PoS PROCEEDINGS OF SCIENCE

Lambda Grid developments, History - Present - Future.

Cees de Laat[1]
University of Amsterdam
Kruislaan 403, Netherlands
E-mail: delaat@science.uva.nl

Paola Grosso
University of Amsterdam
Kruislaan 403, Netherlands
E-mail: grosso@science.uva.nl

About 6 years ago the first baby-steps were made on opening up dark fiber and DWDM infrastructure for direct use by ISP's after the transformation of the old style Telecom sector into a market driven business. Since then Lambda workshops, community groups like GLIF and a number of experiments have led to many implementations of hybrid national research and education networks and lightpath-based circuit exchanges as pioneered by SURFnet in GigaPort and NetherLight in collaboration with StarLight in Chicago and Canarie in Canada. This article looks back on those developments, describes some current open issues and research developments and proposes a concept of terabit networking.

Lighting the Blue Touchpaper for UK e-Science - Closing Conference of ESLEA Project
The George Hotel, Edinburgh, UK
26-28 March, 2007

[1] Speaker

PoS(ESLEA)027

1. From Lambda’s to Hybrid Networks

In the late 90’s new laws slashed the Telco monopolies in many countries. The previously state-owned providers were forced to enter the competitive market. That, combined with the very favourable financial situation for telecom, Internet and ICT led to a huge investment in fiber infrastructure, which became also available to end customers if one had the vision to ask for it. A group of NREN’s and scientists gathered in September 2001 at the TERENA offices in Amsterdam to discuss a new approach to networking that would use circuits and colors on fibers. The goal was to start experiments with a first dedicated Lambda between Chicago and Amsterdam, and to use that for tests of new applications and demonstrations at the iGrid2002 [1] event in the Science Park in Amsterdam. The idea was to get a new order of scalability for the Internet by owning the dark fibers and the optical equipment to handle the traffic and, therefore, incur in just small incremental costs for extra capacity later on. The model up to that point was to go to the Telecom provider and ask for a managed service that would typically always be at full cost.

At that time the concept of Lightpaths was introduced [2]. A lightpath is a dedicated connection in an optical network that gives a guaranteed L1 or L2 service to the end user. Nowadays research and education networks around the world offer such connections to e-Science applications that have large bandwidth requirements or are sensitive to network delays. These applications cannot properly function in the traditional shared IP environment where the network behavior is unpredictable due to the large number of concurrent and competing users, and where routing protocols determine the path followed by the traffic. To name a few, in the high-energy physics community, the upcoming CERN experiments in the LHC (http://cern.ch/lhc) use lightpaths for their wide-area data transfers. Also, the visualization applications developed in the OptIPuter [3] transfer images worldwide using SAGE as shared terapixel display workspaces on dedicated lightpaths.

Inversely these huge streams destroy the normal operation of the routed Internet and would require huge investments in routers. However the big streams usually are very long lived flows from same source to same destination and may therefore be mapped to their own circuit and avoid routers. The unit of bandwidth in these applications is Gb/s up to whatever one color can transport, typically 10 Gb/s. Optical photonic switches have huge capacity for much lower prices than routers: a wavelength selective switch that can handle 72 Lambda’s in five fibers costs as much as one 10 Gb/s router port on a full fledged backbone router. A network that provides both the routed IP packet-switched services and the lightpath optical circuit-switched services is called a hybrid network [2]. The Dutch SURFnet6 network is one of the first designed with the above ideas in mind. It is basically a nation wide dark fiber network where a mix of DWDM and packet encapsulation equipment takes care of the transport of the traffic on multiple colors in the fibers. Routing is done only in two places in the network. The traffic from all the Universities and connected organizations that needs to be routed is transported as 1 and 10 Gb/s streams over the network to Amsterdam, where the routers are located. On the same basic infrastructure organizations can obtain dedicated lightpaths to form their own optical private networks or exchange traffic with other organizations for heavily demanding applications. Such lightpaths can be extended via the hybrid exchange NetherLight in Amsterdam to international locations [4]. Current users of this service are notably the Particle Physics Community for distribution of data from CERN and the Astronomers for correlation of radio telescopes. The SURFnet6 hybrid model enables the network to scale to much higher capacities for the same costs compared to a full routed approach.

