COMPUTATIONAL METHODS IN SCIENCE AND TECHNOLOGY 15(1), 21-30 (2009)

Integrating Instruments in the Grid for On-line and Off-line Processing in a Synchrotron Radiation Facility

Roberto Pugliese, Milan Prica, George Kourousias Andea Del Linz and Alessio Curri

Sincrotrone Trieste SCpA, Strada Statale 14 - km 163,5 in AREA Science Park 34012 Basovizza, Trieste, Italy e-mail: {roberto.pugliese/milan.prica/george.kourousias/andrea.dellinz/alessio.curri}@elettra.trieste.it

(Received: 8 November 2008; published online: 25 March 2009)

Abstract: Experimental stations in facilities like Synchrotrons and Free Electron Lasers produce very large data sets. With access to distributed computing paradigms and Grid infrastructures, the data are processed on-line. This requires considerable computing power and storage requirements. Moreover, the constant advances in sensors, detectors, and light source technologies yield frequent upgrades of the instrumentation and proportional increase of the computing requirements. Other than the computing needs, a key-point for successful experiments is the efficient collaboration between the involved scientists and their convenient access to the research facilities. A viable solution is developing suitable infrastructures that allow remote operation of the experimental facilities. We focus on the importance of the transparent integration of instruments in the traditional Grid. Such capability boosts further the concept of remote operation, on-line & off-line processing, and builds collaborative and distributed virtual research environments. Building on the foundations of previous European projects, the aim of DORII (Deployment of Remote Instrumentation Infrastructure) [1] is to develop and operate an infrastructure on remote operations and on-line & off-line processing in a synchrotron radiation facility.

Key words: grid computing, middleware, remote instrumentation, grid portals, online processing, HPC

I. INTRODUCTION

The volume of data that will be collected in the imminent future by scientific instruments may soon dwarf all the technical and scientific data that have been collected to date.

The terms 'e-science' and 'e-research' are used to encapulate the notion of global collaborations of people and the virtual sharing of the electronic resources needed to tackle new challenges in science, engineering, medicine, arts and humanities. e-Research has been enabled by the huge improvements in computer networks, creating the ability to harness the power of interconnected computers and shared storage. The IT infrastructure that makes such a collaboration possible in a secure and transparent manner is often referred to as the 'Grid' and more recently as 'e-Infrastructure'.

The scale of collaborative research ranges from local to global, and crosses discipline boundaries. While many computational problems can be satisfied by local deployment of inexpensive networked computers at university and lab sites, an important class of HPC problems require largescale specialised facilities located in dedicated international computing centers. The data itself is an increasingly complex mix of numerical information, text and images, creating correspondingly complex technical problems for its access, use and curation.

Complex calculations are required to take diffraction images and convert them in a 3D protein structure [2]. Similarly, complex calculations are required by Computed Tomography (CT) and then perform qualitative and quantitative analysis of the CT data [3].

The visualisation and individual processing of the results requires extensive interactivity. Moreover, the process and the data collection strategy may be concurrent and performed by a group of scientists that are geographically dispersed. The data from instruments and sensors are saved in distributed repositories, then computational models are executed, and finally an interactive data mining process is used to extract useful knowledge.

This kind of application requires both the support of a standard Grid computing environment, that is a virtual organization, a set of distributed storage and computing resources and a resource brokering mechanism, a workflow definition and execution environment and the capability to integrate instruments (the detectors) and interactively collaborate in the data analysis process. A QoS handling mechanism is necessary to effectively utilise the available network.

Such infrastructures could have impact on the financial aspect of the operation. The funds to support trans-national access to research facilities are limited and there is constant effort for reduced travel expenses. Efficient remote operations and access on experimental facilities could be a possible solution to the problem.

