Working Group Report on Weather, Climate and solid Earth Sciences

CONTRACT NO	EESI 261513
INSTRUMENT	CSA (Support and Collaborative Action)
THEMATIC	INFRASTRUCTURE

Due date of deliverable: 31.05.2011

Actual submission date: 01.07.2011

Publication date:

Start date of project: 1 JUNE 2010 Duration: 18 months

Name of lead contractor for this deliverable: CINECA

Authors: G. Aloisio (ENES-CMCC, chair), M. Cocco (INGV, co-chair), J. Baldasano (BSC), J. Biercamp (DKRZ), R. Budich (MPI), M.A. Foujols (IPSL), M. Hamrud(ECMWF), H. Igel (LMU), C. Jones (SMHI), S. Joussaume (IPSL), M. Juckes (BADC), B. Lawrence (BADC), T. Palmer (ECMWF), N. Pinardi (CMCC), G. Riley (Manchester Univ.), J. Silen (FMI), S. Valcke (CERFACS), G.P. Vilotte (IPGP).

Name of reviewers for this deliverable:

Abstract: This document provides a roadmap towards Exascale computing from the Weather, Climate and solid Earth Science (WCES) communities perspective. It highlights specific needs, requirements and expectations coming from these application domains in the Exascale timeframe.

Revision R2.0

Project co-funded by the European Commission within the Seventh Framework Programme (FP7/2007-2013)		
Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Table of Contents

Glossary

Abbreviation / acronym	Description
WCES	Weather, Climate and solid Earth Sciences
ESM	Earth System Model (climate)
WCRP	World Climate Research Program
SIMD	Single Instruction Multiple Data
PGAS	Partitioned Global Address Space
PDE	Partial Differential Equations
ESMF	Earth System Modeling Framework
CPL7	Source code associated with both the driver aspects and the coupler component aspects in NCAR ESM.
FMS	Flexible Modeling System
MPDM	Multiple Program Multiple Data
ENKF	Ensemble Kalman Filter
GHG	Greenhouse Gas
GCM	Global Climate Model
BPEL	Business Process Execution Language
YAWL	Yet Another Workflow Language
CSML	Climate Science Modelling Language
GML	Geography Markup Language
KML	Keyhole Markup Language
AQM	Air Quality Model
IESP	International Exascale Software Project
ESGF	Earth System Grid Federation
NWP	Numerical Weather Prediction
ENES	European Network for Earth System modelling
EPOS	European Plate Observing System
CMIP5	Coupled Model Intercomparison Project Phase 5
MPMD	Multiple Program Multiple Data
REST	Representational State Transfer
AQM	Air Quality Model
ESG	Earth System Grid

1. Executive summary

This document provides a roadmap towards Exascale computing from the Weather, Climate and solid Earth Science (WCES) community perspective. It highlights specific needs, requirements and expectations coming from these application (driver) domains in the Exascale timeframe.

The WCES roadmap has been organized according to the following main sections:

- Earth Science Models;
- Data Intensive applications in Earth Science;
- Cross-cutting dimensions;
- WCES impacts and findings at European Level.

The first two sections of the WCES roadmap include, among others, scientific challenges related to the physics involved, the coupling of codes/models, the scientific data management, the pre-/post-processing, and the data analysis and visualisation.

Cross-cutting issues like performance, programmability, workflows and data management are also investigated in the third section and represent a fundamental part due to their orthogonal impacts and role.

Along with the scientific part, the roadmap highlights, in the fourth section, strengths and weaknesses of Europe in regard to these challenging topics; societal, environmental and economical impacts of Exascale computing; needs for education and training,; potential collaborations outside Europe; and a tentative agenda and needs in terms of costs and allocation of human resources.

A key point highlighted by this document is the need to move towards Co-design centers, in order to set up interdisciplinary groups of experts able to deal with a wide range of complex and multifaceted aspects related to the exascale hardware and software stack.

2. Introduction

The next generation of highly parallel computing systems composed of millions of cores is expected by the end of this decade. These machines will deliver top-level capabilities. Sustain petaflop performance has already been achieved for half a dozen or so codes, multi-petaflop performances will follow soon, and exaflop performances should be achieved by 2020. This scenario leads to outstanding technological breakthrough possibilities, opening unprecedened possibilities for achieving scientific breakthroughs, designing new products and optimizing existing ones benefiting all society domains. Such a complex scenario needs a coordinated effort and planned strategies at several levels: hardware, operating system, programming environments and **applications** are just some of the most relevant in the software stack.

