COMPUTER SCIENCE • VOL. 12 • 2011

MICHAEL SCHIFFERS*, DIETER KRANZLMÜLLER*,
ANDREA CLEMATIS**, DANIELE D'AGOSTINO**,
ANTONELLA GALIZIA**, ALFONSO QUARATI**, ANTONIO PARODI***,
MARINA MORANDO***, NICOLA REBORA***, EVA TRASFORINI***,
LUCA MOLINI***, FRANCO SICCARDI***, GEORGE C. CRAIG****,
ARNOLD TAFFERNER****

TOWARDS A GRID INFRASTRUCTURE FOR HYDRO-METEOROLOGICAL RESEARCH

The Distributed Research Infrastructure for Hydro-Meteorological Study (DRIHMS) is a coordinated action co-funded by the European Commission. DRIHMS analyzes the main issues that arise when designing and setting up a pan-European Grid-based e-Infrastructure for research activities in the hydrologic and meteorological fields. The main outcome of the project is represented first by a set of Grid usage patterns to support innovative hydro-meteorological research activities, and second by the implications that such patterns define for a dedicated Grid infrastructure and the respective Grid architecture.

Keywords: *grid computing, high performance computing, hydro-meteorology, e-science*

ZASTOSOWANIE INFRASTRUKTURY GRIDOWEJ DO BADAŃ HYDROMETEOROLOGICZNYCH

Rozproszona infrastruktura naukowa przeznaczona do badań hydrometeorologicznych (Distributed Research Infrastructure for Hydro-Meteorological Study – DRIHMS) stanowi element skoordynowanej akcji współfinansowanej przez Komisję Europejską. Celem DRIHMS jest analiza głównych problemów spotykanych w dziedzinie hydrologii i meteorologii. Głównym wynikiem projektu będzie zestaw wzorców użytkowania środowisk gridowych w celu wspomagania nowoczesnych badań hydrometeorologicznych oraz wnioski wynikające z powyższego zastosowania, mogące mieć wpływ na dalszy rozwój dedykowanych rozwiązań gridowych.

Słowa kluczowe: *obliczenia gridowe, obliczenia wysokiej wydajności, hydrometeorologia, e-science*

---

* Ludwig-Maximilians-Universität, Munich, Germany, michael.schiffers@ifi.lmu.de
** Institute for Applied Mathematics and Information Technologies, Genoa, Italy
*** CIMA Research Foundation, Savona, Italy
**** German Aerospace Center (DLR), Oberpfaffenhofen, Germany

1. Introduction

Although international agencies like the United Nations (UN) or the European Commission (EC), national government organizations, and local authorities increasingly ask for globally certified management tools to deal with extreme events of precipitation and floods, the scientific community is still reduced to the ability to communicate *scenes* to urban, regional and national decision makers. Unfortunately, *scenes* are typically both information-light and emotion-heavy with a *degraded* view at scientific data. A *scenario*, on the other hand, is a time series of data for studying the causes of observed effects or for predicting future ramifications. Consequently, as the construction of a *scenario* must consider the temporal dynamics and the interactions between processes in order to mitigate risks and disastrous vulnerabilities, a reliable knowledge of the underlying phenomena is required, where *risk* – defined as the probability that a foreseen event, despite the actions taken to contrast it – instigates a certain degree of impact on the population and the environment.

For agencies and governmental authorities it is therefore of utmost importance to be able to issue fact-based warnings in order to reduce the amount of damage people, animals and properties may be exposed to. Important prerequisites of such abilities are observed data, formally sound models to supplement them, the reliable access to distributed data archives, and interoperable computational technologies. Without the support of an adequate Information and Communication (IC) infrastructure such an undertaking will be difficult if not impossible. For Hydro-Meteorological Research (HMR) such an infrastructure is not available today.

Against this background the EC-funded project *Distributed Research Infrastructure for Hydro-Meteorological Study (DRIHMS)* [40] has been set up with three main objectives: Firstly, to better understand the mindsets of both the hydro-meteorological and the Information and Communication Technologies (ICT) communities; secondly, to identify perceived and real gaps; and thirdly, to propose a way of closing the gaps using Grid technologies.

In order to better understand the technical issues behind these objectives questionnaire techniques were applied. While the HMR part of the survey clearly showed the paramount importance of solving probabilistic forecasting issues and of providing formally sound model verification metrics [11], the ICT part emphasized the importance of data management, workflow management and high performance solutions. We will briefly summarize the methodology, the outcome, and the implications of the surveys in Section 2.

Based on the survey results, a rationale for an HMR e-Science environment can be formulated and mapped onto Grid infrastructures provided in the context of European e-Infrastructures like the European Grid Infrastructure (EGI), the Partnership for Advanced Computing in Europe (PRACE) and the European national Grid initiatives (NGI). We will explore this in Section 3 in more detail.

Several research questions are related to such an endeavor. In Section 4 we will briefly address four of them.

In Section 5 we will relate DRIHMS to similar projects before concluding the paper in Section 6.

2. Gap Analysis

2.1. Methodology

Identifying the Grid-related gaps for hydro-meteorological research is a cumbersome task as it requires an understanding of both the real HMR needs and the strengths and weaknesses of modern Grid technologies.

The DRIHMS gap analysis followed a multi stage Delphi-like assessment process (see Figure 1). Based on the input of field experts two questionnaires were defined: One for the hydro-meteorological research community and one for the ICT Grid community. The results of both surveys were consolidated, augmented with requirements known from related projects (see section 5), reviewed, and prioritized at a joint public conference. In (the rare) cases of conflicts additional expert interviews were conducted.

