A strategic vision for the PRISM software infrastructure

This document sets out the key software infrastructure challenges to be addressed by the European Earth System Modelling community over the next 5 years and proposes the organisation of a sustained support to address them. We first list the main drivers of Earth system modelling before describing the concept of the "PRISM software infrastructure", ands its key challenges. The current status of developments is reviewed, the structure and organisation of sustained support for PRISM is proposed, and the document concludes with an implementation plan devised for effective delivery of the objectives.

1. Background and drivers

1.1 Earth system modelling: the challenges ahead.

The continued development and refinement of computational models that simulate the evolution of the Earth system is critical for climate research and prediction. Those models are invaluable tools for the scientific understanding of climate and some of them are also routinely used for numerical forecast purposes. They are critical for environmental impact assessment and provision of adequate information to decision makers. Significant progress has been made in the development of Earth system models (ESMs) over the last decade. The European scene of distributed research centres has been very successful in using high-end modelling as a tool for discovery-driven research. This resulted in diverse and creative approaches and contributed to Europe's leadership in the field. In the meantime however, the software environment of climate modelling has become exceedingly intricate because of the complexity of both the natural systems they describe and the computational platforms they use. If the diversity of approaches has ensured continued improvements in basic understanding of the natural processes, it has proven not to be equally beneficial for efficient software systems development. Modellers and climate scientists around the world realise today that greater software standardisation and isolation from scientific cores would facilitate their research [1], [2], [3], [4]. Likewise, standard data formats, a web-accessible data network and Grid adapted access software would accelerate progress by facilitating broader exchange, dissemination and shared analysis of model results. Staying at the forefront of scientific research in Earth system sciences requires to address these new challenges.

To ensure an adequate level of service to the research community, the development of such software standardisation requires a multi-institute product-driven organisational structure, i.e. a jointly funded team providing maintenance, user support and best practices substantially different from those used in discovery-driven research activities.

1.2 The PRISM concept

Today, Earth System modellers share very little beyond what is provided by computer manufacturers: compilers, message passing libraries, etc. However, for efficiency, they should share a lot more. The PRISM concept, initiated by a Euroclivar recommendation [1], is to share the development, maintenance and support of a comprehensive Earth System Modelling software environment. This is key to facilitate assembling, running, archiving and post-processing of ESM based on state-of-the-art component models developed in the different climate research centres in Europe and elsewhere. The PRISM software environment includes a number of technical standards for existing and future ESM internal composition and external control. It supports model diversity and promotes scientific progress of the community by lowering technical development efforts of ESM teams.

The extensive use of the OASIS coupler illustrates the benefits of a successfully shared software infrastructure [5]. In 1991, CERFACS was commissioned to realise a software for coupling different geophysical component models developed independently by different research groups. Flexibility, modularity, and adaptability naturally emerged as key design concepts. The OASIS

development team also strongly focussed on efficient user support and constant integration of the developments fed back by the users. This interaction snowballed and resulted in a constantly growing community. Today, OASIS capitalises about 20 person years (py) of mutual developments and fulfils the coupling needs of about 15 climate research group around the world. The effort invested in OASIS therefore represents, on a first order, 20 py/15 groups = 1,3 py/group, which is certainly much less than the effort that would have been required by each group to develop its own coupler. Today, OASIS is at the core of the PRISM software (see section 2.2).

A first version of the "PRISM software" is currently being developed under European Framework Programme V funding and is now emerging as a core strategic infrastructure for building the European Research Area in Earth System sciences. PRISM represents the first major collective effort, at the European level, to develop ESM supporting software in a shared and coherent way, as is recognised by the Joint Scientific Committee (JSC) of the World Climate Research Programme (WCRP) that recently endorsed its value as a "key European infrastructure project" [6].

1.3 The link with operational forecasting and other climate-related fields of expertise.

A strength of the climate research community is its links with operational forecasting. This also comes with constraints as several component models share similar infrastructure for both climate and forecasting deployments. Today, the technical and service levels required by operations have higher standards, with priority given to security and stability aspects. Climate modelling, on the other hand, first requires flexibility and extensibility to adapt to the needs of new models, platforms or scientific approaches. Recognising this implies that the convergence of software infrastructure at the community level can only be a long-term goal. In the short term, key parts of the PRISM software can be used in operations. This possibility will be explored in several upcoming projects (like the FP6 project GEMS [7]) paving the way to a future long-term convergence. Climate modelling also more and more interacts with other climate-related fields of expertise, like impact assessment or economic studies, and similar convergence is also desirable with the related software infrastructure.