This document describes (at the **applications** level) issues, requirements, expectations, an agenda and impacts related to the Weather, Climate and Earth Sciences domains in the Exascale timeframe. It has been prepared by the EESI WCES Working group, which includes experts with complementary expertise and knowledge from several European countries (Italy, France, Germany, UK, Spain, Sweden and Finland) operating in these challenging application domains.

The complementarity of the team represents a key point for the reader to have a clear and complete understanding about the future scenarios towards Exascale computing in the WCES domains.

The main structure of the WCES roadmap has been defined according to a few topics ("*Earth Science Models*", "*Data Intensive applications in Earth Science*" and "*Cross-cutting dimensions*") that have been recognized by all of the experts as the most relevant ones in these domains.

The scientific topics have been analyzed and discussed in the three sections by the working group experts; they are reported in this document in a coherent and uniform manner.

A key part of the document (and complementary with regard to the scientific one) is the "WCES impacts and findings at European Level," which highlights societal, environmental and economical impacts of Exascale computing, needs for education and training, potential collaborations outside Europe and a tentative agenda and needs in terms of costs and allocation of human resources.

3. Goals of the WCES Working Group

The main goal of the WCES Working Group is to provide a coherent and integrated vision/roadmap towards Exascale, focusing on the main issues, requirements and expectations coming from the Weather, Climate and Earth Sciences communities.

To achieve this goal the working group composition has taken into account the multifaceted and heterogeneous aspects connected with the WCES domains.

The impact of the roadmap is expected to be wide, which means that it has to present and discuss not only scientific aspects relevant to the WCES communities at exascale, but also the societal impacts of exascale, the needs for education and training and economic aspects in terms of human resources costs, gains or losses connected with the achievement of exascale targets (Exaflop machines) in the timeframe 2010-2020.

4. WCES Scientific Roadmap

4.1 Earth Science Models

4.1.1 Introduction

The field of Earth Sciences or geosciences concerns the scientific study of the planet Earth. It encompasses a wide range of disciplines, from the study of the atmosphere, the oceans, the biosphere to issues related to the solid part of the planet. Earth Sciences address many important societal issues, from weather prediction to air quality, ocean prediction, climate change and to natural hazards such as seismic and volcanic hazards, for which the development and the use of high-performance computing plays a crucial role. In the following, two main domains are considered, on one hand weather and climate that share some similarities, and on the other hand solid Earth. These do not include all the modelling aspects of Earth Sciences but include the two domains mainly concerned by exascale computing issues.

This section provides the scientific and technical vision for the following topics: “Coupled Weather and Climate Models”, “Solid Earth”, “Multidisciplinary data assimilation” and “Uncertainty and predictability”.

The first two sub-sections are domain-specific (Weather and Climate for the first one and Solid Earth for the second one) and present different and relevant aspects organized into sub-paragraphs. The last two sub-sections (“Multidisciplinary data assimilation” and “Uncertainty and predictability”) span across all the WCES domains instead.

4.1.2 Scientific and technical vision

4.1.2.1 Coupled Weather and Climate Models

General circulation models of the atmosphere are the common to both weather and climate models. However, whereas weather forecast models are highly constrained by initialization and run over ten days, climate models require simulations of a hundred years to represent the climate of the last century and estimate possible future changes. Moreover, climate models need to represent the coupled atmosphere-ocean system, and more and more encompass more components of the Earth system such as the biosphere over land and in the ocean to represent the carbon cycle perturbed by greenhouse gas emissions. Climate models are now becoming models of the Earth system, often called Earth system models (ESM). In Europe, the community working on climate models is gathered into the European Network for Earth System modeling (ENES).

Atmosphere and ocean models are also used to investigate other scientific and societal issues. Air quality models aim at forecasting the atmospheric chemistry composition of short-lived species. Ocean models are extensively used to understand ocean processes but also to forecast the evolution of oceans some weeks in advance, either globally or for specific regions.

