Advances in Geosciences

Preliminary deployment of Grid-assisted oceanographic applications

G. Bolzon[1], K. Bylec[2], A. Cheptsov[3], A. Del Linz[4], E. Mauri[1], P.-M. Poulain[1], M. Prica[4], and S. Salon[1]

[1]Istituto Nazionale di Oceanografia e di Geofisica Sperimentale – OGS, Sgonico (TS), Italy
[2]Poznan Supercomputing and Networking Center, IBCh PAS, Poznan, Poland
[3]HLRS, University of Stuttgart, Stuttgart, Germany
[4]ELETTRA, Sincrotrone Trieste SCpA, Trieste, Italy

Abstract. Grid integration of OGS oceanographic remote instruments and coupled physical-biogeochemical model has been explored in the framework of the EC-FP7 DORII project. We discuss here the first preliminary results achieved, describing the different tools developed with the support of the project consortium. A general background of the Grid technology for the e-Science is also provided.

1 Introduction

Observations from oceanographic remote instruments and nowcast/forecast/hindcast fields provided by numerical model simulations are the basis of operational oceanography, a constantly growing and powerful tool to monitor, analyse and predict the state of marine resources as well as the sustainable development of coastal areas (Flemming et al., 2002). In such a framework, OGS, as a partner of the FP7-project DORII (Deployment Of Remote Instrumentation Infrastructure, http://www.dorii.eu/), has collaborated with the DORII consortium to integrate oceanographic instruments and numerical models in a Grid e-Infrastructure (Salon et al., 2008; Cheptsov et al., 2008). The main goal of DORII is the deployment of an e-Infrastructure for those scientific communities where ICT technology and the concept of Grid is perceived as an opportunity to boost their research but that, at present, appears not fully exploited. The oceanographic community may greatly benefit from the resources being offered by the Grid technology, in particular concerning the data driven control/interaction workflow and the post-processing/visualization issues that characterize both the remote instruments and the numerical outputs.

Near real time (NRT) observations at the sea surface and in the water column, such as temperature and salinity profiles, are of paramount importance for the operational forecasting system in the Mediterranean Sea, and the management of the network of Argo floats deployed in the Mediterranean Sea, and handled by the MedArgo Centre at OGS (Poulain et al., 2007), constitutes a complex task that may be greatly supported by the fast developing e-Infrastructures. On the other hand, the advanced visualization of the OPATM-BFM model outputs produced for the forecasting system of the Mediterranean Sea biogeochemistry, pre-operationally deployed within the MERSEA-IP FP6-project (http://www.mersea.eu.org/, Crise et al., 2008, and Lazzari et al., 2009) and currently among the V0 services of the MyOcean FP7-project (http://www.myocean.eu.org/), represents a crucial issue for the post-processing phase of the operational system.

The implementation, in a Grid e-Infrastructure, of the complete workflow that originates from the remote observations and concludes with the 3-D visualization of the model's outputs, passing through the measured data downloading, processing and quality control, and the model running, and integrated with collaborative work among researchers, security access and guaranteed quality of service, represents an interesting challenge for which the DORII consortium is currently having the first preliminary results. In particular, interactive applications can support the remote control of autonomous vehicles (e.g. gliders or bi-directional floats), and handle the presumptive technical problems bound to appear intermittently or the NRT corrections and/or adjustments of the sensors.

Grid integration of oceanographic instruments, computing and data has been already explored and is being developed in the USA within the Ocean Observing Initiative CyberInfrastructure (Edgington, 2008; Chave et al., 2009; see also http://www.oceanleadership.org/programs-and-partnerships/ocean-observing/ooi/ for more details), while in Europe it is worth to mention the NERC

DataGrid effort (Woolf et al., 2009; http://ndg.nerc.ac.uk/).

