Hybrid optical switching for data-intensive media grid applications

Jurgen Baert, Marc De Leenheer, Bruno Volckaert, Tim Wauters, Pieter Thysebaert, Filip De Turck, Bart Dhoedt, Piet Demeester
Dept. of Information Technology, Ghent University – IBBT – IMEC
Gaston Crommenlaan 8 bus 201, 9050 Gent, Belgium
Tel: +32 (0)9 33 14900 Fax: +32 (0)9 33 14899
E-mail: <jurgen.baert@intec.ugent.be>

ABSTRACT
Emerging service grids are being challenged by an increasing number of jobs, larger job sizes, the demand for more storage and computational resources, and ever-rising bandwidth and latency requirements. A specific case of these grids, designed to administer highly demanding media applications (including media production environments) are the so-called media grids. Even though optical networking is the apparent technology of choice for high-bandwidth applications, the question arises whether current burst/packet and circuit switching solutions are flexible enough for the design and development of these next generation optical service grids. This paper introduces the concept of a media grid, deduces the requirements for such a grid. Furthermore, this paper proposes hybrid optical switching for the underlying network architecture, which is expected to substantially increase the overall system performance, compared to single-technology solutions.

Keywords: optical burst switching, optical circuit switching, hybrid optical switching, media grids

1. INTRODUCTION

During the past decennia, computing resources such as processor power and storage capacity, as well as network speeds have known a remarkable increase. Despite this evolution, there are applications where deployment of dedicated hardware is not a profitable approach. For these applications, the grid concept offers a solution: resources can be combined and shared between several clients, thereby reducing the per-client infrastructure cost. This flexibility is achieved by using an underlying network to transfer jobs from client to resource.

First generation grid networks comprise mostly eScience applications: particle- and astrophysics and biomedical research are typical examples. In these grids, the number of users is low, but the computational and storage demands are enormous. An evolution exists however towards service-oriented grids with a larger amount of users and more diverse types of tasks, opposed to the smaller number of tasks and users a traditional grid is limited to. Another distinctive characteristic of service grids is the increased (and near real-time) interactivity with the users, again opposed to eScience grids where a launched job can run for several days, weeks or even months. Service grids typically exhibit smaller job sizes, and the network delay, reliability and QoS demands become more important. Figure 1 shows an overview of the different grid types and their respective job requirements. As can be seen on the figure, the group of service grids can be divided into two subcategories, one for residential users and one for enterprise customers.

Future prospects regarding grids include an ever-increasing bandwidth usage. The choice for optical carriers therefore seems quite obvious, given their relatively low pricing and the high bandwidths they provide.

The remainder of this paper is structured as follows: first consumer grids and media grids are discussed and their requirements are considered. Then a general hybrid node design for use in media grids is proposed.

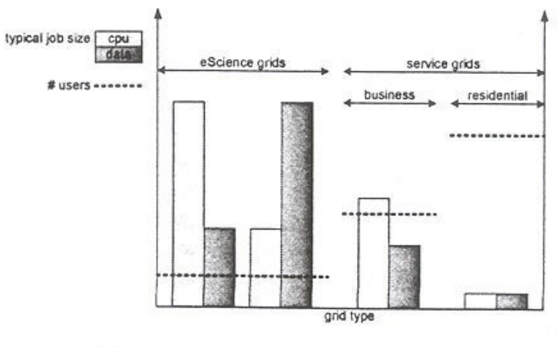


Figure 1: Indicative comparison of the number of users, the storage and processing requirements for eScience and service grids

2. CONSUMER GRIDS

The increasing popularity of grids in the research community has led to an emerging interest in grid computing for the enterprise and consumer markets. Service grids for these types of application will need to handle a large amount of individual jobs, each having only modest resource requirements. Specifically grids for the consumer market, with a potentially vast number of users, will have extreme scalability demands.