Computing power strongly constrains the spatial and time scales resolved for both atmosphere and ocean models. For example, climate models nowadays commonly run at resolutions on the order of 150-200 kms, whereas in order to resolve convective clouds in the tropics, resolutions on the order of ~1 km should be reached, which represents a grand challenge for the climate modeling community (ENES foresight, http://is.enes.org). Computing power also constraint the time length of simulations, the level of complexity resolved and also the number of experiments required to deal with internal variability and uncertainties in model parameters.

4.1.2.1.1 General Strategy for coupled earth system models optimization

Technology and Science Drivers

During the last decades, advances in climate and earth-system modeling were strongly related to the increasing performance of supercomputers. This trend will continue in the foreseeable future. Socially important scientific goals (like prediction of regional climate changes or better simulation of clouds and precipitation to reduce the uncertainty of climate simulation) necessitate massive model improvements, many of which place very high demands on computational capability. This has been made very clear at the World Modelling Summit for Climate Prediction, held in May 2008. The summit statement calls for "*computing capability acceleration, leading to systems at least a thousand times more powerful than the currently available computers*"[1]. Similar reasoning is found in the report Challenges in Climate Change Science and the Role of Computing at the Extreme Scale, prepared for the U.S Department of Energy[2].

Access to increased computational capabilities for climate modelling is in particular required to meet the need for higher spatial and temporal resolution, better physical process representation, explicit modeling or more biogeochemical processes, much longer runs and larger ensembles.

Alternative R&D Strategies

The strategy to adapt and optimize coupled climate earth system models for exascale systems depends on the envisaged application scenario. One goal is the explicit modeling of processes, which today have to be described by parameterizations – the latter being potentially incomplete, inadequate or erroneous. In order to achieve this, the spatial and consequently temporal resolution of the individual component models has to be increased by orders of magnitude. In this case, so called weak scaling – increasing computational concurrency by increasing the problem size – is a key issue.

The other extreme is given by very long simulation times for example the simulation of glacial cycles. Since time is serial by definition this implies the need to sustain or better increase the number of simulated years per real time period. This can either be achieved by strong scaling of the codes (i.e use more processors for a given problem size, aiming to exploit functional parallelism in addition to data parallelism, thereby increasing throughput) and/or to increase the serial performance of the code.

Other science cases lying in between the above extremes include ensemble simulation with medium resolution models or iterative model-data synthesis, the latter implying a large number of (serial) iterations.

Recommended Research Agenda

As it has been pointed out above, a central task is to increase the scalability of earth-system models at the same time trying to optimize the serial performance. The most important and probably also hardest problems are:

- efficient and scalable high volume I/O;
- efficient scaling of individual component models to, at least, tens of thousands of cores;
- optimization of the serial performance of low resolution component models;
- scalable coupling of the individual component models.

Today's climate models are the result of decades of work adding up to hundreds of person-years and millions of lines of code. It would be extremely time consuming and expensive to completely rewrite these models. Instead, the models need to be modularized and restructured to be more scalable, portable and maintainable. Additionally there is the need to develop or improve scalable algorithms for many of these modules. Where possible, the modules should be reused by different groups in different configurations to avoid duplication of work. The organization of the European climate modeling community through ENES can and should foster the development of such common components. One prominent and important example for such a module is the dynamical cores of circulation models, which are currently being rewritten by several groups in order to use more flexible and better scaling computational grids. Dynamical cores are being dealt with in more detail in section 4.1.2.1.2.

---

[1] Workshop Report, World Modelling Summit for Climate Prediction, World Climate Research Programme, WCRP No 31, WMO/TD No. 1468, Jan. 2009.

[2] [Scientific Grand Challenges: Challenges in Climate Change Science and the Role of Computing at the Extreme Scale, Report from the Workshop held November 6-7, 2008, for the U.S Department of Energy under contract DE-AC05-76RLO1830, http://science.energy.gov/~/media/ber/pdf/climatereport.pdf].