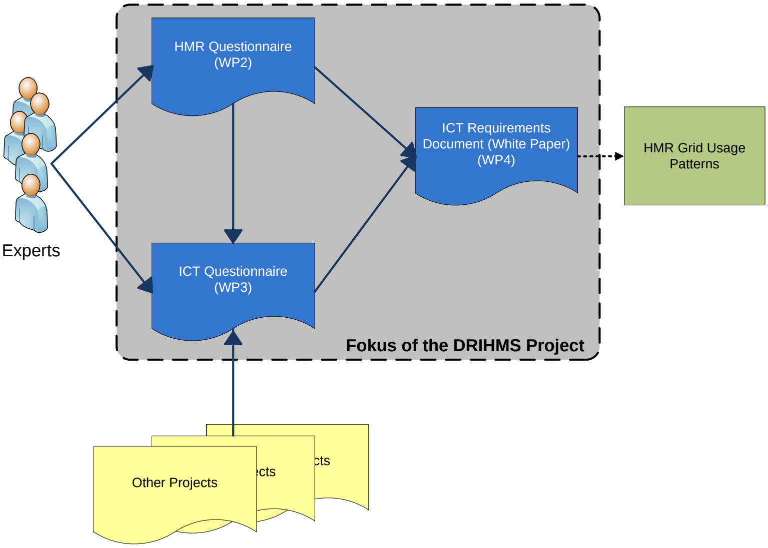

Fig. 1. The DRIHMS ICT requirements analysis process [11]

The purpose of the ICT questionnaire was twofold: Firstly, it should help to understand the Grid implications of the most important HMR requirements as derived from the HMR questionnaire; secondly, it should support exploring Grid opportunities for HMR. Consequently, it polled opinions related to those achievements in Grid Computing that would most probably represent a benefit for HMR.

HMR in Europe is increasingly challenged by the ability to exploit computational resources for online operations and by the ability to retrieve and access data from different sources/countries, stored in different formats, and to be used in various hydrological/meteorological models. Not surprisingly, the HMR questionnaire identified as the most important requirements for HMR Grid infrastructures the availability of objective methodologies for downscaling model outputs to local conditions, the availability of objective comparison methods of model performance, the ability to store data from field campaigns in commonly agreed standard formats, and the facilitation of operational forecasting.

Both the HMR and the ICT questionnaires were published on the project's web site [12] and distributed via various national and international mailing lists, personal contacts, and to conference participants. The HMR questionnaire was returned by more than 200 respondents, the ICT questionnaire by 81 respondents from most European countries, Ukraine, Brazil, Korea, Taiwan and the US. It should be noticed that the answers are based on a non-random sample. The results are thus not projectable to any population other than the experts selected for this sample. Nonetheless, there is a high degree of representativeness since the variances in the answers are small.

2.2. Gaps and Expectations

A detailed analysis of the questionnaire results is given in [43]. The most important findings are summarized in Figure 2 which relates the weighted mean of the (perceived) importance to the (perceived) maturity/availability of the HMR hot topics (identified by the HMR questionnaire) for the next three years (higher numbers denote higher importance and higher maturity, resp.).

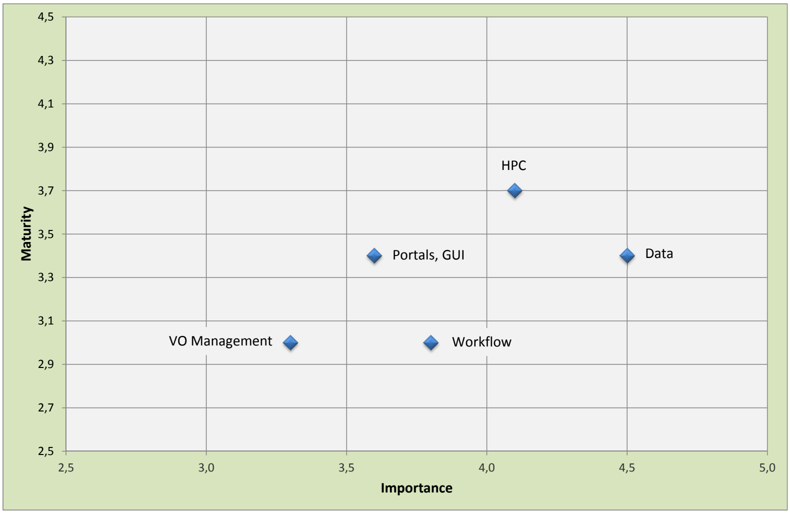

Fig. 2. Correlation between (perceived) importance and (perceived) maturity of HMR hot topics [43]

From Figure 2 we can draw some general conclusions:

1. Data management is very important. The respondents, however, do not see significant progress in the next years.
2. High Performance Computing (HPC) is seen very important and significant progress within the next years is expected in hardware, middleware, and access technology.
3. Workflow management is considered important but no significant progress is expected – even short term.
4. Portals and user interfaces are considered important and existing solutions seem to fulfill most of the requirements. Nonetheless, the respondents expect further progress in this field.
5. Managing Virtual Organizations (VO), a core concept in Grids [19], is seen less mature but also less important. It seems that for HMR Grids the current solutions are operationally sufficient. There is also a tendency to look at VOs from a broader perspective like the Virtual Research Communities (VRC) in the sense of the e-Infrastructure Reflection Group's (e-IRG) 2011 White Paper [14].

Analyzing the responses in a bit more detail reveals an expectation of a relatively high maturity level within the next three years for efficient data transfer services (66% of the respondents), for data access and monitoring services (47%), for data format conversion services (46%), and for High Throughput Computing (HTC) infrastructure interoperability (45%). On the other hand, there are critical services where the respondents do not anticipate any progress within the next three years. Examples are data validation and quality assurance services (52% of the respondents), services for Quality of Service (QoS) provisioning (46%), services to support quality assurance of scientific data (38%), metadata sharing across organizations (36%), services to support scientific data validation (36%), and workflow validation services (35%).