We present here some of the outcomes achieved so far in the framework of the DORII project focusing on the current state of the application proposed by OGS. In the next section a brief background of the technology employed is given, then we describe the different components used in the prototypical version of the application also discussing some of the most important results achieved so far.

2   Grid technology for e-Science

Being a constantly evolving technology, it is quite difficult to include all the characteristics of what a Grid is in a single definition. According to Foster and Kesselmann (1998) "a computational Grid is a hardware and software infrastructure that provides dependable, consistent, pervasive, and inexpensive access to high-end computational capabilities". However, the developments in Grid technology during the last decade have substantially enriched this definition, including many aspects that refer to concepts like resources sharing, secure access, standard and open protocols and interfaces, quality of service (for a general discussion, see also http://www.gridcafe.org). As a result, current Grid infrastructures share a virtually unlimited computational power and storage capacity. In particular, the computing resources, intended both as server farms and multi-core machines, are usually referred to as "computing element" (CE), while the shared space where data are stored represents "storage element" (SE). The community that has access to the resources, either as providers or as users, is defined as "virtual organization" (VO).

Such a complex and distributed application system (High Level Architecture) requires a software layer that connects separate applications and exchanges data among them. The connecting software is usually referred to as "middleware", and allows interaction among applications running on different operating systems, making easier the implementation, maintenance and customisation of single applications. Within the Grid, this service is carried out by the middleware, e.g. gLite, Globus or UNICORE.

In recent years the interest of experimental science communities using the Grid infrastructure has moved towards the scientific instruments and sensors, in order to create an environment able to link the data collection with their analysis and visualization (for more details and an overview on this issue see Davoli et al., 2008). Consequently, a new component has been added to the CE-SE-VO original architecture, the "instrument element" (IE), that integrates the instrument into the Grid. The IE is defined as an abstraction of the instrument (or group of instruments) into a standard interface. It was initially developed by the FP6-GRIDCC project (http://www.gridcc.org/cms/), and then successively re-implemented for the objectives of DORII project (Prica et al., 2008; Pugliese et al., 2009). In this context, the term "instrument" refers to a piece of equipment that needs to be initialised, configured, operated (start, stop, standby, resume, application specific commands), monitored or reset. The IE adopted by DORII represents another level of middleware designed to supply the remote control of the real sensors and to perform a fast interaction with SE and CE within the e-Infrastructure, using the gLite/GSI[1] security model. For example, for the oceanographic community, the instrument could be a float or a glider, and the CE/SE can be then used to run numerical models that assimilate the measured data and store their outputs.

3   The OGS application in the DORII e-Infrastructure

OGS participates in the DORII project with the OCOM-MOON (Oceanographic and coastal observation and modelling Mediterranean ocean observing network) application, that involves the Argo floats managed by the OGS MedArgo Centre, with Sea-Bird sensors on board to measure vertical profiles of temperature, salinity and pressure, the OGS glider Slocum Trieste-1, which measures temperature, salinity, pressure, dissolved oxygen, fluorescence chlorophyll and turbidity in the upper water column, and the OPATM-BFM parallel off-line coupled 3-D physical-biogeochemical model for the Mediterranean Sea. The goal of such application is to provide an integration of the instruments and numerical model within the DORII e-Infrastructure. This is the first step towards the realization of a prototype of a Grid portal, based on the Virtual Control Room developed in the framework of GRIDCC project (Ranon et al., 2008), which will offer a direct and customized access to the Grid resources and services for the OGS researchers involved in the instruments or model management and use. Requirements from this group of users have been examined during the evaluation processes carried out within DORII.

In the following sections, we briefly describe the components of the application and expected impact for the users. The most relevant outcomes achieved so far are also discussed.