The operations that have to be made remotely in experimental facilities like Accelerators, Synchrotrons and Free Electron Lasers, but also experimental stations, include the planning phase, maintenance, troubleshooting, repair of delicate equipment, optimising performance, commissioning new hardware and setting up routine tasks. All these activities are based on large amounts of information that are usually accessible only on-location at the site of the facility. The key for success of remote operations and online and off-line processing of experimental data is the capability to integrate instruments in the traditional Grid to build collaborative and distributed virtual research environments. Built on the foundations of the previous European projects, the aim of DORII (Deployment of Remote Instrumentation Infrastructure) is to build and operate an infrastructure addressing this issue to support different areas of eScience. The following sections provide details on the way we are utilising the DORII infrastructure for remote operations and on-line and off-line processing in the synchrotron radiation facility named ELETTRA. We begin by introducing the architecture of DORII, then the concepts of Instrument Element (IE) and Virtual Control Room (VCR), and then we describe use cases, remote operations and on-line processing that is under development at Sincrotrone Trieste SCpA (Fig. 1). Future work and research challenges will be presented in the final section.



Fig. 1. The ELETTRA Synchrotron Radiation Facility and FERMI@Elettra Free Electron Laser located in Trieste

ELETTRA is a state-of-the-art (2-2.4) GeV electron storage ring and about 30 synchrotron radiation beam lines with 13 insertion devices. The facility provides (4 700--5 100) hours/year of user operation on each beamline. ELETTRA covers the needs of a wide variety of experimental techniques and scientific fields, including photoemission and spectro-microscopy, macromolecular crystallography, low-angle scattering, dichronic absorption spectroscopy, and X-ray imaging serving the communities of materials science, surface science, solid-state chemistry, atomic and molecular physics, structural biology, and medicine.

ELETTRA is part of the Trieste Area Science Park, and the Trieste scientific environment which includes worldlevel institutions such as the International Center for Theoretical Physics and the International Centre for Genetic Engineering and Biotechnologies. ELETTRA is building a new light source called FERMI@Elettra which is a single-pass FEL [4] user-facility covering the wavelength range from 100 nm (12 eV) to 10 nm (124 eV). The advent of femtosecond lasers has revolutionized many areas of science from solid state physics to biology. This new research frontier of ultrafast VUV and X-ray science drives the development of a novel source for the generation of femtosecond pulses.

II. THE DORII PROJECT

Building on the experience gained in previous European projects (notably, among others, GRIDCC [5], and RINGrid [6], on Remote Instrumentation, int.eu.grid [7], on interactivity, and g-Eclipse [8], on software frameworks for application developers), the DORII (Deployment of Remote Instrumentation Infrastructure) project is currently designing and setting up an extended infrastructure [9] with the direct involvement and cooperation of users in three main areas:

- Earthquake community (with various sensor networks),
- Environmental science community,
- Experimental science community (with synchrotron and free electron laser facilities).

DORII applications span a significant range of e-Science domains, each presenting some challenges for the effective exploitation of Grid and networking services. The goal of DORII is to provide an integrated support environment to such applications, in terms of networking and middleware, by using similar concepts and abstractions, allowing interactivity among scientists and application developers, collaborative working, real-time collection, transfer, manipulation and visualization of data, virtualization and publication of real instrumentation as grid resources.

II.1. DORII Architecture

The DORII eInfrastructure is mainly based on the EGEE [10] infrastructure and its middleware of choice gLite. Extensions of the middleware dealing with the management of remote instrumentation are being built by the DORII project (these are evolutions of services and concepts developed by the GRIDCC project) and will be discussed in section V of this paper. To deal with the interactivity requirements of the applications, the DORII eInfrastructure integrates a selection of services built by the Interactive European Grid Project (int.eu.grid).

The analysis of requirements of the DORII application has resulted in building the architecture shown in Fig. 2.



Fig. 2. The DORII architecture

The gLite middleware offers the main grid services such as information, job management, data management and security control. Information about the resources and services of the infrastructure are provided by the Berkeley Database Information Index (BDII), which uses standard LDAP databases populated by an update process. The Workload Management System (WMS) is the service responsible for the distribution and management of tasks across Grid resources in such a way that applications are conveniently, efficiently and effectively executed. The LCG Computing element (LCG-CE) is responsible for submitting jobs to the underlying local cluster of Worker Nodes (WNs). Storage Elements (SEs) are responsible for data storage and management, while the LCG File catalog (LFC) offers a hierarchical view of files to the users with a UNIX-like client interface. From the security perspective the Virtual Organization Management Service (VOMS) is a full-fledged Attribute Authority whose job is to assign attributes like group membership and role ownership to members of a Virtual Organization (VO), so that other Grid services can make informed decisions based on those attributes, with levels of granularity that range from extremely coarse to extremely fine.

