StarPlane – an Application Controlled Photonic Network

Paola GROSSO, Jan-Philip VELDERS, Li XU, Cees de LAAT
Universiteit van Amsterdam, Kruislaan 403, 1098SJ, Amsterdam, The Netherlands
Tel: +31 20 525.7533, Fax: +31 20 525.7419, Email: P.Grosso@uva.nl

Abstract: The StarPlane research project in the Netherlands aims at building an 'application controlled photonic network'. Applications running on the DAS-3 Grid cluster will request lightpaths on a dedicated portion of the SURFnet6 network. Fast reconfiguration times and topology changes, together with the use of photonic devices in the core of the network are the most innovative aspects of this project.

1. Introduction

e-Science applications started to pose ever-growing demands and requests for more specialized network services. The traditional routed IP infrastructure cannot provide these applications with optimal services; many of these new scientific applications do not perform well when they have to contend for bandwidth with others. For these reasons research and education networks around the world are now offering dedicated connections named lightpaths to end users. They deploy networks that run in *hybrid* mode: routed IP and lightpath services coexist in the same infrastructure [1]. The network operator, either directly or through automated tools, reserves capacity (a wavelength) in the underlying (DWDM) network and assigns it for exclusive use to an application. The application becomes the sole user of this path for the allotted timeslot.

This model works well for applications that have fairly static connectivity constrains and that use a path for relatively long time scales (months); it is not well suited for applications that need more dynamic topologies over shorter time scales. In these cases there is a need for a novel approach that gives the applications themselves control of the network without intermediation from the network administrator. We also need mechanisms to quickly change the underlying network topology so that the application can obtain lightpaths matching the specific communication requirements of the moment.

StarPlane is a research project funded by the Dutch Research Council (NWO) and carried out by researchers at the Universiteit van Amsterdam (UvA) and the Vrije University Amsterdam (VU). In StarPlane Grid applications that run on the Distributed ASCI Supercomputer 3 (DAS-3) have direct access to a portion of the national research and education network SURFnet6. The project researches how to create an 'application controlled photonic network' that is both flexible and dynamic, and can accommodate any request coming from e-Science applications running in the DAS-3 Grid on short notice. The vision is to make the network infrastructure part of the programming model of the application.

In this document we will describe the innovative aspects of the StarPlane project. We aim to provide a clear overview of the differentiating and original aspects of this research with respect to the already existing or under development systems for lightpath provisioning.

2. StarPlane Infrastructure

In StarPlane Grid applications are in control of the photonic paths. The applications run on the DAS-3 clusters and use the paths provided by the SURFnet6 network. In the following two sections we will introduce first the DAS-3 and then SURFnet6.

2.1 DAS-3 Clusters

To create a shared computing infrastructure for their students and researchers a couple of Dutch universities that participate in the Advanced School for Computing and Imaging (ASCI) created DAS – the Distributed ASCI Supercomputer [2].

The DAS is now in its third generation, with the DAS-3 incarnation just having been installed. DAS-3 provides a test-bed for Grid applications, parallel and distributed systems, including communication substrates, programming environments, and applications.

DAS-3 is composed of five clusters with about 270 dual-cpu nodes supercomputers in total, which are integrated into a single large-scale distributed system. The five DAS-3 clusters are located in Leiden, Delft and in Amsterdam at the Vrije University (VU) and at the Universiteit van Amsterdam (UvA). The clusters installed at each location differ slightly in total number of nodes, type of processors, amount of storage and memory: but they all share the same high-level architecture, shown in Figure1.
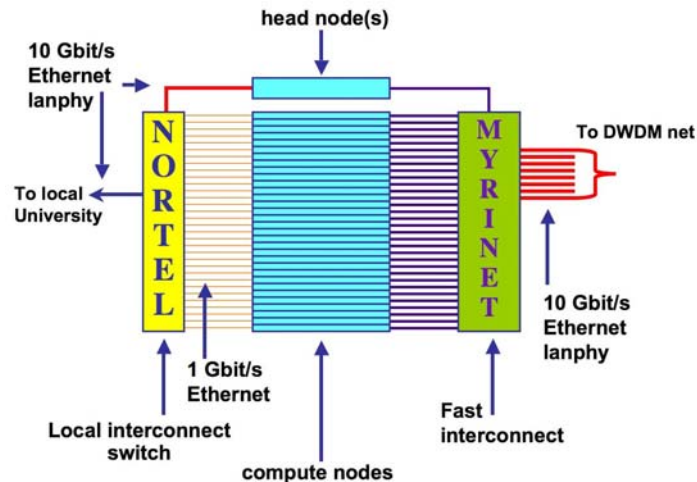

Figure 1 - DAS-3 Cluster Architecture

Each cluster has one head node and a number of compute nodes. The head node connects to the local area network with a 10GE interface while the compute nodes connect with 1GE interfaces. The clusters can use these connections to communicate with local university resources and with each other over the traditional routed part of SURFnet6. At the same time all nodes in the cluster connect to a Myrinet switch at 10Gbps for a fast local interconnect. Parallel computing techniques based on MPI - Message Passing Interface – benefit from and make extensive use of these connections.