3.1   VCR

The Virtual Control Room (VCR) is a web portal that puts together a rich collaborative environment with easy access to e-Infrastructure. The VCR is built on top of the Gridsphere portal (http://www.gridsphere.org/) and the DORII common library, a pure Java client for the gLite middleware (developed by CERN, http://glite.web.cern.ch/glite/). It integrates DORII workflow client (Sect. 3.4) as well as its own Application Manager. The main goals of the VCR are to hide the Grid complexity from the users and to allow them to use the resources securely from any Java-enabled web browser. The latest edition of the VCR includes tags that allow for easy user customisation of the working environment, so that the

_____

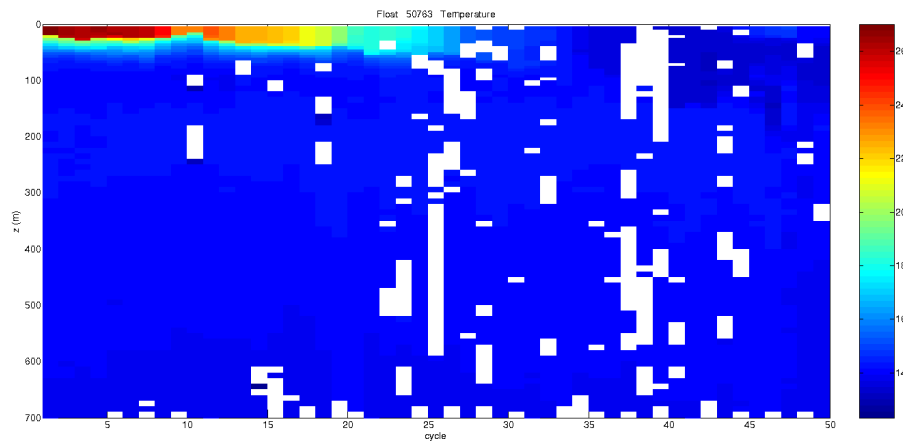[1]Grid Security Infrastructure.

Fig. 1. Snapshot of a retrieved image showing the time evolution of the temperature vertical profiles between two user-specified cycles of surfacing. Mixing at the surface layer can be observed starting from cycle 20.

users once logged into the portal, are presented with only the applications and resources they are entitled to use. In the DORII architecture it is the central building block and it acts as the SSO (single sign-on) entry-point to the Grid for the application users. From the VCR, users may control instruments, process datasets on the Grid computing resources and see both the acquired and the processed data on the Grid storage resources.

3.2 Instrument Managers for the OGS oceanographic instruments

The Instrument Manager (IM) runs inside the IE installation, and is the interface that allows the middleware to communicate with the control system of the remote instrument. The IM is implemented by specifying the description of the actual instrument with: (i) the list of its status, (ii) the values to be monitored (attributes), (iii) the values set by the user (parameters) and (iv) the commands and their parameters. Each of the previous points is integrated in the IE by a Java class that implements the interface with the instrument (the I/O protocol) and executes the tasks (get/set values, send commands and receive outputs). Therefore, the IM implementation results in an extension of the capabilities of the instrument itself, and allows a further customisation. Within DORII we have developed IMs for the MedArgo floats and for the OGS glider.

3.2.1 IM for floats

Every 5–10 days the MedArgo floats transfer via satellite the measured raw data to a control system which stores them in a specified server.

The float IM basically connects to such a server and shows the last measurement data, position, and technical status. Further, the IM allows to explore the float historical database and to generate a graphical output of temperature and salin-

ity vertical profiles for a time-range specified by the user (Fig. 1). This is carried out by showing the attributes on the VCR and by allowing the transfer of the binary files to a specific SE. In this case the extension of the concept of instrument concerns the capability of accessing databases and elaborating data, for example by passing the data to a graphic generator running outside of the Java code.

3.2.2 IM for the glider

The OGS glider is an AUV remotely controlled by a complex system which is composed by the glider's antenna, the Iridium Satellite and a dedicated server (Dock Server), hosted at OGS. The user communicates with the glider via a command-line client interface (Glider Terminal).

