at higher speeds this possibility might be limited because of deployment and cost issues.

- In addition to the CAPEX, there are significant OPEX costs related to site management. NRENs have relatively little experience of this, which is made even more complex by the multi-lingual and multi-administrative environment. However, as described in section 5.3.2, these issues can be avoided with the use of transmission techniques such as Nothing-In-Line (NIL).

Furthermore, from a technical point of view, the line-termination equipment and in-line equipment must be compatible. It is not possible to procure “off-the-shelf” components and put them together.

7.3. Conclusions

Which network architecture to deploy depends on the set of user requirements that need to be fulfilled, network management principles from the operator’s point of view, costs of equipment, and running costs. This is a very complex equation, especially in the pan-European case where costs of connectivity are subject to variations of a factor of 40. For each network topology, and for each set of services that need to be offered, an extremely detailed cost analysis needs to be done in order to evaluate the costs of one solution versus another and this needs to be combined with a detailed technical analysis of the various solutions.

Our view is that all network architectures proposed here can fulfil the foreseen requirements of future networks. In order to decide which is more appropriate, one must carry out the detailed technical and cost evaluations outlined. It is expected that different solutions will apply in different countries, depending on the countries involved, topology, network management principles of the operator, number of networking components required, scale of geographic coverage needed for advanced services, and so on. Another factor to be taken into account is the attitude of carriers that currently own fibre, towards the provisioning of dark fibre or managed fibre services. This attitude might be significantly affected by the current changes in the economic realities of the telecommunications sector.

8. Customer-empowered networks

The technical developments that can be expected in the next five years, dealing with transmission, switching and routing components, have consequences in terms of network architecture, as well as of real-time network management and control. In addition, from a business perspective, NRENs feel ready to look into customer-empowered networks, where user communities purchase optical fibre infrastructures or lease them on a long-term basis, in order to set up their own networks, taking over some functions traditionally provided by telecommunications operators. An analysis of this alternative infrastructure model has been part of a separate SERENATE report[8] and is outside the scope of this equipment study, but there are some relevant implications in terms of fibre and equipment requirements that are important for NRENs.

8.1. Rationale

As explained in section 7.2.3, operation of fibre networks by NRENs constitutes a new working model in the relation between telecommunications operators (carriers) and NRENs. Regardless of whether the network architecture is based on a shared IP-only model or on a hybrid one, without access to dark fibres there is no special reason for NRENs to look after transmission equipment and fibre infrastructure.

Traditionally NRENs build their network by using telecommunications services delivered by operators (carriers) who own or lease the fibre and are responsible for data transmission, fibre maintenance and repair, fibre lighting, Wave Division Multiplexing (WDM), signal amplification, dispersion compensation, signal regeneration etc. In this model, carriers are only responsible for the low-level transmission layer and NRENs are not involved in the acquisition, installation and operation of dark fibres and transmission equipment.

A carrier-managed infrastructure is necessarily a multipurpose transmission layer fitting many customers' requirements, whereas NRENs often require a transmission layer designed for “on-demand” delivery of dedicated connections and experimental lines and services for high-demand applications (see section 7). NRENs' requirements for link capacity are usually lower than those of operators and higher than those of wide-area enterprise networks. Moreover, specific requests for transmission equipment from NRENs are being recognised by equipment vendors as an emerging market segment for long-distance gear. There are expectations that knowledge and prices of transmission equipment will play a very important role in the design and development plans of NRENs. The opportunity for NRENs to acquire long-distance dark fibres is emerging. Additionally, NRENs have the facility to house equipment in university premises.

The development towards customer-empowered networks is of key importance for NRENs, both technically and economically, giving them technical, operational and price independence from traditional carriers. In the traditional model, prices for telecommunication services depend on competition and market conditions, rather than on grounds related to technical costs, like fibre routes and type of equipment used. Moreover, an independent position allows NRENs to choose between buying telecommunication services or using dark fibre. Some NRENs have experienced that the use of dark fibres can save up to 50% compared to buying telecommunication services.

Dark fibres are already being used by NRENs at campus and metropolitan level and now more and more also at the national and regional level (e.g. CESNET, HEANET, PIONIER, SANET, SURFnet, SWITCH; international connection is in place between CESNET, SANET and ACOnet). The first project at a continental level, the National LightRail project, is expected to connect Los Angeles with Seattle and Washington DC in July 2003 using dark fibre.

The use of dark fibres by NRENs should not be seen as a naïve competition with telecommunication services. NRENs are buying professionally provided dark fibre (e.g. leasing and maintenance) and high-quality advanced transmission equipment, which fit closely to their needs and additionally provide services like, for example, transmission with 10-GE interface or single-fibre transmission with FE, GE or 10 GE, which are still missing from the service portfolio of most carriers.

---

[8] SERENATE deliverable D3: Report on the experience of various communities that have experimented with "alternative" models of infrastructures.

8.2. Dark fibre

Dark fibre is optical fibre dedicated to a single customer (or a group of customers) where the customer is responsible for attaching the transmission equipment to "light" it.

Optical network implementations, based on dark fibres, are classified according to different geographical size ranges of installation, e.g. campus, metro, national or regional, and continental or intercontinental. The primary distinction among them is based on:

- **the distance** signals must be transmitted and
- **the density,** i.e. number of wavelengths (signals, colours) in a single dark fibre.

Each of these parameters has different implications for the transmission equipment.

Currently, and in the foreseeable future, the required density of connections by NRENs is lower (e.g. 1 - 64) than in some dark fibres used by carriers (e.g. 32 – 256). In both cases, the density required in edge networks is lower than in the core. Requests for transmission distance is the same in both cases, but for longer distances (e.g. 2,000 km and more) buying lambda services from carriers could be cheaper than using dark fibre. This depends of course on the actual business offer, but in principle a higher level of fibre sharing is necessary to achieve lower expenses, and carriers have better opportunities to organise fibre sharing for longer distances than NRENs. The cost comparison is not easy, because usually it is necessary to take into account the cost of future lambda upgrade (upgrade of number of lambdas or upgrade of lambdas transmission rate).