The users and the instruments interact via the Instrument Element (see ch. II.2).

The VCR (see ch. II.3) is a Grid portal which enables registered users to interactively access the DORII Grid resources. It provides them with additional collaboration services like e-Logbook, video and audio conferencing, etc. The VCR is designed to support the DORII applications with scripts and workflows through an external but integrated workflow management system (Wf-Man System). This is a native client application launcher and a tunnelling technology which can be used to integrate application components and to support interactivity in software for the Grid.

The g-Eclipse framework provides tools [15] that customize Grid user applications, manage Grid resources and support the development cycle of new software applications. g-Eclipse can be used as an access and development platform for the DORII Grid, especially by the application developers.

Other services will be added if needed (in Fig. 2 they are represented as "OtherS").

To access all the resources and services of the DORII e-Infrastructure, a suitable Common Library is under development. This will be a common access layer between integrated middleware components and the infrastructure.

In the following paragraphs we will examine in more detail the IE and the VCR which are critical for remote operations and online processing.

II.2. IE

Instrument Element (IE) represents the virtualization [12] of the scientific data sources, mainly digital instrumentation and sensors. IE is implemented as an open source Java middleware that handles the tasks of remote control and safe access to the instrumentation. In particular, it handles the tasks of user authentication and authoriza-

tion, remote operations, concurrent access to devices, locking and reservation of instrumentation during an experiment, publishing of alarms and events generated by the instrumentation, monitoring of the device variables, and a simple connection to the Glite's Grid resources. A schematic overview of the IE architecture may be seen in Fig. 3.

The IE has been first introduced [13] as a novel grid component during the GRIDCC project, and by the end of the project it has been deployed in a number of pilot applications including the remote operations of the Elettra accelerator (see ch. III.1). After the GRIDCC ended, the IE development has been continued at the Elettra, and a complete rewrite of the middleware has been accomplished in order to overcome certain shortcomings that emerged during tests, and to make the IE deployment and use significantly simpler [16]. The IE framework has been part of the EGEE's RESPECT program since the summer of 2008 and it is among the building blocks of the DORII project.



Fig. 3. The instrument element internals

IE presents to the outside world a web service interface that integrates the Glite's authentication module based on the delegation of the client's X509 proxy certificates. Proxy certificates may have VOMS extensions that allow for a fine grain authorization mechanism based upon the user's role in his virtual organization (VO). Access to the instrument operations is limited only to the authorized users and may be overseen by a super-user with higher level permissions. The delegated proxy is further used to access the traditional grid resources, in particular the Storage Elements (SE).

The actual device, or rather its control system, is linked to the middleware through a protocol adapter called the Instrument Manager (IM). An IM consists of a set of Java classes and a single XML description file that fully describes the device. Making IM development and deployment simple and straightforward was the main motivation behind the middleware rewrite. Any device that may be accessed and controlled locally from a Java class is ready to be attached to the IE framework and thus to the Grid environment. An instrument is described by its state machine, transactions, regular commands, attributes (variables), and operating parameters. The above mentioned XML device descriptor is used by the developer to list the commands, attributes, parameters, initial and final states, input and output types, measurement units, attribute and parameter value boundaries, etc. Compiling such a file properly frees the developer of the IM from performing a number of cumbersome check-ups inside the IM code, greatly simplifying the coding procedure. A template is provided to facilitate the IM development.

The IE offers the support for saving the instrumentation output directly to the remote grid storage or uploading of some parameters to the device from the remote grid storage using the GridFTP protocol. The IM developer calls within a single line of code an apposite IE utility, specifying the data and the target SE where s/he wants his/her data to be moved to. The middleware takes care of all the rest, hiding effectively the complexity of the Grid from the user.

Instrument variables (attributes) may be monitored synchronously by polling or, much more efficiently, asynchronously using the Java Message System (JMS). JMS is also used to publish the alarms generated by the device or the events that occurred during the operations.

IE allows for the plugging of the scientific instrumentation into a complex data processing workflow pipeline consisting of traditional grid resources like storage capacities and processing power in a safe and standardized manner. IE also allows for remote operations and monitoring even of the complex equipment.