The glider IM allows the user to manage the instrument without the command-line interface, thus avoiding the drawbacks related to it, e.g. giving commands with the Glider Terminal's proprietary syntax or expressing variables in the proprietary unity (e.g. angles in radians). Since IM is basically an open source software, it has some advantages that go beyond the concept of a friendly user interface controlling a remote instrument. In particular, sending a command to the glider can be wrapped in a more complex operation, able to run some controls before executing the required command, to check if all the security conditions are satisfied or to even wrap the command in a loop. This improves the instrument control, since the user can avoid wrong sequences of instructions or execution of dangerous commands. Moreover, being an extremely delicate instrument, the glider developers have the necessity to introduce and manage privileges for the glider interaction: this can be fulfilled by using features of the DORII e-Infrastructure, i.e. VO Management System (VOMS) attributes and roles, that can be translated into IE roles. Furthermore, an open source code gives the way to a development of software for a general management of the instrument: pre-processing (set up of the mission), processing
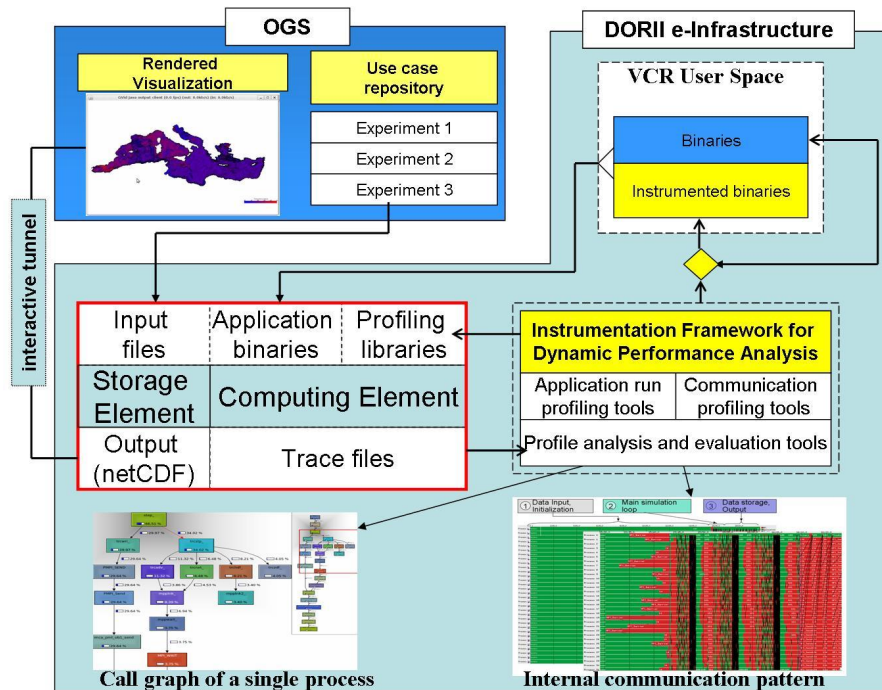
Fig. 2. Framework of analysis of the OPATM-BFM integrated in the DORII e-Infrastructure. A snapshot of the visualization of models output is showed in the top-left corner.

(control during the mission), and post-processing (data management and quality control).

3.3 OPATM-BFM gridification

The complexity of numerical computation is typical for many oceanographic simulation applications, and mainly consists in the high number of prognostic variables to be integrated, the domain's dimension, the parameterisation schemes and many others. These issues also concern the OPATM-BFM model: in this context, the efficient solving of its tasks poses several challenging scenarios for the efficient utilization of the e-Infrastructure's resources.

For the data management, the DORII e-Infrastructure provides an unique access to diverse storage systems, offering a total of 147TB storage space for the applications deployed on it (Adami et al., 2009). For such applications as OPATM-BFM, which produces ∼110GB of output data for the standard operational short-term forecast (10 days of analyses and 7 days of forecast), the e-Infrastructure enables performing experiments with a large amount of data produced. The local user space provided by VCR is also available for storing the light-weight user data containing the basic settings and simulations description data.