A critical aspect is the usability of an HMR Grid. Nearly all respondents (from both communities) claim a significantly more user driven approach ("Most of the developers seem to develop what they want to develop, but not what is really needed. No one really seems to put the user into the focus."). Technically, this is often expressed as a severe difficulty to add new application components (e.g., models) to existing workflows or as a lack of adequate data provenance management.

Finally, there has been reported a set of issues related to non-functional requirements like dependability, fault tolerance, standardization, and interoperability (with special emphasis of the latter on INSPIRE [16], OGF PGI [34], EDGI [15] and OpenMI [22]. In this context the respondents could also present a list of (subjectively) underestimated technologies. Among other they mentioned the UNICORE Grid middleware, Apache Hadoop for scalable data storage, Web-enabled Geographic Information Systems (GIS), the Internet of Things (sensors), semantic content management and ontologies, data mining techniques, and Clouds.

2.3. HMR Grid Usage Patterns

From the questionnaire responses it becomes apparent that simply closing these gaps technically is not enough for a successful application of Grid technologies in hydro-meteorological research. An additional source of requirements is therefore derived from observing the way HMR scientists would use Grid services. The joint HMR/ICT poll revealed four basic Grid usage trends with decreasing importance: The need for probabilistic forecasting using large ensembles of forecasts to deal with uncertainties; the need for model verification metrics with advanced spatial error; the need for advanced data merging/fusion techniques; and the need for precipitation downscaling using statistical and dynamic methods. These trends, however, are accompanied by a dramatic increase in both the quantity and the complexity of the tools and data sets available for hydro-meteorological research. There are basically three reasons for this development: Remote sensing observations from satellites and ground-based radars are increasingly providing complete multi-dimensional views of the atmosphere and land; ensemble forecasting methods that combine multiple numerical weather predictions and hydrological models to quantify the uncertainty in the forecast are multiplying the computational costs; a more thorough recognition of the need to understand the entire forecasting chain, from observations to civil protection responses, will result in complex workflows combining different data sets, models and expertise in a flexible manner.

One result of the DRIHMS project is the outline of such a forecasting chain as schematically depicted in Figure 3 [23]: The *rainfall layer* (top level) pertains to the combination of different Numerical Weather Prediction (NWP) models to build a high-resolution multi-model ensemble to enable the production of quantitative rainfall predictions for severe meteorological events. The *discharge layer* (mid level), however, concerns the fusion/combination of rainfall predictions (from the *rainfall layer*) with corresponding observations in order to enable river discharge predictions. Finally, the *water level, flow and impact layer* (bottom level) addresses the execution of hydraulic model compositions in different modes in order to assess the water levels, flow and impact created by flood events and to compare them against observations using adequate modeling verification metrics.

Given such HMR Grid usage patterns, the major objective of an HMR Grid architecture is to provide the adequate tools and to facilitate their ease of use. While HMR scientists ideally devote their time to accomplish tasks like phenomena exploration, initial analysis, conclusion derivation and report publishing, the current practice looks a bit different: The daily work of HMR scientists is often bounded to preparatory activities like finding data, retrieving huge data sets, learning formats, extracting parameters, identifying data qualities, filtering data, developing tools, and applying for access to HPC resources.

An HMR Grid may help to minimize these "manual" activities and to maximize the automated processes. Minimizing requires solutions at three different levels: Large data sets, whether from remote sensing instruments or from ensembles of numerical

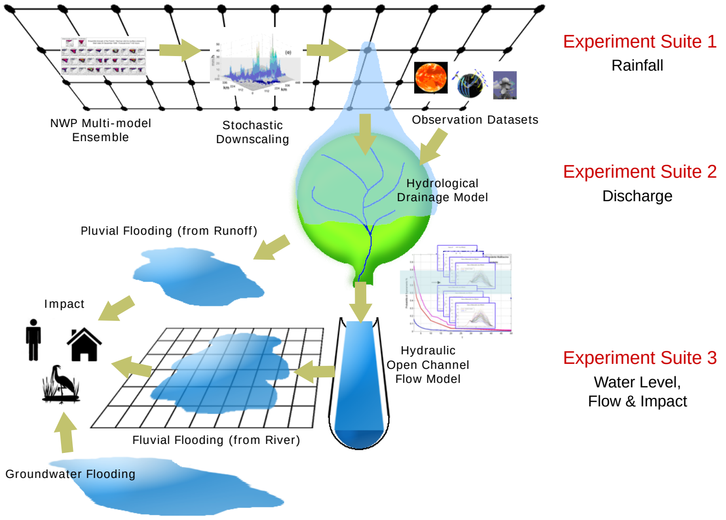

Fig. 3. Graphics rationale of the hydro-meteorology probabilistic forecasting chain [23]

weather forecasts, need to be accessible in a transparent and easy manner; models must be available with minimal effort; as workflows become an issue when treating the hydro-meteorological forecast chain as a whole, model interfaces and run time infrastructures are necessary to not only allow results from alternative tools to be fed into current computations, but also to compare different configurations without the need to continually develop new interfaces and analysis tools.

3. Outline of an HMR e-Science Environment Based on Grids

The DRIHMS surveys suggest a general rationale of an HMR e-Science environment as conceptually depicted in Figure 4 with HMR resources (e.g., hardware, models, simulation methods and tools), integrated services that are aiming at organizing the use of these resources, and dedicated HMR application services accessible via various portals.

As the currently existing European e-Infrastructures are designed to handle the processing of large amounts of distributed data and the sharing of resources in a co-ordinated way, the HMR applications (necessary for example to execute the hydro-meteorological probabilistic forecasting chain of Figure 3) would certainly benefit from such a design and the close interaction with EGI and PRACE as shown in Figure 5.

Figure 4 depicts a *conceptual* view of an HMR e-Science environment. As such, it does not describe the *technical* building blocks and how they interact over a Grid infrastructure.