2) Situation today

2.1 The PRISM project

Recognising the need for shared software infrastructure, the European Network for Earth System Modelling (ENES) organised the PRISM project, which is currently funded for 3 years (starting December 2001) by the European Union under the 5th Framework Programme. The PRISM project gathers 22 partners, including the main European climate modelling institutions and four computer manufacturers. It has an overall budget of 4.8 MEur, corresponding to a total effort close to 80 py.

One main objective of PRISM is to provide a portable, user-friendly, flexible, and standard based infrastructure for assembling, compiling, running, monitoring and post-processing Earth System Models built on state-of-the-art component models developed in the different European modelling groups. Structured data archiving and dissemination is not part of the project.

2.2 PRISM achievements

Today, PRISM provides as standard software:

1. a standard coupler and I/O software, OASIS3 [5]
2. a standard compiling environment (SCE) at the scripting level [8]
3. a standard running environment (SRE) at the scripting level [8]
4. Graphical User Interfaces (GUIs) to configure the SCE and SRE [9]
5. a GUI for an end-to-end monitoring of climate experiments [9]
6. standard diagnostic and visualisation tools [10]

Currently, the PRISM SCE and SRE are configured for 9 European sites, and their respective GUIs are implemented on two sites, at the Max Planck Institute (MPI) in Hamburg and at the European Centre for Medium-range Weather Forecast (ECMWF) in Reading. At the end of the project and as a major delivery, a new coupler and I/O software will be available (OASIS4); thanks to the expertise gathered in PRISM, it will benefit from the new opportunities offered by modern software design.

The PRISM project partners will provide an open source access to all PRISM software, and cost-free access to the monitoring GUI for research purposes in Europe.

Most state-of-the-art European component models have been adapted, or are being adapted, at different levels, to the PRISM standards:

- atmosphere-land models: **ECHAM5** (MPI, Germany), **ARPEGE4** (Météo-France, France), **LMDz** (Institut Pierre Simon Laplace - IPSL, France), and **HadAM3** (Hadley Centre, UK)
- ocean and sea-ice models: **MPI-OM** (MPI, Germany), and **ORCA-LIM** (IPSL, France, and Université Catholique de Louvain – UCL, Belgium)
- marine biogeochemistry models: **HamOCC** (MPI, Germany), and **PISCES** (IPSL, France)
- atmospheric chemistry models: **MOZART** (MPI, Germany), **MOCAGE** (Météo-France, France) and **TM** (Royal Netherlands Meteorological Institute - KNMI, The Netherlands)
- regional atmosphere models: **HIRLAM** (Danmarks Meteorologiske Institut - DMI, Denmark) and **RCA** (Swedish Meteorological and Hydrological Institute – SMHI, Sweden)

A number of combinations of these component models are assembled for the PRISM demonstrations [11] on the main community platforms (NEC-SX, SGI-Origin, Fujitsu-VPP, IBM-Power4): ECHAM5 + MPI-OM; ECHAM5 + ORCA-LIM; ARPEGE4 + ORCA-LIM; ; LMDz-ORCHIDEE + ORCA-LIM; HadAM3 + ORCA-LIM; ORCA-LIM+ PISCES; MPI-OM + PISCES; MPI-OM + HamOCC; ECHAM5 + HIRLAM; ECHAM5 + RCA*;* ECHAM5 + MOZART*;* ECHAM5 + MOCAGE; Full ESM ECHAM5 + MOZART + MPI-OM + HamOCC*;* Full ESM ECHAM5 + MOZART + MPI-OM + PISCES,...

2.3 PRISM: the success of a well co-ordinated network of experts

Although PRISM was designed as a demonstration project, its technical value is already recognised by many European research groups. First user experiences show that using the PRISM system eases the assembly, compilation and running of complex component models via the use of PRISM standards. Some of the ESM configurations described above are starting to be used both for local scientific projects and wider community undertakings (IPCC runs, ENSEMBLES FP6 project, German COSMOS project,...)