The high performance and scalability of computation algorithms is another important aspect where OPATM-BFM can benefit from running on the e-Infrastructure. Unlike dedicated supercomputer architectures, the Grid infrastructure

largely eliminates the limitation on the number of computation power available for the applications. The current DORII infrastructure configuration enables its applications to run on ∼2200 CPU cores (Adami et al., 2009) provided by National Grid Initiatives like German D-Grid, Polish Grid or Greek GRNET ones. Moreover, being an MPI parallel application, OPATM-BFM can capitalise the facilities provided by the e-Infrastructure, including implementation by means of OpenMP (inside a single SMP[2] node), MPI (parallelization across many interconnected SMP nodes) or other technologies. In particular, the DORII infrastructure provides a strong support for the parallel MPI applications, deploying the state-of-the-art tools for the implementation of the MPI standard (as Open MPI library; Graham et al., 2006), the support of inter-cluster (or inter-CE) application execution (as PACX-MPI; Keller et al., 2005) and the improved running MPI applications in Grid environments for different MPI libraries.

Visualization of the output data produced by OPATM-BFM is another crucial point: current single-system software running on commodity machines scale hardly up to the large amount of data generated by OPATM-BFM, especially when produced by long-term simulations for climate scenarios. To address this problem, the DORII e-Infrastructure deploys a flexible and extensible Grid-based visualization framework for time-critical, interactively controlled data set rendering that enables visual browsing of spatially and tem-
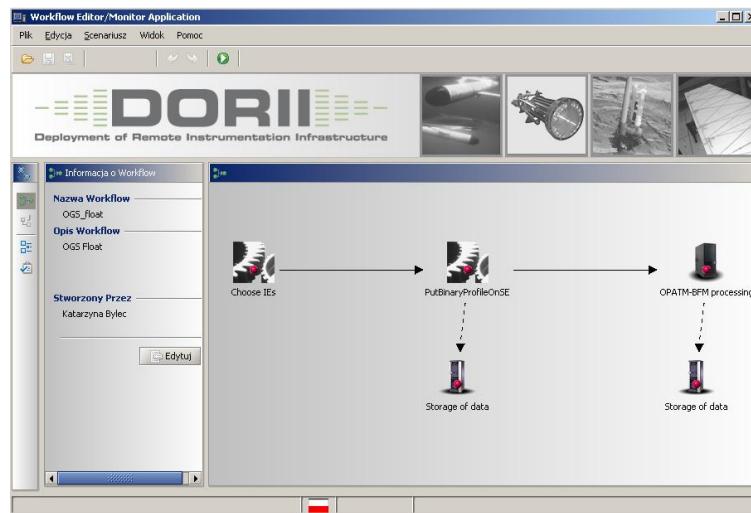
---

[2]Shared-Memory processor.

Fig. 3. Snapshot from the Workflow Editor with workflow relative to OCOM-MOON application.

porally distributed large-scale datasets in a Grid environment (including the netCDF data format used in OPATM-BFM). The framework is based on VTK (http://www.vtk.org/) and GVid (Köckerbauer et al., 2005) toolkits which are the core of the visualization application for OPATM-BFM developed in DORII.

Finally, a critical issue for the majority of oceanographic applications is the ability to *effectively* use computational Grid resources for data and communication intensive applications. The heterogeneous nature of the traditional Grid poses a significant challenge for computationally demanding applications, like OPATM-BFM, especially when porting from dedicated computational platforms. The performance analysis techniques and optimisation strategies are extremely beneficial for coping with this challenge. For this purposes, the DORII e-Infrastructure offers a special performance analysis framework (Fig. 2), that includes both tools for automatic instrumentation of the applications executable files (that aims at collecting time stamps of occurred events in form of traces) and graphical backends for analysing acquired event traces (such as Vampir trace; Knüpfer et al., 2008).