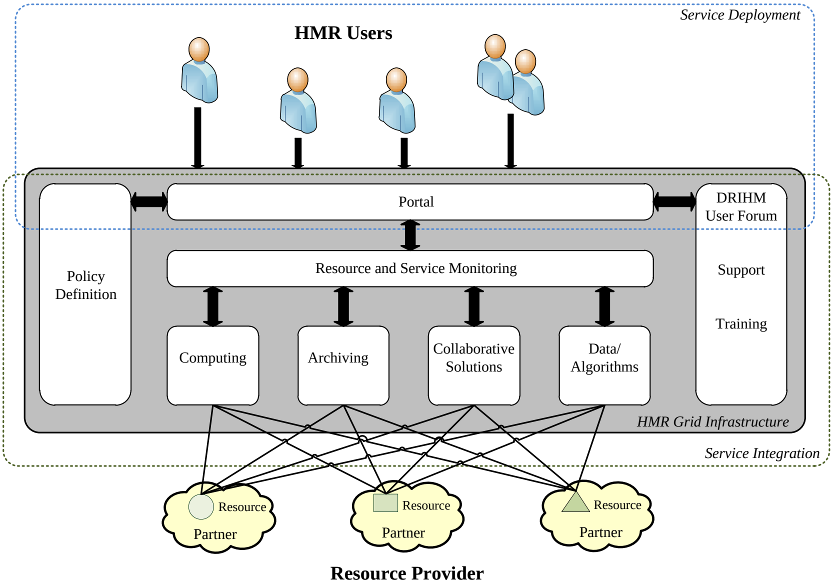

Fig. 4. Conceptual view of the HMR e-Science environment

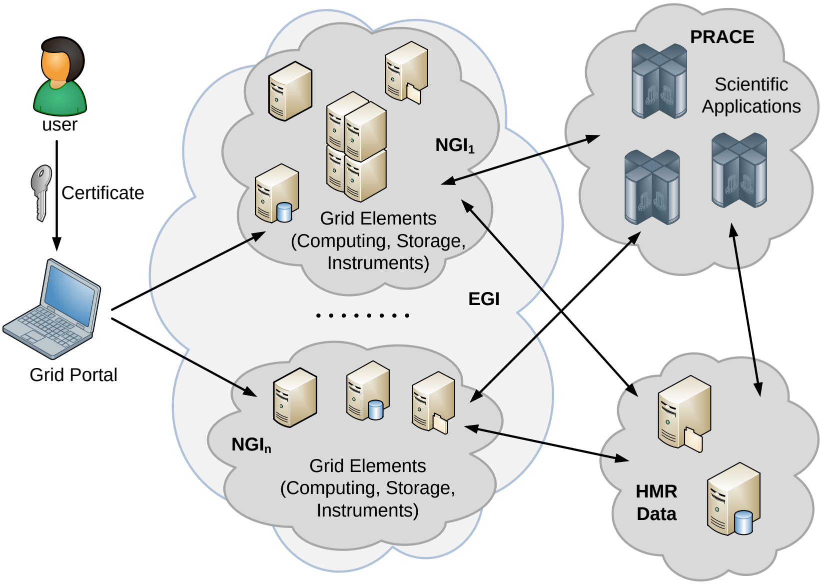

Fig. 5. General HMR e-Science environment

This is the purpose of the *architectural* view shown in Figure 6. It is organized using the well-known layer pattern for implementing the inevitable separation of concerns.

