StarPlane – a national dynamic photonic network controlled by grid applications

Paola Grosso, Li Xu, Jan-Philip Velders and Cees de Laat

SNE Group – University of Amsterdam, Amsterdam, The Netherlands

Abstract

Purpose – This paper aims to look at a research project – StarPlane –a national dynamic photonic network controlled by grid applications.

Design/methodology/approach – The paper examines StarPlane, a research project funded by the Dutch Research Council (NWO) and carried out by researchers at the University of Amsterdam (UvA) and the Vrije University (VU).

Findings – The paper finds that StarPlane provides the management functions needed by applications running on the DAS-3 Grid clusters to interact with the network control plane. The most unique features of StarPlane are the use of photonic equipment in the network core and the fast reconfiguration times of lightpath topologies.

Originality/value – This paper was written with the expertise of the authors and will be of interest to those people involved in computing and e-science.

Keywords Wavelengths, Computer networks, Sciences, Computer applications

Paper type Research paper

Application-controlled dynamic lightpaths

Lightpaths are dedicated end-to-end optical connections in use by a single application. Lightpaths are the answer to the demand for a guaranteed performance from high-end e-science applications; such paths provide a guaranteed set of services to the applications, for example, a predetermined amount of bandwidth or a predictable latency (de Laat *et al.*, 2003). This type of quality of service cannot easily be guaranteed in a regular, routed IP network, where multiple flows share or fight for the same network resources.

There are many ways to actually implement a lightpath: the differences are in the technology chosen to provide the path in the providers' network, the duration of the connection and in its dynamicity.

In some networks lightpaths are dedicated Ethernet connections, in some others they are TDM timeslots, and in others wavelengths over a DWDM infrastructure. Applications or providers control a lightpath; when the user has the possibility of specifying the network topology, or any other of the service's characteristics, at setup or run time we talk about user-controlled lightpaths. The lifespan of lightpaths also varies. In their most simple form lightpaths are fairly static; once configured they remain in place for weeks or months assuming the role of private lines for the user. In

more dynamic environments lightpath setup and teardown happen in shorter timeframes.

StarPlane is a research project funded by the Dutch Research Council (NWO) and carried out by researchers at the University of Amsterdam (UvA) and the Vrije University (VU). In StarPlane Grid, applications that run on the third generation of the Distributed ASCI Supercomputer (DAS-3) have direct access to a portion of the national research and education network SURFnet6. DAS-3 applications use StarPlane to dynamically request lightpaths through the SURFnet6 photonic network.

In the StarPlane case, lightpaths are dedicated colors (wavelengths or lambdas) in the SURFnet DWDM network. A differentiating aspect from other lightpath implementations in similar DWDM networks is that the core equipment carrying the lambdas is uniquely photonic; no OEO conversion takes place within the provider's network (that is, the light never gets converted to electricity for regeneration or switching purposes). StarPlane is an application/user-driven network aiming at second or sub-second reconfiguration times, so that applications can tap the full flexibility of the network.

In the following sections we will give an overview of DAS-3, SURFnet6 and the StarPlane middleware, focusing on its innovative aspects. We will finally outline which DAS-3 applications will most benefit from using StarPlane.

Related work

There are several projects that try to address the management and control of switched lightpaths. In Japan the G-Lambda project (www.g-lambda.net/) is developing a web services-based interface to set up lightpaths in an optical network. In the USA the NSF-funded "Enlightened" project (www.enlightenedcomputing.org/) has similar goals and is closely collaborating with the DRAGON project (http://dragon.maxgigapop.net/twiki/bin/view/DRAGON/WebHome) that aims to deliver a GMPLS implementation to facilitate the signaling for actual lightpath setup. The European Commission funded the project Phosphorus (www.ist-phosphorus.eu/). That project will enhance and demonstrate solutions that facilitate vertical and horizontal communication among applications middleware, existing Network Resource Provisioning Systems, and the Grid-GMPLS Control Plane. The development team of StarPlane is participating in Phosphorus and closely collaborating with the other projects mentioned. Most of the other projects use optical switches in their networks. That means that the lambdas on the fibers are terminated in SONET switches or in some cases Ethernet devices and all lightpath provisioning is done in the electrical domain. Compared to those, StarPlane is unique in that the reconfiguration is done purely in the photonic domain.

