Chapter 4
A Federated Approach to Information Management in Grids

Mehmet S. Aktas
TUBITAK-Marmara Research Center, Turkey

Geoffrey C. Fox
Indiana University, USA

Marlon Pierce
Indiana University, USA

ABSTRACT

We propose a novel approach to managing information in grids. The proposed approach is an add-on information system that provides unification and federation of grid information services. The system interacts with local information services and assembles their metadata instances under one hybrid architecture to provide a common query/publish interface to different kinds of metadata. The system also supports interoperability of major grid information services by providing federated information management. We present the semantics and architectural design for this system. We introduce a prototype implementation and present its evaluation. As the results indicate, the proposed system achieves unification and federation of custom implementations of grid information services with negligible processing overheads.

1. INTRODUCTION

Independent Grid projects have developed their own solutions to problems associated with Information Services. These solutions target vastly different systems and address diverse sets of requirements (Zanikolas et al., 2005). For example, large-scale Grid applications require management of large amounts of relatively slow and varying metadata, while others such as e-Science Grid applications dynamically assemble modest numbers of distributed services and are designed for specific tasks, tasks that can be as diverse as forecasting earthquakes (Aktas et al., 2004) or managing
audiovisual collaboration sessions (Wu, 2005). These dynamic Grid/Web service collections require specific support for dynamic metadata.

Existing solutions to Grid Information Services present some challenges for metadata services: First, independent Grid applications use customized implementations of Grid Information Services, whose data model and communication language is application specific (Zanikolas et al., 2005). These information services are in need of greater interoperability to enable communication between different grid projects so that they can share and utilize each other’s resources (OGF-GIN, n.d.). Second, previous solutions do not address metadata management requirements of most Grid applications that have both large-scale, static and small-scale, highly dynamic metadata associated with Grid/Web Services (Zanikolas et al., 2005). Third, existing solutions do not provide uniform interfaces for publishing and discovery of both dynamically generated and static information (Zanikolas et al., 2005). The lack of a uniform interface limits clients, who must interact with more than one metadata service. In turn, this necessity increases the complexity of clients and creates fat clients. We therefore see the existing solutions of Grid Information Services as an important area of investigation.

To address these challenges, an ideal Grid Information Service Architecture should meet the following requirements: uniformity: the architecture should support one-to-many information services and their communication protocols; federation: the architecture should present a federation capability where different information services can interoperate with each other; interoperability: the architecture should be compatible with widely used, existing Grid/Web Service standards; performance: the architecture should search/access/store metadata with negligible processing overheads; persistency: the architecture should back-up metadata without degradation of the system performance; and fault tolerance: the architecture should achieve distribution and redundancy of information.

We have previously investigated the design, implementation, and evaluation of two specific data-systems: UDDI XML Metadata Service and WS-Context XML Metadata Service (Aktas et al., 2008a). We designed, implemented, and evaluated centralized versions of these metadata-systems and applied them to different application domains, such as geographical information systems, sensor grids (Aktas et al., 2004), and collaboration grids (Wu, 2005). However, these systems did not fully meet the aforementioned metadata management requirements of these application use domains.

We propose a Hybrid Grid Information Service called Hybrid Service that addresses the challenges of announcing and discovering resources in Grids, as seen in previous work and that improves our own previous work by addressing complete metadata management requirements of a number of application use domains.

In this study, we present the semantics and architectural design of the centralized Hybrid Service. We introduce a prototype implementation of this architecture and present its performance evaluation. As the main focus of this paper is information federation in Grid Information Services, we discuss unification, federation, interoperability, and performance aspects and leave out distribution and fault-tolerance aspects of the system. The main novelty of this study is that it describes an architecture, implementation, and evaluation of a Hybrid Grid Information Service that supports both distributed and centralized paradigms and manages both dynamic, small-scale and quasi-static, large-scale metadata. This novel approach unifies custom implementations of Grid Information Services to provide a common access interface to different kinds of metadata. It also provides federation of information among the Grid Information Services, so that they can share or exchange metadata with each other. This study should inspire the design of other information