SensorML for Grid Sensor Networks

Giovanni Aloisio[††], Dario Conte[†,††], Cosimo Elefante[†], Italo Epicoco[††], Gian Paolo Marra[†,††], Giangiuseppe Mastrantonio[†] and Gianvito Quarta[†,††]
[†)] Institute of Atmospheric Sciences and Climate of the Italian National Research Council
via per Monteroni, 73100 Lecce – Italy
[††)] Center for Advanced Computational Technologies/ISUFI, University of Lecce,
at NNL/INFM&CNR Lecce, Italy
{d.conte, c.elefante, gp.marra, g.mastrantonio, g.quarta}@ isac.cnr.it, {giovanni.aloisio, dario.conte, gianpaolo.marra, gianvito.quarta }@unile.it

Abstract - *This paper describes an approach based on Globus toolkit for developing grid sensor networks. The key aspect is also represented on the use of a novel information service based on a relational data model, namely iGrid developed within the European GridLab project. iGrid is used to integrate sensor networks in Grid environments by means of the design of an information structure based on Sensor Modeling Language (SensorML). A case study is also presented in order to provide some relevant details about our approach and to verify the concrete applicability of the proposed methodology.*

Keywords: Grid Sensor Networks, Grid Computing Environment, SensorML.

1 Introduction

A sensor network [1] is a computer network of many spatially distributed devices using sensors to monitor conditions at different locations. Such distributed devices are typically used to monitor temperature, sound, vibrations, pressure, motion, pollutants and so on. Usually these devices are small and inexpensive, so that they can be produced and deployed in large numbers therefore, their resources in terms of energy, memory, computational speed and bandwidth are severely constrained. Due to the increased use of this technique in several applicative domains, the requirements, in terms of processing, data gathering, data storage and mining were rapidly increased. At the same time, Grid computing [2,3] has evolved as standard-based approach for heterogeneous and geographically spread resources integration and access. In fact, computational Grids are spreading as technology that allows the sharing and utilizing of computing resources, software, knowledge, scientific tools and so on, in an efficient and coordinated manner.

Several developments in Grid computing have focused on compute and data grids as suitable solutions for putting together computational power and provides a seamless access to very large amounts of storage resources. Another domain in which the grid paradigm is beginning to be used is the sensor networks. In fact, sensor grids extend the grid approach to the sharing of sensor resources in wired and wireless networks. As demonstrated by several works and mentioned in [4,5,6], sensor grids are suitable for improving the sensor network with regard to the following issues:

- sensors can be different in terms of providing information and its data format;
- sensors can be owned by different organizations which would like to share the sensors in a controlled manner;
- sensors can continually produce a very large amount of data;
- particular prominent information can be found if a powerful tool to query the sources is provided.

Grid approach promises to efficiently solve these issues providing a suitable technology through which large amounts of data can be collected, stored and processed, using a data grid from one side, and a computational grid from the other. Moreover, a very prominent feature of grid approach is: several Virtual Organizations (VOs) [7] can by easily defined, in order to provide fine grained access control to geographically spread sensors.
In this paper, we propose an approach based on iGrid information service to integrate sensors in grid environments. In fact, sensors can be considered as grid resources that provide specific information about some phenomena. The sensors are characterized by a strong heterogeneity, a variable state, are often geographically spread and can be shared among many organizations.
In this paper we also describe Ligris, a case study in which our approach used, to put a SODAR (SOnic Detection And Ranging) sensor into a grid environment.
In the next section, we describe some relevant related

works, thus, in section 3 we illustrate the designed approach and its implementative details. In section 4 we describe the case study and finally, in section 5, we conclude the paper.

2 Related Works

Hingne et al in [4] identifies some research issues and challenges for building infrastructures that enable computational abilities of Grid to answer queries on sensor data. In this work the tradeoffs relating to sensor accuracy consumption are addressed and experimental results are reported.

Hock et al [5] addresses several challenges in sensor grid designs. In fact, the authors propose SPRING, a proxy-based sensor grid architecture. They also developed a sensor grid test-bed which aims to study the design issues of sensor grids and to improve their sensor grid architectural design.

Other works deal with the definition of middleware architecture for sensor devices, able to facilitate the integration with the grid. Bramley et al [6] presents in fact, the Common Instruments Middleware Architecture (CIMA), which aims to enable instruments and sensors as real-time data sources and to facilitate their integration with the grid. The implementation was based on OGSA. In our work, we developed a SensorML interface to Globus toolkit, which aims to allow the complete abstraction from specific features of sensors. In fact, Sensor Model Language (SensorML) [8] proposed by Open Geospatial Consortium (OGC) [9], provides standard models and an XML encoding for describing any sensing process. This description includes the process of measurement by sensors, the instructions for deriving higher-level information from observations.