DAS-3 clusters

A number of Dutch Universities participate in ASCI, the "Advanced School for Computing and Imaging". A few of these ASCI members, namely the University of Amsterdam (UvA), the Vrije University (VU) in Amsterdam, Leiden University and Delft University designed and contributed to the latest Distributed ASCI Supercomputer, DAS-3 (Bal *et al.*, 2000). DAS-3 is dedicated to research on wide-area distributed and parallel applications. Five clusters totaling 270 dual-CPU supercomputers are located at the four participating universities, with two distinct clusters housed at the UvA.

At each location the clusters differ slightly in total number of nodes, type of processors and amount of storage and memory, but they all share the same high-level architecture, shown in Figure 1.

Each cluster has one head node and a number of computing nodes. The head node connects to the local area network with a 10GE interface while the computing nodes have 1GE interfaces. The clusters can use these connections to communicate with local university resources and with each other over the traditional routed Internet. At the same time all nodes in the cluster connect to a Myrinet switch at 10Gbps for fast local interconnection. From the Myrinet switch there are eight 10GE interfaces towards the SURFnet6 network: these interfaces provide the dedicated connections to the photonic network that is used in the StarPlane research. Figure 2 shows in more detail the optical wide-area setup in more detail.

The eight connections from the Myrinet switch go to a SURFnet Nortel switch, an OME 6500. It is the OME switch that takes the Ethernet interfaces in input and translates each one into a specific wavelength. These eight wavelengths enter the CPL – Common Photonic Layer – at one location and travel as such to another DAS-3 location, where they will be mapped again in Ethernet interfaces by the corresponding remote OME.

CPL – common photonic layer

SURFnet6 is one of the first nation-wide hybrid networks in the world, providing both IP and lightpath services to end-users. SURFnet6 entered production in 2006 and connects universities and research centres with five network rings running through the Netherlands. Every ring can carry up to 36 colours, each with a network bandwidth of 10 Gbit/s packed in a spectral bandwidth of 100 GHz. In the future the number of possible wavelengths will grow to 72 by replacing filters to allow for a spectral bandwidth of 50 GHz. At the same time new coding mechanisms may allow for more
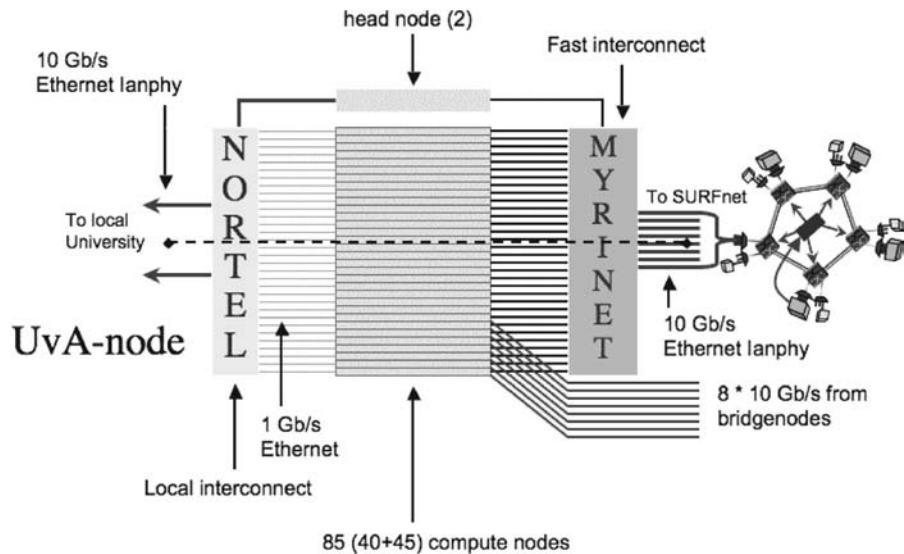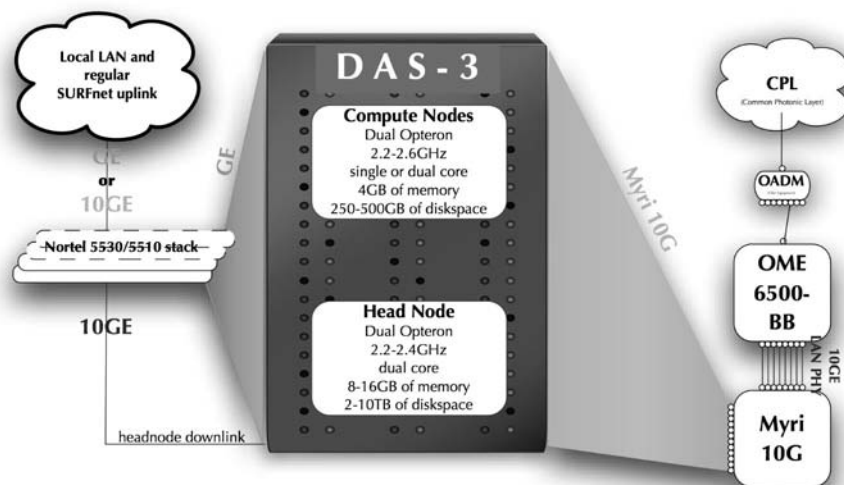