Professional companies who specialise in dark fibre systems take care of the installation of the fibre and also maintain it on behalf of the customer. In many cases, there is no additional management complexity or overhead associated with customer-empowered dark fibre, and this is more reliable than traditional telecommunication services. Some additional management complexity only appears in long-distance fibre lines (e.g. longer than 250 km), when in-line optical signal amplification or regeneration is necessary. In such cases, the customer takes responsibility for remote management of in-line equipment and the company maintaining fibre provides some local assistance.

Technically, companies actually own the fibre, but sell IRUs (Indefeasible Rights of Use) for up to 20 years for unrestricted use. The key thing to note is that the up-front cost for the purchase of a 20-year IRU is a one-time investment. Usually, the IRU can be considered as a physical asset, which can be resold, traded or used as collateral. As such, the cost of an IRU can be depreciated over its 20-year lifetime, which results in a monthly cost substantially lower than the cost of traditional telecommunication services.

In some countries, the legal system does not consider IRUs as assets (with known tax consequences). For this reason or for reasons of competitiveness some fibre owners offer long-term lease with a regular monthly fee and, in some cases, an additional one-time installation fee. NRENs had very positive experiences with this, having the possibility to decide which dark-fibre line to lease and for how many years. Paying a monthly fee for the fibre is similar to paying monthly telecommunication services, but often is cheaper. The choice for longer-term lease is usually determined by price, but there might be reasons for choosing a shorter time, including:

- need for moving end points in the future,
- need for different types of fibre in the future,
- need for changing network topology in the future,
- expected future better offers resulting from installation of new fibre trunks with shorter length or advanced fibre types, etc.

A number of next-generation service providers in metro or regional areas are now installing fibre networks and will sell strands of fibre to any organisation that wishes to purchase and manage its own dark fibres. Many of these new fibre networks are built along the same model as a condominium apartment building (condominium fibre network). For NRENs, the role of customer usually belongs to a national group of universities and research institutions. For the pan-European network the role of customer belongs to a grouping of NRENs (where responsibility may be delegated to an organisation established for this purpose).

Customer-empowered dark fibre can be more reliable than traditional carrier services, particularly if the customer deploys a diverse or redundant dark-fibre route. Dark fibre is a very simple technology. Traditionally managed services usually have many devices in the network path such as SONET multiplexers, Add/Drop multiplexers, switches, routers, etc. Each of these devices is susceptible to failure and that is why traditional carriers have to deploy complex networks and systems to insure reliability and redundancy.

The use of dark fibre brings, in some sense, simplicity into network design by avoiding the use of complex telecommunication services, which are not transparent and which are difficult or impossible to manage from the customer point of view. Wide-area network design is today closer to computer system design. One can read slogans as “the network is the supercomputer” or “the supercomputer is the network” and one can see the practical consequences.

When discussing future optical network architecture, it is important to take into account the trends in optical transmission system design. Results in research and development of new fibre types, lasers, optical amplifiers, switches, integrated circuits etc. will have a big influence on network design, and have been considered in this report (see also section 9). However, in planning and deploying optical transmission for research and education networks, we have to look at transmission equipment that is currently available or will be on the market in the near future. From an engineering point of view, we need a kit of transmission systems, which are available for building research and education networks at national and international level and are proven by experimental deployment. The current demand is characterised by:

- transmission of single lambda or few lambdas,
- transmission rate from 100 Mb/s to 40 Gb/s and more per lambda,
- transmission over long distance,
- scalability and cost of (reliable) equipment,
- single fibre instead of fibre pair is preferred for sufficiently short distances (now up to 125 km), if overall expenses are lower.

A narrow band of a few lambdas is suitable for optical amplification rather than a wide band of few or many lambdas, so demand for long distance means in some sense demand for narrow band. For metropolitan research and education networks, the demand is the same with the exception of long distance. This means that narrow band, in-line amplification or DCF are usually not necessary in metropolitan networks up to transmission rates of 10 Gb/s per lambda. Transmission equipment for metropolitan networks is closer to research and education networks’ needs, because low-shared links are common in some parts of metropolitan areas (usually in peripheral parts of large cities that have a low density of users) and they are also common in some parts of NRENs.

9. Future and emerging technologies

The purpose of this section is to explore future and emerging technologies that may have a positive impact on the development of more advanced networking services.

The reason for the need to look at new developments is that some existing technology is reaching its physical limits of speed and reliability. As an example, the laws of physics suggest that the maximum practical electrical signal path-length possible on a transponder working at 10 Gb/s is around 30 cm. The expectation is that at 40 Gb/s the maximum possible path-length will be 7 cm, or even less. This makes the implementation of a 40 Gb/s electro-optical device difficult.

9.1. Fibre technologies

Fibre optic cables have been in use since the 1970s and are the underlying physical technology that has permitted the explosive growth in the availability of communications bandwidth. Fibre behaviour is now well understood and manufacture techniques have improved radically. The consequences of this are that fibre installed today has much better dispersion characteristics than that installed in the early days. We have been told that the fibre installed pre-1992 is unlikely to be able to provide support for high-density WDM systems, over long distance, without the need for expensive dispersion compensation equipment, and that, even with this equipment, it may not work reliably. Old fibre will not support DWDM at 40 Gb/s and above over the long haul. Quite apart from the need for the new termination equipment to support 40 Gb/s, the pre-1992 fibres will need to be replaced (or used for slower transmission rates). Fibres manufactured according to the ITU recommendations G.652 (for standard single-mode fibres) and G.655 (for non-zero dispersion shifted fibres) should be able to support such links.