3 Sensor Networks Integration into the Grid

3.1 The Challenge

Integrating sensor networks into Grid environments are a problem and a challenge because different sensors are built using non standard protocols.

In order to provide a standard description of sensors, the OGC has recently provided the Sensor Model Language mentioned above. In fact, processes described in SensorML are discoverable and executable; indeed, all processes define their inputs, outputs, parameters, and methods, as well as provide relevant metadata. SensorML models are encoded in XML.

On the other side, the Open Grid Services Architecture (OGSA) [10] represents an evolution towards a Grid system architecture based on Web services concepts and technologies. The OGSA standards are the emerging API for grid applications. Web Services provide standard methods for discovery and invocation services. Web services (as well as web applications) use HTTP protocol for communication and they are thought and modeled as web applications that produce HTTP responses for HTTP requests, other web services provide SOAP responses to HTTP/SOAP requests (SOAP - Simple Object Access Protocol- is a protocol built over XML language, it allows to invoke web functions and to get the responses). Web services specification information are provided through their WSDL (Web Service Description Language) that codify their invocation modalities and related responses.

Since the release of the Globus Toolkit 3.0, the Globus Project offers an open source collection of Grid services [11] that follow OGSA architectural principles. The Globus Toolkit also offers a development environment for producing new Grid services that follow OGSA principles.

Consequently, taking into account the OGC approach which describes sensors from one side and OGSA specification from the other side, a good approach for integrate sensors into grid environments can exploit the Globus Toolkit and SensorML standards.

3.2 Resources description in Grid environments

In grid environments, the description of resources is an important aspect. The involved resources can be computing elements, software objects, data repository but also instruments and sensors. It is important to recall here the key role of information: a so called grid-aware application can not even be designed and implemented if information about the execution environment is lacking.

A well known approach for managing information can be found in the information system developed in the Globus project [12]. In Globus Toolkit version 2.x a LDAP based Information Service, called Metadata Directory Service (MDS-2) [13] was provided. MDS-2 provides a large distributed collection of generic information providers that extract information from local resources. The gathered information is structured in terms of a standard LDAP data model. MDS-2 also provides a set of higher level services that collect. manage and index information provided by one or more information providers.

Due to the non- satisfactory performances provided by this solution, a new version of Monitoring and Discovery Service (MDS-3) [14] was provided by Globus Toolkit version 3.x. MDS-3 and was developed by using lava and the Grid Services framework. Finally, in Globus Toolkit version 4 a new version (MDS-4) will be based on the emerging Web Service Resource Framework (WSRF) specification [15].

In each version of the Globus's MDS, the information schema can be easily extended and additional information providers can be developed in order to manage and publish extended or a new set of information. A complementary approach for information management is implemented in iGrid information service developed within the European GridLab project [20]. iGrid is based on a relational data model and presents a GSI-enabled web service interface by means of the gsi-plugin for gSOAP. iGrid has a hierarchical and distributed architecture that ensure a decentralized control and a distributed access to information related to grid resources, it is fault tolerant, scalable and easy to extend with new information. The relational data schema allow also a user to use the SQL language and its expressiveness to extract and retrieve any kind of information from the relationship among handled data.

3.3 Architecture

In our approach, we think about sensors in a network as distributed grid resources that provide information about one or more objects or phenomena. In fact, a sensor can be viewed as a hard implementation of such an information provider (see Figure l). Moreover, we can think about networks as a physical organization that can be composed in order to define new ones as VOs.

Considering the mentioned iGrid, it allows to solve the resources discovery and monitoring issues. In particular, with respect to sensors, the monitoring can be viewed as the process in which the information is extracted.
Moreover the sensors in one or several networks can be geographically distributed, can have a variable status, can be variably grouped.

For these reasons, we have designed an information structure able to contain all relevant information about sensors and the values observed by it. More in depth, during the design phase, we have dealt with the following aspects:

- we have chosen which information is relevant to publish;
- we have analyzed how the information must be represented in a relational data model;
- we have designed a standard interface that allows sensors to publish its data;
- we have addressed how the communication can be done.