In previous DAS installations there were two ways to exchange data: communication between sites using the Internet, or between cluster nodes at the same location using a fast local (Myrinet) interconnect. In addition to that, DAS-3 sites now have a dedicated wide area dark fiber and DWDM technology based network over the hybrid all-optical SURFnet6 network. This constitutes the innovative aspect in the DAS-3 network design. From each 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.

2.2 SURFnet6

SURFnet, the Dutch Research and Education Network, is now running its 6th generation network. SURFnet6 entered production in 2006 as one of the most innovative research and education network in the world, fully implementing the hybrid network paradigm.

SURFnet6 deploys multiple fiber optic rings that connect the various academic and research locations around The Netherlands. Each ring consists of 36 lambdas available for either production IP services or lightpaths. The number of wavelengths on each ring will actually increase in a second phase to 76. This adaptive, all-optical, intelligent network was built using Nortel Networks equipment. The network equipment used is a combination of the Common Photonic Layer (CPL), the Optical Multiservice Edge (OME) 6500 and the Optical Metro 5000. The Metro Ethernet Routing Switch 8600 is deployed to aggregate and backhaul Ethernet connections towards the routed IP core.

One of these photonic rings connects the universities in Leiden, Delft and Amsterdam, the locations of the DAS-3 clusters. Eight wavelengths in one of the nine bands on this ring, basically a frequency band, constitute the StarPlane lightpaths. The same underlying physical infrastructure simultaneously carries both production services for the various universities and research institutes that connect to the ring, as well as the experimental dynamic lightpath services for the DAS-3 locations.

3. The StarPlane Technology

The use of lightpaths, as we indicated in the Introduction, is a current international trend in research networks, which provides the best functionality and services to applications. StarPlane delivers these 'coloured' connections to DAS-3 applications by using two innovative technologies: at the hardware level using Wavelength Selective Switches (WSS) and electronic dispersion compensating modems, and at the software level by creating Web Services interfaces that provide the interaction mechanisms between the application and the network.

3.1 WSS - Wavelength Selective Switch

What enables fast and dynamic reconfiguration of coloured paths in the SURFnet6 network are the Nortel WSSs. These devices route wavelengths in and out of fibers: given a series of colours coming into such a device, they can send subsets of colours to different outputs. A WSS has incoming & outgoing fibers, and Micro Electro Mechanical mirrors, that tilt in 10ms to reflect appropriate colours to output ports. Figure 2 shows the principle of WSS operation.

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 into electric signals takes place and the whole switching of wavelengths happens at the photonic level. It therefore saves the cost of transponders and can easily scale.

DRAC [3] is the software suite developed by Nortel that controls the CPL equipment. DRAC configures the WSSs for StarPlane and creates the paths between the OMEs at the DAS-3 sites. It provides Web Services interfaces to the StarPlane middleware.
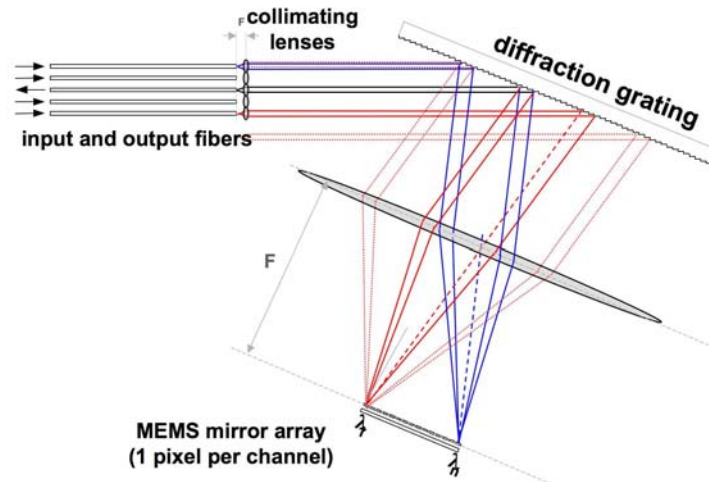
Figure 2 - Principle of Operation of WSS

4. StarPlane Software Design

In StarPlane we look at two kinds of dynamic behaviours: a fast initial setup time and a fast lightpath service change. It is expected in fact that also during runtime DAS-3 applications will require changes in the underlying network topology to access data and computing nodes at different clusters than the ones they had originally started outwith. In both cases we strive to reach (re-)configuration times of the order of seconds and ideally sub-second: this requires fast reaction times of the WSS, CPL, DRAC and obviously the StarPlane middleware.

To achieve these goals the overall software architecture of StarPlane is composed of:

- a *management plane*, that handles the requests from the applications and translates them into instructions for the control layer; the development of this management plane is the central component of the research project;

- a *control plane* that receives the management requests and interacts with the network devices. This control plane is effectively the DRAC and the CPL provided by Nortel.

