Architecture of a Network Monitoring Element

Augusto Ciuffoletti\textsuperscript{1} and Michalis Polychronakis\textsuperscript{2}

\textsuperscript{1} CNAF-INFN, Bologna, Italy
\textsuperscript{2} FORTH-ICS, Heraklio, Greece

\textbf{Abstract.} A Network Monitoring system is a vital component of a Grid; however, its scalability is a challenge. We propose a network monitoring approach that combines passive monitoring, a domain oriented overlay network, and an attitude for demand driven monitoring sessions. In order to keep into account the demand for extreme scalability, we introduce a solution for two problems that are inherent to the proposed approach: security and group membership maintenance.

\textbf{Keywords:} network monitoring, passive network monitoring, on demand network monitoring, network monitoring element, scalability issues, security issues.

1 Introduction

Monitoring the network infrastructure of a Grid has a vital role in the management and the utilization of the Grid itself. The Global Grid Forum (GGF) schema \cite{GGF}, splits this activity into three distinct phases: \textit{production}, \textit{publication}, and \textit{utilization} of measurements. Here we focus on the production and publication, with a special concern for scalability: for measurement production we address the usage of passive network monitoring techniques, while for the publication activity we adopt a domain-oriented overlay network which reduces the complexity of the task.

The challenge comes from the fact that a Grid oriented network monitoring should address network routes, not single links, since this is the kind of information needed to optimize distributed applications. Since each pair of Grid Services should be individually monitored, this makes an $O(n^2)$ complexity for many aspects of network monitoring: from the size of the database containing the results, to the number of pings that probe the system.

We combined a number of ideas in order to limit the complexity of our solution: \textit{i}) monitoring shouldn’t address single Grid resources, but pools with similar connectivity; \textit{ii}) monitoring tools shouldn’t inject traffic, but observe existing traffic; \textit{iii}) monitoring activity should be tailored on application needs. Only the integration of above ideas can effectively control the problem size, and, in some sense, the first two open the way for the application of the third one.

We observe that a Grid topology is made of pools of resources reachable through dedicated ingress points: the accessibility of such pools depends on ingress points connectivity, and local administration avoids internal bottlenecks.
Therefore the monitoring topology can be simplified by monitoring *Network Elements* between ingress points of distinct pools.

One Network Monitoring architecture, called GlueDomains [3], has been recently designed and prototyped according to a two levels hierarchical overlay; the purpose of such experiment was mainly the assessment of a number of design principles. A Grid-wide deployment of GlueDomains was carried out during the summer of 2006, as part of the Italian branch of the Large Hadron Collider Computing Grid Project (LCG). Apart from the statistics collected (usual packet loss and roundtrip time, together with an experimental one way jitter measurement tool, published through the GridICE Grid Information Service [2]), the most relevant results from the GlueDomains experiment concern the ease of deployment, as well as the resilience, and stability of the architecture, which were assessed during a one month trial. GlueDomains is included in the current release of the Italian branch of LCG.

GlueDomains architecture centers around a number of specialized units hosting the agents in charge of monitoring the network. Such agents are able to autonomously (re)configure their activity based on a dynamic description of the network monitoring topology, available from a relational database. The monitoring activity was based on a domain partitioning of Grid resources: the target of such monitoring is the *Network Element*, which abstracts the network infrastructure in charge of interconnecting two domains.

One relevant lesson learned from GlueDomains experience is the identification of the role played by the agent that concentrates the network monitoring activity for a domain. This role corresponds to a new resource in the Grid architecture, which is mainly dedicated to network monitoring. In the architecture proposed in this paper we call such agent a *Network Monitoring (NM) Element*: its activity is organized into *Network Monitoring Sessions*.

Another cornerstone concept in our architecture is *passive monitoring*, which is non-intrusive by nature. The internal architecture of NM Elements adopts specific hardware and software solutions to address passive network monitoring.

A third concept that cuts down network monitoring complexity is an *application driven* configuration: this is feasible in a Grid, where applications negotiate computing resources with resource brokers, which can configure *Network Monitoring Sessions* on the fly, providing adequate credentials to NM Elements. The relevant conclusion is that, if network monitoring activity is bound to applications, it will increase linearly with system throughput, not with the square of system size.

A relevant aspect of an *application driven* approach is the interface that a NM Element should offer to the outside. Currently brokers find resource characteristics in the *Grid Information System* (GIS), automatically collected by *preconfigured* network monitoring sessions. This attitude is inappropriate in an *application driven* scenario for its limited scalability, and it would be preferable to connect NM Elements to brokers through a publish-subscribe system. Given that this aspect is still a research topic, we indicate a composite interface, which decouples the input, consisting of monitoring requests, from the output, consisting of observation records.