II.3. VCR

We are convinced that hiding the complexity of the Grid from the end users greatly helps the adoption of the Grid computing paradigm. The Virtual Control Room (VCR) [14] is an open source Grid portal based on Gridsphere and Web 2.0 techologies. Easy discovering, browsing and using GLite Grid resources are its most significant features (Fig. 4). Upon registration, the VCR allows its users to safely create a long term proxy certificates on their VO's MyProxy server. From that moment on, the VCR becomes a SSO portal to the Grid. As users log in with their username and password, short term (usually 12-hour) proxy certificates are created. Such certificates are then used for authentication at VO's computing and storage resources. From the VCR, the users may browse the LFC catalogues and storage elements, and

copy files in both directions. They may submit jobs on the computing elements and monitor its outcome. Preparing a job to be launched on a computing element normally requires the knowledge of the JDL language. On the other hand, VCR's application forms allow the user to prepare his/her job by filling a set of web forms and never have to manage the JDL. The Jython scripting environment allows to create and run simple workflows.

000	Multipurpose Cooperative E	nvironment	
Apple (S1) * Amazon eBay Yahoo! News (93)	 Apple 	gavascript-enabled 😡 🔍 📿 Coogle	
Multipurpose Collaborative E a Virtual Control Roc	nvironment om over the Web	Welcome Milan Prica - Logout	Sep 18, 2008 - 12:35 CEST
Home Administration mceinstruments			
Resources Control Instrument Control Script Manager A	pplication Manager		
? Resource Browser		Storage Element	
Update Resources List 🚳 Geo Visua	Host: gsiftp://gridit003	.pd.infn.it:2811//	
Instrument Elements	HEF/ Listing successful - 48 alice alice bin bin bin bin bin bin bin bin bin bin	2 (11) Ne(=)	D
	?	Computing Element	
Resource Brokers https://prod-wms- 01.pd.infn.it:7443/glite_wms_wmproxy_serve	CE:		
LFC Catalogs	Create New Job, Update Jo	bs List	
Portal File Systems	Job ID Status Submis	sion Date Op. Output	
VCR shared space	?	Multiple Instrument Control	

Fig. 4. The VCR at work

VCR is a front end to the Instrument Element (IE) thus allowing for remote control of the scientific instrumentation, adopting the same security scheme (proxy delegation) as for the storage and computing resources. Each instrument (manager) appears in a separate portlet as showed in Fig. 5. SSL tunelling allows users to safely access the not yet gridified devices or complex legacy instruments.