Since the number of computational grid points in most science cases will remain rather limited due to the required long simulation periods, communication overhead in message passing seriously limits the possible number of parallel partitions. Therefore, in addition to techniques to overlap communication with computation and ensure the 'loosening' of synchronization between processes, where possible, hybrid programming concepts using SIMD type programming within one partition or within one computational node will have to be implemented. Within a data-parallel partition, there is an opportunity to exploit functional parallelism; for example, updates to the prognostic variables can often be computed concurrently, leading to small but potentially useful performance multiplication. Opportunities for parallelism in the vertical dimension will also need to be explored, both in the dynamics – for example, through the use of (thread-based) parallel implicit solvers, where applicable, – and also in the physics routines. Again, hybrid programming concepts are needed to address the (primarily software engineering) limitations of the current MPI-OpenMP approach. At the same time, climate modelers should be open to new programming concepts like PGAS and the use of accelerators through specific programming constructs. However, this will only be feasible if these concepts are efficiently supported at high abstraction levels in the programming interfaces, e.g. through directives in FORTRAN code.

Another serious bottleneck in highly parallel climate models is efficient and scalable output of very large volumes of data. The data challenge is not only limited to the fast and comfortable parallel output during production of the data, but extends to storage and curation. This is dealt with in the Data Management section of this report.

Global Climate models consist of models of many different components describing very different realms. (atmosphere, ocean, sea ice, land biosphere, atmospheric chemistry, land usage are some notable examples) on potentially different grids. These models cover processes on many length and time scales. One way to couple these systems is to combine them into one very large executable. This is probably the most efficient way but imposes strong coding constraints, which are not easy to apply to modelling groups not working under hierarchical management structures (as in the European climate modelling community). Considering the variety of modeling groups in Europe and the above mentioned structural organization through ENES, a better way for Europe would probably be the usage of one common software framework which allows combination and coupling of component models with minimal intrusion into or restructuring of existing legacy codes. The efficient implementation of such systems on exascale systems, however, is a challenge on its own and is dealt with in more detail in sub-section 4.1.2.1.3.

Cross-Cutting Considerations

On the way to exascale, there will be a big increase in complexity, both in the architecture and programming of high performance computing systems as well as in the deployed coupled climate models. The only alternative to cope with this is a considerably strengthening of the cooperation of climate modelers with software engineers.

As many of the sub problems solved in climate models are common to other domains as well, the development of algorithms appropriate for exascale is a cross cutting issue that should involve domain scientists, mathematicians and computer scientists.

Some of the issues arising from the inefficiency of parallel I/O at high abstraction levels are specific to climate models, in that they are due to the complex structure of the data and the high output frequency. However, underlying software- and hardware layers for parallel I/O are common to many other fields as well. It is hoped that climate models will be able to leverage general tools and methods provided by hardware vendors and software engineers.

4.1.2.1.2 Dynamical cores

The dynamical cores are often seen as the most challenging problem within geo-physical applications for future Exascale computing. In today's ESM, the dynamical cores are the main scalability bottle-necks. The fundamental problem is the sequential nature of the time-stepping algorithm used to solve the PDEs involved. As the resolution of the models increases, the time-step used normally needs to shorten in order to control errors caused by the time-discretization and to avoid instabilities. Each time-step involves communication: possibly local or quasi-local for the advection but often global if any form of implicit time-stepping is used. Abandoning the currently used implicit (or semi-implicit) time-stepping schemes to avoid the need for global communication may be counterproductive if this in turn leads to a vastly decreased size of the time-step.

In the timeframe of Exascale computing we will enter the cloud-resolving scales for global atmospheric models (~1km). This will necessitate moving to non-hydrostatic dynamical cores.

The definition of the representation used for the physical quantities is a crucial part of the design of a dynamical core. Some schemes currently used, like the spectral method with its associated Gaussian grid, and the latitude-longitude grid seem to present special problems when going into the Exascale regime. The spectral method requires global communications within the transforms between the spectral representation and the Gaussian grid, which may scale poorly at very large core counts. The latitude-longitude grid instead suffers from the convergence of the longitudes when approaching the poles. More promising seem to be various quasi-uniform grids (i.e., cubed sphere, icosahedral, Yin-Yang, Fibonacci) currently being developed.

