Chapter 16

A Mathematical Analysis of a Disaster Management Data-Grid Push Service

Nik Bessis
University of Bedfordshire, UK

Antony Brown
University of Bedfordshire, UK

Eleana Asimakopoulou
University of Bedfordshire, UK

ABSTRACT

Much work is under way within the Grid technology community on issues associated with the development of services fostering the integration and exploitation of multiple autonomous, distributed data sources through a seamless and flexible virtualized interface. These developments involve fluid and dynamic, ad hoc based interactions between dispersed service providers and consumers. However, several obstacles arise in the design and implementation of such services. In this chapter, the authors examine a notable obstacle, namely how to keep service consumers informed of relevant changes about data committed in multiple and distributed service provider levels, and most importantly, when these changes can affect others’ well-being. To achieve this, the authors use aggregated case scenarios to demonstrate the need for a data-Grid push service in a disaster management situation. In this regard, the chapter describes in detail the service architecture, as well as its mathematical analysis for keeping interested stakeholders informed automatically about relevant and critical data changes.

INTRODUCTION

Data integration has long been discussed in other literature reviews. Many concerns have been encountered, as most of the datasets addressed by individual applications are very often heterogeneous and geographically distributed. Hence, the ability to make data stores interoperable remains a crucial factor for the development of these types of systems (Wohrer et al., 2004). Clearly, one of the challenges for such facilitation is that of data integration; these challenges have been widely discussed (Calvanese et al., 1998; Reinoso et al., 2008). Moreover, Foster et al. (2001) explain that
the combination of large dataset size, geographic distribution of users and resources, and computationally intensive analysis results in complex and stringent performance demands that, until recently, have not been satisfied by any existing computational and data management infrastructure. Recent advances in computer networking and digital resource integration resulted in the concept of Grid technology. In particular, Grid computing addresses the issue of collaboration, data and resource sharing (Kodeboyina, 2004). It has been described as the infrastructure and set of protocols to enable the integrated, collaborative use of distributed heterogeneous resources including high-end computers, networks, databases, and scientific instruments owned and managed by multiple organizations, referred to as Virtual Organizations (Foster, 2002). A Virtual Organization (VO) is formed when different organizations come together to share resources and collaborate in order to achieve a common goal (Foster et al., 2002).

The need to integrate databases into the Grid has also been recognized (Nieto-Santisteban, 2004) in order to support science and business database applications (Antonioletti et al., 2005). Significant effort has gone into defining requirements, protocols and implementing the OGSA-DAIS (Open Grid Services Architecture – Data, Access and Integration Services) specification as the means for users to develop relevant data Grids to conveniently control the sharing, accessing and management of large amounts of distributed data in Grid environments (Antonioletti et al., 2005; Atkinson et al., 2003). Ideally, OGSA-DAIS as a data integration specification aims to allow users to specify ‘what’ information is needed without having to provide detailed instructions on ‘how’ or ‘from where’ to obtain the information (Reinoso Castillo et al., 2004).

On the other hand, working with obsolete data yields to an information gap that in turn may well compromise decision-making. Bessis (2009) and Bessis and Asimakopoulou (2008) explain that it is value creation for collaborators to automatically stay informed of data that may change over time. Repeatedly searching data sources for the latest relevant information on a specific topic of interest can be both time-consuming and frustrating. A set of technologies collectively referred to as ‘Push’, ‘NetCasting’ or ‘WebCasting’ was introduced in late 90s. This set of technologies allowed the automation of search and retrieval functions. Ten years on, Web Services have overtaken most of Push technology functionality and become a standard supporting recent developments in Grid computing with state-of-the-art technology for data and resource integration.

However, if a Grid is a system to enable flexible, secure, coordinated resource sharing among dynamic collections of individuals, institutions and resources (Foster, 2002) then it should be all about designing a dynamic service that is inherent in a VO (Weishaupl & Schikuta, 2004). That is to say, updates within a distributed data environment are much more frequent and can happen from within any data source in the network. Hence, there is a need for updates to be migrated to other sites in the network so that all the copies of the latest, relevant and up-to-date data are synchronized and communicated to maintain a consistency and homogeneity across the VO. Several authors have highlighted the need from different viewpoints, including Foster (2002), Bessis (2003), Magowan (2003), Raman (2003), Watson (2003), and Venugopal et al. (2005). On this basis, OGSA-DAIS as a data integration specification should ideally address the ability to allow users to specify ‘what’ information is needed without having to provide detailed instructions on ‘how’ or ‘from where’ to obtain the information, as well as to automatically ‘keep’ users ‘informed’ of latest, relevant, specific changes about data in a single or multiple autonomous distributed database(s) and/or data source(s) that are registered within the VO. The requirement is widely regarded as a highly important service for individual and collaborative decision-making, as it will sustain
Related Content

Dynamic Task Assignment with Load Balancing in Cloud Platform
www.igi-global.com/chapter/dynamic-task-assignment-with-load-balancing-in-cloud-platform/139853?camid=4v1a

Grid INFN Virtual Laboratory for Dissemination Activities (GILDA)
www.igi-global.com/chapter/grid-infn-virtual-laboratory-dissemination/20526?camid=4v1a

Service Providers Indexing Using P2P Systems
www.igi-global.com/chapter/service-providers-indexing-using-p2p/40831?camid=4v1a

Replica Placement Strategy for Data Grid Environment
Mohammed K. Madi, Yuhanis Yusof and Suhaidi Hassan (2013). International Journal of Grid and High Performance Computing (pp. 70-81).
www.igi-global.com/article/replica-placement-strategy-for-data-grid-environment/78736?camid=4v1a