Further developments in the design and fabrication of fibres are already on the horizon. Companies are pioneering the development of hollow photonic crystal fibres (hollow-PCF) that may push the bandwidth and link-length possibilities to a much higher level than today. In a conventional fibre optic cable, the light is guided from end-to-end by a process of total internal reflection. The light travels through a glass medium for all of its passage and in doing so interacts with solid glass. The result of these interactions is that many non-linear effects such as dispersion lead to signal degradation.

As the prefix "hollow" implies, at the centre of the new hollow-PCF fibres is a minute hole (approximately 50 μm in diameter) through which the light is guided. The researchers developing the technology claim that the hollow-PCF approach provides several benefits, including the ability of the fibre to attenuate unwanted multi-modal signals in its walls (by a process known as modal sieving) and the photonic band-gap effect which provides a mechanism by which the fibre is selective about the wavelengths it transmits.

The main benefit of hollow-PCF technology appears to be near-lossless transmission of light, without the introduction of dispersive, non-linear effects into the signal. In addition, because the fibre is hollow and the light does not interact with the glass, higher laser powers could be used. This will further increase transmission distances without the need for intermediate signal conditioning.

Researchers are also working on the development of Polymer Optical Fibres (POF). The benefits seen for POFs is that they can be manufactured more cheaply than glass/silica fibres and have mechanical characteristics more suited to working in the end-user environment, e.g. flexibility and strength. A major problem, however, is that the signal attenuation in a POF is an order of magnitude larger than for its equivalent silica/glass counterpart. It may be that development of hollow POFs alleviates this problem and polymers might find a place in the manufacture of long distance fibre optic links. In the mean time, it is likely that they will be confined to short links of less than a few hundred meters.

9.2. Optical switching

As indicated in the fibre section above, dealing with opto-electrical conversion at speeds of 40 Gb/s and higher will be problematic through the limits imposed by the laws of physics. A more practical approach to optical switching at high speeds would be to undertake all required signal processing fully in the optical domain. Currently, true optical switching cannot be undertaken, due to the lack of availability of optical logic components.

Although rudimentary optical logic components already exist, they cannot support the complex processing needed to support the architectures of networks that real users would like to build. Although optical cross connects are currently available, they require external electronics to provide the switching function. Indeed, some describe the currently available range of "optical switches" as smart patch panels.

With the future availability of mature MPLS and G-MPLS, the motivation for fully-optical switching will strengthen. The reason for this is that a combination of IP packets, coupled with G-MPLS, could be transported directly on the DWDM infrastructure, without the need for additional intermediate layers, each of which requires their own management and control. The concept of Optical Packet Switching (OPS) has been discussed in many forums. It may be implemented as slotted OPS for fixed length packets, or in an un-slotted fashion for variable length packets. A critical requirement, needed to implement OPS, will be the availability of optical delay buffers. These buffers are needed to temporarily store the payload data while the optical packet header is processed and the appropriate optical route is set up within the switch fabric. Although some information on prototype demonstrators is available, it is likely that it will be many years before practical OPS systems are available in the market place.

A hybrid approach, known as Optical Burst Switching (OBS), has been proposed. In an OBS transmission system, the optical payload data is kept separate from the control signal, which itself can be in the optical domain. The control plane is most conveniently transmitted on a separate path, or wavelength, from the payload data. The control signal is transmitted from source switch to destination switch, slightly in advance of the payload data, so that it arrives sufficiently in time to set the required optical path through the switch before the payload data arrives. As the optical control plane is required to transmit comparatively small amounts of data, it can provide signalling for many payload data links. It is the only transmission link in the system that will require opto-electrical conversion at the intermediate nodes, the payload data remaining in the optical domain. The fact that the payload data can stay in the optical domain will dramatically reduce the cost of intermediate node switches, compared with those that require opto-electrical conversion of all streams. Additionally, since the payload data will pass through a fully optical path from end-to-end, it will experience no delays and will provide a predictable transmission system that will be well suited to support Grid-style applications that demand this level of service.

9.3. Error correction techniques

Current fibre systems incur some degradation of signal as the light traverses the fibre. In future, it may be possible to devise and construct fibre systems that do not impose noticeable levels of dispersion, and hollow-PCF may obtain this result. Even if such advanced fibres become common, unless there is total replacement of the installed fibre infrastructure (which is extremely unlikely), there will always be some links on which dispersion will continue to be a problem. It will therefore be necessary to compensate for dispersion and correct errors introduced into the data stream.

An approach to dealing with signal degradation that has been used for many years, is based on detecting errors and correcting them by using check bits. In order to achieve this, the value of the check bits has to be computed before transmission and analysed at reception. This technique is known as Forward Error Correction (FEC). The simplest form of such error correction can be achieved by using Hamming codes, however this adds about 36% overhead in order to correct corrupt data. The G.709 recommendation specifies the use of Reed-Solomon error correction, which is much more efficient, requiring adding around 6% overhead. When implemented as weak error correction, this might even be accommodated in unused bits of the transmission stream and therefore impose no overall loss of payload bandwidth (it should however be noted that a danger with using in-band FEC is the use of bytes that maybe be used in the future for other functions).

As transmission speeds approach and exceed 10 Gb/s, particularly in DWDM systems with fine spacing, optical degradation, through non-linearities, becomes increasingly problematic. If end-to-end systems are to reach anywhere near their full transmission potential, then errors introduced by dispersion and other non-linear effects have to be corrected. Systems such as Reed-Solomon alone will not be sufficient, so enhanced approaches known as Super-FEC will have to be developed. One supplier has told us that the availability of Super-FEC is crucial to the support of 40 Gb/s, even over the newest of installed fibre.