hirdRobot - on		E 16
E ld: http://140.105.3.224:8080/minilE	/services/IEService	
ontext:		
M Id: ThirdRobot		
CONTROL		
Commande history:		
Mar 14, 2008 9:24:59 AM GET [milan]	status:	
execCommand: "GoStraight" (backward	(100) -> OK on	
Mar 14, 2008 9:25:07 AM GET [milan]	Translations 1 late	
execCommand: "StopStraight" -> OK	I ransitions List:	
Mar 14, 2008 9:25:11 AM GET [milan]	Choose Trans	<u> </u>
execCommand: "Rotate" (RIGHT;30) -> OK		
	No Transition Only	hotod
	No Transition Set	ected
COMMANDS Commands List: Choose Comm	ents:	
COMMANDS Commands List: Choose Comm I No Corr MONITORING	ents:	
COMMANDS Commands List: Choose Comm I No Com MONITORING	ents: imand Selected	
COMMANDS Commands List: Choose Comm MONITORING Name Min Max Value	enta: imand Selected Unit Op.	
COMMANDS Commands List: Choose Comm I No Com MONITORING Name Min Max Value Distance 0 -	ents: mmand Selected Unit Op. 58 - Set	
COMMANDS Commands List: Choose Comm MontoRiNG Name Min Max Distance 0	enta: imand Selected Unit Op. 58 ~ Set	
COMMANDS Commands List: Choose Comm I No Com MONITORING Name Min Max Value Distance 0 Add Time chart, Add Sar chart, Add Sar	ents: mand Selected Unit Op. 58 - Set atter plot Add Time chart (with JMS),	
COMMANDS Commands List: Choose Comm Image: No Com MONTORING Name Min Max Value Distance 0 - Add Time chart, Add Bar chart, Add Sc worldthe Research.	ents: mmand Selected Unit Op. 58 - Sector atter plot Add Time chart (with JMS),	
COMMANDS Commands List: Choose Comm MontoRiNG Name Min Max Distance 0 Value Oistance 0 Add Time chart, Add Bar chart, Add Sc ow/Hide Parameter Parameter	ents: mand Selected 58 Set_ atter plot Add Time chart (with JMS),	
COMMANDS Commands List: Choose Comm Image: No Common MONITORING Name Min Max Value Distance 0 - Add Time chart, Add Bar chart, Add Sc ow/Hde Parameter	ents: immand Selected Unit Op. 58 - Selected atter plot Add Time chart (with JMS),	
COMMANDS Commands List: Choose Comm MonTORING Name Min Max Distance 0 Add Time chart, Add Bar chart, Add Sc cw/Hide Parameter Monto Charts	enta: imand Selected Unit Op. 58 Set. atter plot Add Time chart (with JMS),	
COMMANDS Commands List: Choose Comm I No Corr MONITORING Name Min Max Value Distance 0 Add Time chart, Add Bar chart, Add Sc ww/Hide Parameter	enta: imand Selected Unit Op. 58 - Set atter plot Add Time chart (with JMS),	
COMMANDS Commands List: Choose Comm Image: No Com NoNTORING Name Min Max Value Didatance 0 Conv/Hide Parameter Conv/Hide Parameter	enta: mand Selected 58 - Set_ atter plot Add Time chart (with JMS),	

Fig. 5. The VCR Instrument Control

VCR is also a collaborative environment that offers a set of groupware tools in support of a scientific teamwork. Tools include an e-logbook, an integrated chat, a wiki-like help system and people, and resource browsers. VCR integrates with the third party tools like Skype. VCR may be adopted as a ready-to-use virtual collaboratory or as an extendible framework for developing advanced, application-specific collaboratories. A set of libraries are provided with the VCR software development kit that supports the implementation of application-specific plug-ins.

The VCR portal has been developed during the GRIDCC project. The development continued at Elettra after the termination of the GRIDCC, and currently the VCR is adopted by the DORII project as the front-end to the Grid. Since summer 2008, the VCR makes part of the EGEE's RESPECT program.

III. REMOTE OPERATIONS AND ONLINE AND OFFLINE PROCESSING OF EXPERIMENTAL DATA

In this chapter we will examine the use cases and scenarios of interest in our laboratory for photon science, the Synchrotron Radiation Facility named Elettra.

III.1. Remote Operations Use Case

Scientific facilities like Accelerators, Synchrotrons, Free Electron Lasers, and experimental stations can greatly boost their efficiency with Remote Operations. These involve planning operations, maintenance, troubleshooting, repairing delicate equipment, performance profiling, optimisations, commissioning of new set-ups and various other routine tasks. All these activities are based on large amounts of information stored locally. These big data sets, raw or processed, are usually on-site, and this is against most of the distributed computing models.

The Virtual Control Room is a system that enables operators, scientists and support personnel to use remotely the experimental setup and collaborate in a virtual but efficient way. This can be used in both routine and troubleshooting scenarios.

Our belief is that most of the end-user software applications and workflows are suitable to be migrated to Grid technologies. One example is the implementation of feedback workflows where a high level of instrument integration of and traditional Grid resources is required.

In the feedback parameters of the setup are sent via Grid to a remote computing center where the feedback correction code is executed and the corrections are fed back to the equipment, thus improving some working parameters of the controlled equipment.

The Remote Control of an experimental facility has the potential of substantially altering positively the mode of operation and the degree of exploitation of large experimental physics facilities. Grid technologies facilitate the sharing of these resources across a distributed computing environment. In particular, the Grid handles issues of authentication, authorization, resource description and location, data transfer and resource accounting. Moreover, Grid technologies can be used to integrate operations with computing farms where complex models of the experimental facility can run.