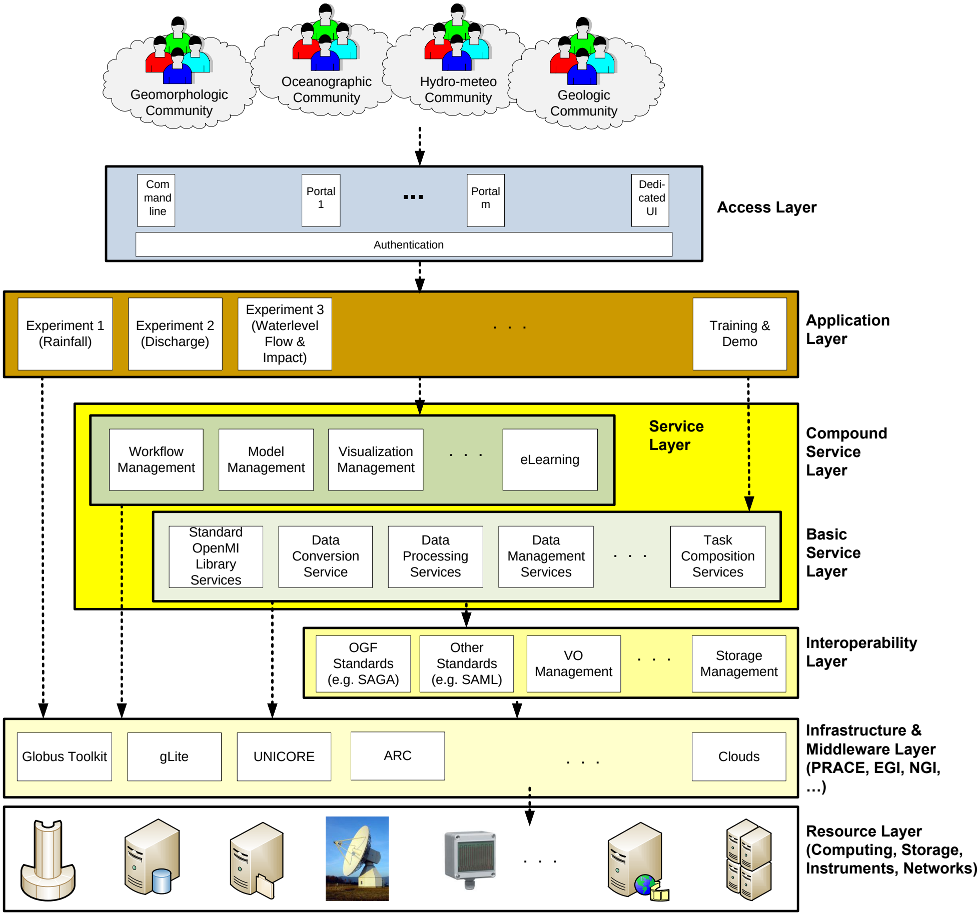

Fig. 6. Architecture of an HMR Grid

The *Resource Layer* contains all physical and logical resources (computing elements, storage elements, instrument elements, sensor networks, communication networks, licenses) as provided and operated by various Grid Resource Providers over the e-Infrastructures like PRACE, EGI or the NGIs. For accessing them, a respective Grid middleware system is required. Some of the resources may also be available in a Cloud (public or private). The middleware and the access services are part of the *Infrastructure and Middleware Layer*.

In order to cope with the inherent heterogeneity of Grid resources and Grid middleware, an *Interoperability Layer* assists in transparently accessing the resources. The interoperability layer is based on standards like those defined in the Open Grid Forum, OGF [33], or in the Organization for the Advancement of Structured Information Standards, OASIS [35].

The HMR related services comprise the *Service Layer* which itself is separated in a *Basic Service Layer* and a *Compound Service Layer*. The Service Layer typically interacts with the Interoperability Layer for transparently accessing the HMR resources. However, it may also interact with the Middleware Layer directly. While the Basic Service Layer provides all fundamental services (for example data conversion, model access or task composition), the Compound Service Layer deals with more complex services assembled from the basic ones. Typical examples of the latter are the creation and the management of workflows, the management of model sets, or the visualization of simulation results.

The *Application Layer* uses services from the Service Layer for specific HMR applications like simulations, experiments and e-Learning activities. The Application Layer is accessed via the *Access Layer* which provides the required capabilities to authorize the access to the HMR applications via various portals or clients and through an authentication interface.

4. Outline of a Derived Research Agenda

4.1. Guiding Roadmap

A roadmap towards HMR Grids requires to distinguish first between technical and administrative issues; second between infrastructure/data operations and their management in the context of the Grid basic architecture (see Figure 6); and third between basic hydro-meteorological *production* applications and HMR *experimental* setups. It is obvious that such a roadmap defines a long-term process with short- and mid-term milestones. The following HMR Grid roadmap has been triggered on the one hand by the survey responses and on the other hand by similar efforts, especially Cossu at al. [9]. However, compared to these works, the focus here is on using Grid infrastructures rather than on Grid deployments per se.

The starting point for the roadmap is implicitly given by the respective questionnaire results, whereas the direction of the roadmap follows the gap analysis (see section 2).

One of the major objectives to be addressed *short-term* is community building and standardization work via bodies like the OGF and the Outreach and Community Adoption Program which is part of the Open Geospatial Consortium [36]. As community building exhibits strong governance components, many organizations and research groups need to participate (by granting access to data or other scientific equipment or by installing Grid nodes, Grid access middleware, tools, datasets, etc.). In close relationship to community building activities, specialized Grid training programs are needed with special focus on HMR issues and on exploring the technical and operational benefits of Grid infrastructures (e.g., single sign-on authentication, policy-based authorization). For a successful deployment of HMR services at large, the identification and discovery of data and models across communities is needed. A short-term goal will thus be a standardized way not only to describe, address and discover such entities but also to describe, address and discover HMR related competence centers and specialized Resource and Service Providers.

The main concern of the *mid-term* period will be to overcome the blocking issues for HMR applications when deploying them on a Grid infrastructure, the porting of existing measurement and sensing applications from different disciplines, and the easy access to HMR Grid services. Other issues that need a thorough mid-term investigation refer to data retrieval, access to storage in near real-time and fast data transfer. Especially the latter is important as there is the necessity to transfer huge amounts of data in a great variety of formats using heterogeneous access protocols.

The *long-term* objectives are fundamental for the real exploitation of HMR Grids. They are mainly related to a sophisticated end user engagement rather than to the deployment of a particular technology. Thus, special emphasis should be given to the ease of adoption, the ease of use, the hiding of complexity, and the support of computationally expensive scientific workflows. Due to the time range involved long-term objectives can only be recommendations which heavily depend on the achievement of critical mid-term objectives.

When defining a roadmap it is important to note that it will not be a static sequence of fixed steps. Rather it should be considered as a set of research and development activities that may benefit from scientific progresses in various areas. In the following we briefly discuss four example research areas.

4.2. Example Research Area: Mashups

The main gaps identified by the ICT questionnaire are associated with data-related issues such as interoperability, availability, and extensiveness (amount and size of data). Moreover, the survey also pointed out a lack of commonly accepted tools for discovering, interchanging, and merging scientific data from different sources (satellites, private meteo-stations) for processing and visualization. As a matter of fact, a number of applications and projects have already been emerging to retrieve and visualize vast amounts of earth science data from remote sensing observations and models [28, 29].

In the last decade, however, the exploitation of Web Services technologies along with new software engineering paradigms like Service Oriented Architectures (SOA) fertilized the development of a huge amount of services delivered through and accessed via the Web. The use of Web technologies to explore geo-scientific data is now well-established [2] and public and private sources of hydro-meteorological data are freely available from weather services such as NOOA [32] or YahooWeather [46].

What is missing, however, is the ability to aggregate and combine this heterogeneous information in some flexible way according to the continuously changing HM scientist needs.

A promising approach is based on the "mashup" concept [6] which permits to combine data and/or programs published by external online sources into an integrated experience. In other words, a mashup application may grab data from one place in the Web, augment it with relevant information from other places and present it as a single application – including mechanisms for estimating the uncertainty and the bias of data. Mashups have a number of advantages and disadvantages. From the HM researcher's perspective a well-designed mashup system may represent an almost ideal solution – not least due to the ability to reuse existing data and services [4]. Future research results in developing mashups may also boost the design of data oriented Grids.