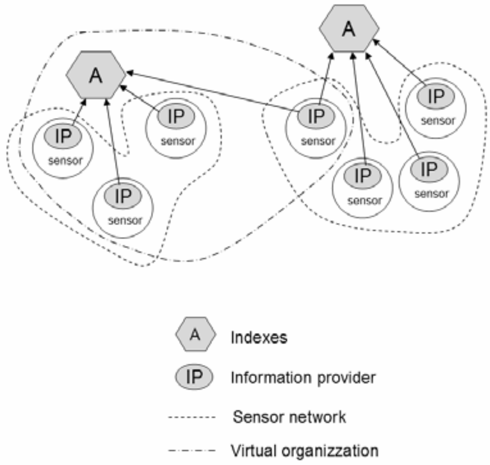

Figure 1 – Grid’s point of view of sensors networks

Thus, we have defined, according to SensorML guidelines, an *objectclass* named *GridSensor* that contains the following information:

- *SENSOR-ID:* contains the sensor’s unique id;
- *SENSOR-HOSTNAME:* describes the hostname of the sensor or the host name of resource in which it is installed;
- *SENSOR-LOCATION:* contains the geographic coordinate of the sensor on Earth;
- *SENSOR-STATE:* describes the sensor state and can contain one of these values:
    - *ready*
    - *busy*
    - *uninstalled*
    - *unknown*
- *SENSOR-TYPE:* describes the kind of sensor and the kind of supplied information;
- *SENSORML-VALUE:* contains the instantaneous value (or values) of observed phenomena.

This object is modeled by an XML schema that will be published through an ad-hoc provider (outputting XML), written in a language supported by the sensor.

Security issues are also addressed by iGrid and in particular by the Globus Security Infrastructure (GSI) [16] that support a wide range of access control policies. We assume that an information provider may specify the credentials that must be presented to access that information. GSI public-key security mechanisms are used to verify credentials and to achieve mutual authentication between information consumers and sensor.

Finally, the front-end was provided by a web application.

3.4 Implementation Details

As previously described we adopted iGrid as information service and in particular, we have extended the relational information schema in order to add all the above mentioned information related to sensors. The iGrid information schema has been extended with the following table:

Table 1

Table Name: GridSensor		
Field Name	Type	Dimension
Sensor-ID	Integer	
Sensor-Hostname	String	255
Sensor-Location	Real	
Sensor-State	Enum	“ready”, “busy”, “uninstalled”, “unknown”
Sensor-Type	String	255
SensorML-Values	Integer	Foreign key to the Sensor-Values table

Moreover, in the *igrid.conf* file we have specified how the sensor's information provider can be contacted and the data supplied by itself.

The implementation of each information provider depends on the hardware and software architecture of the device in which it must be installed. In the next section, we describe a case study in which we have developed a simple provider implemented through a shell script.

Finally, we have developed a web application as a front-end. This application, developed with PHP, allows users to easily access the needed information, to manage the resources with respect to VOs and to sensor specific configurations. The web application was hosted by Apache 2 web server.

4 The Ligris Sensor: a Case Study

4.1 The atmospheric boundary layer

The dynamics or in general the phenomena which take place in the atmospheric boundary layer (ABL) are very complex because of the great number of variables, such as the feedback processes, the highly not linear interactions between different scale motions and, above all, the turbulence that characterizes the exchange processes in this part of the atmosphere.

The ABL plays a crucial role for several reasons: it is the layer in which energy (in the forms of, water vapour, heat and momentum) is transferred from the surface to the atmosphere and vice versa; it is also a place in which a large amount of human activity residuals are released.

Knowledge of the ABL structure is therefore basic for studying the atmospheric dynamics and the transport of pollutants and its diffusion in the atmosphere.

To study these topics, many different measures are needed which are usually gathered by sensors that are based on different technologies and principles. It is often necessary to assimilate these measures into mathematical models in order to be able to describe the phenomenon and its evolution. Therefore numerical models may need to automatically access the sensors to get the data.

4.2 The SODAR

The SODAR (SOnic Detection And Ranging) is a powerful tool for investigating the phenomena that takes place in the ABL. It provides a “snapshot” of the thermal structure of the ABL by displaying the echo intensity in facsimile format: in this format each pulse produces a vertical trace whose darkness varies with the echo strength from a particular height in the atmosphere, a height proportional to the ordinate on the chart. Since the traces from successive pulses are closely spaced along the abscissa, the eye can easily integrate corresponding features and detect the evolution of the atmospheric structure with time. Moreover, through the harmonic analysis of the signal received, it is possible to derive the radial wind profile, usually in a range of 1000 m, and to visualize the dynamic evolution of the airflow crossing the antenna beams. The SODAR relies, in its monostatic configuration (scattering angle = 180°), on backscatter from a field of turbulence-driven temperature fluctuations. If three antennas are used, one pointed vertically and the others tilted at right angles from one another one, the SODAR can map the entire three-dimensional wind field at all heights within the echosonde range.

