Future Generation Computer Systems 26 (2010) 576–589

Contents lists available at ScienceDirect

Future Generation Computer Systems

journal homepage: www.elsevier.com/locate/fgcs

Joint resource and network scheduling with adaptive offset determination for optical burst switched grids

Mehmet Koseoglu *, Ezhan Karasan

Department of Electrical and Electronics Engineering, Bilkent University, TR-06800, Bilkent, Ankara, Turkey

ARTICLE INFO

Article history:
Received 26 February 2009
Received in revised form
3 November 2009
Accepted 16 November 2009
Available online 24 November 2009

Keywords:
Grid computing
Optical burst switching
Grid resource provisioning
Grid network provisioning

ABSTRACT

Optical burst switching (OBS) is a promising technology for optical grids with short-lived and interactive data communication requirements. On the other hand, burst losses are in the nature of the OBS protocol and these losses severely affect the grid job completion times. This paper first proposes a joint grid resource and network provisioning method to avoid congestion in the network in order to minimize grid job completion times. Simulations show that joint provisioning significantly reduces completion times in comparison to other methods that perform network provisioning after grid scheduling. An adaptive extra offset based quality of service (QoS) mechanism is also proposed in order to reduce grid burst losses in case of network congestion. Results show that this adaptive mechanism significantly reduces grid completion times by exploiting the trade-off between decreasing loss probability and increasing delay introduced by the extra offset time.

1. Introduction

Computational requirements of scientific problems are growing dramatically and this increase requires collaboration between remote institutions in order to solve these problems. For example, the Large Hadron Collider at CERN will produce 15 petabytes of data in a year which has to be processed to determine useful observations [1]. Since it is not possible to process this enormous amount of data at a single processing site, geographically distributed resources have to collaborate to solve such problems. Grid computing introduces a new paradigm in which independent remote institutions collaborate through a uniform interface to solve complex problems [2].

In addition to scientific problems, grid computing is envisioned to be used for consumer applications in the future [3]. In this case, consumers rent computational resources from remote servers and pay as they get service. Consumer grids are expected to lower the initial and maintenance costs of expansive resources that may be required for running computationally expensive applications. Real-time rendering for video games or interactive high-definition TV are possible applications for this type of grid. These applications also require high bandwidth similar to science grids but the network infrastructure has to support more interactive traffic, and delay must be kept low for supporting real-time applications.

Although the natures of these two usages of grid computing are different, optical networks provide a suitable infrastructure for both [4]. In the scientific grid, the amount of data transferred is very high but the frequency of these transfers is low. On the other hand, the duration of data transfers is much smaller in the consumer grid but the interactions among entities are more frequent. For that reason, both applications require high bandwidth, which can be provided by optical networks.

Although optical networking is suitable for both applications, the switching method to be used in the optical grid has to be selected based on the specific grid application. For example, wavelength switching is more appropriate for data transfers that are long-lasting and whose bandwidth demands are not fluctuating. Meanwhile, optical burst switching (OBS) performs better when the data traffic is short-lived and dynamic [5]. Hence, wavelength switching is more suitable for science grids whereas OBS is more suitable for consumer grids.

Even though OBS is suitable for consumer grids, burst losses have to be considered when running the grid over OBS networks. In the OBS protocol, a control packet is sent before the data burst in order to reserve network resources. The data burst is sent after a predetermined duration without waiting for an acknowledgment so that the delay is kept at a minimum. Because of this one-way reservation mechanism, a data burst can be lost if its control packet could not reserve network resources. Lost bursts carrying grid jobs need to be retransmitted, resulting in an increase in the completion times of grid jobs.

The consumers in a grid environment have flexibility in both resource selection and network path selection. Since the consumers in a grid environment can request service from various providers,

* Corresponding author.
E-mail addresses: kmehmet@ee.bilkent.edu.tr (M. Koseoglu), ezhan@ee.bilkent.edu.tr (E. Karasan).

0167-739X/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.future.2009.11.002

they can perform resource and network scheduling jointly to improve performance. There are several proposals describing how joint scheduling can be done for optical grids running over wavelength routed networks [6,7]. Our study investigates the joint resource and network scheduling problem in OBS grid networks. In the OBS grid, if the paths to a resource which offers a short processing delay are severely congested, a consumer may decide to select another resource with slower processing but with less congested paths. Beside resource selection, a consumer also has the option to select among the paths destined to the resource to send the grid job with less delay. We show that joint selection of computational resources and network paths reduces grid job completion times significantly by lowering loss rates when the transmission times are comparable with the grid computation times, which is the case for consumer grids.