2. Research and Development in Hybrid networks

The operation of the hybrid networks is two-fold. The routed part is operated in the same way as is current practice in the Internet. The lower layer part for the lightpaths is still mostly done by hand using phone and email as the service layer. A number of pilot implementations exist to aide users and network managers to control the setup of the switches, but they require detailed insight of the local and global setup, of the topology, the interface properties and equipment specifications. They usually also require privileged access to the network devices; any errors could kill someone else's traffic through the same box. Some of the main issues in a multi domain hybrid networking architecture are:

- virtualisation of the network elements by a Network Resource Provisioning System
- authorization (and associated authentication, cost accounting, etc.)
- intra and inter domain topology for path-finding
- addressing schemes of endpoints on the lower layers
- web services versus in band signalling
- connecting together domains with different NRPS's and signalling systems
- applicability of new photonic devices

International projects as Phosphorus, GN2, DRAGON, G-Lambda, Enlightened and organizations as CANARIE and Internet2 try to alleviate several of these issues by developing protocols, interfaces, middleware and tools. In Europe the Phosphorus project is addressing the inter domain lightpath setup signalling where different domains may use different NRPS'es. In the rest of this section we delve into two specific research tracks at the University of Amsterdam. Those are the dynamical DWDM project named StarPlane and the development of a resource description framework for describing the lower layer networks and their interconnections in a format that can be exchanged between domains.

2.1 The StarPlane project

The StarPlane project is designed to add a new level of flexibility to the photonic networking on the dark fiber infrastructure in SURFnet6. The current Lambda based networks usually implement their LightPath services on top of SONET switches that use rather statically configured DWDM infrastructure requiring transponders at switching points. In our opinion a higher level of flexibility in a dark fiber infrastructure of modest size can be achieved operating directly on the photonic level. The challenge in StarPlane is to develop the technology that allows applications running on a grid infrastructure connected to a portion of the SURFnet6 to autonomously switch colors on the photonic layer. In StarPlane applications can optimize the interconnection network topology as function of the computational needs of the application.

PoS(ESLEA)027

The current photonic networks are static for several reasons. Lambda's in fibers can influence each other; e.g. in an amplifier section. Photonic level switching causes the distance, hence dispersion, to change. In the StarPlane project the flexibility is added by introducing NORTEL's wavelength selective switches (WSS) and electronic dispersion compensating transponders (eDCO). The WWS's are functionally a combination of Dense Wavelength Division Multiplexers and micro electro mechanical switches. The switches are capable of selecting any combination of wavelengths out of a number of fibers (in our case 72 colors in 4 fibers) and put them in the outgoing fiber. The actual switching is done by the MEMS part of the WSS and can in principle be very fast. The eDCO works by trying to figure out what the dispersion effect of the fiber is and then modulating the sending laser to generate a kind of inversely deformed signal so that when that signal arrives at the receiver the dispersion is cancelled out. The eDCO can adjust to changing paths. These switches and transponders are now installed in several locations in SURFnet6 and interconnect the DAS-3 [5] computer clusters. The middleware on the clusters is being adapted to enable it to signal to the network its desired network topology.

2.2 Semantic Web Resource Description Framework applied for lightpaths

When applications are in need of a lightpath service the current best practice is to contact the local NREN by phone or email and let them use their knowledge of the worldwide GLIF to construct a connection. In order to get to an automatic setup system it is imperative to somehow distribute the topology of the existing infrastructure. Based on ideas from the semantic web world we constructed an ontology and associated schemas named NDL, the Network Description Language, to exchange, distribute and store this topology information. NDL is an under-development RDF-based ontology that provides a well-defined vocabulary and meaning for the description of optical networks topologies [6].