4.3. Example Research Area: Optimization

The accuracy in spatio-temporal modeling of atmospheric scenarios inducing intense rainfall processes is very important both from a research and an operational (Civil Protection, early-warning activities) point of view. The importance has evidenced as extreme weather conditions strongly influence human activities and are damage prone. Many government institutions invest a lot of resources in guaranteeing safety for the population but efforts in this direction not always provide the desired results. This is due to both the significant time requirements to alert civilians (at least six hours before the expected injurious event) and to the ability to accurately forecast future weather scenarios.

To tackle the second issue, hydro-meteorological science has made strong progress over the last decade as new modeling tools, post processing methodologies and observational data have emerged. The available weather forecast tools include numerical weather prediction models and meteorological observations. The former are necessary to predict the state of the atmosphere whereas the latter are typically used to initialize the models themselves.

However, strong scientific and technical challenges are still associated with using such tools. One such challenge is the actual connection between hydro-meteorological models and the real observations, i.e. the possibility of using different computational modules [27]. Another is the lack of expertise to efficiently exploit HPC resources [7, 37].

A first step to fill this gap would be running high-resolution weather forecast models and comparing the results with observed data on HPC facilities. Exploiting HPC and Grid resources would significantly improve the space-time predictability of extreme convective events by a formally sound characterization of precipitation forecasts in terms of kinematics, thermodynamics and microphysics properties, and by the evaluation of the uncertainties related to high-resolution modeling [20]. Understanding the potentials and the limitations of space-time predictability of extreme convective events through the combined use of high-resolution meteorological models and hydro-meteorological observations by means of Grid and HPC resources is an open research field.

4.4. Example Research Area: Software Components for Model Coupling and Software Adaptivity

The demand for model coupling tools, post processing tools and other software modules in dynamic workflows gives evidence to the necessity of wrapping software modules into suitable workflow components. Component model design for scientific computing has been an active research area since several years [1, 3]. Several research challenges are related to an efficient and effective component model design: the trade-off between flexibility and performance requiring a thorough understanding of interface definitions and coupling mechanisms between components; the trade-off between local and distributed execution of components; the appropriateness of the overall architecture of a component model; the adequacy of the framework that implements such an architecture; or the suitability of the underlying infrastructure. In a recent survey [27], Jagers reviewed different component models (including models particularly suited for hydrological or environmental studies such as OpenMI [22] and The Earth System Modeling Framework, ESMF [44]).

Component model design is still a huge research area which also covers the field of adaptive software implementation where different implementations are associated with a single interface. Adaptive software behavior is obtained by dynamically triggering the "best suited" implementation – based on specific application parameters or on the availability of computing resources. An example is calling a component via a single interface while the execution environment is dynamically selected from a sequential, a cluster based parallel implementation, or an implementation based on General Purpose computing on Graphics Processing Units (GPGPU), depending on the size of the input data set and/or on the availability of computing nodes. A specific example is described below in the context of visualization.

4.5. Example Research Area: Remote and High Performance Visualization

Within the HMR community interactive visualization of 3D data is of paramount importance for the understanding and conveyance of both data and simulation processes [45, 42].

As for scientific workflows in general, visualization instruments also have to deal with an increasing amount of data generated by modeling tools, post-processing methodologies, and more sophisticated acquisition instruments. A successful visualization system for HMR should therefore address two aspects: The capability of quickly combining heterogeneous building blocks into workflows in a dynamic way while considering application and resource characteristics; the capability of exploiting the concept of adaptive implementation of components, in order to reduce the computing time and to more efficiently support interactive applications. For both aspects, an adoption of component-based frameworks [5] with dedicated core services [41] and formal descriptions of data and models [24] seems promising.

Traditionally, the efficient processing of large amounts of data was performed by parallel HPC clusters. However, the present evolution of multi- and many-core architectures allows turning smaller workstations into massively parallel-computing devices. As such, they represent a cost-effective instrument for processing and visualization of volumetric data. For example, isosurface extraction is a basic operation that allows querying and visualization of 3D data in HMR workflows. Depending on the volume size, this operation may require a considerable amount of computing power on different architectures [13].

5. Related Projects

A number of hydro-meteorology research projects have been financed by the European Commission in the last years. Although some of these projects overlap in parts with the idea of an HMR Grid, their focus was completely different. However, all of these projects contributed to the surveys described above.

- HyMeX (Hydrological cycle in the Mediterranean eXperiment, [25]) aims at better understanding and quantifying the hydrological cycles and processes in the Mediterranean, with special emphasis on high-impact weather events, inter-annual to decadal variability of the Mediterranean system, and associated trends in the context of global change;
- COST 731 (Propagation of uncertainty in advanced meteo-hydrological forecast systems, [10]) addresses issues associated with the quality and uncertainty of meteorological observations from remote sensing and other potentially valuable instrumentation, as well as their impacts on hydro-meteorological outputs from advanced forecasting systems;
- MAP-D-PHASE (Mesoscale Alpine Programme – Demonstration of Probabilistic Hydrological and Atmospheric Simulation of flood Events, [30]) demonstrates some of the many achievements of the Mesoscale Alpine Programme (MAP), in particular the ability of forecasting heavy precipitation and related flooding events in the Alpine region. It addressed the entire forecasting chain ranging from limited-area ensemble forecasting, high-resolution atmospheric modeling (km-scale), hydrological modeling, and nowcasting to decision making by the end users, by setting up an end-to-end forecasting system.
- MEDEX (MEDiterranean Experiment, [31]) was designed to contribute to the better understanding and short-range forecasting of high impact weather events in the Mediterranean, mainly heavy rain and strong winds;
- FP6 PREVIEW (Prevention Information and Early Warning, [38]) was looking for new techniques to better protect European citizens against environmental risks and to reduce their consequences;
- CLIVAR (CLImate VARiability and predictability: a programme of the World Climate Research Programme, [8]) develops a better understanding of climate variability and applies this to provide useful prediction of climate variability and change through the use of improved climate models;

