Computer Systems Group
Institute of Computer Science
AGH-UST

Knowledge supported data access optimization for grid environments

Darin Nikolow[1], Renata Słota[1], Jacek Kitowski[1,2]

[1] Institute of Computer Science, AGH-UST, Cracow, Poland

[2] Academic Computer Centre CYFRONET-AGH, Cracow, Poland

Introduction

As the grid technologies get more matured and widely adopted demands for new functionalities appear. Clients are expecting that they will be served with the desired level of quality of service negotiated with the server. Efficient management of data storage resources in SLA (Service Level Agreement) respecting system is not a trivial task. Extended monitoring of the SLA fulfillment is needed in order to proactively prevent the lost of performance for the given client request. The monitoring of SLA fulfillment may not be possible without some low level storage monitoring services.

In this paper our idea of using services for monitoring and data access time estimation for data access optimization and the supporting of SLA in grid environment will be presented. These services could be used by the data management middleware and especially by the data replication module to effectively manage the storage resources fulfilling the performance and cost goals presented in the SLA.

The problem

There are many different storage systems available on the market today. Providing monitoring data for all kind of storage systems used in a given environment requires knowledge of details which are specific for the given storage system. The obvious approach is to have specialized storage monitoring services for every storage system used. From the other hand storage systems have some common or similar functionalities and characteristics which can be exploited by a more general service. The more intelligent approach proposed in the paper is creating a knowledge base addressing monitoring storage systems and then using this knowledge to create workflow which will gather and provide the necessary performance parameters of a given storage system.

Since the storage system monitoring service in production environment should be available at any time a functionality of non-disruptive updating of the low level services and the knowledge base should be present. Another problem which should be addressed in any monitoring system is the impact of the monitoring itself on the actual performance of the storage system. Special care should be taken to minimize the influence of the storage monitoring service on the performance of the storage system.

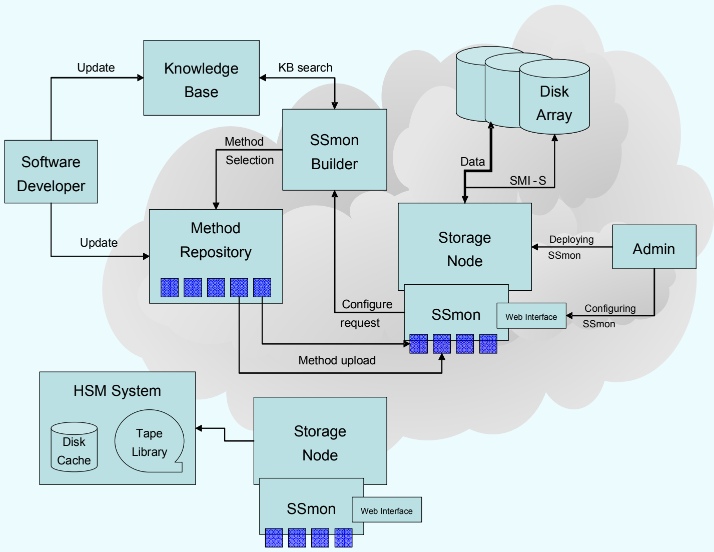
System Architecture

The conceptual diagram is shown on the right. The system consists of the following modules:

- KB - knowledge base. It keeps the ontology descriptions of the available methods for storage systems monitoring.
- SSmon builder - this module is responsible for the automatic building and configuring of the SSmon service.
- MR - Method Repository. It contains methods for monitoring different parameters for different kinds of storage systems.
- SSmon service - this service is installed on every storage node being monitored.

The installation and configuration process of storage system monitoring facilities looks as follows: First the base SSmon service is deployed on the storage system which has to be monitored by the administrator. Next the administrator configures SSmon by specifying the identity of the storage system, its access points and methods and the storage system parameters which needs to be monitored. After that a configure request for the specified storage system's type, model and version is sent to the SSmon builder, which in turn does a KB search and select the most suitable service methods. Next, the selected methods are uploaded Ssmon. At this point Ssmon is ready to serve for monitoring requests from the other grid components.

When the developer team is updating the method repository or the knowledge base a notification is sent to the Ssmon services and new versions get uploaded on the fly if necessary.



State of the Art

The Service Negotiation and Acquisition Protocol (SNAP) [snap] has been designed to enable automatic negotiating of SLA. It mainly addresses task submission and resource usage used in traditional computational grids. Sahai at al. in [spec-sla] propose method for specyfying and monitoring SLA fullfillment in commercial grids.

In our previous work we have proposed a graybox based system for estimating data access time for HSM systems [ppam03-gray]. Some parts of this system can be also used in this project.

The SOKU (Service Oriented Knowledge Utilities) vision presented by the Next Generation Grids Expert Group identifies a semantic approach for creating and operating commercial and scientific grids [soku].

The Storage Management Initiative-Specification (SMI-S) is a standard that has been designed with the purpose of standardizing and streamlining storage management functions and features into a common set of tools that address the day-to-day tasks of the IT environment [smi-s]. SMI-S is based on the Common Information Model (CIM) which is an open standard designed to enable vendors to exchange semantically rich management information between systems throughout the network [cim].

Bibilography

[snap] Czajkowski, K., Foster, I., Kesselman, C., Sander, V., and Tuecke, S., "SNAP: A Protocol for Negotiating Service Level Agreements and Coordinating Resource Management in Distributed Systems", 8th Workshop on Job Scheduling Strategies for Parallel Processing, Edinburgh, Scotland, July 2002.

[spec-sla] Sahai, A., Graupner, Sven., Machiraju, V., Moorsel, A., "Specifying and Monitoring Guarantees in Commercial Grids through SLA", Proceedings of the 3rd IEEE/ACM International Symposium on Cluster Computing and the Grid (CCGRID'03), IEEE, 2003.

[ppam03-gray] Nikolow, D., Slota, R., Kitowski, J., "Gray Box Based Data Access Time Estimation for Tertiary Storage in Grid Environment", in: Wyrzykowski, R., Dongarra, J., Paprzycki, M., Waśniewski, J. (Eds.), Parallel Processing and Applied Mathematics. 5th International Conference, PPAM 2003, Czestochowa, Poland, September 2003, LNCS, no. 3019, Springer, 2004, pp. 182-188.

[soku] NGG Expert Group, "Future for European Grids: GRIDs and Service Oriented Knowledge Utilities - Vision and Research Directions 2010 and Beyond", http://www.semanticgrid.org/NGG3/

[smi-s] SMI-S: The Storage Management Initiative Specification, http://snia.org/forums/smi/tech_programs/smis_home/

[cim] DTMF Common Information Model (CIM), http://www.dmtf.org/standards/cim/

CGW'07 Cracow, Poland