In this paper, we first propose a joint resource and path selection algorithm and it is shown to outperform other algorithms that make resource and path selection separately. We then extend this scheme with the adaptive extra offset mechanism which adaptively selects the offset of the grid bursts depending on the congestion in the network. The extra offset mechanism for OBS bursts decreases the loss probability of bursts at the expense of increasing the latency. Since the completion time of an OBS grid job is a function of both loss probability and delay, reducing the loss probability of a burst using extra offset requires fewer retransmissions, so the average completion time may decrease although the transmission delay is increased by the extra offset. The proposed mechanism finds the optimum extra offset which minimizes the average completion time by exploiting this tradeoff between the delay and loss probability.

We improve the previous version of this work [8] by extending the numerical results and by providing several mathematical analyses. The numerical results are extended to include the effects of changing best-effort traffic load and burstiness as well as the effects of computational resource parameters on the average completion time. In addition to these, the control plane load created by the proposed mechanism is analyzed. We also provide a mathematical method for calculating the optimum extra offset.

The remainder of the paper is structured as follows. In the next section, we discuss how the OBS grid architecture addresses latency problems with the contemporary grid architecture. In Section 3, the optical grid architectures in the literature are presented. Analysis of a grid job lifetime is given in Section 4. In Sections 5 and 6, we present the proposed algorithms, joint resource-path selection and adaptive offset based QoS mechanism, respectively. The extra OBS control plane load generated by the proposed algorithms is analyzed in Section 7. Performance evaluation results for the algorithms are presented in Section 8.

2. The need for a low latency grid architecture

Since current grid implementations are generally designed for long-lasting jobs, the overhead caused by resource and network scheduling does not significantly affect the whole duration of the computation. However, as grid computing is starting to be used by highly interactive applications where the job lengths are shorter, these overheads become significant. For that reason, grid computing models have to be revisited to perform scheduling operations faster.

There are two types of models proposed for scheduling network and grid resources in wavelength switched grids [9]. The first is the overlay model, where applications request resource scheduling from the grid middleware. Once resource allocation is completed, applications ask for network connectivity to the selected resources. Network provisioning is also performed by the grid middleware, which communicates with the optical control plane to reserve end-to-end lightpaths between the user and the resource.

The second solution is to establish a unified control plane for both grid and network resource provisioning. A network layer protocol such as GMPLS [10] can be used for grid resource reservation as well as network provisioning. Delays caused by the middleware can be reduced to some extent by integrating the network and resource reservations. However, since wavelength switching requires connection set-up, this approach is still not suitable for small data transfers.

The OBS grid architecture is proposed for addressing the delay problems in the current grid practices and it makes the grid more suitable for small-sized real-time jobs. Currently, grid resource allocation is done in a centralized manner. However, centralized allocation is not feasible for an increasing number of users with highly dynamic requests [3]. For that reason, OBS grid architecture enables a distributed way of resource reservation where users interact directly with resources. In addition to distributed resource reservation, low latency data transmission provided by the OBS protocol makes this architecture suitable for dynamic jobs.

3. OBS grid architectures

In this section, the consumer grid architecture based on OBS is discussed. OBS offers sub-wavelength granularity for optical networks [11]. In OBS, a burst control packet (BCP) is sent before the optical data to configure switches between the source and the destination. The optical burst is delayed at the source node for an offset time waiting for the control packet to configure an all-optical path and the BCP is sent without waiting for an acknowledgment. If the BCP fails to find an available wavelength for a link, the optical burst is dropped at that node.

OBS performs better than wavelength switching for grids with dynamic and short-lived data transmission requirements. Since each lightpath has a bandwidth of one wavelength, the granularity of wavelength switching is very coarse. Besides, setting up a lightpath takes in the order of hundred milliseconds so it is not suitable for short-lived connections. As shown in [5], OBS performs better than wavelength switching for grids with scarce wavelength resources.

There are several proposals for consumer grid architectures running over OBS networks [12,13]. In a consumer grid, jobs are small in contrast to science grids, and assigning a single job to multiple resources increases the communication time overhead. For that reason, a grid networking architecture where each job is assigned to a single resource is considered, as in [14]. In the OBS grid, an optical burst carries a single grid job. The BCP carries the grid job information which is sent to the network without a specific destination address (anycasting). The BCP is routed to a suitable grid resource by intelligent routers which have grid layer information in addition to network layer information. After the offset time, the burst containing the grid job is sent to the network without waiting for an acknowledgment.

Instead of leaving the resource selection to intelligent routers in the network, another reservation mechanism is presented in [13], where the resource provisioning and wavelength reservation are performed in two separate steps. In this mechanism, which is called explicit reservation, a BCP containing the job description is sent to the network for resource discovery. After receiving these discovery probes, intermediate routers perform resource discovery and send acknowledgments to the consumer if they find a suitable resource. Fully aware of all options, the consumer selects a resource and sends the job as an optical burst to the selected resource.

In our work, we use explicit reservation architecture because we find this architecture more practical. In the case of anycasting, the routers need to perform on-the-fly routing based on resource states, and resource information has to be distributed to routers in the network. This approach significantly increases the complexity of routers and the cost of the infrastructure. Instead, we use a