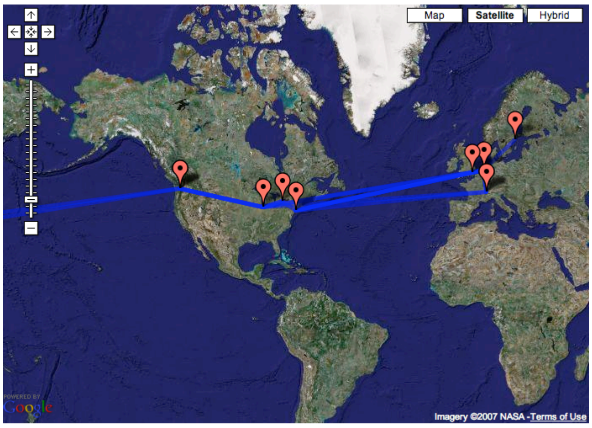

Fig 1 - A Google maps based graph of GLIF lightpaths.

RDF, the Resource Description Framework provides the mechanisms to organize the data. RDF uses a triple-based model to represent resources and the relationship between them. The predicate defines a property of the subject; and the object is the value of such property. RDF identifies subject, object and predicate with URIs - Universal Resource Identifiers. NDL uses RDF/XML to define the ontologies. Currently NDL provides five schemas: topology, layers, capability, domains and physical schemas. Applications for NDL in hybrid networks are: tools to generate up-to-date network maps, to provide input data to path finding algorithms, and to detect errors and faults in existing lightpaths' setup. We have used the mapping tool to model the GLIF network as can be seen in fig 1.

3. The future.

The developments in Networking technology outpace those in computing and storage. In principle the interconnecting networks do not limit anymore the development of e-Science applications. Events like iGrid2002 and iGrid2005 and projects as GigaPort-NG, CineGrid and OptIPuter demonstrate a completely new wealth of applications if effort is spend to carefully engineer all the systems involved. The e-Science and high quality media applications are now approaching the Terascale. We are currently conceptually thinking about Terabit networking. The Lambda networks have given access to an order of magnitude more bandwidth. The unit of bandwidth has shifted from packets to complete photonic channels of 10 Gb/s. However for application to be able to use the Lambdas effectively the programming models need to be adapted. The Lambdas must be treated in the same way as multiple cores and processors in parallel programs. Middleware as MPI and the use of multithreading turns 1000 computers of 1 GigaFlop each into a TeraFlop machine where one application can get the equivalent of a tera number of multiplications for its own use. We need to think of the equivalent of MPI in Networking such that applications can program an equivalent of a Terabit network for the benefit of one application. Examples of such applications are LOFAR, LOOKING, OPTIPUTER, CINEGRID. The solution is to work at removing or lessening the factors that inhibit the optimal use of the emerging new local and wide are photonic networks and to design systems, protocols, and grid middleware that empower applications to optimally allocate and use the infrastructure. This way applications can become location and distance independent except for the unavoidable limit of the speed of light.

References

[1] Thomas A. DeFanti, Maxine D. Brown, Cees de Laat, "editorial: Grid 2002: The International Virtual Laboratory", iGrid2002 special issue, FGCS, volume 19 issue 6 (2003).

[2] Cees de Laat, Erik Radius, Steven Wallace, "The Rationale of the Current Optical Networking Initiatives", iGrid2002 special issue, FGCS, volume 19 issue 6 (2003).

[3] Larry L. Smarr, Andrew A. Chien, Tom DeFanti, Jason Leigh, Philip M. Papadopoulos, "The OptIPuter," Communications of the ACM, Volume 46, Issue 11, November 2003, pp. 58-67.

[4] Leon Gommans, Freek Dijkstra, Cees de Laat, Arie Taal, Alfred Wan, Bas van Oudenaarde, Tal Lavian, Inder Monga, Franco Travostino, "Applications Drive Secure Lightpath Creation across Heterogeneous Domains", IEEE Communications Magazine, vol. 44, no. 3, March 2006

[5] H.E. Bal et al.: "The distributed ASCI supercomputer project", ACM Special Interest Group, Operating Systems Review, Vol. 34, No. 4, p 76-96, October 2000.

[6] J..J. van der Ham, F. Dijkstra, F. Travostino, Hubertus M.A. Andree and C.T.A.M. de Laat, "Using RDF to Describe Networks", iGrid2005 special issue, FGCS, Vol. 22 issue 8, pp. 862-867 (2006).

PoS(ESLEA)027