Applications running on the DAS-3 clusters request topologies from the management plane of StarPlane via Web Services interfaces. Web Services interfaces are also used in the interaction between the management and the control plane. This separates the development of the two and makes them independent from one another as long as the interfaces are well defined at the beginning of the development.

The management plane of StarPlane does not only interface with the application; it also maintains the schedule of present and upcoming network reservations, and it interacts with the authorization servers that verify the credentials of the end users.

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.

4.1 Network Topologies

Network topology information can be expressed using semantic models based on the ideas of the Semantic Web. In StarPlane we intend to make use of NDL – the Network Description Language [4] - 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 its

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 colors and paths at the specified times are exposed. At the same time applications can use NDL to express their requests for paths and their desired complex topology that they would like to be delivered to them.

4.2 *Lightpath Planning*

The path provisioning times of CPL and DRAC are currently not short enough to guarantee the desired reconfiguration speed of (sub-)seconds. StarPlane middleware will research the concept of 'lightpath portfolios'. The management plane of StarPlane will reserve a series of lightpaths to DRAC independently from any applications requests. It will try to optimize this set based on information from previous usage. By having these preconfigured paths ready it will be possible to satisfy requests for new paths in real time. It will also be possible to use these pools of paths to modify the lightpath's service characteristics, i.e. increase bandwidth or change the underlying topology of existing connections, without the application noticing.

5.   Applications Using StarPlane

In this dynamically (re)configurable network DAS-3 applications can find and experiment with new ways to share, consume and distribute data. Such a network represents an extension of the Grid into the realm of wavelengths. Lambdas in this photonic network become a resource that can be scheduled, allocated and aggregated, as we have been doing for years with computing resources; we essentially create a 'lambda grid' that cooperates with the computing Grid.

The ultimate goal for building a lightpath management plane like Starplane is to offer Grid applications ways to optimize 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 endpoints, for example the data backup between computing centers. Large file stage in/out is needed to move data between clusters where applications will at a later moment perform computation. The movement of this data is a pre-requisite to the actual processing, during runtime the application does not move data and has minimal communication requirements. Another use-case is applications that have large (dynamic) bandwidth requirements at runtime. The bandwidth might be needed continuously or intermittently, with periods of large usage followed by periods of computation and little network usage. In this case it is interesting to research whether the paths need to be kept allocated to the applications during periods of quiescence or can be relinquished and re-instantiated at need.

6.   Business Benefits

The StarPlane project provides feedback to the academic community on the optimal design of provisioning systems for lightpaths devoted to e-Science applications. The whole idea of hybrid networks, and dedicated connections stems from the observation that scientific applications are changing in nature: they are becoming more bandwidth intensive and also more geographically distributed. SURFnet with its SURFnet6 design addresses this need. StarPlane takes the idea further placing control closer to the user.

But apart from the academic background of such a project, there are also clear implications for the business community. StarPlane creates in fact a synergy between the research world and commercial world, in this case through close collaboration with Nortel Networks. The development of products such as DRAC could be shaped by the needs of the R&E community brought forward by StarPlane applications and *modus operandi*. Use of DRAC could then transfer this know-how to the wider networking world outside the academic and research institutions, and spur on new insights and novel uses.

6.1 Business Potential

From the current trend in research networks and increased business interest it is clear that lightpath services will become more readily available to the general public through the commercial providers, given the booming of service sensitive applications such as digital cinema distribution, relocatable virtualized machines or even data center disaster recovery. To support these kind of services we will certainly witness the emergence of new business models that integrates lightpath usage for and by the customer, be they an end user or business. Initiatives such as the IPSphere forum [5], whose mission is to deliver an enhanced commercial framework for IP services and other new technologies, could address this new paradigm.

The whole networking and computing industry will certainly incorporate some of the ideas coming out of StarPlane and other similar research projects in their new product designs. Making networks an integral part of an application end-to-end architecture provides added value to the service to clients.

7. Conclusions

StarPlane develops an application controlled photonic network. It controls a dedicated portion of the SURFnet6 network, allowing deployment and experimentation with new services. StarPlane gives applications complete control of a photonic lambda grid, empowering end users to create multiple concurrent network topologies that satisfy their computational requirements in real time. Medical applications, visualization applications and high data volume applications running on the DAS-3 cluster will all take advantage of this infrastructure.

Acknowledgements

References

[1] Cees de Laat, Erik Radius, Steven Wallace - "The rationale of the current optical networking initiatives", In: Future Generation Computer Systems, Vol.22, Issue 6, pp. 999-1008, 2003

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

[3] DRAC- Dynamic Resource Allocation Controller, URL: http://www.nortel.com/drac

[4] J.van der Ham, F. Dijkstra, F. Travostino, Hubertus M.A. Andree and C.T.A.M. de Laat, "Using RDF to Describe Networks", In: iGrid2005 special issue, Future Generation Computer Systems, volume 22 issue 8, pp. 862-867 (2006).

[5] IPSphere forum, URL: http://www.ipsphereforum.org/