Although the processing load to implement FEC and Super-FEC are not excessive, the algorithms have to be executed continuously and in real-time at line speed. In practical systems, FEC and Super-FEC have to be implemented on custom chips (ASICs), which will be built into interface cards.

There are products currently on the market that implement a Super-FEC capable of working at line speeds of up to 10 Gb/s and 40 Gb/s. These products work by concatenating a range of FEC techniques (Bose, Chaudhuri, Hocquenghem and Reed-Solomon) to give a more robust error correction capability. It should be noted that current devices are implemented in a proprietary way. The cost of increased robustness is an increased overhead in terms of the redundancy data that must be carried. However, it can be demonstrated that, even with the addition of this extra data, there is a net coding gain. It looks likely that FEC and Super-FEC will be needed until such times as error-free and dispersionless fibres are ubiquitous.

The short-term future for 40 Gb/s (particularly on the long haul) looks inextricably tied to the availability of custom ASICs that will deliver the required throughput of error correction. From an availability point of view, we have been told that, due to the reluctance of chip manufacturers to make capital investments in the turbulent telecommunications sector, most 40-Gb/s Super-FEC developments have been put on hold. This could be one of the largest factors delaying the availability of 40-Gb/s DWDM systems, at least on the long haul. It may suggest that 40-Gb/s systems for the metropolitan area will appear first.

9.4. Hyperfine DWDM

There are two opposing views on how the best commercial utilisation of fibre infrastructure can be achieved. The simple view, shared by many in the industry, is that bandwidth per channel is increased until the physical limits of the fibre are reached. Supporters of this view reasonably expect that a 2.5-Gb/s system would be replaced by 10 Gb/s, 40 Gb/s and eventually even greater bandwidth per channel (or wavelength).

The opposite view is that this is not an effective way for the development to proceed and hyperfine DWDM is a better solution. The proponents of hyperfine DWDM say that, as bandwidth per wavelength is increased, so do the non-linearity problems and the need for more expensive signal correction and interference between channels. They suggest that the majority of the end-user market is not looking for single 40-Gb/s or 80-Gb/s channels. Investing in equipment that delivers such large single channels will only result in the need for additional multiplexers to break the single channels into user-sized chunks. They suggest that a better approach would be to limit the bandwidth transmitted, on each wavelength, resulting in much cheaper equipment being required. Moreover, this will reduce, or remove, the need for additional multiplexers. Since, at lower bandwidths per wavelength, the interfering non-linearities are very much reduced, the inter-wavelength spacing can be decreased to as low as 50 Mhz. The consequence of this is that, perhaps, more than 1,000 different wavelengths could be supported per fibre. It is expected that hyperfine DWDM will find application first in the metro area close to the users.

10. General conclusions

GÉANT, the European interconnect, is amongst the most advanced research networks in the world, with nine core links at 10Gp/s and fourteen links at 2.5 Gb/s. Several research and education networks in Europe and in other continents have country-wide DWDM-based network infrastructures in place. The next generation of GÉANT and research networks in Europe will require higher capacity and speed in order to support particular users and applications such as astrophysical or geophysical observation and modelling, as well as high-energy physics Grids. Technical developments for the introduction of higher speeds lead to different roadmaps, depending on specific characteristics of transmission, switching and routing components. But in all cases technology developments and deployment are affected by the current lack of customer demand and market conditions.

10.1. Transmission capacities

10-Gb/s transmission capacity is supported by equipment available today. Several vendors interviewed shared the opinion that carriers still do not require more than 2.5 Gb/s and the expected lack of user demand for higher transmission capacity will last for a couple of years. Even recognising that NRENs are an important factor in driving the introduction of networking technologies and services, none of the vendors seemed to consider their demand to be a sufficient case for the investments required in providing 40 Gb/s soon. The solution, which will likely be the next step, will be to load-balance 10 Gb/s across multiple links. There are limitations to this, due to router interfaces (see section 3.1.2.), which will probably be the drivers for providing single-channel 40-Gb/s interfaces and favour wide deployment of 40-Gb/s transmission systems. Some vendors have expressed serious concerns about the economical viability of having single-channel transmission at speeds higher than 10 Gb/s, whereas others were much more optimistic. On the other hand, one of the manufacturers observed that, in some cases, 10 Gb/s is still more expensive than four times 2.5 Gb/s. Even if many operators plan their networks to operate at higher bit rates (e.g. 10-Gb/s wavelengths), they end up installing lower-rate components (e.g. 2.5 Gb/s wavelengths).

There is a trend to increase the bit rates per wavelength, as well as the number of wavelengths supported in a single optical fibre, but, as discussed in section 9.4, non-linearity problems, the need for expensive signal correction, and interference between channels suggest that to limit the bandwidth transmitted on each channel and increase the number of wavelengths transmitted down a single fibre, would be a more efficient and cost-effective way of providing higher bandwidth.

As far as technology is concerned, limitations in the availability of custom interfaces affect the introduction of 40 Gb/s. However, all necessary subsystems have been already developed. But the current high production costs can only be lowered by mass production and there does not seem to be demand emerging soon. 40-Gb/s circuits will become more cost-effective only when their price will be less than four times 10 Gb/s; this is generally the case for 10 Gb/s and 2.5 Gb/s today. There are claims that the price factor is as high as 10, and some believe there will be no cost-effective solutions soon to justify rollout plans. But some DWDM equipment vendors have demonstrated that cost reduction can already be achieved today. Many optimistically think that 40 Gb/s would become price-competitive by the end of 2003. Thus, we would start seeing some initial deployment in 2004 with a replacement of 10 Gb/s in core networks within the next 5 years.