Besides those technical achievements, one important success of PRISM is that it has brought the different partners of the European Earth system research community to interact and work closely together. This led to invaluable trust-building, naturally opening up into scientific co-operation and co-ordination. Today, this closely co-ordinated network of experts (IT specialists, climate scientists, computer manufacturers,...) is ready to go one step further by providing a routinely maintained state-of-the-art software infrastructure for the Earth System Modelling community.

3) A proposal for a sustained PRISM team

3.1 Why go for sustained support ?

The PRISM software described above needs to be maintained and constantly improved to fit the evolving needs of the Earth system modelling community. The system is also intended to integrate progressively more component models and data archive infrastructures, to be implemented on additional sites, and serve a wider community. Sustained staff and financial support is therefore needed to ensure the continued and co-ordinated maintenance of PRISM, together with an adequate level of user support, both required to guarantee a growing community buy-in and trust-building. Without sustained support, it seems inevitable that the PRISM software will diverge over time.

Local support teams are unlikely to be able to develop the same level of expertise across the whole software system and any result is likely to be more expensive and/or of lower standard than would be available by an expert co-ordinated team, as proposed here.
Sustained support is also key to

- attract highly qualified experts,
- draw additional EU and other temporary funding (the CAPRI project submitted to the FP6 showed that sustainability is a key evaluation criterion),
- work towards convergence of software infrastructures used in climate research and related fields of expertise, such as impact studies, scientific data assimilation and operational forecasting (see section 1.3).

The proposed goals for such a team and for the next 3 years are:

1. **Co-ordinate improvement, maintenance and user support of current PRISM Software** (see section 2.2);
2. **Support adaptation** of other component models to PRISM technical standards;
3. **Install** PRISM software environment at additional computing sites;
4. **Prepare for the future** by seeking additional funding and proposing development strategies.

3.2 Management and technical tasks, and associated costs

Different skills needed to fulfil these goals, and a first list is given below:

The management tasks are key in a distributed network of experts:

- Project management (including internal communication, meeting organisation)
- Technical development co-ordination (see below)
- User support (help-desk, tutorial, FAQ, documentation, workshops, etc.)
- Communication and dissemination
- Additional fund raising

The technical tasks are needed to maintain and develop the system:

1. Maintenance, improvement and quality control of existing software:
    - General architecture and system integration
    - Coupler and I/O systems
    - SCE and SRE
    - GUIs (SCE and SRE)
    - Data archive environment
    - Diagnostic and visualisation tools
2. Provision of services:
    - Central software and model repository management
    - User support[+] and best practice for
        a. new component model adaptation to the PRISM standards
        b. PRISM software environment installation at new computing sites
        c. new coupled combination assembling

Looking at existing support efforts for similar tasks, the PRISM Steering Group estimates that a minimum of 7 full-time persons[*] (for management, technical and support tasks) plus associated costs of 60 KEuro/year (travel, equipment, workshops, etc.) is needed to compose a first PRISM sustained team for the next 3 years.

[+] Even if active user support will be provided, some effort will still be required from the institutes to adapt and integrate a new component model in the PRISM environment, install the PRISM software at a new site and assemble a new coupled combination, as routinely done today.
[*] or 7 py per year

No sustained European infrastructure funding exists outside inter-governmental set-ups (as for ECMWF, for instance). Furthermore, EU Framework Projects structures have shown their limits in managing such complex and service-oriented infrastructures (lack of flexibility, difficulty to attract skilled experts, limitation in time,...). Therefore the structure and organisation described in the next section are proposed for the sustained PRISM team.

3.3 Organisation and structure

We propose to organize the PRISM sustained team via a Consortium formed by a number of funding institutions. Several types of contributions can be expected:

1. Major Earth System institutions substantially contributing to sustained team (one or more full-time person or equivalence in financial means);
2. Other Earth System institutions contributing less to sustained team;
3. Other model development teams contributing adapted models;
4. IT companies contributing software expertise (see 3.4);
5. Computing centres contributing support to install the PRISM environment.

The Executive committee (with representatives from the contributors – details to be defined) is in charge of the PRISM development strategy. The Executive committee appoints the PRISM team management. The procedure to appoint the PRISM technical team needs to be defined. The duration of the Consortium will be three years, renegotiated for three years every year. An advisory group, composed of internationally recognised experts will advise the Executive committee on an annual basis. A user group will be established to ensure that proper feedback is received from the community.