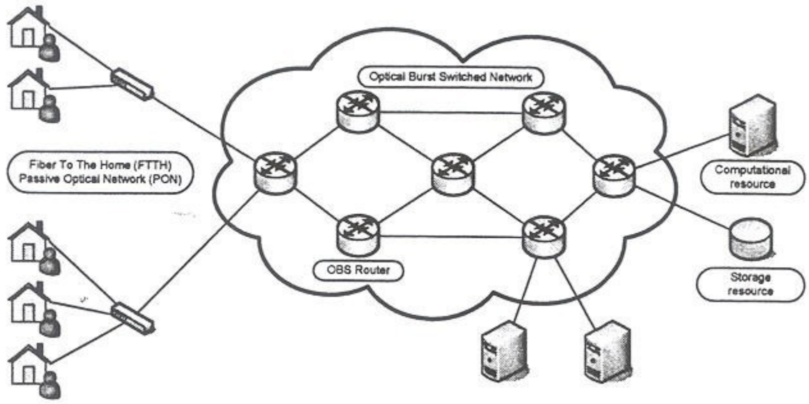


Figure 2: Example consumer grid configuration

2.1 Use case

One application example of consumer grids is the online visualization of a virtual environment [4]. A virtual environment is typically made up of various objects, described by their shape, size, location, the applied texture, etc.. Usually, the description of a scene can be realized in a limited storage space, typically around a few megabytes. Rendering the scene however, requires substantial computational power: a demand of 300 million polygons per second results in required capacities as large as 10.000GFlops. With a frame rate of 25fps, latency below 40ms per frame is required. Assuming a 2.5MB scene description size (transmission time of 8ms on a 2.5Gbps link), about 30ms is available for processing a frame on the remote resource, which is feasible.

2.2 Job characteristics

Several important properties of the job generation process in consumer grids can be identified:

- A large amount of jobs is generated.
- Individual jobs have fairly modest resource requirements.
- Job sizes are quite small, usually in the megabyte range. This means network holding times are short compared to switching times. For instance, a 1MB burst is transmitted in 0.2ms over a 40Gbps link, while current photonic switches have setup times in the order of several milliseconds.
- Job generation is bursty both in time and in space: not only the time when, but also the location where a job is generated, is highly unpredictable and has a bursty nature. This can lead to frequent mismatches in available versus generated load, indicating remote execution of jobs is a necessity.
- Several application types will be interactive in nature, and as such, a strict deadline must be met for successful operation. Other applications have different degrees of tolerance for various system parameters, which implies the existence of several QoS classes.

2.3 General requirements

The architecture as a whole, and the control plane in particular, should have the following properties:

- Because a large number of users and resources must be supported, *scalability* of the architecture is essential. This becomes even more relevant when considering the dynamic nature of job submission: peaks in generated load should be handled seamlessly.
- It is economically not justifiable to build a dedicated network for each application. Consequently, the basic infrastructure should be able to support all application types, each with its own access and resource usage patterns.
- Support for real-time applications is only possible through adequate levels of speed and flexibility in the control plane, implying low latency signaling is required. Also, the overhead imposed by the control plane on the involved entities should be minimized.

3. MEDIA GRIDS

Grids for the enterprise market fill the void between the extremes that eScience and consumer grids pose. In terms of the number of users, average job size, bandwidth and processing requirements, these grids exhibit widely differing characteristics. One specific case of such a grid, the *media grid*, is considered. As the name suggests, the primary application of such a grid network is transferring and processing media (audio and video) streams, in a production environment, such as a news broadcaster.