- FLOODsite (Integrated Flood Risk Analysis and Management Methodologies, [18]) covers the physical, environmental, ecological and socio-economic aspects of floods from rivers, estuaries and the sea. It considers flood risk as a combination of hazard sources, pathways and the consequences of flooding on people, properties and the environment;
- GMES (Global Monitoring for Environment and Security, [21]) is dedicated to the monitoring and forecasting of the Earth's subsystems affected by climate change. GMES services also address emergency response (e.g. in case of natural disaster, technological accidents or humanitarian crises) and security-related issues (e.g. maritime surveillance, border control);
- IMPRINTS (Improving Preparedness and Risk maNagemenT for flash floods and debriS flow events, [26]) contributes to the reduction of loss of life and economic damage through the improvement of the preparedness and the operational risk management of flash flood and debris flow generating events, as well as contributing to sustainable development through reducing damages to the environment;
- FloodProBE (Flood Protection of the Built Environment – Technologies for Improved Safety of the Built Environment in Relation to Flood Events, [17]) develops technologies, methods and tools for flood risk assessment and for the practical adaptation of new and existing buildings, infrastructure and flood defenses leading to a better understanding of vulnerability, flood resilience and defense performance;
- PRISM (Partnership for Research Infrastructures in earth System Modeling, [39]) supports sharing the development, maintenance and support of standards and state-of-the-art software tools to assemble, run, and analyze the results of Earth System Models based on component models (ocean, atmosphere, land surface, etc.) developed in the different climate research centers in Europe and elsewhere.

6. Conclusion

In this paper we presented the results and the implications of a recent survey conducted to poll the specific requirements for an HMR Grid, a Grid infrastructure devoted solely to the scientific activities in the hydro-meteorological research community. From the survey we could not only derive the hot topics, we could also identify the main gaps to fill, a simple roadmap towards an HMR Grid, and a first draft of an architectural pattern. We also presented a list of related European projects which contributed to the findings. In addition, we drafted a dedicated research agenda to be tackled over the next years.

Acknowledgements

This work has been funded by the Seventh Framework Program of the European Commission (Grant 246703).

References

[1] Allan B. A., Armstrong R., Bernholdt D. E., Bertrand F., Chiu K., Dahlgren T. L., Damevski K., Elwasif W. R., Epperly T. G. W., Govindaraju M., Katz D. S., Kohl J. A., Krishnan M., Kumfert G., Larson J. W., Lefantzi S., Lewis M. J., Malony A. D., McInnes L. C., Nieplocha J., Norris B., Parker S. G., Ray J., Shende S., Windus T. L., Zhou S. A.: *Component Architecture for High-Performance Scientific Computing.* International Journal of High Performance Computing Applications, 20(2), 2006, pp. 163–202.

[2] Amagasa T., Kitagawa H., Komano T.: *Constructing a Web Service System for Large-scale Meteorological Grid Data.* [in:] Proceedings of the 3rd IEEE International Conference on e-Science and Grid Computing (e-Science 2007), 2007, pp. 118–124.

[3] Armstrong C. W., Ford R. W., Riley G. D.: *Coupling Integrated Earth System Model Components with BFG2.* Concurrency and Computation: Practice & Experience, 21(6), 2009, pp. 767–791.

[4] Bedrina T., Quarati A., Clematis A., Parodi A., Rudari R.: *Using of Mashup Technology for Hydro-Meteorological Research.* Geophysical Research Abstracts, 13, EGU2011-13733, 2011. European Geosciences Union General Assembly 2011, Vienna, Austria.

[5] Castronova A. M., Goodall J. L: *A Generic Approach for Developing Process-Level Hydrologic Modeling Components. Environmental Modelling & Software*, 25(7), 2010, pp. 819–825.

[6] Chow S.-W: *PHP Web 2.0 Mashup Projects.* Packt Publishing, ISBN 978-1-84719-088-8, 2007.

[7] Clematis A., DAgostino D., Galizia A., Quarati A., Parodi A., Rebora N., Kranzlmller D., Schiffers M.: *Selecting High Performance Computing and High Throughput Computing Capabilities for Hydro Meteo Research e-Instrastructures.* Invited Presentation at the IEEE International Workshop on High Performance Computing on Complex Environment – HPCCE 2010, Heraklion, Greece, 2010.

[8] *CLIVAR*: `http://www.clivar.org/`. (Last checked July 27, 2011).

[9] Cossu R., Petitdidier M., Linford J., Badoux V., Fusco L., Gotab B., Hluchy L., Lecca G., Murgia F., Plevier C., Renard P., Schwichtenberg H., Som de Cerff W., Tran V., Vetois G.: *A Roadmap for a Dedicated Earth Science Grid Platform.* Earth Science Informatics, 3(3), 2010, pp. 135–148.

[10] COST 731: `http://cost731.bafg.de/servlet/is/Entry.9691.Display/`. (Last checked July 27, 2011).

[11] Craig G. C., Parodi A.: *Results of the Questionnaires HMR.* Deliverable D2.2 of the DRIHMS Project; `http://www.drihms.eu/project/material/Deliverable_D2.2_revised_DRIHMS.pdf`, 2010. (Last checked July 27, 2011).

[12] *DRIHMS Project Web Site*: `http://www.drihms.eu`. (Last checked July 27, 2011).

[13] D'Agostino D., Seinstra F. J.: *An Efficient Isosurface Extraction Component for Visualization Pipelines Based on the CUDA Architecture.* Technical Report IR-CS-64-2010, Vrije Universiteit, Amsterdam, The Netherlands, 2010.