Figure 1.
DAS-3 architecture

bits per Hertz, thus increasing the bandwidth per lambda. Nortel Networks provides the equipment used in SURFnet6 and the infrastructure carrying the various colours is called CPL – Common Photonic Layer. Figure 3 shows the different dark fibre rings throughout the country.

StarPlane is deployed on ring number 1, the ring running through the cities of Amsterdam, Leiden and Delft where the DAS-3 clusters are located. Of the colours available, eight are dedicated for StarPlane use.

One of the components more relevant for StarPlane in the whole CPL architecture is the Nortel Wavelength Selective Switch – WSS. These switches route wavelengths in and out of fibres: given a series of colours coming into such a device, they can send subsets of colours to different outputs. A WSS has incoming and outgoing fibres and Micro Electro Mechanical mirrors that can tilt in about 10 ms to reflect the appropriate colours to the output ports. Figure 4 shows the principle of operation of such device.

The use of these devices differentiates StarPlane from other provisioning projects; a WSS allows for faster reconfiguration and creation of a more dynamic network topology. No conversion to electric signal takes place and the whole switching of wavelengths happens at the photonic level. It therefore saves the cost of transponders and can easily scale. It is the WSS that allows for the creation of dynamic topologies that best suit the application. Another essential component in the system is the use of the Electronically Dispersion Compensating Modems (Transponders) developed by NORTEL. These modems are able to compensate for the varying dispersion conditions when the different colours are switched to different destinations. Since dispersion is dependent on the distance travelled in the fibre and the switching can cause a lambda to travel on a very different distance, the dynamic compensation is essential.

