Seminar May 16, 2003

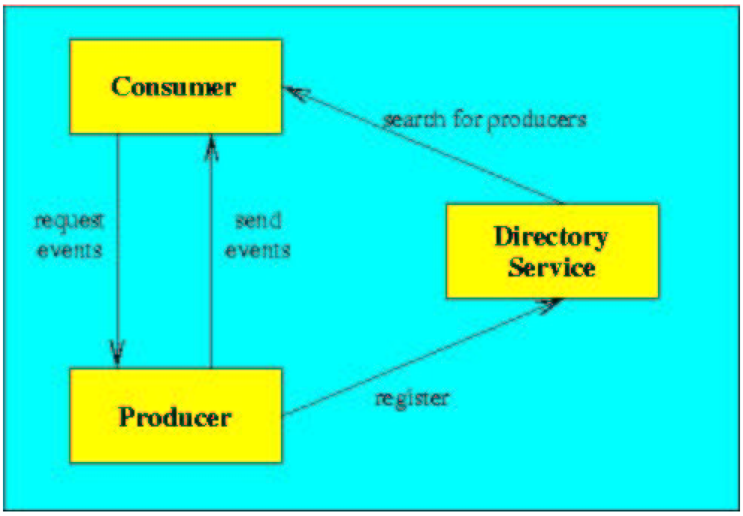
Resource Monitoring in GRID computing

Augusto Ciuffoletti
Dipartimento di Informatica - Univ. di Pisa

next: Network Monitoring Architecture

index

Network Monitoring Architecture



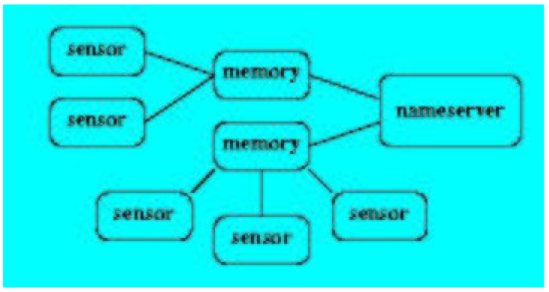
- controls a complex monitoring system;
- is fault tolerant;
- minimizes overhead (over monitored resource):
- makes available the results of the monitoring activities, also implementing security;
- integrates different monitoring tools;
- scales well with system size and probe frequency.

next: Network Weather Service

up: Seminar May 16, 2003

index

Network Weather Service



The components of a NWS resource monitoring system:

- a **nameserver**, a centralized controller, that keeps a registry of all components and monitoring activities;
- **sensors**, that produce resource observations;
- **memories**, that store resource observations;
- **forecasters**, that process resource observations;

Limits of the NWS architecture, bound to the presence of the **nameserver**:

- a communication bottleneck;
- a single point of failure;
- no security;

next: Store observations into LDAP directories

previous: Network Monitoring Architecture

up: Seminar May 16, 2003

index

Store observations into LDAP directories

One limit of NWS is centralized control: in order to obtain data, one has to address (possibly indirectly) the nameserver, specifiying the id of the sensor.

The distributed nature of the LDAP architecture is appealing to solve this problem: information is organized hierarchically, and the resulting tree (might be) distributed over different servers.

The hierachy often reflects the organization of the Grid: the first level under the root is composed of continental networks, below are national networks and next local networks. Leaves are single resources, like clusters of computers, single computers or storage elements.

The well known Globus toolkit is based on LDAP.

However, the LDAP archtecture is appropriate to store static information, like the number of processors in a cluster, or the size of a disk partition.

When data are more volatile, the LDAP architecture is the wrong choice: it is unsuitable to support frequent `write` operations.

Examples of characteristics that induce frequent `writes`:

- available computing power on a computing resource (e.g. nummber of idle nodes in a cluster);
- percent of free space in a partition;
- average roundtrip time on a link

In all these cases the LDAP architecture suffers serious performance and scalability problems.

next: Store observations in a SQL database

previous: Network Weather Service

up: Seminar May 16, 2003

index

Store observations in a SQL database

The SQL architecture is a distributed relational database. There are several excellent implementations of this architecture, that are designed to support extremely large databases.

They offer a good compromise between the distributedness, and the cost of `query` operations.

In particular the database can be replicated in order to improve scalability and fault tolerance, as long as the number of `queries` operations dominates the number of `writes`.

The **R-GMA** architecture is being developed to address this problem: the scalability of the architecture is further improved introducing components that combine data from the database and cache the results (similar to NWS forecasters).

A modular architecture for the Grid

In a SQL scenario, the structure of the Grid is not bound to the hierarchical tree induced by an LDAP tree.

Therefore we need to introduce a structure, that will be used to index the database and improve performance.

The case of resources that are bound to a node is quite different from the case of resources that represent communication between nodes: the former grow with the number of nodes, the latter with its square.

As a consequence, representing communication resources using end-to-end characteristics may limit the scalability of a monitoring architecture.

In addition, availability of end-to-end mesurements implies that:

- nodes that host Grid resources also support network monitoring protocols and
- each pair of nodes hosting Grid resources should generate network monitoring traffic, with an increment of network load that is a square of the nodes.

Therefore it would be appropriate to partition the nodes in the Grid so that the number of partitions is remarkably smaller than the number of nodes: we call **domains** these partitions.