4.3 The Ligris Sensors Network

Usually in a sensors network there are sensors to monitor conditions at different locations: these devices are typically small and inexpensive. Instead, in this case study, the type of sensor used is the SODAR developed by Mastrantonio, Fiocco and Argentini [17,18]. This powerful tool for investigating ABL is large in dimensions and expensive, but however, it is a sensor, which if integrated into a network, would permit the measuring of the atmosphere's samples distributed on the territory. Such integration would in fact bring an elevated scientific value, creating a presupposition for a vast scale study of the ABL. The implementation of the proposed approach realized in this case study is denominated LiGriS Network (Linux Grid Sodar). It is composed of a network of sensors of which, each is formed by the following components (see figure 2):

- Three antennas for the transmission and reception of acoustic signals;
- A Personal Computer (with operative system GNU Linux Debian) equipped with acquisition card Daqboard 2001 IOTECH (containing 16 analogical inputs, 4 digital output and 40 I/O digital lines);
- Globus Toolkit 3.2 preogsi, with active configuration of GRIS (Grid Resource Information Service) for this particular use;
- Software developed ad-hoc for the management and control of the SODAR;
- Information Provider capable of interfacing with the controlling and managing software and providing the reading in accordance with the object GridSensor.

In each sensor the software for the SODAR is developed in C language exploiting the Daqboard's API independently and allows the management of each acquisition channel. The control software of the SODAR directly interfaces with the Information Provider (IP). Instead, the IP was realized through script shell (specifically GNU Bourne Again Shell). In particular this provider receives the parameters at the start of the acquisition and returns the information acquired. The initialization phase is necessary because the sensor works in function of some parameters, such as the frequency of the signal to send, the duration of the acquisition and the frequency pattern of the echo received.

The realized network is composed by three nodes of equal characteristic (*{cassandra, icaro, icaro1}.le.isac.cnr.it*) organized in a single VO (*le.isac.cnr.it*). The GRIS service is running in each node. Instead, the GIIS service is active in only one node (*cassandra.le.isac.cnr.it).*

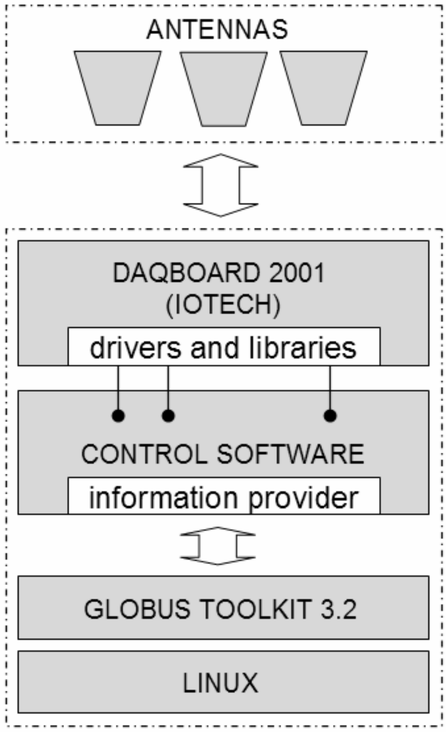

Figure 2 – Sensor's architecture in Ligris

In order to carry out the test on the behavior of the realized network, a Web Application was implemented, which permits controlled access to the information gathered from sensors.

In detail, the Web Application allows to call some scripts, which permit the execution of the basic operations, such as: searching, reading, contacting and so on. For example, the search script is based on *grid-info-search* command as:

/grid-info-search -x -b 'Mds-Vo-name=le.isac.cnr.it, o=Grid' objectclass=GridSensor | grep GridSensor | awk '{printf ("%s\n", $2)}'| sed "/^uninstalled/d" > /sensorml/tmp/sensorhost.txt

Therefore, a My-Proxy Server [19] was utilized as a container of the proxy in which the Web Application makes preference to obtain valid user's credentials.

5 Conclusions and Future Works

In this paper, the authors presented the extensions made on the iGrid Information Service for supporting the grid sensors network information. iGrid information schema has been easily extended in order to support different information sources and able to handle information related to sensors by means of an ad-hoc information provider. Moreover, the Sensor Modeling Language (SensorML) proposed by OGC has been also considered, in order to implement a standard description of sensors.