Based on the experience for the pilot DORII applications, the strategy of the performance analysis was proposed, allowing application developers not only to identify the main performance shortcomings and bottlenecks in the application internal message-passing and I/O communication patterns, but also to work out several optimisation proposals minimizing the shortcomings and increasing the scalability. Exemplary use of the strategy for OPATM-BFM allowed OGS to improve the OPATM-BFM performance for a standard short-term forecast on a generic cluster of workstations by 5% and scalability on increasing number of cores by 15% (Cheptsov et al., 2009).

3.4 Workflow

Introducing scientific applications to the Grid results in increase of performance, but this in turn raises the complexity and entrance level for new users. Although tools such as the VCR offer access for e-Infrastructure, the user is still required to perform a sequence of many Grid operations to achieve the desired result and outcome of scientific applications. Even for most experienced Grid users going through those repetitive steps in daily work is inconvenient. Therefore porting applications to Grid may be mistakenly seen rather as an encumbrance than an improvement. This is the point when a need for a tool which would encapsulate Grid operations appears and therefore we adopt the concept of workflows for DORII.

The idea of workflow is not new neither in scientific nor in Grid environment (Taylor et al., 2007; Ludäscher et al., 2008). The term itself is self-explanatory: workflow is created for managing the flow of work seen as a sequence of operations. Abstraction of those operations is always the same, in submission of workflow they differ only with well-defined parameters. Users are not required to be aware of technical implementation of workflow steps. As long as they understand the abstraction of each operation they can carry out their daily work with scientific applications.

Workflow Management System (WfMS) is the implementation of workflows for the DORII project. It is composed with a client-side Workflow Editor and a server destined Workflow Manager. Workflow Editor is a graphical tool for submission and monitoring of workflows (Fig. 3), adopted from VLab project (Lawenda et al., 2004). At this level three abstract Grid operations are defined as building blocks of workflow, corresponding to e-Infrastructure components. CE is represented with Grid jobs, SE with I/O operations as well as with data transfers and IE with Instrument Elements

Table 1. Details of the OCOM-MOON workflow.

Workflow step	Block type	Grid operation
Choose IE	IE	Query IE for available Argo float devices
PutBinaryProfileOnSE	IE	Turn on chosen float, fetch data from it, switch it off
Storage of data (1)	SE	Browse chosen SE, chose remote location to which data from float will be transferred
OPATM-BFM processing	CE	Submit OPATM-BFM job file to the Grid
Storage of data (2)	SE	Browse chosen SE, store output of OPATM-BFM in remote location chosen by user

interaction. The user of scientific application is provided with a predefined workflow built with blocks presenting each step of the application. User action is narrowed down to customising the workflow providing values for operations' parameters. The workflow is then submitted to Workflow Manager which acts as an interface for Grid actions. It translates workflow blocks to Grid operations and performs them in the right order.

In case of the OCOM-MOON application five workflow's steps were identified together with block types and corresponding Grid operations: they are illustrated in Fig. 3 and described in Table 1.

Once the user has properly chosen the Argo float and the remote locations for the data transfer, workflow submission is readily available by clicking on the "submit" green button (top of Fig. 3). Then, it is possible to monitor the job execution via the logbook tool included in the Workflow Editor.

4 Conclusions

As already demonstrated by US initiatives, the operational oceanography scientific community can greatly benefit from the integration with the Grid of the oceanographic remote instruments and forecast numerical models. The possibility to offer a single, certified access to a collaborative environment able to manage the flow of information between observations and forecast simulations, embedded in the OGS research activities, has been explored within the DORII project. The DORII consortium has strongly supported OGS and its scientific application OCOM-MOON to implement prototypical, efficient management tools fit to cope with the researchers' requirements.