The flow in simulations with recent climate ocean models is essentially laminar. Mesoscale eddies, the ocean equivalent to atmospheric weather systems, are not explicitly resolved, but parameterized. The trend in global ocean modelling is towards eddy-resolving resolutions (10 km and higher). On the other hand, the physical time scales of important ocean processes require integration times of 100-1000 years. An approach to meet both requirements - high-spatial resolution and long integration times - is to develop models on non-uniform grids, i.e., use locally varying resolution. The resolution may vary statically, as many interesting/critical regions in the ocean are set by the topography, or it may vary dynamically. In any case, this implies a fundamental challenge in the formulation of the dynamical core, which in its extreme case must resolve hydrostatic and non-hydrostatic regimes, i.e., solve different PDEs. An important aspect of this challenge in order to get efficient codes concerns the development of appropriate time integration schemes that are able to handle grid cells of largely varying size. A second research area is the development of scale-sensitive parametrizations.

4.1.2.1.3 Flexible couplers and coupling frameworks

Technology and Science Drivers

As stated before, global climate models consist of different component models describing very different realms and usually developed independently by different modelling groups. These components potentially use different numerical grids and cover extremely varying inherent time scales. Increasing the parallelism of global climate models by increasing the number of concurrent components is probably essential to exploit the exascale platforms expected before the end of the decade. To model the whole climate system, these components have to interact and exchange information at their boundaries; the efficiency of the coupling between these components is therefore crucial for exascale computing.

Alternative R&D Strategies

The current coupling technologies can roughly be split into two main categories, each one presenting advantages and drawbacks. On one hand, direct coupling using a fully concurrent multiple executable approach, such as with the OASIS coupling software currently widely used in the European climate modelling community, is somewhat less flexible, but it is relatively straight-forward to implement. In this case, the original components are run as separate concurrent executables, and their main characteristics, such as memory management or internal parallelisation, remain practically untouched with respect to their standalone mode. The exchange of coupling data is performed through "put" and "get" instructions implemented in the component codes. It is the user’s responsibility to ensure that the component models coherently define some global parameters such as the total run duration, the calendar, etc. The main advantage of this approach is that it requires minimal intrusion into, or restructuring of, existing legacy codes. The drawback is that it is less flexible, and in some cases, less efficient than the coupling exchanges necessarily imply some data transfer.

On the other hand, coupling via top-level interfaces within one integrated application (such as with American couplers ESMF, CPL7 or FMS) requires some standardization around high level interfaces, design, and datatypes, but provides opportunities to run models in more flexible and more efficient configurations. In this case, the model source code is decomposed into init, run, and finalize units and the interface must match the standard expected by the coupling layer. The coupling data are made available as input and/or output at each calling interface. This approach forces the components to expose both data and control interfaces; it is the coupling layer that executes the units and this "inversion of control" automatically ensures consistency of global parameters across component models (e.g., run duration). This approach is more flexible and in some case more efficient as the component models can be executed concurrently, sequentially, or in some hybrid mode and coupling exchanges can be in some cases optimized as shared memory accesses. A drawback to this approach is that it potentially requires more modification to legacy codes (e.g., splitting them into *init*, *run*, and *finalize* units, which might be difficult to achieve) and imposes the use of the coupling layer standard-calling interface and data structure.

Recommended Research Agenda

In the short and mid term, developments and tests have to be done with the existing coupling tools to evaluate if the multiple executable approach will still be an option to consider for exascale computations. In all cases, as this approach may be the only realistic one for some legacy codes for still many years to come, it is important to ensure that future operating systems support the MPMD (multiple program multiple data) mode.

In parallel, the integrated approach should be seriously considered for further development of coupled climate models in Europe to address exascale challenges. In the longer term, the new languages and the expected re-write of the models should be considered as an opportunity to widely agree, at the European level but maybe even at the international level, on coding and coupling standards that will make possible the assembling of different climate components into integrated applications.

Regarding the regridding functionality, which is essential in the coupling, efficient 2D and 3D conservative algorithms have to developed to complement more traditional ones (2D linear, higher order and user-defined). The type of grids supported should include all logically-rectangular grids (latitude-longitude, stretched, rotated, etc.) but also unstructured grids such as icosahedral grids which are becoming more and more popular. Coupling of components with adaptive grids is also foreseen in the near or mid- term. In this case, the coupler should be able to recalculate the regridding neighbours and weights during the run and efficiently manage the impact on the communication.