The PRISM technical team will be physically distributed among few "PRISM competence centres" (to be defined) to 1) benefit from existing distributed expertise and 2) allow a better integration in the scientific community, where as the PRISM management will be grouped for efficiency. The PRISM management structure can either be created via a new legal structure (maybe a mid term goal) or use an existing one (easier to set up).

A base level of service will be available to the Earth System modelling community (user support, adaptation support, help-desk, tutorials, workshops, etc...). The priority given to any additional level of service required by a group/institution will be a function of its level of contribution to the sustained team. If a development is seen as urgent for an institution, but not by the Consortium, the institution can provide temporary additional resources to the PRISM team to achieve it.

3.4 Detailed expected benefits

The establishment of such a sustained PRISM team in charge of a shared software infrastructure is a long term commitment. The expected mid and long term benefits are multiple and a few are listed here:

- Each institution benefits, at relatively low cost, from a high-performance ESM software infrastructure developed by dedicated IT experts. The efforts saved on technical developments can be invested in scientific work, ensuring 1) the survival of smaller scientific expertise groups, hence the key scientific diversity of models and 2) a continuous European leadership in climate sciences;
- Assembling ESMs based on community models coming from the different institutions is easier as component models are adapted to the same supporting software;
- Sharing a common software infrastructure naturally leads to increased interaction at all levels between the institutions. Experience shows that this facilitates co-ordination of climate science and paves the way to integration of efforts for supercomputing resources deployment at the European level;
- Interaction and co-ordination with equivalent international projects, such as ESMF in the USA [3] for instance, may be done officially at PRISM level;

- Computer manufacturers are inclined to work within or with this team, offering IT solutions for climate modelling and providing expertise for porting, optimising, and maintaining PRISM software on their platforms. Conversely, the platforms of the next generation can then be optimised for the ESM community needs, leading to significantly reduced computing costs.

4) Implementation plan

4.1 What is expected from institutions now ?

Institutions considering contributing to the proposed PRISM sustained team are invited to 1) provide feedback to this document to the PRISM Steering group (via ericg@met.rdg.ac.uk) and 2) attend a set-up meeting (in summer 2004, details to follow) where outstanding issues and estimated contributions will be discussed, and where a Letter of Intent (LoI) to work towards a Consortium Agreement will be prepared.

4.2 Additional documents planned

Two additional documents are planned in the months preceding the set-up of the Consortium: 1) a detailed work plan for the first year of the PRISM sustained team and 2) an external audit of the feasibility of the project, including risk assessment.

4.3 Milestones

- Set-up meeting to discuss details of project and finalize LoI:	summer 2004
- LoI signed by participating institutions	September 2004
- Work plan for first year available:	September 2004
- External audit of project available:	October 2004
- Draft Consortium Agreement available:	November 2004
- Signature of Consortium Agreement:	December 2004
- Start of project:	January 2005

REFERENCES

1- Euroclivar (1998). Climate Variability and Predictability in Europe, 1999-2004. Euroclivar recommendations, ISBN 90-369-2146-5. xxiv + 120 pp.
2- USGCRP (2000). High-End Climate Science: Development of modelling and related computing capabilities. http://www.usgcrp.gov/usgcrp/Library/models2001
3- ESMF (2002).ESMF: A High-Performance Framework for Earth Science Modeling and Data Assimilation, Washington D.C. http://www.esmf.ucar.edu/
4- Guilyardi E., Budich R., Komen G., Brasseur G. (2003). PRISM System Specification Handbook, Version 1. PRISM Report series No. 1, 230 pp.
http://prism.enes.org/Results/Documents/Handbook/handbook.v.1.0.3.pdf
5- Valcke S., Caubel A., Declat D., Terray L. (2003) OASIS3 Ocean Atmosphere Sea Ice Soil User's Guide. PRISM Report series No. 2, 54 pp.
http://prism.enes.org/Results/Documents/PRISM_Reports/Report2_oasis3_UserGuide.pdf
6- WCRP (2002). JSC 23rd session, Hobart, Australia, WMO/TD – No.1137
7- http://www.ecmwf.int/research/EU_projects/GEMS/
8- http://prism.dkrz.de/Workpackages/WP3i
9- http://prism.enes.org/WPs/WP4b/
10- http://prism.enes.org/WPs/WP4a/
11- http://prism.enes.org/WPs/WP5/