Low Layer Network Technologies Research Areas for the Next Generation of the Research and Education Networks.

The ASTON Experience

WhitePaper

by:	M.Campanella,	(editor)	GARR	mauro.campanella@garr.it
	M. Przybylski,	(editor)	PSNC	michalp@man.pozanan.pl
	V. Cavalli,		TERENA	valentino.cavalli@terena.nl
	V. Reijs,		HEAnet	victor.reijs@heanet.ie
	R. Sabatino,		DANTE	roberto.sabatino@dante.org.uk
	S. Sima		CESnet	stanislav.sima@cesnet.cz

Date: April 24th 2003
Version 1.0
Status *Draft*

1. Introduction

The creation of networks devoted to Research, Academic and Education Institutes in Europe has been a key element for the development of research itself and society as a whole. These networks have made possible an unprecedented availability of knowledge and information through the use of IP and its applications, now part of the lives of millions of people. The technical developments need to be kept on going to allow an even greater distribution, quality and efficiency of knowledge sharing and use.

Research users' needs in term of capacity, outreach and services fostered rapid advances in network deployment. Today most of European NRENs feature in each country a multi-Gigabit per second backbone and offer advanced services to their community, as multicast, Quality of Service and soon IPv6 native connectivity.

The GÉANT network [1] is the European research backbone. It links more than 32 NRENs located in the European area and provides access to all other research networks world-wide. It features a meshed core made by 10 Gbit/s links, with access up to 2,5 Gbit/s from the various countries. The access links from various countries are expected to be upgraded soon to 10 Gb/s, which requires a core backbone upgrade to 40Gb/s links or multiple parallel 10 Gb/s links.

Historical data in various NRENs, see e.g. the online statistics of GARR [2], shows a traffic growing at an average rate of 1,5 – 2 times every year.

The increase springs from the natural growth of capacity requirement from connected institutions, the increase in number of institutions and researchers connected and also from the surge of new ways of using the network infrastructure, like GRID application and the Peer to Peer model.

The scenario sketches the steady exponential growth needs for research networking infrastructure. The challenge is now moving into unknown territory, as the networks need to scale not only in capacity, but also to face a much higher level of physical internal and external meshing.

The rapid evolution of optical technologies and general public interest for Internet access has timely developed the availability of dark fibres and high capacity channels both on long distance and metro area.

Whilst optical fibre may provide the ideal medium for carrying the traffic, this white paper elaborates on the techniques the NREN environment find important to light the fibre and on the architectures needed to manage such highly-meshed, multidomain environments. The techniques span multiple layers of the data transmission protocol stack. The white paper mainly focuses on the lower layers, up to the network one.

The paper feeds and elaborates further on the discussions and experiences of participants to the ASTON project; see the ASTON Expression of Interest [3].

The Serenate IST project [6] is a strategic study aiming at providing input to the European Commission on initiatives targeted to keep the evolution of the European research networking at the forefront of worldwide development.

Research and testing on new architectures are strongly suggested in the immediate future, coordinated at the European or continental scale.

2. Current Trends in Networking

2.1 Customer empowered networks

The availability of fibre and large diversity of customers allowed the deployment of technologies enabling the construction of networks owned and managed by its customers. The mechanisms at the foundation of such networks are the WDM technology, optical switching with VPNs and virtual routers.

These technologies are also very well suited to support the new applications types, like the GRID environment, which constitute a fundamental test case.

2.2 High capacity

Gbit/s speed is the standard in core backbones. 10 Gbit/s SDH is already deployed. 10 Gbit/s Ethernet has just been standardised and it's not yet widely available on the WAN market.

40 Gbit/s has been tested but it is not yet used by any NREN, nor it is easily available on the market. Higher transmission speeds have been tested in laboratories, but not deployed on large testbeds.

The increase in speed of the single link is proceeding in parallel with the increase in link (wavelength) density for each fibre. The combination of the two can provide an efficient solution on the mid-term to the capacity increase requirements.