Cross-Cutting Consideration

As for any highly parallel computation, efficient communication, i.e. efficient coupling exchanges between the different components of the coupled system, is a central issue to address. The communication costs should be reduced as much as possible. In particular, this will have an impact on the localization of the components on the different processors and on the individual component model decomposition. Overlay of calculation and communication and redundant calculations are options that should also be explored. The (possibly dynamic, i.e. evolving during the simulation) load balancing of the different processes (here the components) will also have to be addressed.

Code coupling per se is in fact not specific to climate modelling. In many other disciplines, such as computational fluid dynamics or electromagnetism, multi-physics and multi-scale coupling is a central issue. Therefore, interaction with other communities should be considered essential to progress in addressing the challenges raised by exascale computing.

4.1.2.2 Solid Earth

Technology and Science Drivers

Computational challenges in solid earth sciences span a wide range of scales and disciplines, and address fundamental problems in understanding the Earth system - evolution and structure - in its near surface environment. A rich panoply of societal applications has emerged from basic research. The community plays today a central role in natural hazard mitigation (seismic, volcanic, tsunami and landslides), hydrocarbon and energy resource exploration, containment of underground wastes and carbon sequestration, and national security (nuclear test monitoring and treaty verification). Emerging new applications address Earth's environmental changes and hazards: paleo climate evolution, dynamic interactions with the near surface environment: glaciology, ocean wave and atmosphere with a whole set of new natural seismic sources and seismic waves. Solid Earth computational challenges are observation driven and rely on the effective analysis and modelling of the increasing wealth of multi-attribute observational data continuously transmitted by global and regional monitoring systems.

Understanding and mitigating earthquakes risk critically depends on physics-based prediction of ground motions using realistic dynamic rupture and wave propagation models in 3D geological structures. This requires extreme-scale computing capabilities to adddress the actual spatio-temporal scales governing these processes, and to improve our physical modelling of the earthquake rupture physical processes.

Imaging accurately the Earth from the near surface to the deep interior, on land or below the sea floor, is a challenging problem today, with important implications in terms of energy resources and environmental hazard management. New imaging (correlation-based migration) and inversion methods exploit the spatial and time coherency of continuous seismic waveforms, within increasingly dense acquisition systems. These new deterministic (local) and probabilistic (global) methods require extreme scale computing and data capabilities.

Understanding the patterns of mantle flow today and in the Earth's past is a challenge confronting many Earth science disciplines: seismology, geodynamics, and mineral physics. Issues ranging from thermal history of the planet to the driving-forces of present-day tectonics are intimately linked to this topic. High resolution, three dimensional thermo-chemical mantle convection models, including melt-induced differentiation, solid-solid phase changes and self-consistent near surface plate tectonics behaviour require extreme scale computing capabilities

Understanding turbulent flow in the Earth's fluid core and how the geodynamo generates the magnetic field - and why it undergoes spontaneous reversals - is a challenge in solid Earth and in astrophysics. Today none of the Earth's core magneto hydrodynamics simulation models are able to simulate the Earth's core parameter regime and dimensions due to the extreme time and space resolution required. Next generation geodynamo models need to simulate regimes where the thermal and viscous (eddy) diffusivities are no larger than the actual magnetic diffusivity of the Earth's fluid core, while using the core's dimensions, mass, rotation and heat flow. In addition, many realizations are necessary to perform ensemble modelling and forecasting in the case of such highly nonlinear systems, with strong dependence on initial and boundary conditions.

Understanding and simulating the multi-scale and multi-physics processes governing the dynamics of the interior Earth systems and their near surface interactions, improving spatial and temporal resolution, testing hypotheses, rely to an increasing extent on extreme scale computing capabilities with new computation models and algorithms that can cope with integration and manipulation of massive observational and synthetics data sets.

Solid Earth systems have intrinsically limited observational resolutions due to their wide range of spatio-temporal scales. Fusion of observational data with computational modelling provides the only means to gain the required understanding of the driving processes, and to implement physically based monitoring and warning systems. Computational modelling is increasingly used in the evaluation and the prediction of natural hazards, and to inform policy-making and risk mitigation strategies. A critical issue is to better identify and quantify uncertainties, and to estimate the probability of extreme events through simulation of scenarios and exploration of parameter spaces.