In terms of deployment, 40 Gb/s poses additional problems, requiring greater control of physical effects (e.g. dispersion and non-linear effects). In fact, even at 10-Gb/s capacity there are strong requirements for the availability of adequate fibre infrastructure. Chromatic dispersion becomes a critical factor, especially on long haul networks and compensation can be required at any distances over 4 km. Fibres deployed more than ten years ago present very high Polarisation Mode Dispersion (PMD) and require more sophisticated/smart (and expensive) equipment for compensation and regeneration. Of the three range applications (ultra long haul, long haul, metro and VSR range), 40 Gb/s is best suited to the longer reach. With the present technology and amplifiers (EDFA), transmission systems for 2.5 Gb/s and 10 Gb/s could reach 1,000-1,500 km (LH) and 3,000-4,000km (ULH) without regeneration. Only a couple of the vendors interviewed have introduced Raman amplifiers in their production systems. One of them was able to reach up to 3,200 km at 2.5 Gb/s. 40 Gb/s unregenerated systems are claimed to be capable of reaching about 1,000 km, on the same common equipment.

The majority of transmission equipment vendors think that, due to cost-effectiveness, 40 Gb/s will first emerge over DWDM long haul systems. However, there was no consistent view and particularly router vendors seemed to estimate that it will first emerge in the metro area in 2003.

80 Gb/s has been demonstrated by several vendors in the laboratories. The four-fold increase of bandwidth has proved to be effective, in terms of the cost reduction, which enables the introduction of a new technology. The next likely step beyond 40 Gb/s should be 160-Gb/s line rate. However, even if available in the laboratories, this is still a very expensive and immature technology, needing a much higher quality of the necessary components. Higher line rates have also been demonstrated. An experiment was done in Japan with single-channel transmission at 600 Gb/s, which is very far beyond the capabilities of near-future router interfaces and requires a new generation of fibres.

In the Ethernet domain, the next step would be to go from the 10 GE, currently available, to 100 GE. There was some interesting debate whether the Ethernet-like growth factor (10) would not be more appropriate in the optical domain than traditional SONET/SDH-like factor (4), but in general that was seen as too big a step and 100 Gb/s equipment was not expected to become available soon.

Alternative solutions to SONET/SDH, for transport and bandwidth grooming, will be available in the short term, including support for a variety of mapping techniques developed inside the SONET/SDH world, like Generic Framing Procedure (GFP), Virtual Concatenation, LCAS as well as outside, like G-MPLS and various flavours of Ethernet interfaces. Transparent wavelength services, based on G.709 interfaces, were expected to extend the management and monitoring features inherent into the optical layer and will become widely available in a few years' time. However, carriers seem to favour GFP with virtual concatenation and eventually LCAS to effectively transport different types of payloads (i.e. non-SDH framed payloads such as for instance Ethernet packets).

10.2. Routers

Router interfaces seem to be the main driver for the adoption of (single-channel) 40 Gb/s, due to problems associated with load balancing across n x 10 Gb/s.

Latest generation routers already support 40 Gb/s. However, line-speed interfaces are not yet ready and are likely to be very costly. In the near future, new generations of routers will certainly support interfaces to SONET/SDH, 1 GE, 10 Gb/s with different range lasers and to a limited extent, WDM interfaces. The study did not reveal any significant plan to support "coloured" G.709 interfaces at the moment.

Router vendors broadly support the development of G-MPLS and contribute to its standardisation. First implementations of G-MPLS support are available and will soon become common on routers from many different vendors. The same situation can be expected regarding protocols as Constrained Shorted Path First (CSPF) and Link Management Protocol (LMP), the latter only in a static version initially. Some vendors have already done interoperability tests with various OXC manufacturers.

In terms of scalability, it is clear that Terabit routers are becoming available but only as multi-chassis implementations (e.g. four 320-Gb/s full-duplex chassis interconnected by a switch fabric chassis). This involves very complex configuration of component parts, which is not necessarily easily done by customers.

Virtual routers, separating a single physical device into multiple independent logical partitions, have been in use in edge nodes for a few years and are now being adopted on core routers. By allowing operators of large IP networks to utilise virtually private router resources, and to manage them autonomously, virtual routers may aid support for Grid projects so that Virtual Private Routed Networks can be implemented on common packet-switched IP networks.

10.3. Intelligent optical networking

Optical fibre transmission and WDM equipment have been available for several years now. However, in the traditional IP networks, running on top of SONET/SDH, over WDM, the transmission layer was totally transparent to the network control plane. What is really new with the concept of optical networking, is the integration of the transport layer and the network control plane, introducing network intelligence at the lower layers and allowing users to manage and control wavelengths (colours or Lambdas). However, coloured interfaces are still offered today as SONET/SDH. The main technical and operational issue in offering lambdas to users is agreeing on the signal format. Different WDM vendors utilise different optical channel parameters so that current systems are not standardised enough to have a single coloured interface compliant with all WDM platforms. The only "standard" really available today is the ITU wavelength grid with digital wrapper overhead, but it allows vendors enough freedom to implement it in a proprietary fashion so that they can ensure the highest performance or link length. As a consequence, interoperability can nowadays be guaranteed only within a single-vendor system.

The ability for users to signal-high bandwidth "channels" and hence to switch/route wavelength "on-demand" is available now, but technologies remain immature and vendor-interoperability will require some time in coming. OIF UNI1.0 has been implemented by some vendors, but it has limited functionality. OIF UNI2.0 will soon be finalised. G-MPLS is still rather immature; many protocols are available but they are put together in vendor proprietary fashions and in many cases there is still a lot of centralised network management needed. Intelligent optical networking is already available today, but only as single vendor solution. However, multi-vendor laboratory trials being carried out at the moment show that interoperable solutions may start to become available in a couple of years.