2.3 Optical technology

Many advances have been done in the optical technology – starting from intensification of fibre use with WDM systems, through optical protection to optical switching and GMPLS. Below we try to outline the major development areas.

Intensification of fibre use – WDM systems

The intensification of fibre use (by the means of multiplication of available optical channels in single fibre with Wavelength Division Multiplexing techniques) is now in the focal centre of the attention of many NRENs. This is usually caused by the need to provide dedicated, high capacity channels for advanced research projects – namely HEP (High Energy Physics), Virtual Laboratories, radio-astronomy etc.

Dense WDM now is meant to be used for long haul networks – only in NRENs in geographically large countries or within pan-European deployments, such as possible future GEANT network, where it is economically justified.

Characteristics of DWDM:

range (w/o regeneration)	hundreds to thousands of km
speed	up to 40 Gbit/s
number of channels	up to 160 – 190

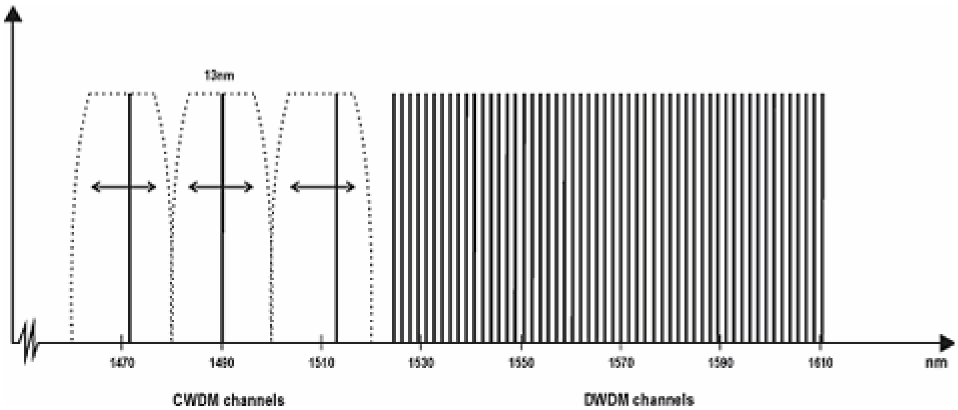

Figure. Comparison of bandwidth needed for Coarse WDM and Dense WDM (only part of DWDM wavelength grid shown) [5].

Many small countries will use short haul WDM in the form of Coarse WDM or Wide WDM systems that allow for lower number of lambda's to be used, on shorter distances and usually at lower speed. These devices can be much cheaper but also less flexible than long haul DWDM systems. Metro systems are usually CWDM.

Characteristics of CWDM:

range (w/o regeneration)	tens to hundreds of km
speed	up to 10 Gbit/s
number of channels	17-18 at most

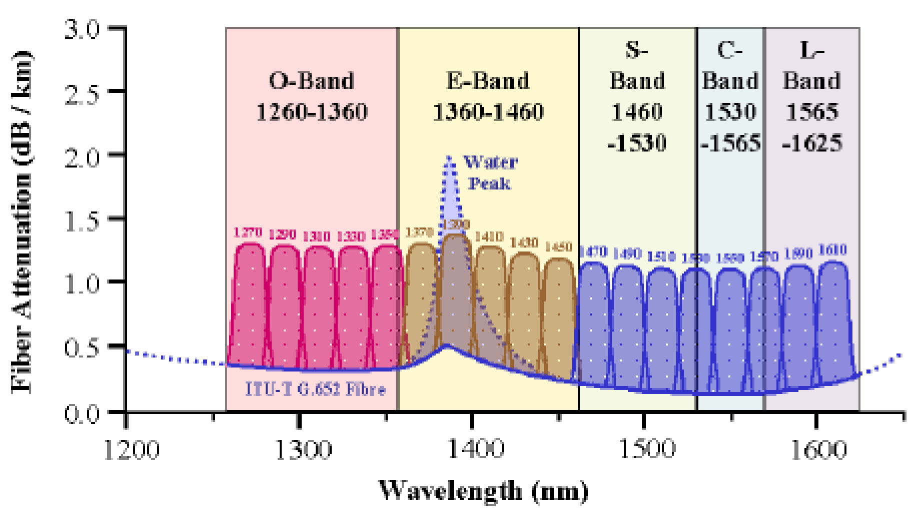

Figure. Metro CWDM Wavelength Grid as specified by ITU-T G.694.2 [5]