StarPlane middleware architecture

StarPlane middleware provides a web services interface to the Grid scheduler on DAS-3 and interacts with the network control plane that manages the configuration of
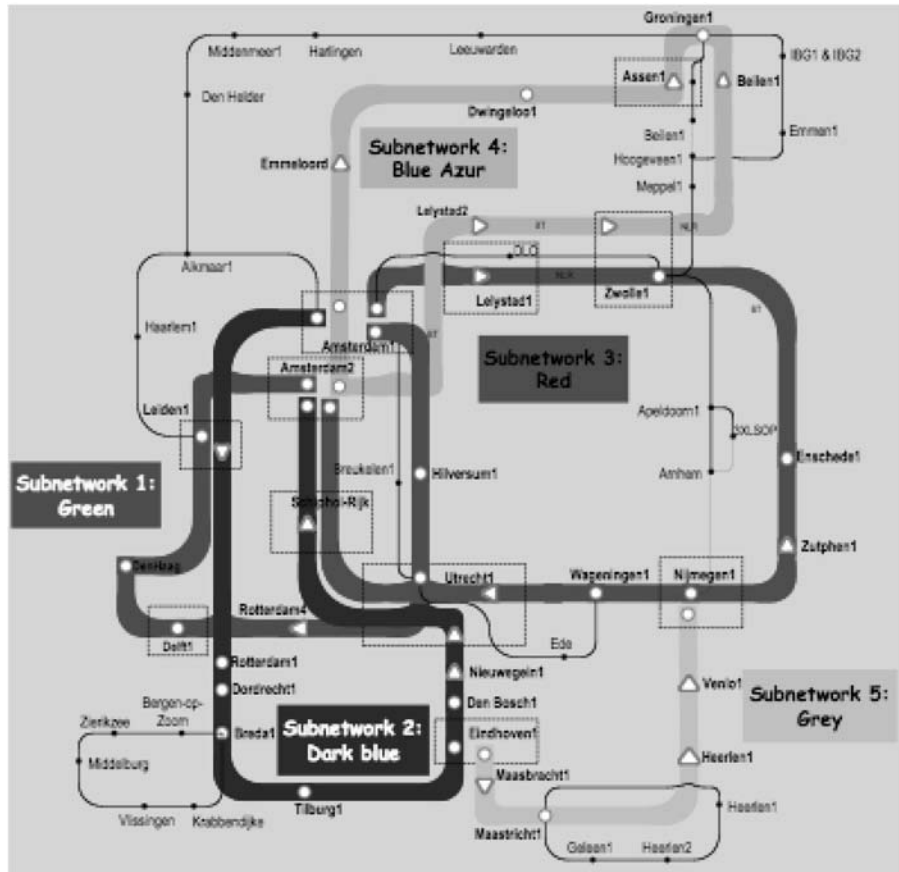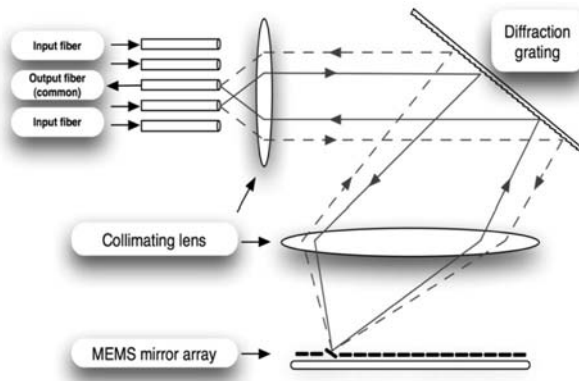
550

Figure 3.
SURFnet6 CPL
architecture

Figure 4.
Wavelength selective
switch

the network devices. Nortel provides DRAC – Dynamic Resource Allocation Controller – for the control plane function. DRAC (www.nortel.com/drac) is a software suite to control the WSS and to create the paths between the OMEs at the DAS-3 sites. Interaction with DRAC also happens through web services.

Two interesting research aspects we are currently incorporating in the design of the system are: the use of semantic models for topology information exchange between applications and middleware; and the concept of a "lightpath portfolio" to achieve the fast configuration and runtime changes that are an essential component of the project.

Network topologies

Network information can be expressed using semantic models based on the ideas of the semantic web. In StarPlane we intend to make use of NDL – Network Description Language (www.science.uva.nl/research/sne/ndl/; van der Ham *et al.* (2006)) as a vehicle for communication between the application and the middleware. NDL provides a series of semantically rich schemas that StarPlane can use to advertise the network topology to applications. The topology presented to the user can be "time-invariant", showing all the lambdas and paths available or "time-dependent" in which only available colours and paths at specified times are exposed. At the same time applications can use NDL to request paths and the desired complex topology that can deliver them.

Lightpath planning

CPL and DRAC provisioning times are currently not short enough to guarantee the desired reconfiguration speed. The configuration times are in the order of minutes depending on the size and complexity of the CPL network. The reason for this is that switching on and of lambdas in principle also influences other lambdas in the same fibre. Amplifier sections where energy is shared by the lambdas and if one vanishes, the others get some extra cause this. Also, non-linear effects can cause crosstalk. Currently, the CPL system takes care for this by very carefully tuning up and down lambdas, monitoring them and regulating filters and amplifiers in the process. However, the system has quite some tolerance, so research is going on to determine if fast (sub-second) switching a limited number of lambdas can be tolerated. StarPlane middleware will research the concept of "lightpath portfolios". The management plane of StarPlane will request a series of lightpaths to DRAC independently from applications requests. It will try to optimise this set based on information on previous use. By having these preconfigured paths ready, it will be possible to satisfy requests for new path in real time, to increase bandwidth and change the underlying topology.