10.4. All-optical networking

The drivers for all-optical network solutions will be scalability (the number of ports on electrical cross connects will be reduced when the line rate increases, whereas all-optical cross connects interfaces are independent of the bit rate) and cost-benefit. O-E-O conversion is required in OADM or switching nodes where signals in a traditional architecture are forced to be all terminated, while only some are effectively reaching their destination in a given site. The other signals are fed back into another DWDM terminal to continue their journey. In true “all-optical” networks O-E-O conversions are eliminated, representing a significant cost saving. They are meant to become less expensive solutions in the long run than non-transparent networks, especially for networks at high capacity.

O-E-O conversion is still required in many cases to supervise the signal and being able to offer, monitor and enforce Service Level Agreements. However on an “all-optical” network, one needs to monitor only the ends of the circuit and the analogue characteristics (such as signal-to-noise ratio) along the journey. Long Haul and Ultra Long Haul transmission technologies allow building large-size networks without the need for regeneration. In small-size cases, where regeneration is not needed, optical switches and solutions with Nothing-In-Line have smaller footprints and require less power than regenerative solutions.

In offering 40-Gb/s optical interfaces to users, serious difficulties are associated with the availability of the appropriate fibre plant to be able to handle such high bit rate propagation. Many innovations currently being developed will allow the operation of completely all-optical networks but their deployment will require “Greenfield” installations.

Some transparent optical interfaces are available in switches today (for example, framing and bit rate do not matter up to certain levels, e.g. 2.5 Gb/s or 10 Gb/s), but the main challenges for transparent optical switching are posed by all-optical wavelength conversion with optical 3R regeneration. In practice there is no need for wavelength conversion unless extreme capacities are required. Moreover, since all-optical wavelength conversion with optical 3R regeneration requires proper network planning and design as well as well-developed routing schemes, it is not expected to become widely available until 5-10 years from now.

Tuneable lasers should be able to provide savings on component inventories, but although available now, they still are very immature and expensive.

11. List of acronyms

3R	Restoration Reshaping Retiming
ACL	Access Control List
ANSI	American National Standards Institute
API	Application Programming Interface
APS/MPS	Advanced Photon Source/Machine Protection System
ASIC	Application Specific Integrated Circuit
ASTN	Advanced Switched Transport Network
ATM	Asynchronous Transfer Mode
BER	Bit Error Rate
BGP	Boarder Gateway Protocol
BLSR	Bi-directional Line-Switched Ring
BML	Business Management Layer
CAPEX	Capital Expenditure
CLI	Command Line Interface
CMIP	Common Management Information Protocol
CORBA	Common Object Request Broker Architecture
CoS	Class of Service
CR-LDP	Constraint-based Routing Label Distribution Protocol
CSPF	Constrained Shortest Path First
CS-RZ	Carrier Suppressed Return to Zero
CWDM	Coarse Wavelength Division Multiplexing
DCF	Dispersion Compensating Fibre
DGF	Dynamic Gain Flattening
DWDM	Dense Wavelength Division Multiplexing
DXC	Digital Cross Connect
EDFA	Erbium Doped Fibre Amplifier
EMS	Element Management System
E-NNI	Exterior Network-Node Interface
EXC	Electronic Cross Connect
FE	Fast Ethernet
FEC	Forward Error Correction
FWM	Four Wave Mixing
GBIC	Gigabit Interface Converter
GDMO	Guidelines for the Definition of Managed Objects
GE	Gigabit Ethernet
GFP	Generic Framing Procedure
G-MPLS	Generalised Multi Protocol Label Switching
GUI	Graphical User Interface
IEEE	Institute of Electrical and Electronics Engineers, Inc.
IETF	Internet Engineering Task Force
IGP	Interior Gateway Protocol
IP/PPP	Internet Protocol/Point-to-point Protocol
IPv6	Internet Protocol (version 6)
IRU	Indefeasible Right of Use
ISI	Inter-Symbol Interference
IS-IS	Intermediate System to Intermediate System Protocol
ITU-T	Telecommunication Standardisation Sector of the International Telecommunication Union
LCAS	Link Capacity Adjustment Scheme
LD	Laser Diod
LH	Long Haul
LMP	Link Management Protocol
LSP	Label Switched Path
MBS	Managed Bandwidth Service
MEM	Micro Electromechanical Mirror
MPLS	Multi Protocol Label Switching
NEL	Network Element Layer
NIL	Nothing In Line
NMS	Network Management System

NZDSF	Non-Zero Dispersion Shifted Fibre
OA	Optical Amplification
OADM	Optical Add Drop Multiplexer
OBS	Optical Burst Switching
OC-n	Optical Carrier
O-E-O	Optical-Electrical-Optical
OIF	Optical Internetworking Forum
OMG	Object Management Group
O-O-O	Optical-Optical-Optical
OPEX	Operational Expenditure
OPS	Optical Packet Switching
OSPF	Open Shortest Path First
OTN	Optical Transport Network
OXC	Optical Cross Connect
PCF	Photonic Crystal Fibre
PMD	Polarisation Mode Dispersion
PNNI	Private Network-to-Network Interface
POF	Polymer Optical Fibre
POS	Packet over SONET
PXC	Photonic Cross Connect
QoS	Quality of Service
RFC	Request For Comments
RSVP-TE	Resource Reservation Protocol – Traffic Engineering
SDH	Synchronous Digital Hierarchy
SFM	Single Mode Fibre
SNCP	Sub-Network Connection Protection
SNMP	Simple Network Management Protocol
SNR	Signal-to-Noise Ratio
SONET	Synchronous Optical NETwork
SPM	Self-Phase Modulation
STM-n	Synchronous Transfer Module
SVC	Switched Virtual Circuit
TDM	Time Division Multiplexing
TMN	Telecommunications Management Network
ULH	Ultra Long Haul
UNI	User Network Interface
UPSR	Uni-directional Path-Switched Ring
VC-n	Virtual Concatenation
VLAN	Virtual Local Area Network
VPN	Virtual Private Network
VSR	Very Short Reach
WDM	Wavelength Division Multiplexing
XPM	Cross Phase Modulation