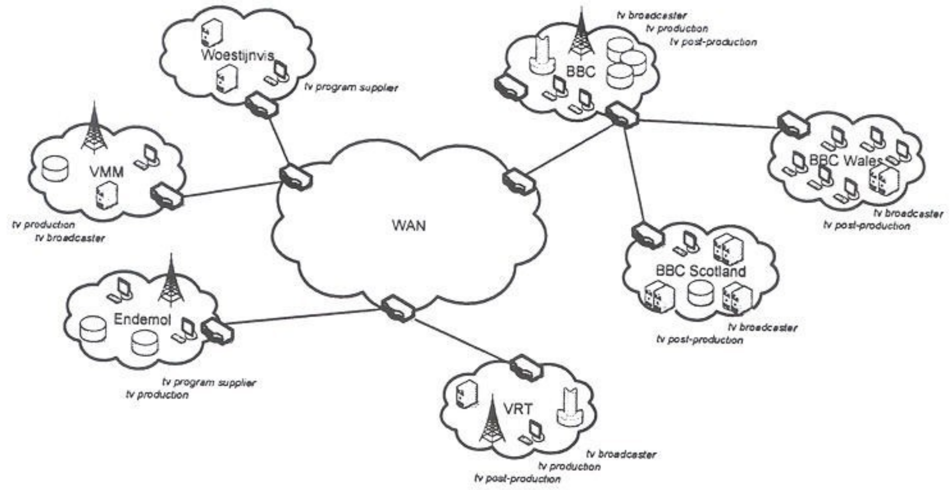


Figure 3: Example media grid configuration

3.1 Use case

Consider a national television production site. In such an organization, hundreds of journalists gather video material that has to be ingested, edited, archived and played out. Journalists, craft editors, producers and directors need to be able to access and perform operations on the media streams. To this end, current facilities include computational clusters and centralized data storage and archive infrastructure. The capacity of these resources has to be large enough to account for all possible scenarios, and is therefore often overdimensioned. In a situation as depicted in Figure 3, several media sites are interconnected using a network. Peak loads from one site can then be adequately handled by executing part of the job load at remote sites with idle resources. This allows sites to have smaller resource pools, reducing the overall cost and increasing the resource usage efficiency.

3.2 Job characteristics

In a media grid environment, task profiles exhibit a huge diversity in their requirements. Operations performed on low resolution data streams require little bandwidth, but some amount of computational power for the conversion of the stored high resolution video to the desired lower resolution. Video craft editing requires both high bandwidths (several high resolution streams) and large amounts of computational resources. Rendering on the other hand requires less bandwidth (often a mere scene description), but a lot of cpu power.

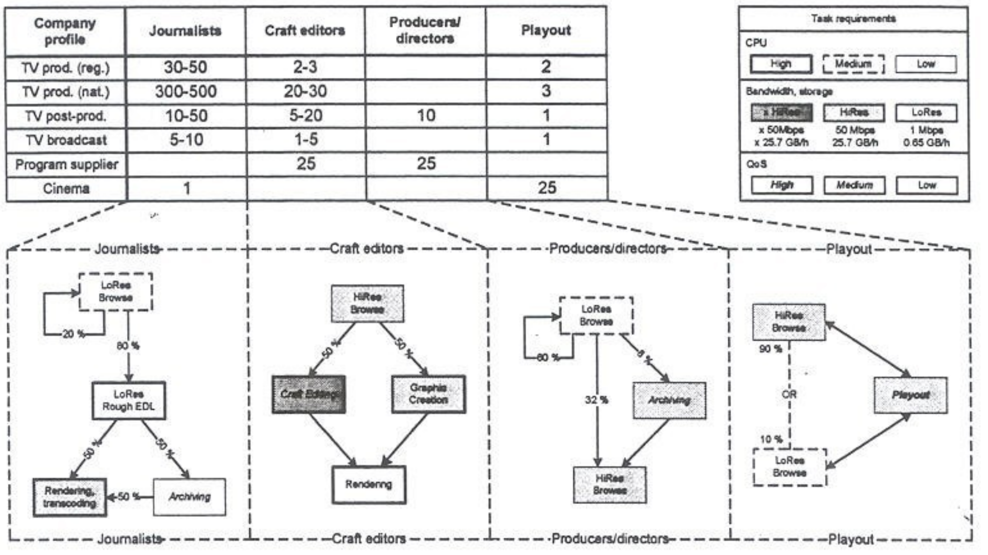
The different types of media sites (television production, post-production and broadcast, program suppliers, cinema, etc.) each have different user classes, such as journalists, craft editors and directors [1]. Figure 4 contains a non-exhaustive overview, along with the tasks associated with each of the displayed profiles and their respective job requirements. As can be seen, each user will have a more or less fixed pattern of job groups, with varying lengths and varying requirements, and will periodically go from one type of job to another (in a Markov-like process).