E-science applications on StarPlane

The ultimate goal for building a lightpath management plane like Starplane is to offer to Grid applications ways to optimise their performance and implement new and hopefully more efficient operational models. There are primarily four types of DAS-3 applications in need of StarPlane: large "stand-alone" file transfers, large file stage-in/stage-out, applications with static bandwidth requirements and finally applications with dynamic bandwidth requirements.

The large stand-alone file transfers constitute the most classical application in need of lightpaths. The primary service required is sufficiently large bandwidth to move

data between clusters, for example, backups of data between computing centres. Large file stage in/out is needed to move data on clusters where applications will at a later moment perform computation. The movement of this data is a pre-requisite to the actual processing; during run time the application does not move data and has minimal communication requirements. Another use-case is applications that have large (dynamic) bandwidth requirements at run time. The bandwidth might be needed continuously or intermittently, with periods of large usage followed by period of computation and little network usage. In this case, it is interesting to research if paths need to be kept allocated to the applications during periods of quiescence or can be relinquished and re-instantiated when needed.

Conclusions

The StarPlane project in The Netherlands is developing the middleware that will enable applications to request lightpaths and to make changes to the underlying network topology in near real-time. StarPlane provides the management functions needed by applications running on the DAS-3 Grid clusters to interact with the network control plane. The most unique features of StarPlane are the use of photonic equipment in the network core and the fast reconfiguration times of lightpath topologies. StarPlane fits nicely into developments in the Global Lambda Integrated Facility (GLIF) (http://glif.is/) enabling the emerging network-hungry applications to program and use the Photonic networks as needed (de Laat and Grosso, 2007).

References

Bal, H.E., Epema, D.H.J., van Gemund, A.J.C., Radulescu, A., van Reeuwijk, C. and Sips, A.J.C. (2000), "The distributed ASCI supercomputer project", *Operating Systems Review*, Vol. 34 No. 4, pp. 76-96.

de Laat, C. and Grosso, P. (2007), "Lambda grid developments, history – present – future", *Lighting the Blue Touchpaper for UK E-science*, Closing Conference of ESLEA Project, Proceeding of Science, June, available at: http://pos.sissa.it//archive/conferences/041/027/ESLEA_027.pdf

de Laat, C., Radius, E. and Wallace, S. (2003), "The rationale of the current optical networking initiatives", *Future Generation Computer Systems*, Vol. 22 No. 6, pp. 999-1008.

van der Ham, J., Dijkstra, F., Travostino, F., Andree, H. and de Laat, C.T.A.M. (2006), "Using RDF to describe networks", *Future Generation Computer Systems*, Vol. 22 No. 8, pp. 862-7.

About the authors

Paola Grosso received her PhD in physics from the University of Turin – Italy. She has worked at the Stanford Linear Accelerator Center and since 2004 she is a member of the SNE (System and Network Engineering) group of the University of Amsterdam. She is leading the group effort in the fields of optical networking and her current research interests are dynamic lightpath provisioning and semantic network models. Dr Grosso is co-chair of the OGF Network Mark-up Language Working Group (NML-WG).

Li Xu is PhD student on the StarPlane project at the University of Amsterdam. Li Xu received his Master's degree from the University Leiden in 2006. In 2003 he received a BSc in computer software engineering from Tsinghua University, Beijing, China and in 2001 a BSc in management information systems from Fudan University, Shanghai, China.

Jan-Philip Velders is System Administrator in the ICT Advanced Technologies Group at the University of Amsterdam. He is working part-time on the traffic engineering, control and setup of the StarPlane project. He is member of the SURFnet CERT team.

Cees de Laat is associate professor in the Informatics Institute at the University of Amsterdam. Current research in his group includes optical/switched networking to optimize Internet transport of massive amounts of data for TeraScale eScience applications, RDF to describe networks and associated resources, distributed cross-organisation authorisation architectures and systems security. With SURFnet he implements projects in the GigaPort Research on Networks. He collaborates in the NSF – OptIPuter project. He serves in the Open Grid Forum as Grid Forum Steering Group (GFSG), and is Infrastructure Area Director and IETF Liaison. He is co-founder and organiser of several of the past meetings of the Global Lambda Integrated Facility (GLIF). Cees de Laat is the corresponding author and can be contacted at: www.science.uva.nl/ ~ delaat