In this use case the key roles are held primarily by the Grid software developers, the scientists of the experiment and the operators of the scientific equipment. From the point of view of the Information System, the developers have access to all the system options available to the scientists in addition to the option of deploying software and defining the Collective Behavior with suitable workflows and automation scripts. The scientist can perform all the activities of the operators plus control and monitor the equipment. The operators can define the operations parameters, run the workflows involved in the operations, and monitor the evolution of the workflows. The users can access and visualize historical data in case it is needed to reprocess them with better algorithms or different parameters. In specific situations like in troubleshooting, they need to collaborate with experts to understand where the problem is and how to solve it. They can also run simulations and then compare the simulated results with historical data and possibly understand how to improve the performance of the equipment.

The resources involved in the remote operations are instruments, sensors and detectors interfaced with the equipment. Such resources are accessed via Instrument Elements (IE). In a remote operation scenario, these IEs are generally co-located but sometimes can be distributed geographically. The data acquired need to be stored in the Grid Storage resources and processed by Grid Computing resources. We will refer to them as Storage Elements (SE) and Computing Elements (CE), respectively. Brokers of resources both storage and computing and other support systems like information, security, and user management, are also involved. The user interacts with the involved resources and collaborates with the other users via the Virtual Control Room (VCR). VCR provides the tools in support of the collaboration effort.

The diagram shown in Fig. 6 shows the remote operations of the Elettra accelerator. The devices involved are beam position sensors and correction magnets interfaced with the rest of the instrumentation. Both may be remotely controlled via Instrument Element (IE). Running the pictured orbit correction workflow (i.e. a feedback via Grid) is a higly demanding task. In order to be effective, the orbit correction loop needs to have frequency not more than 1 Hz, which has a challenging implication in the middleware.



Fig. 6. Orbit-correction Feedback Workflow

An appropriate AAA-mechanism is required. Normal VO management systems can provide a solution in such cases.

Remote operations demand a high quality of service and performance on the whole pipeline that includes instrument control, data acquisition, data transfer to storage resources, fast computation, and data transfer back to storage and instruments. Receiving feedback via Grid to perform orbit correction is a typical example where a response close to real-time is strongly required. This can be achieved by locking and pre-fetching all the required resources including bandwidth. Interactivity is also required when interacting with instruments and visual inspection and judgment from the user is required.

III.2. On-line and Off-line Processing

On-line processing refers to the class of processing that is performed in Computing Elements (CE) accessed through a Grid infrastructure. Off-line processing tends to be performed in local computers or in remote P2P computing resources (i.e. submitting a processing job in a cluster's queue). On-line and off-line processing of experimental data actually includes a group of applications since each beamline and experimental station represents a completely different data collection process, with specific processing, storage, analysis and sharing and visualization requirements.

The individuals involved are grouped in three categories: 1) software developers, 2) beamline scientists, and 3) BL users. Their access policy is hierarchical in the form of inclusive sets of actions. The developers can perform all the activities of the scientists and also deploy software and define the collective behavior. The beamline scientist can perform all the activities of the users plus control the beamline instrumentation, the detectors and the rest of the instrumentation. Finally, the user can access and visualize data even remotely, collaborate with other users, beamline scientists and developers, define the data collection strategy, start and stop data acquisition and monitor the acquisition and online processing phase. The user can also run scripts and workflows and monitor their execution and perform offline processing of the available data.

The resources involved in the online processing related to experimental science are instruments, sensors and detectors. An example are X-Ray Diffraction detectors, and other imaging detectors mounted in beamlines and experimental stations. These are the data sources and in our system analysis they are considered to be Instrument Elements (IE) as described in chapter II.2. A processing workflow for online processing is shown in Fig. 7.