The transition from conventional deterministic to probabilistic/stochastic methods and models of extreme events relies on extreme scale computing and storage capabilities. The economic benefit to society of quantifying the uncertainty and impact of prediction scenarios, on whatever time scales, is potentially enormous for earthquake, volcanoes and tsunami hazard mitigation but also for the potential impact on our global communication infrastructures from the Earth's magnetic field time reversals. Probabilistic inverse methods, large stochastic simulation, and 4D data assimilation will increase by several order of magnitudes the complexity of the algorithms and the required capabilities and capacities of these models.

The solid Earth science community is facing a fundamental paradigm shift. New-era data-intensive applications are shaping requirements for extreme scale architectures and framework integrating data and computing capabilities. This has been made clear within EPOS, the European Plate Boundary Observatory System, the ESFRI initiative of the solid Earth community, and in the related European projects: NERA, SHARE & VERCE in infrastructure; QUEST in Research and training. The community is also closely collaborating with other similar initiatives in the US and Japan.

Alternative R&D strategies

Computation models based on communicating sequential process parallelism (CSP) are no longer capable of supporting effective exploitation of current and future generations of massive multicore architectures foreseen for the extreme exascale computing. New instructions set architectures and accelerators (GPGPUs and attached processors) will require hybrid parallel programming models that can cope with different programming languages, including Fortran. These changes in the foundation model of computation will have a dramatic impact on algorithms and software engineering - that bridge solid Earth application characteristics and hardware characteristics – with conflicting goals in the management of concurrency and locality (tasks and data). Lessons learned when moving HPC solid

Earth applications to petascale will hopefully help in establishing future exascale applications footprints.

The foreseen critical issues in the refactoring and adaptation of the HPC solid Earth applications are:

- Efficient scaling toward high-level task concurrency with a billion or more threads: this implies an increased level of asynchronous task parallelism lowering global barriers, hiding latency and handling variability among cores
- Explicit locality model: locality is key to efficient parallelism and will require a portable and abstract hierarchy of locality with collocated distributed tasks and data, a higher degree of vertical data locality – temporal locality (i.e. reuse) – exploiting nodes' memory hierarchy; a higher degree of horizontal data locality through adaptive domain decomposition; and load balancing exploiting nodes' cache and new instruction set architectures (GPGPUs and attached array processors).
- Scalable coordination and synchronization: this implies improving point –to-point synchronization, collective operations (map-reduce) and lowering contention and overhead issues.
- Tipping the balance to data: for data-intensive applications a key issue is an efficient data-crawling strategy to hide the latency gap beyond the main memory. This needs to efficiently exploit the memory fabrics; node'ss' memory hierarchy and multi-level caching; scalable and efficient parallel I/O; high I/O bandwidth.

The explosion of observational data, and the emergence of distributed data infrastructures integrated at the European and international levels in the solid Earth Science community, has created a major challenge for cutting-edge scientific projects in solid Earth Science. Solid Earth science is facing today, with other observational communities, a data intensive paradigm shift and very strong requirements for extreme exascale computing and data capabilities.

Emerging new data intensive applications, related to innovative combinatorial data analysis of massive data sets and model-to-data matching (theory-to-observation) methods, lead to enormous challenges both in terms of computational models and in terms of architecture models integrating efficiently data and HPC infrastructures. Data volumes in solid Earth science are expected to increase 1000x in 10 years while I/O bandwidth improve ~3x in 10 years. Such a mismatch will create massive data management challenges together with parallel data mining challenges.

Ahmdal has established important laws for building a balanced computer model in the context of the explosion of observational and synthetics data. A balanced system needs one bit of I/O and one byte of memory for each CPU cycle. This clearly underlines the fact that today's computer systems' I/O subsystems are badly lagging CPU cycles.

Hiding this gap will require innovative algorithmic developments to balance parallel mining of data and mining of parallel data, and improved layered architectures. Having multiple tiers provides a system with a certain amount of hierarchical spread of memory and disk storage. The low level data can be spread evenly among server nodes on the lowest tier, all running in parallel, while query aggregations are done on more powerful servers in the higher tiers.