Annex I. Questionnaire used in the interviews

NEP-2002-068
R. Sabatino 27 August 2002
V. Cavalli, M. Campanella 09 September 2002
M. Enrico, 01 November 2002

SERENATE WP3
Questionnaire for Equipment Vendors

Background

The GÉANT network is a multi-gigabit IP backbone that provides pan-European transit between European National Research & Education Networks (NRENs) and connectivity to R&E networks in other parts of the world (e.g. Abilene in the US, NACSIS in Japan, etc). The NREN in a given country is usually unique and typically interconnects universities and other institutes of higher or further education and public research institutes. GÉANT has been implemented and is operated on behalf of a consortium of European NRENs by DANTE, which is a not-for-profit limited company based in Cambridge, UK.

The GÉANT network consists of a single POP in each member country housing one or more carrier-class IP routers (currently mostly Juniper M160s) that are connected by unprotected international leased lines that range in capacity from 155Mbit/s to 10Gbit/s. Most of the trunk circuits are 2.5Gbps (currently eleven in number) or 10Gbps (nine). These circuits (although mostly implemented by the carriers as single DWDM channels) are presented on G.957 or G.691 compliant interfaces for direct connection to OC48c/STM-16 or OC192c/STM-64 packet-over-SONET/SDH (POS) interfaces on the routers.

Many of the NREN backbones are now implemented along similar lines although there are increasingly those with access to their own fibre infrastructure who are using long reach gigabit Ethernet and 10GE and experimenting with operating DWDM transmission equipment. Access to this type of infrastructure is still not really possible on the pan-European scale although DANTE is now starting to discuss the notion of long-distance "managed fibre" services with some of its present suppliers.

Most of the production traffic carried by the European R&E networks is currently "best-effort" IP and the utilisation levels of the 2.5/10G trunk circuits are often below 10%. The development of other services is as follows. Multicast is more or less ubiquitous. A DiffServ-based Premium IP service has been trialled and will soon become available as an operational service on the GÉANT backbone and some of the European NRENs. In addition, a DiffServ-based "less-than-best-effort" or "scavenger" service is currently being tested. An MPLS-based L2 VPN service is also currently being developed for the GÉANT backbone and should soon be available for geographically dispersed research projects that require relatively low capacities of layer-2 separated connectivity.

IPv6 is at a relatively advanced stage of deployment within the research networking community as a whole. Some of the European NRENs are already offering a pre-production IPv6 service as a result of operating their IP backbones in "dual stack" mode and it is planned to have a production IPv6 service on GÉANT by mid 2003.

Apart from the inevitable increases in volume of production traffic, it is also foreseen that there will be an increasing requirement to be able to offer a guaranteed "bandwidth-on-demand" service with wide-ranging granularity to accommodate the needs of short or medium term research projects and GRID initiatives. The granularities may range from STM-1 to STM-64 levels and the time scales involved may be of the order of days or weeks (short term) up to periods of about 1 year (medium term). In addition, the nature of the connectivity required by these projects may range from simple L2 or L3 VPN-style connectivity to dedicated point-to-point wavelengths.

SERENATE is the name of a series of strategic studies into the future of research and education networking in Europe, addressing the local (campus networks), national (national research & education networks), European and intercontinental levels. The SERENATE studies, which are organised by TERENA, DANTE, the Academia Europaea, the European Science Foundation and the Technical

University of Denmark, bring together the research and education networks of Europe, national governments and funding bodies, the European Commission, network operators, equipment manufacturers and, last but not least, the scientific and education community as the users of networks and services. More details can be found on the project web site given below.

This questionnaire is intended to be the framework for a series of bilateral meetings with leading vendors of networking equipment that will provide the primary input for a report assessing the availability and characteristics of the necessary equipment for large-scale, long-distance networks operating at 100Gbps or above.

NOTE: not all of the questions below are relevant to all vendors. Nor, indeed, would we expect all of those that are relevant to be fully answered during the course of the bilateral meetings. Rather they are more of a framework to guide the discussion during the meetings.

Useful URLs:

SERENATE	http://www.serenate.org
DANTE/GÉANT	http://www.dante.org.uk
TERENA	http://www.terena.nl
European NRENs (selection)	http://www.heanet.ie (Ireland)
	http://www.garr.it (Italy)
	http://www.psnc.pl (Poland)
	http://www.renater.fr (France)
	http://www.switch.ch (Switzerland)
	http://www.ja.net (UK)
	http://www.dfn.de (Germany)
	http://www.surfnet.nl (Netherlands)
	http://www.cesnet.cz (Czech Republic)

1. 40+Gbps interfaces, what is the framing?

1.1 What is your view of the market drivers for 40+Gbps (40Gbps, 80Gbps, 160Gbps…) transmission capacities over the next 2 and 5 years?

1.2 Which types of interface do you plan to use as a function of the equipment, SDH/SONET, WDM, G.709, 1GbE, 10GbE and so on..

1.3 Will signalling be supported on such high-speed interfaces, and if so, which kind of signalling protocols?

1.4 When will you offer 40+Gbps interfaces? What are the requirements and implications for the fibre infrastructure?

1.5 What equipment have you developed / are you developing that is capable of supporting 40+Gbps interfaces and what is the status of its deployment (please give answers for the next 2 years and next 5 years)?