[14] *e-Infrastructure Reflection Group*: e-IRG White Paper 2011. Report of the e-Infrastructure Reflection Group, 2011.

[15] *EDGI Project*: `http://edgi-project.eu/`. (Last checked July 27, 2011).

[16] *European Commission, INSPIRE Directive*: `http://inspire.jrc.ec.europa.eu/`. (Last checked July 27, 2011).

[17] *FloodProBE*: `http://www.floodprobe.eu/`. (Last checked July 27, 2011).

[18] *FLOODsite*: `http://www.floodsite.net/`. (Last checked July 27, 2011).

[19] Foster I., Kesselman C., Tuecke S.: *The anatomy of the grid: Enabling scalable virtual organizations.* International Journal of High Performance Computing Applications, 15(3), 2001, pp. 200–222.

[20] Galizia A., Queijo C. C.: *Exploiting the Finis Terrae Supercomputer to Improve the Characterization of Spatio-Temporal Predictability of Severe Weather Events.* Geophysical Research Abstracts, 13, EGU2011-13791, 2011. European Geosciences Union General Assembly 2011, Vienna, Austria.

[21] *GMES*: `http://www.gmes.info/`. (Last checked July 27, 2011).

[22] Gregersen J. B., Gijsbers P. J. A., Westen S. J. P.: *OpenMI: Open Modelling Interface. Journal of Hydroinformatics*, 9(3), 2007, pp. 175–191.

[23] Harpham Q: *Private Communication.* 2011.

[24] Horsburgh J. S., Tarboton D. G., Maidment D. R., Zaslavsky I.: *A Relational Model for Environmental and Water Resources Data. Water Resources Research*, 44(5), 2008.

[25] *HyMeX*: `http://www.hymex.org/`. (Last checked July 27, 2011).

[26] *IMPRINTS*: `http://www.imprints-fp7.eu/`. (Last checked July 27, 2011).

[27] Jagers H. R. A.: *Linking Data, Models and Tools: An Overview.* [in:] Proceedings of the 2010 International Congress on Environmental Modelling and Software Modelling for Environments Sake, Ottawa, Canada. 2010.

[28] Keiichirou U., Toshiyuki A., Hiroyuki K.: *A FUSE-Based Tool for Accessing Meteorological Data in Remote Servers.* [in:] Proceedings of the 20th International Conference on Scientific and Statistical Database Management, volume 5069 of *Springer Lecture Notes in Computer Science*, Springer, Berlin, 2008, pp. 592–597.

[29] Le Phuoc D., Parreira J. X., Hausenblas M., Han Y., Hauswirth M.: *Live Linked Open Sensor Database.* [in:] Proceedings of the 6th International Conference on Semantic Systems (I-SEMANTICS 10), Graz, Austria. 2010.

[30] *MAP–D–PHASE*: `http://www.map.meteoswiss.ch/map-doc/dphase/dphase_info.htm`. (Last checked July 27, 2011).

[31] *MEDEX*: `http://medex.aemet.uib.es/`. (Last checked July 27, 2011).

[32] *National Oceanic and Atmospheric Administration*: `http://www.opengeospatial.org/ogc/programs/ocap`. (Last checked July 27, 2011).

[33] *Open Grid Forum*: `http://www.ogf.org/`. (Last checked July 27, 2011).

[34] *Open Grid Forum Production Grid Infrastructure*: `http://www.ogf.org/gf/group_info/view.php?group=pgi-wg`. (Last checked July 27, 2011).

[35] *Organization for the Advancement of Structured Information Standards*: `http://www.oasis-open.org/`. (Last checked July 27, 2011).

[36] *Outreach and Community Adoption Program, Open Geospatial Consortium*: `http://www.opengeospatial.org/ogc/programs/ocap`. (Last checked July 27, 2011).

[37] Ploski J., Scherp G., Petroliagis T. I., Büchner O., Hasselbring W.: *Grid-Based Deployment and Performance Measurement of the Weather Research & Forecasting Model.* Future Generation Computer Systems, 25(3), 2009, pp. 346–350.

[38] PREVIEW: `http://www.preview-risk.com`. (Last checked July 27, 2011).

[39] PRISM: `http://www.prism.enes.org`. (Last checked July 27, 2011).

[40] Quarati A., Rebora N., Schiffers M.: *Filling the Gap Between ICT and HydroMeteo Research Communities: The DRIHMS Project. ERCIM News*, (85), 2011, pp. 55–56.

[41] Rao I., Huh E., Lim S.: *An Adaptive and Efficient Design and Implementation for Meteorology Data Grid Using Grid Technology.* [in:] Proceedings of the 15th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE 06), Manchester, United Kingdom, 2006, pp. 239–246.

[42] Sanyal J., Zhang S., Dyer J., Mercer A., Amburn P., Moorhead R. J.: *Noodles: A Tool for Visualization of Numerical Weather Model Ensemble Uncertainty. IEEE Transactions on Visualization and Computer Graphics*, 16(6), 2010, pp. 1421–1430.

[43] Schiffers M., Parodi A.: *Results of the Questionnaires Grid.* Deliverable D3.2 of the DRIHMS Project; `http://www.drihms.eu/project/material/Deliverable_D3.2_revised_DRIHMS.pdf`, 2010. (Last checked July 27,2011).

[44] *The Earth System Modeling Framework*: `http://www.earthsystemmodeling.org/`. (Last checked July 27, 2011).

[45] Van der Z. J., Vickery R. J., Moorhead R.: *Vortex Detection Through the Visualization Toolkit.* [in:] Proceedings of the 2009 DoD High Performance Computing Modernization Program User Group Conference, San Diego, USA, IEEE Computer Society, 2009, pp. 395–397.

[46] YahooWeather: `http://weather.yahoo.com`. (Last checked July 27, 2011).