Optical switching

Optical switching includes two types of devices

- OXC (Optical Cross Connects, OEO switches) which in fact are electronic cross-connects with optical interfaces
- PXC (Photonic Cross Connects, OOO switches), which process signal only in optical form.

Both technologies have some advantages and disadvantages. To indicate some: OXC are more flexible, perform 3R (amplification, reshaping and retiming) and use a mature technology, although expensive and not very scalable. PXC is a fairly new, immature technology, with high signal attenuation, no wavelength conversion, low granularity and per-port cost. In addition buffering is quite complex, if feasible at all, in all optical domain. The emergence of PXC devices raised hopes for All Optical Networking, but it does not look real for the near future.

Optical switching also co-exists with the GMPLS architecture, which should allow for more flexibility into optical networking.

2.4 Routing

OSPF and IS-IS are usual choices as Interior Gateway Protocols. BGPv4 is the only choice for interdomain routing. There are various enhancement proposed to BGP to cope with the routing table growth. But still BGP is a protocol based on static link costs and it needs manual policy configuration and accurate planning to load balance even few parallel peering links.

The mass appearance of IPv6 in the global Internet will put an additional burden on the routing protocols and hardware platforms. The research networks will face as first the challenge.

The idea of customer empowered networks requires change in the router’s architecture. The new mechanisms, such as virtual routers will help to create multiple overlay networks on single infrastructure, shifting the control from the network operator to end-user of the service. The inclusion in routing algorithms computation of dynamic parameters like link load is an additional step required to enable advanced services.

2.5 Multidomain

More and more often customer networks reach over the single network boundary, usually the path for customer data spans multiple domains. There is a need to provide mechanisms to establish and operate virtual networks with predictable services build over multiple, independently managed networks.

The need is particularly live for the research environment, which is intrinsically made by world-wide research groups. The constant exchange of information between researchers, the sharing of resources and the mobility requirements create the vision of a “virtual” research network which spans all the physical NRENs in a seamless way.

2.6 Network engineering

Multi Protocol Label Switching is becoming the only “de facto” architecture to provide techniques to perform network engineering at network layer three or two, or both, on a single domain. It has to be coupled with Quality of Service techniques to ensure traffic guarantees. MPLS coupled with dynamic signalling (RSVP) may provide a platform for managing a complex, heterogeneous domain.

Extension and techniques for fast provisioning, failure recovery and resiliency are being engineered now. See for example the ATRIUM project [4].

The mandatory extension to multidomain environment demands advances on authentication and authorisation protocols and additional improvement to network engineering algorithms to take into account the long physical delays.

2.7 Mobile networking

High-speed mobile networking is now available at affordable prices using 802.1X or similar wireless technologies. Users need the technology to decouple the task to be performed from the physical location.

3. NREN networks scaling issues

The national research networks face various scaling issues:

1. *network capacity*. Users' capacity demand grows due to the local network development in users' premise. Users are adopting the new optical, high-speed technologies. Gigabit Ethernet LANs are now common and affordable.
2. *number of users.* The number of users accessing the network is increasing, due to the network access granted to classical faculties and the whole university student population. A major, abrupt leap in capacity requirement will be the connection of schools of all grades to the infrastructure. A significant number of NRENs in Europe is already proceeding along this path.
3. *higher meshing level*. The deployment of fibre infrastructure at the MAN and WAN level creates the case for multiple connection at an affordable price. The fibre infrastructure may be in common with other metropolitan or regional networks and can be shared at the optical, data link or IP layer. The increase in the meshing level is also due to higher expectation for network redundancy at all sites and also between countries and all connected entities (Universities for example).
4. *variation in traffic pattern on the network*. Peer to Peer application and meshing increase bring a significant variation in the traffic patterns in every network. Users stay connected longer in time and transfer large amount of traffic uniformly on the networks, causing an almost "flat" distribution between peers. The classical "sinks" and "sources" of traffic are less important and therefore the network engineering must develop new topological models to sustain the constant load in time and geography. Beside this enough headroom must be kept available to provide bandwidth for new applications.
5. *ubiquitous and mobile computing*. Although not immediate, the demand for ubiquitous and mobile computing is rising fast. As an example, the evolution of

mobile phones to PDA's, with serious computing and networking capabilities, can easily scale by orders of magnitude the number of concurrent mobile flows.

6. *diverse user requirements*. Different user communities are requesting from the network very different behaviours. For example high-energy physicists and radio astronomers require a high capacity, high throughput GRID. Videoconferencing requires a resilient and predictable behaviour.
7. *geographical dispersion of users.* Users of single network services are dispersed in multiple countries. There is a need to provide services with predictable quality in multi-vendor, multi-domain environment. Also dedicated networks.
8. *last mile.* Connecting the network backbone to the users' location is still a non-trivial task, if fibres are not available. New techniques make available an intermediate solution at medium speeds, in the range of Megabits per second. It will be possible in the near future connect at "broadband" speed many sites, previously connected only with a single low speed connection. In addition techniques to lay fibres using existing infrastructures, like gas pipes, water ducts or power lines, allow a more economical large scale deployment of a fibre infrastructure.

4. NREN networks research areas

The NREN have to work closely with research bodies, network providers and manufactures to develop a new generation of research networks capable of sustaining the quantum leap, scaling factor and challenges outlined above.

Key research areas may be identified:

1. *optical infrastructure*. To enable the best possible transmission infrastructure for high speed, long and medium and last mile distance links. The research is not aimed at creating new optical devices, which lies outside the scope and knowledge of NRENs, but rather to closely collaborate with network providers and optical manufacturers to ensure that the infrastructure will be able to support a transport of at least 40-80 Gb/s for each single channel in the core. Switching at the optical layer should be considered as a possible evolution of current linking of channels.
2. *Bandwidth on Demand*. In the broad sense, how to face the challenge for providing a clear channel end to end, in multi-domain environment, where "end" is considered the customer device, exploiting different technologies at layer three, two or one. MPLS is a good example of a possible implementation of this type of network engineering. This will also include using a proper Authorisation, Authentication and Accounting environment.
3. *40Gb/s and beyond***.** Upgrading core links to 40 Gb/s or higher speed represents the simplest upgradepath for the current network infrastructure to satisfy the foreseen capacity requirements. Such speed impose particular requirements not only to the optical physical infrastructure, but also to the router hardware which must be capable of offering standard and advanced services at such speeds.
4. *Simpler data link transport Layer*. SDH at the current stage of deployment is seen as too rigid to accommodate the requirements of modern networks. A more flexible data link, capable of managing the whole link capacity in variable, non fixed by

specification, capacity slices, without multiplexing and demultiplexing the whole transport hierarchy, is needed. An evolution of the present SDH technology or the evolution of Generic Optical Framing can both provide a step in the right direction. Advances in the Ethernet protocol family should also be investigated closely, up to its maximum speed.

5. *routing model*. Routing on a highly meshed network, whether it is composed by many physical or virtual channels, created by MPLS for example, may demonstrate itself a hard task. This both for internal routing and for external routing, in the case where the number of peerings increases in the ten’s range. A shift from the present static, manual configuration of routing cost to the use of more dynamic cost and link usage is an additional research issue.
6. *monitoring and network management*. The high speed, high number of nodes network poses a real challenge on the current monitoring architectures. The quality of monitoring reflects itself closely on the quality of the network response. Active and passive measurements, sampling versus precise monitoring should find the correct balance. Instrumentation to analyse and provide data in real time to Network Operation centres and. possibly’ GRID bandwidth brokers has to be developed. Multidomain management and monitoring needs to be addressed.