1.6 To what extent do these developments follow standards as opposed to proprietary solutions?

1.7 What are the key features of this equipment? How does it differ from other solutions?

1.8 To what extent do you foresee 40+Gbps being offered as a single channel or multi-channel with few channels (eg., 4 or 8) to users, as opposed to being used as a transport system between DWDM terminals? In this case, will you have parallel processing inside the interface?

1.9 What are the difficulties in offering 40+Gbps to users and when do you foresee them being solved?

1.10 What interoperability issues are there with other vendor’s solutions?

1.11 Presentation to users: Colours vs. SONET/SDH framing vs. Other Data Links

    1.11.1 Today interfaces to user equipment are based on SONET/SDH or GE framing. What are the technical and operational issues for presenting a specific colour/wavelength to users?

1.11.2 Given the context we explained, what are your views on offering colours as a service?
1.11.3 Standardisation is certainly a key factor in developing colour presentation of services. What contribution have you made to the relevant standards activities?
1.11.4 Are you developing proprietary solutions/proposals?
1.11.5 Do you have any prototype available?
1.11.6 Do you foresee implementations in 2, 5 years time?

2 Resilience and range

2.1 Research networks rely more on multiple diversely routed connections to a location to guarantee high availability rather than use protected services. What are your plans of offering resilience, at the optical layer / routing layer?

2.2 What are the requirements for regeneration and amplification for your 40+Gbps transmission? What are the maximum unregenerated transmission distances?

2.3 What is the capacity of your equipment? (E.g. how many ports/card and cards/chassis? What is the maximum number of chassis in multi-chassis setup? What is the maximum, non-blocking switching speed? Does the size of the system affect the resilience?)

3 Bandwidth grooming capabilities (spanning one wavelength or multiple wavelengths)

The predecessors to the GÉANT network (TEN-34 and TEN-155) were based on managed ATM connectivity. At the same time many of the NRENs had ATM backbones as did the connected institutes. This allowed ad-hoc pan-European ATM VPNs (based on end-to-end ATM VCs and VPs) to be built for end-user community groups to fulfil the needs of particular short and medium term projects. This was called the Managed Bandwidth Service (MBS) and offered layer 2 separated connectivity with bandwidth and quality guarantees. The GÉANT backbone and now most of the European NRENs no longer have ATM in their networks yet there is still the requirement from end-user communities (e.g. GRIDs) to have access to an MBS-like service and with ever increasing bandwidth requirements (e.g. 2.5Gbps per link).

3.1 Do you have smart solutions based on SONET/SDH? And what are the alternatives to SDH? What techniques do you have to inject different signals and present them as separate interface to a user/client?

3.2 What techniques do you foresee available in 2 and 5 years time?

3.2 What is the granularity of these techniques?

3.4 Can these techniques operate across multiple wavelengths?

3.5 Do you participate in any related standardisation activity?

4 Circuit (lambda) switching

As the bandwidth requirements of GRIDs increase then it may be that the best way to support them and the production IP services is to implement a pan-European research networking infrastructure that is capable of supporting "lambda switching" (possibly even "on demand"). Discussions on BoD are currently underway within the European research networking community.

4.1 What are the drivers for an all-optical solution?

4.2 To what extent are we tied up with O-E-O, and what are the real technical and operational implications of this? For how long do you see this situation lasting?

4.3 What developments do you have in this area, what products do you have for the next 2 years and what do you foresee to have in the next 5 years?

4.4 How much of your activity is related to standardisation?

4.5 What impact will the recent announcement regarding the Lucent LambdaRouter (http://www.lightreading.com/document.asp?doc_id=19801) have on the development and deployment of optical switches?

5 Network management issues

Traditionally each networking layer and its resources are managed by dedicated management entities each with a view of their own resources. G-MPLS is working towards integrated management of network resources, as this is seen as a more effective way of utilising resources. In terms of managing an IP-over-optical network there are two trends: the overlay and the peer model.

5.1 Which network management model do you favour and why (peer or overlay)? Which network management tools do you foresee?

5.2 Do you have G-MPLS implementations? Are these compliant to the standards?

5.3 To what extent can this model enable true end-to-end, interdomain, as opposed to edge-to-edge, intra-domain network management and allocation of resources.

5.4 What working experience do you have with G-MPLS

5.5 What limitations do you see in G-MPLS, what activities have you undertaken to address these?

5.6 Would you allow SNMP access, and provide accounting management?

6 Service model: provider/(N)RENs

In today's optical networks there is a clear service boundary between providers and Research Networks, in terms of presentation and management. We expect that the ongoing technological developments will facilitate the establishment of different service models. Today only a handful of NRENs engage in the procurement and management of wide area fibre networks (examples being SURFnet (NL), SWITCH (CH), PSNC (PL) and CESNET (CZ)).

6.1 Which kind of tools would you provide for the management of a mixed infrastructure for (N)RENs?

6.2 To what extent is it feasible that (N)RENs procure "managed fibre" from carriers. By managed fibre we mean that amplification and regeneration are performed by a service provider whilst (de)multipexing, switching and Add/Drop functions are performed by an (N)REN. Are you going to provide the tools, for example AAA, that allow (N)RENs to do this?

6.3 Managed fibre as opposed to managed wavelength connectivity represent extreme opposites in terms of service models between providers and (N)RENs. What other models do you foresee being enabled by your technological developments?

7 Cost Distribution

7.1 In an IP-over-optical network, composed of IP routers, DWDM, optical switches, optical add/drop multiplexers, amplifiers, regenerators etc, what is the distribution of the capital cost (percentage–wise) between all these components?

7.2 What impact does an "all-optical" as opposed to O-E-O solution have on this capital cost distribution?

7.3 And what is the typical impact on the operational costs (including man-power)?