Instrument Managers have been developed to manage technical and observational data coming from the MedArgo Argo floats in the Mediterranean Sea and to interactively communicate with the OGS Slocum glider Trieste-1, whereas the OPATM-BFM model for the short-term forecasts of the Mediterranean biogeochemistry has been integrated in the DORII infrastructure. Moreover, VCR and Workflow, even if at a very preliminary stage, have been appreciated by

not-experienced users as useful tools to interact with the new Grid environment.

Future planned development includes efforts oriented to the implementation of instrument geo-reference and improved graphic utilities integrated within the OGS-customised VCR, IM for bi-directional floats, fast and easy access to remote visualization tool based on GVid, and targeted dissemination within wider communities such as the National and European working groups involved in operational oceanography, i.e. the Gruppo Nazionale di Oceanografia Operativa (National Group for Operational Oceanography, GNOO) and the Mediterranean Operational Oceanography Network (MOON).

Edited by: G. M. R. Manzella and S. Nativi
Reviewed by: two anonymous referees

References

Adami, D., Cheptsov, A., Davoli, F., Liabotis, I., Pugliese, R., and Zafeiropoulos, A.: The DORII Project Test Bed: Distributed eScience Applications at Work, in: Proceedings of the 5th International Conference on Testbeds and Research Infrastructures for the Development of Networks & Communities and Workshops (TridentCom 2009), Washington DC, USA, 6–8 April 2009, 1–4, doi:10.1109/TRIDENTCOM.2009.4976247, 2009.

Chave, A. D., Arrott, M., Farcas, E., Krueger, I., Meisinger, M., Orcutt, J. A., Vernon, F. L., Peach, C., Shofield, O. and Kleinert, J. E.: Cyberinfrastructure for the US Ocean Observatories Initiative: Enabling Interactive Observation in the Ocean, IEEE Oceans 2009, Bremen, Germany, 11–14 May 2009, OS22A-02, 2009.

Cheptsov, A., Dichev, K., Keller, R., Lazzari, P. and Salon, S.: Porting the OPATM-BFM Application to a Grid e-Infrastructure – Optimization of Communication and I/O Patterns, Comp. Met. Sci. Tech., 15, 9–19, 2009.

Crise, A., Lazzari, P., Salon, S. and Teruzzi, A.: Final report on the BFM OGS OPA Transport module, MERSEA deliverable D11.2.1.3, 21 pp., 2008.

Davoli, F., Meyer, N., Pugliese, R., and Zappatore, S.: Grid Enabled Remote Instrumentation, series: Signals and Communication Technologies, Springer New York NY, USA, 591 pp., 2008.

Edgington, D. R., Davis, D. and O'Reilly, T.: Ocean observing system instrument network infrastructure, in: Proceedings of the Marine Technology Society/Institute of Electrical and Electronics Engineers Oceans Conference (OCEANS 2006), Boston MA, USA, 18–21 September 2006, 1–6, doi:10.1109/OCEANS.2006.307043, 2006.

Flemming, N. C., Vallerga, S., Pinardi, N., Behrens, H. W. A., Manzella, G., Prandle, D., and Stel, J. H.: Operational Oceanography - Implementation at the European and Regional Scales, Elsevier Oceanography 66, Elsevier, Amsterdam, Netherlands, 572 pp., 2002.

Foster, I. and Kesselman, C.: The Grid: Blueprint for a New Computing Infrastructure, Morgan Kaufmann Publishers, San Francisco CA, USA, 675 pp., 1998.

Graham, R. L., Shipman, G. M., Barrett, B. W., Castain, R. H., Bosilca, G.,and Lumsdaine, A.: Open MPI: A High-Performance, Heterogeneous MPI, Proceeding of the 2006 IEEE International Conference on cluster Computing, 25–28 September 2006, Barcelona, Spain, doi: 10.1109/CLUSTR.2006.311904, 2006.

Keller, R. and Liebing, M.: Using PACX-MPI in metacomputing applications, 18th Symposium Simulationstechnique, Erlangen, Germany, 12–15 September, 2005.