The proposed approach has been implemented in a case study in which small SODAR network has been integrated into a grid environment (the Ligris Sensors Network).

Now we are investigate an approach based on the emergent Web Service Resource Framework (WSRF) specification. However, we must also investigate the applicability of the proposed approach on small sensors, in which the computing power and communication capability are very limited.

4 References

[1] Akyildiz, I.F., W. Su, Y. Sankarasubramaniam, E. Cayirci, "A Survey on Sensor Networks", *IEEE Communications Magazine*, pp. 102–114, August 2002.

[2] I. Foster et al., “The Physiology of the Grid: An Open Grid Services Architecture for Distributed Systems Integration”, Grid Forum white paper, 2003.

[3] Foster and C. Kesselman, *The Grid 2: Blueprint for a New Computing Infrastructure*, Morgan Kaufmann , 2003.

[4] Vipul Hingne, Anupam Joshi, Elias Houstis, John Michopoulos. "On the Grid and Sensor Networks", Fourth International Workshop on Grid Computing, grid, pp. 166, 2003.

[5] Hock Beng Lim, Yong Meng Teo, Protik Mukherjee, Vinh The Lam, Weng Fai Wong, Simon See. "Sensor Grid: Integration ofWireless Sensor Networks and the Grid," The IEEE Conference on Local Computer Networks 30th Anniversary (LCN'05)l, lcn, pp. 91-99, 2005.

[6] R. Bramley, et. al., *Instruments and sensors network services: Making instruments first class member of the grid.* Technical Report 588, Indiana University, CS Department, December 2003.

[7] Foster, I., C. Kesselman, and S. Tuecke: “The Anatomy of the Grid: Enabling Scalable Virtual Organizations”, *International Journal of High Performance Computing Applications*, Vol 15, No. 3 pp. 200-222, 2001.

[8] SensorML core specification. [*http://vast.nsstc.uah.edu/SensorML/docs/OGC-05-086r2_SensorML.doc*].

[9] Open Geospatial Consortium. [*http://www.opengeospatial.org*].

[10] Foster, I., Kesselmann, C., Nick, J., Tuecke, “The Physiology of the Grid: An Open Grid Services Architecture for Distributed System Integration”. Technical Report for the Globus project. [*http://www.globus.org/research/papers/ogsa.pdf*].

[11] Foster, I., Kesselmann, C., Nick, J., Tuecke, “Grid Services for Distributed System Integration”. Computer, Vol. 35, No. 6, pp. 37–46, 2002.

[12] The Globus Project. [*http://www.globus.org*].

[13] K. Czajkowski, S. Fitzgerald, I. Foster, C. Kesselman “Grid Information Services for Distributed Resource Sharing”, Proceedings of the Tenth IEEE International Symposium on High-Performance Distributed Computing (HPDC-10), IEEE Press, August 2001.

[14] Czajkowski K., Fitzgerald S., Foster I., Kesselman C., “Grid Information Services for Distributed Resource Sharing,” Proceedings of the Tenth IEEE International Symposium on High-Performance Distributed Computing (HPDC-10), IEEE Press, August 2001.

[15] The WSRF specification. [*http://www.globus.org/wsrf/specs/ws-wsrf.pdf*].

[16] Tuecke S., Grid Security Infrastructure (GSI) Roadmap. Internet Draft 2001 (2001). [*www.gridforum.org/security/ggf1-200103/drafts/draft-ggf-gsi-roadmap-02.pdf*].

[17] Mastrantonio G., and Fiocco G., “Accuracy of wind velocity determinations with Doppler Sodar”, J.Appl.Met., Vol 21, pp. 820-830, 1982.

[18] Mastrantonio G., Argentini S., *A modular PC-Based multiband sodar system. Acoustic Sounding and Applications. S.P. Singal. Ed.*, Narosa Publishing House, New Delhi, India, pp. 105-116, 1997.

[19] J. Novotny, S. Tuecke, and V. Welch. “An Online Credential Repository for the Grid: MyProxy,” Proceedings of the Tenth International Symposium on High Performance Distributed Computing (HPDC-10), IEEE Press, pp 104-111, August 2001.

[20] “LNCS 3470, Lecture Notes in Computer Science,” iGrid, a Novel Grid Information Service Proceedings of the First European Grid Conference (EGC) 2005, Springer-Verlag, pp. 506-515, P.M.A. Sloot et al. (Eds.), 2005.