3.3 General requirements

A media grid architecture should provide adequate support in the following areas:

- The grid needs to be able to handle a broad spectrum of tasks, as well in storage size and bandwidth requirements, as in processing power.
- More than traditional grids, media grids should be able to handle peak loads: important news events draw a lot of attention and result in an avalanche of job requests to a specific location, requiring media files related to this event. Data replication and load balancing strategies for handling these peak loads are required to ensure adequate functionality at all times.

- As is the case in consumer grids, the large number of users leads to the need for scalability. For example, the link and resource state information in larger networks cannot be flooded over the entire network, so localized scheduling and routing decisions might need to be made by the switches.
- The media grid will require the presence of QoS-classes, to be able to guarantee real-time interaction between clients and resources for applications that need this feature.
- Reliability is an extremely important characteristic of media grids. Tasks such as broadcasting can never be allowed to be compromised by (single and multiple) link or node failures.
- Security and admission control is an obvious requirement when interconnecting several corporate sites.



Company profile	Journalists	Craft editors	Producers/ directors	Playout
TV prod. (reg.)	30-50	2-3		2
TV prod. (nat.)	300-500	20-30		3
TV post-prod.	10-50	5-20	10	1
TV broadcast	5-10	1-5		1
Program supplier		25	25	
Cinema	1			25

Figure 4: Company, task, job profiles and workflows inside a media grid

4. NETWORK ARCHITECTURE

Given the nature of a grid, the importance of the underlying network is apparent. Therefore, network architecture choices influence the overall performance of the grid in a substantial way. When dealing with optical carriers, the employed switching technology is one of the most influencing architectural decisions. Obviously, the implementation of the control plane is another important factor.

Traditionally, a distinction is made between two types of switching paradigms: optical burst/packet switching (OPS, OBS [2]) and optical circuit switching (OCS). Circuit switched technologies set up a lightpath from source to destination to reserve bandwidth before sending data, while burst (and packet) switched networks do not reserve a channel for communication. In such a configuration, data is sent into the network accompanied by a control packet – which may be separated in space and time from the data packet.

Traditional eScience grids with high bandwidth requirements, long job lengths and a small amount of clients are known to perform well in an OCS configuration. Consumer grids on the other hand, with their much larger user base and smaller job sizes, can only achieve an acceptable efficiency in an OBS network, as stated in [3].

Media grids, as mentioned before, can be found between these two extremes. The large number of users in these grids makes OCS a less attractive strategy (given the limited amount of available wavelengths). Using OBS for all high bandwidth communication however significantly increases the cost of the network without outperforming OCS. A hybrid approach, combining the advantages of the OBS and OCS switching paradigms, is expected to obtain the best price/performance ratio.

The control plane of the envisaged network will need to be able to deal with both burst and circuit switching in an adaptive way. However, timing is a much more important issue in OBS networks than it is for OCS. Therefore, in the control plane, priority handling will require special attention.

5. HYBRID SWITCHING NODE DESIGN

Burst switching requires faster hardware, which comes at the direct cost of a higher price. Even though this expensive hardware can be used for circuit switching as well, OCS does not require such fast switch reconfiguration. Therefore, a hybrid switch can be made up of a substantial portion of slower, cheaper switching equipment without significantly reducing the overall performance.