Fig. 7. Online Processing Workflow

Important to the development of a Grid infrastructure is the deployment of suitable applications. These applications, although similar to those targeting clusters and other parallel systems, have to be designed with the Grid architecture in mind. This means that such topics as interactivity, latency, accessibility, and performance have to be considered in advance by the software engineer. Moreover, due to the continuous advances in Grid technology, many core parts are different across various Grid infrastructures. One of the main considerations is the middleware that is used and the software installed on the Computing Elements (CE). Often, when the application has already been developed for a different computing paradigm, it may require porting or substantial changes. In other cases, the laboratories may be cautious with new paradigms as that of the Grid, and may ask for an application that could be deployed both on-line and offline. One of the applications that are being developed for the Grid is a system for Computed Tomography (CT) with a synchrotron radiation source (Fig. 8). Both input and output data sets are very large and special issues regarding the SE and bandwidth have to be considered. The computationally



Fig. 8. Parallel reconstruction from test sinogram generated from Radon transformation of a reference matrix

demanding task is that of reconstructing the slices that apart the 3D volume of the sample. The initial stage requires suitable parallelization of the reconstruction algorithm [17, 18] and then careful design to enable full Grid utilization. For example, MPI programming [19] may utilize a CE with a cluster of WN but does not allow parallelisation in multiple CEs. In the system we are developing by initiating an asynchronous controller, the computation can be performed in cross Grid heterogeneous CE. Additional parts of our CT system require steps that have to be computed off-line and in high-level numerical computing environments like GNU Octave and Matlab.

IV. SUMMARY & FUTURE DIRECTION

Remote operations and online processing are currently under experimental deployment at Sincrotrone Trieste SCpA. There is an active team working on the R&D issues and challenges of the topic. A concept that has matured and was further developed in our lab is that of the Instrument Element. We are providing a suitable e-Infrastructure that can be used to effectively support remote operations and both online and offline processing of experimental data and hence a future-proof solution to the data deluge problem. One of the most critical points is the possibility to integrate in the e-Infrastructure the real sources of data: sensors, instruments and detectors. Another critical success factor is the capability to support effectively interactive applications and their access to scientific data. Our goal is the full implementation of the metaphor of a virtual computer with the scalable capabilities, where the VCR is the display, keyboard and the human computer interaction device (a scaled down Grid-OS), the storage resources and brokers are the disk, the computing elements and brokers are the CPUs, the IE represents the I/O of sci. data. The future direction is towards optimising the above technologies and working towards a Grid paradigm as transparent as possible to the end-user.

Acknowledgments

DORII is a project funded by the European Union under the Seventh Framework Programme (FP7 DORII RI-213110)

APPENDIX: SHORT GLOSSARY

The following is a list of terms and acronyms that may not be widely known due to the continuously updated Grid infrastructure. Some terms from the field of Physics are briefly stated. too. This is an incomplete list. Longer glossaries are available [20] from completed Grid projects. But again, due to the nature of the topic, these may be incomplete, too.

AAA	Authentication, Authorization and Accounting			
Apps Man	application manager			
beamline (BL)	refers to experimental endstation/lab			
CE	computing element			
СТ	Computed tomography			
DB	database element			
DORII	Deployment of Remote Instrumentation			
	Infrastructure			
eScience	reference to computationally intensive scienc			
FEL	free-electron laser			
gEclipse	developer workbench and IDE			
GridFTP	an extended File Transfer Protocol (FTP) for			
	use in the Grid			
НРС	High-performance computing			
	(supercomputers & computer clusters)			
IE	instrument element			
IM	instrument manager			
IS	information system			
JDL	Job Description Language			
LCG	the Computing Grid of the Large Hadror			
	Collider, Cern			
LFC	location file catalog (SE broker)			
OtherS	other services (e.g. OGC)			
QoS	Quality of service			
Script Man	script manager			
SE	storage element			
SSO	Single sign-on			
Synchrotron	a type of particle accelerator			
VCR	virtual control room			
VO	Virtual Organisation (aka Application			
	Community)			
VOMS	VO management system			
VUV	Vacuum Ultraviolet radiation			
Wf-Editor	workflow editor			
Wf-Man	workflow management system			
Wf-Monitor	workflow monitor			
WMS	workload management system (CE broker)			
WN	Worker Node			

References

- [1] DORII project website, http://www.dorii.org
- [2] D. E. McRee, (1993), *Practical Protein Crystallography*. San Diego.
- [3] A. Kak and M. Slaney (1988), *Principles of Computerized Tomographic Imaging*. IEEE Press.
- [4] Ch. Brau, (1990), *Free-Electron Lasers*. Boston: Academic Press, Inc.
- [5] GRIDCC project website, http://www.gridcc.org.
- [6] RINGrid project website, http://www.ringrid.eu.
- [7] Interactive Europran Grid project website, http://www.i2g.eu/.