A must is that all these research items are to be developed in the context of a multidomain environment. In addition, the end to end model must be applicable to all the architectures and techniques developed, at least as the final outcome of a transparent collation of different approaches in each domain.

Particular attention has to be applied not only to interoperability testing of equipment of various vendors, but also to the preliminary harmonisation of techniques and ideas, which can provide a solution to the above-mentioned issues. Some of the key research areas may actually have different technical solutions, which may fit to a greater or lesser extent the research environment. It is considered fundamental to start the discussion with vendors and standardisation bodies as soon as possible and to foster inter-vendor technical discussion through the standard bodies.

5. Connection to non low link layer technology research

Other supporting architectures must be engineered and deployed to allow the deployment of some of the advanced techniques aforementioned.

5.1 Authentication, Authorisation and Accounting

Some of the advanced research areas mentioned above require a secure communication between entities. Nodes or domains have to exchange commands or request and the network must ensure that proper authentication and authorisation is performed.

This is especially true if a dynamic provisioning system for resources is pursued. Signalling protocols between node in the same domain or different domain for resource reservation employ strong authentication.

In addition to the authentication functions, a secure connection between certification authorities must be in place, to allow, for example, authentication of mobile users or access to remote resources. A secure method for information exchange between different authorities and different authentication techniques has to be set-up.

5.2 Measurement and monitoring infrastructure

Given the link speeds and the size of the network a quantum leap is expected also from the monitoring and measurement infrastructure. Issues related to sampling efficiency, effective capabilities in the node to keep track of usual SNMP variable for each circuit (physical or virtual) have to be actively addressed and resolved.

5.3 Testbeds, test cases and test applications

All the techniques to be developed aim at a rapid prototyping and implementation cycle in the production network. The creation and operation of multidomain testbeds is thus considered essential for the validation and developments of the proposed solutions.

Although the solution will not be engineered having in mind the requirements of a particular application, the GRID paradigm can provide a quite complete and challenging test case at the application level. GRID requirement range already today from very high capacity virtual networks to diffuse, on demand, cooperative computing.

A testbed, which spans multiple NRENs with different user base requirements and GRID application, can provide a good environment where to validate and harmonise the proposed solutions.

5.4 Service Models

New user services will be developed, like Bandwidth on Demand and user empowered networking. The NRENs will have to develop new architectures to operate and offer these services, which are a notable extension to the classical Best Effort transport service. In particular the coexistence of the different services may require a different, more constant and diffuse, management of the network, which has to balance complexity and simplicity in operation and cost.

In addition to user service models, an NREN will probably manage a network made of privately owned fibres, leased infrastructure and leased capacity.

This service model is similar to the model traditionally employed by large carriers, but NRENs have a different set of requirements and constraints. An NREN network must seamlessly integrate itself to a greater extent with other research networks, provide faster services and it is based on a not for profit nature.

It is thus foreseen that, in addition to new techniques the NRENs will have to develop a different model for the creation and management of the network.

REFERENCES

[1] GÉANT: http://www.dante.net/geant

[2] GARR – Italian Research Network – Yearly graph for various destination at http://www.garr.it/mappagarr/garr-b-mappagarr-engl.shtml

[3] “A Step Towards the provision of Optical Networking”, Editor V. Reijs, HEAnet, http://www.terena.nl/tech/projects/testbed/

[4] ATRIUM IST-1999-20675 - http://world.alcatel.be/atrium/index.htm

[5] http://www.broadcastpapers.com/sigdis/NetworkElectronicsWavelengthDM01.htm

[6] SERENATE, IST-2001-34952 - http://www.serenate.org