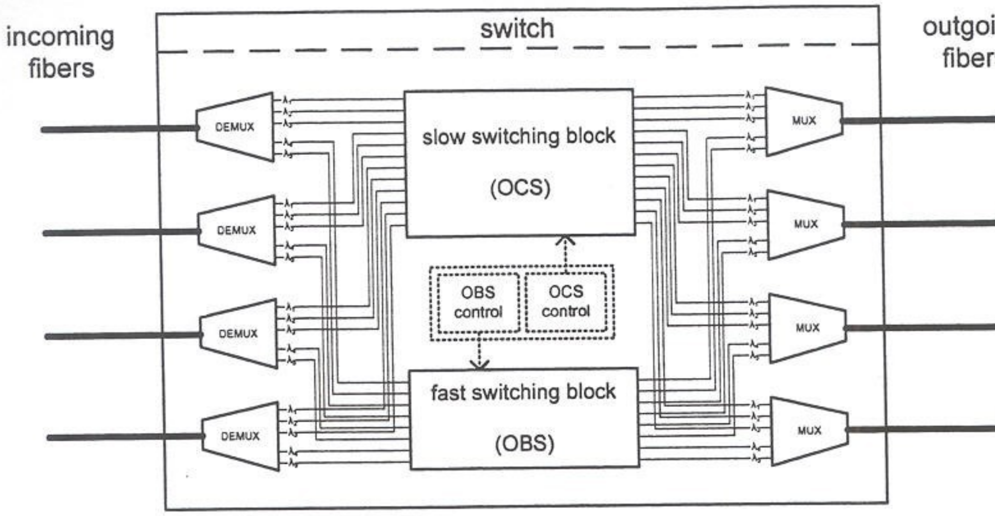


Figure 5: Simple hybrid node architecture

Figure 5 shows the basic layout of a hybrid node capable of OBS and OCS. In this simplified example, the available wavelengths have been pre-divided into the two categories, one for fast switching using OBS, and one for slower OCS handling.

More complicated schemes without a preconfigured division (such as lambda switching, OBS tunneling, etc.) could dynamically divide the wavelengths. The generic node design however will remain similar: a block of expensive hardware, capable of fast burst switching, combined with a block of slower OCS hardware to keep the total cost at a reasonable level.

The control units for these respective switching blocks have obvious similarities (such as switch fabric interface) as well as differences: the network processor for the OBS part includes a policy scheduler and queue management entities, which are not needed for circuit switching.

6. CONCLUSIONS AND FUTURE WORK

We have introduced the media grid as a case study of data-intensive service grids and deducted the general requirements for such a grid. Hybrid optical switching was proposed as an efficient technology to fulfill these requirements, and a hybrid node design combining speed and cost-efficiency as needed was suggested.

Future work will include research on efficient routing and scheduling algorithms with intelligent dynamic switching technology selection. Cost-effective network dimensioning and a more extensive study and implementation of the control plane are other topics of future research.

7. ACKNOWLEDGEMENTS

Part of this work has been funded by the IBBT FIPA project – https://projects.ibbt.be/fipa.

M. De Leenheer and B. Volckaert are supported by the Institute for the Promotion of Innovation by Science and Technology in Flanders (IWT-Vlaanderen).

P Thysebaert is a Research Assistant by the Fund for Scientific Research – Flanders (F.W.O.-Vlaanderen).

F. De Turck is a postdoctoral Fellow of the Fund for Scientific Research – Flanders (F.W.O.-Vlaanderen).

REFERENCES

[1] FIPA – File based Integrated Production Architecture Project. https://projects.ibbt.be/fipa.

[2] C. Qiao and M. Yoo. Optical Burst Switching (OBS) – A New Paradigm for an Optical Internet. *Journal of High Speed Networks*, 8(1):69-84, Jan 1999.

[3] M. De Leenheer, E. Van Breusegem, P. Thysebaert, B. Volckaert, F. De Turck, B. Dhoedt, P. Demeester, D. Simeonidou, M.J. O'Mahoney, R. Nejabati, A. Tzanakaki and I. Tomkos. An OBS-based Grid Architecture. *Proc. of the Globecom 2004 Workshop on High-Performance Global Grid Networks*, Nov 2004.

[4] D. Simeonidou (ed.) and R. Nejabati (ed.). Optical network infrastructure for grid, Mar 2004. https://forge.gridforum.org/projects/ghpn-rg/document/draft-ggf-ghpn-opticalnets-1/en/1/.

Workshop on

Design of Next Generation Optical Networks:
from the Physical up to the Network Level Perspective

Ghent, Belgium

February 6th, 2006

Organization by COST291 & IWT GBOU ONNA

COST

INTEC

IWT
GBOU

Table of Contents