Available online at www.sciencedirect.com
SCIENCE DIRECT®
Future Generation Computer Systems 21 (2005) 559–571
FGCS FUTURE GENERATION COMPUTER SYSTEMS
www.elsevier.com/locate/fgcs

GridICE: a monitoring service for Grid systems

Sergio Andreozzi[a,*], Natascia De Bortoli[b], Sergio Fantinel[c], Antonia Ghiselli[a], Gian Luca Rubini[a], Gennaro Tortone[b], Maria Cristina Vistoli[a]

[a] *INFN, CNAF, Viale Berti Pichat 6/2, 40126 Bologna, Italy*
[b] *INFN, Napoli, Via Cintia ed. G, 80126 Napoli, Italy*
[c] *INFN, Lab. Naz. di Legnaro, Via Romea 4, 35020 Legnaro, Italy*

Available online 7 December 2004

Abstract

Grid systems follow a new paradigm of distributed computing that enables the coordination of resources and services that are not subject to centralized control, can dynamically join and leave virtual pools, and are assigned to users by means of an explicit assignment functionality. The monitoring of a Grid is a multi-institutional and Virtual Organization (VO)-oriented service. It must deal with the dynamics, diversity, and geographical distribution of the resources available to Virtual Organizations, and the various levels of abstraction for modeling them. This paper presents the requirements, architecture and implementation of GridICE, a monitoring service for Grid systems. The suitability of this tool in real-life scenarios is analyzed and discussed.

Keywords: Grid computing; Distributed systems; Grid monitoring

1. Introduction

The advent of high-speed networks enabled the definition of new paradigms for the coordination of geographically dispersed resources. In this context, Grid systems offer distributed computing capabilities on resources not subject to centralized control, heterogeneous, and provided by different organizations. These resources and services are grouped into virtual pools that are accessible to users with the right credentials.

According to Németh and Sunderam [1], the key properties that make Grid systems different from conventional distributed systems are: (1) computational and storage resources are abstracted and form virtual pools; (2) the users who require access to resources of the virtual pools are different from the users who have valid accounts and login rights to physical resources; (3) mapping functions are available for entities translating from virtual to physical resources and users; (4)

* Corresponding author.
E-mail addresses: sergio.andreozzi@cnaf.infn.it (S. Andreozzi), natascia.debortoli@na.infn.it (N. De Bortoli), sergio.fantinel@lnl.infn.it (S. Fantinel), antonia.ghiselli@cnaf.infn.it (A. Ghiselli), gianluca.rubini@cnaf.infn.it (G. Luca Rubini), gennaro.tortone@na.infn.it (G. Tortone), cristina.vistoli@cnaf.infn.it (M. Cristina Vistoli).

0167-739X/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.future.2004.10.005

the number of resources in the pool is in the order of thousands or greater.

At the virtual level, users form Virtual Organizations (VOs). Even though this concept is widely used in the Grid community, there is still no agreement on a common definition in the literature. In this paper, we adopt Travica's definition: "a Virtual Organization is a new organizational form which manifests itself as a temporary or permanent collection of geographically dispersed individuals, groups or organizational units, either belonging or not belonging to the same organization, or entire organizations that depend on electronic links in order to complete the production process" [2].

These characteristics of a Grid system must be taken into account when designing a multi-institutional and VO-oriented service. One aspect of particular importance is Grid monitoring, that is, the activity of measuring significant Grid resource-related parameters to analyze usage, behavior, and performance of a Grid system, and to detect and notify fault situations, contract violations, and user-defined events. This is an open area of research where several research teams attempt to define requirements and designs for working solutions [3].

In this paper, a number of requirements are identified while exploring three real-life scenarios (Section 2). They are evaluated against existing tools and the need for a new solution is substantiated (Section 3). Then, GridICE is presented as a monitoring service architecture satisfying these requirements (Section 4). Particular attention has been paid to the different categories of consumers of monitoring information. Aggregation dimensions such as the Grid Operations Center (GOC), the Virtual Organization, and the Site have been considered in the design of the proposed system. Next, implementation details (Section 5) and experience results are reported (Section 6). Finally, conclusions are presented together with directions for future work (Section 7).

2. Use cases and requirements

The characteristics of Grid systems raise new requirements for a monitoring service. In order to identify them, three use cases are presented here. They refer to three different levels of abstraction in a Grid system: the VO level, the site level, and the operations domain level. On the basis of these use cases, a set of requirements is inferred; they form the basis for the design of the presented work.

The first use case is the usage of a Grid system from the VO viewpoint. Virtual Organizations consider Grid systems an appealing solution for sharing resources among geographically dispersed members. They make agreements with organizations providing resources and create virtual pools. Their main requirements are fourfold: (1) visualize at various aggregation levels the actual set of resources accessible to its members; (2) assess how Grid mapping functionalities from virtual to physical resources and users meet the members' demands; (3) evaluate how the global service offered by resource providers fulfills the service level agreement; (4) analyze data retrospectively to understand how to improve the effectiveness of VO applications running in a Grid, as the target machine for different executions of the same application can vary over time.

The second use case concerns local site administration. Organizations participating in a Grid may have different characteristics. Some of them may contribute a large set of hardware resources monitored by their own systems and managed by highly qualified staff working in 24×7. They may want to activate Grid monitoring only on their control and access machines. Other organizations may offer a limited set of machines and a limited number of people devoted to local control and management. These people may not always have the right skills or enough time for managing the local resources based on Grid middleware. They should have the ability to perform fine-grained monitoring activities without having to set up and manage local monitoring systems. A Grid monitoring service should be able to support both situations in order to collect the data of interest at the Grid level.

The third use case concerns the operations domain of a Grid. In production Grid environments, an organization needs to be formally in charge of coordinating and monitoring the operation of a Grid infrastructure, devising and managing mechanisms that encourage its optimal operation, offering support to Grid users, and interacting with local support groups to help them to provide the best possible service while their equipment is connected to a Grid.

In this paper, this organization is called the Grid Operations Center. It is responsible for an entire Grid infrastructure, e.g., a national Grid or the Large Hadron