- [8] g-Eclipse project website, http://www.geclipse.org/.[9] DORII Project Deliverable DSA1.1, *Analysis of Applica*tions and Network Requirements for Remote Instrumentation Infrastructure.
- [10] Enabling Grids for E-sciencE, http://www.eu-egee.org.
- [11] gLite, http://glite.web.cern.ch/glite.
- [12] D. F. McMullen, R. Bramley, K. Chiu, H. Davis, T. Devadithya, J. C. Huffman, K. Huffman and T. Reichherzer, The Common Instrument Middleware Architecture. In: Grid Enabled Remote Instrumentation,. Springer, New York, NY, 2008, pp. 393-407.
- [13] F. Lelli, E. Frizziero, M. Gulmini, G. Maron, S. Orlando, A. Petrucci and S. Squizzato, The many faces of the integration of instruments and the grid. International Journal of Web and Grid Services 3 (3), 239-266 (2007).
- [14] R. Ranon, L. De Marco, A. Senerchia, S. Gabrielli, L. Chittaro, R. Pugliese, L. Del Cano, F. Asnicar and M. Prica, A Webbased Tool for Collaborative Access to Scientific Instruments in Cyberinfrastructures. Proceedings of INGRID 2007 - Instrumenting the Grid: 2nd International Workshop on Distributed Cooperative Laboratories, Springer, Berlin, 2007, in press.

- [15] C. Klausecker, T. Köckerbauer, R. Preissl and D. Kranzlmüller, Debugging MPI Programs on the Grid using g-Eclipse. 2nd Parallel Tools Workshop, Stuttgart, Germany, July 7-8, 2008.
- [16] M. Prica, R. Pugliese, A. Del Linz and A. Curri, Adapting the Instrument Element to Support a Remote Instrumentation Infrastructure. INGRID 2008 Conference Proceedings (to be published).
- [17] Deans, Stanley R. (1983). The Radon Transform and Some of Its Applications. New York: John Wiley & Sons.
- [18] J. B. T. M. Roerdink, M. A. Westenberg, Data-parallel tomographic reconstruction: a comparison of filtered backprojection and direct Fourier reconstruction, Parallel Computing 24 (14) 2129-2142 (1998).
- [19] J. J. Dongarra, S. W. Otto, M. Snir and D. Walker. A message passing standard for MPP and workstations. Communications of the ACM. 39 (7), 84-90 (1996).
- [20] EGEE-II Glossary, http://egee-technical.web.cern.ch/egeetechnical/ documents/glossary.htm



DR. ROBERTO PUGLIESE is a research coordinator at Sincrotrone Trieste S.C.p.A. where he is leading the Scientific Computing Group. Since October 2002 he has also been Professor of E-Commerce at the University of Udine. His research interests include Web Based Virtual Collaborations and Grid technologies. Roberto Pugliese was the technical coordinator of the GRIDCC project and is currently coordinating the Applications workpackage of the DORII project.



MILAN PRICA is a senior software engineer working for the Software for Measurements Group of Sincrotrone Trieste. His background is in Computer Science, Information Retrieval and Search Engine technologies. His interests include Grid Systems, Data Management, Reliable Systems and Web 2.0. Major projects that he has participated in include ADAPT, GRIDCC and DORII.



GEORGE KOUROUSIAS is a computational mathematician at Sincrotrone Trieste. His background is in Numerical Methods and Artificial Intelligence. His current interests include signal processing, medical imaging, HPC, Grid & Cloud computing.



ALESSIO CURRI is a technologist at Sincrotrone Trieste. His professional skills include system analysis and design of Grids and distributing computing environments. He has participated in EGEE I&II, GRIDCC, BIOXHIT projects and is currently involved in DORII and the Italian Grid Infrastructure.



ANDREA DEL LINZ is a software engineer specialised in OO programming and Web development in Sincrotrone Trieste. His interests are in Robotics, Grid Infrastructures, and Remote Instrumentation. He has participated in the GRIDCC and DORII projects.

COMPUTATIONAL METHODS IN SCIENCE AND TECHNOLOGY 15(1), 21-30 (2009)