Knüpfer, A., Brunst, H., Doleschal, J., Jurenz, M., Lieber, M., Mickler, H., Müller, M. S. and Nagel, W. E.: The Vampir Performance Analysis Tool-Set. Tools for High Performance Computing, in: Proceeding of the 2nd International Workshop on Parallel Tools for High Performance Computing, July 2008, HLRS, Stuttgart, edited by: Resch, M., Keller, R., Himmler, V., Krammer, B. and Schulz, A., Springer, Berlin Heidelberg, 139–156, 2008.

Köckerbauer, T., Polak, M., Stütz, T., and Uhl, A.: GVid – Video Coding and Encryption for Advanced Grid Visualization, in Proceedings of the 1st Austrian Grid Symposium, Linz, pp. 204–218, 2005.

Lawenda, M., Meyer, N., Rajtar, T., Oko, M., Stokosa, D., Stroiski, M., Popenda, L., Gdaniec, Z. and Adamiak, R. W.: General Conception of the Virtual Laboratory, in Proceedings of the International Conference on Computational Science 2004, LNCS 3038, Cracow, Poland, 6–9 June, 1013–1016, 2004.

Lazzari, P., Teruzzi, A., Salon, S., Campagna, S., Calonaci, C., Colella, S., Tonani, M., and Crise, A.: Pre-operational short-term forecasts for Mediterranean Sea biogeochemistry, Ocean Sci., 6, 25–39, doi:10.5194/os-6-25-2010, 2010.

Ludäscher, B., Bowers, S. and McPhillips, T.: Scientific Workflows, in T. Özsu and L. Liu, Eds., Encyclopedia of Database Systems. Springer, 2008.

Poulain, P.-M., Barbanti, R., Font, J., Cruzado, A., Millot, C., Gertman, I., Griffa, A., Molcard, A., Rupolo, V., Le Bras, S., and Petit de la Villeon, L.: MedArgo: a drifting profiler program in the Mediterranean Sea, Ocean Sci., 3, 379–395, doi:10.5194/os-3-379-2007, 2007.

Prica, M., Pugliese, R., Del Linz, A., and Curri, A.: Adapting The Instrument Element To Support A Remote Instrumentation Infrastructure, Proceedings of INGRID'08 Conference, Ischia, April 9–11, 2008.

Pugliese, R., Prica, M., Kourosias, G., Del Linz, A., and Curri, A.: Integrating Instruments in the Grid for On-line and Off-line Processing in a Synchrotron Radiation Facility, Comp. Met. Sci. Tech., 15, 21–30, 2009.

Ranon, R., De Marco, L., Senerchia, A., Gabrielli, S., Chittaro, L., Pugliese, R., Del Cano, L., Asnicar, F., and Prica, M.: A Web-based Tool for Collaborative Access to Scientific Instruments in Cyberinfrastructures, in: Grid Enabled Remote Instrumentation, edited by: Davoli, F., Meyer, N., Pugliese, R., Zappatore, S., Springer, New York NY, USA, pp. 237–251, 2008.

Salon, S., Poulain, P.-M., Mauri, E., Gerin, R., Adami, D., and Davoli, F.: Remote Oceanographic Instrumentation Integrated in a GRID Environment, Comp. Met. Sci. Tech., 15, 49–55, 2009.

Taylor, I. J., Deelman, E., Gannon, D. B., and Shields, M.: Workflows for e-Science: Scientific Workflows for Grids, Springer, 530 pp., 2007.

Woolf, A., Cramer, R., Gutierrez, M., Kleese van Dam, K., Kondapalli, S., Latham, S., Lawrence, B. N., Lowry, R., and O'Neill, K.: Data Virtualisation in the NERC DataGrid, in: Proceedings of the U.K. e-science All Hands Meeting, edited by: Cox, S. J., 2003.