The monitoring activity should be limited to communications between domains.

next: GlueDomains

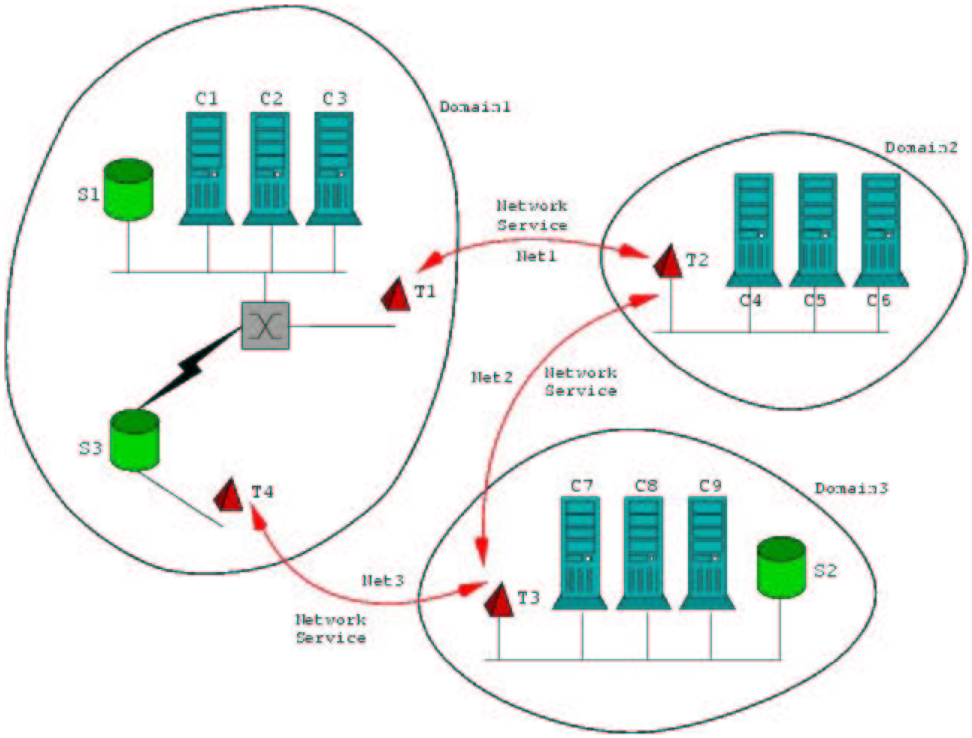
previous: Store observations in a SQL database

up: Seminar May 16, 2003

index

GlueDomains

Together with A.Ghiselli and C.Vistoli of CNAF-Bologna, we are currently implementing a prototype architecture for Grid partitioning, in the mainframe of the GLUE collaboration: we call it **GLUEDOMAINS**.



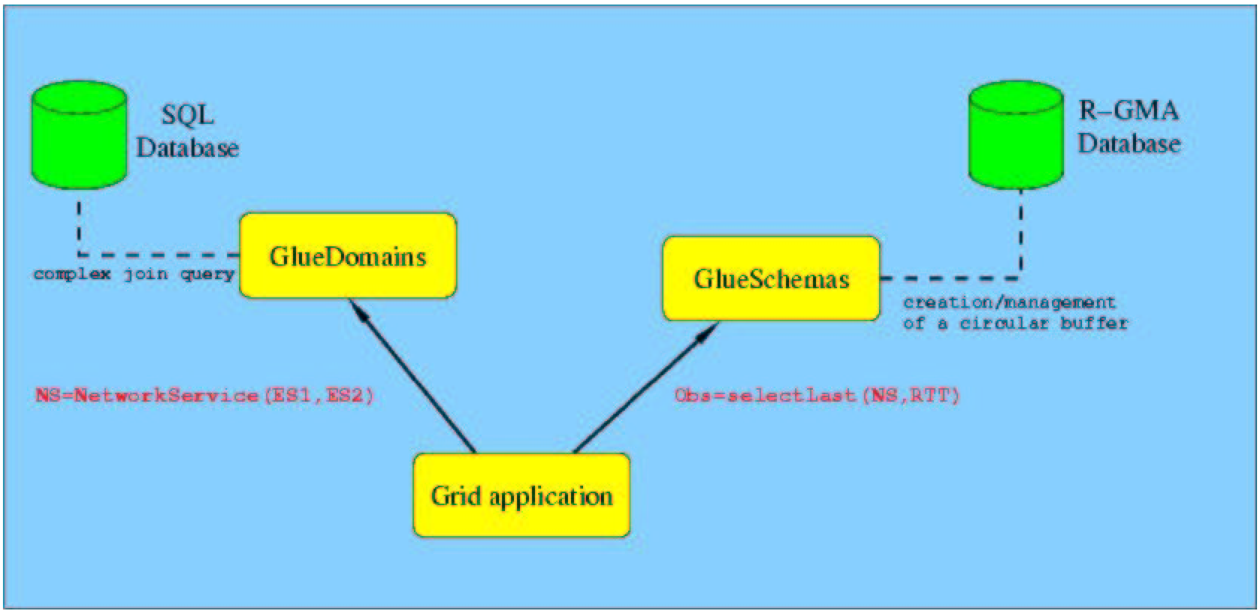
next: GlueDomains architecture

previous: A modular architecture for the Grid

up: Seminar May 16, 2003

index

GlueDomains architecture



next: Database contents

previous: GlueDomains

up: Seminar May 16, 2003

index

Database contents

The content of the R-GMA database, concerning Network Services, is represented by a collection of measuremnts, commonly called **observations**.

Observations fall into different categories, each representing a **characteristic** of the Network Service.

For instance, a characteristic is the **round trip delay** measured between two monitoring hosts.

There is presently much effort on designing a standard representation for these characteristics, much like has been done in the Internet with the MIB specifications.

This standardization is mandatory to guarantee the usability of database contents by developed separately applications: the GGF organizion works in this direction.

During last GGF meeting (number 7, Tokio), the working group about network monitoring schema definition was presented with the problem of characterizing wireless links and devices: the chair’s answer was at best vague.

There might be some work for a group tah focuses on wireless networks.

next: Indice

previous: GlueDomains architecture

up: Seminar May 16, 2003

index