Another issue regards the EPOS solid Earth research infrastructure with the development of an efficient e-Science environment built upon an architecture and framework of tools that allow the integration of data and HPC infrastructures, together the provision of high sequential bandwidth, the support of most e-Science patterns, efficient database design allowing seamless movement and efficient data ingest mechanisms for massive data sets and applications across those components.

Recommended Research Agenda

Solid Earth science spans a broad range of disciplines and physical and numerical models. Not all the solid Earth applications that operate at sustained terascale today will scale out easily to petascale and exascale, given current application and architecture characteristics. Scaling strategies are application specific and all face barriers in the areas of algorithms, software engineering, computational models and data management.

The main strategy, shared among the solid Earth applications, is related to the so-called weak scaling model. Weak scaling refers to the concept of adding work as an application is run on more processors/cores to ensure that starvation, overhead and latency do not destroy performance.

The classical weak scaling model is related to higher resolution in space and time in simulating 3-D solid Earth systems for a fixed system size (global Earth systems) or for larger system sizes (regional systems). When solving larger system size problems, the memory scales nearly proportionally with work. Solving higher resolution problems requires smaller grid size and time steps, both for the explicit methods used in wave simulation problems or semi-implicit and implicit problems used in CFD Earth mantle and core simulation problems. The relative ratio of work versus memory increases. Structured and unstructured mesh discretization of complex 3-D geometries of heterogeneous and anisotropic geological media are a critical issue that can ultimately dominate the overall runtime of the simulation. Parallel generation of meshes of many billion of elements on massively parallel architectures are required, as well as effective partitioning of adaptively defined unstructured mesh with dynamic load balancing.

Transformative new weak scaling models are emerging today in solid Earth applications. Compared to traditional weak scaling of those HPC applications, it is however more difficult to predict applications footprints for the new-era solid Earth applications mentioned above.

Flexible model coupling related to the study of the interactions between solid Earth systems that were previously model and simulated independently. Coupling physical and numerical models with different dynamical regimes and spatio-temporal scales is becoming critical, in particular for simulations of the interactions between the core and the fluid inner core convection, and of paleo climate evolution, and of solid-fluid interactions during slow deformation and wave propagation. This may involve sophisticated and demanding coupling strategies, well-standardized interfaces implying some modification to the existing codes, as well as data structure issues …

Multi-scale simulation related to the evolution of systems requiring the resolution of physical processes or material heterogeneities spanning space and time scales of several order of magnitude differences within one model. A related issue is the coupling between different level of physical description, like continuum and first principles (molecular or granular) levels descriptions. This is becoming critical in order to overcome the intrinsic limitations of empirical models based on a parametric description. This requires parallel adaptive non-uniform and/or non-conforming mesh strategies. New methodological and algorithmic developments allow today the construction of temporal and spatial hierarchical coarse graining models - through theoretical or numerical homogenization, slow and fast phase variable separation, adaptive time and space resolution, multi-iterate integration methods that allow different time steps for different part of the model – spanning the different scales. Key issues here are asynchronous task parallelism, dynamic adaptive load balancing, and abstract hierarchy of task and distributed data locality, memory fabrics and latency. Typical applications are related to wave propagation in complex geological media including fluid coupling, earthquake rupture dynamics, simulation of seismic cycles along fault systems, dynamics of the fluid inner core and rapid magnetic time reversals, convective fluid dynamics in porous media, avalanches and debris flows, erosion dynamics, etc.

Large-scaled optimization-based 3-D inversion related to problems with high-dimensional parameter spaces. Large 3-D inversion based on adjoint methods, associated with a local linearization of the problem, open today new perspectives through iterative multi-scale optimization strategies. Because large scale 3-D forward and adjoint problems need to be simulated and stored for each of the optimization iterations, this kind of applications is extremely demanding in terms on computational and data capabilities. The key issues here are the memory fabrics and the latency gap beyond main memory, efficient and scalable parallel I/O, and aggregated shared memory. Typical applications are those related with full waveform inversion and high-resolution imaging in seismology including differential time seismic imaging…

Probabilistic and stochastic models account for intrinsic observation uncertainties and resolution limitation, and for quantifying probabilistic uncertainties in inferred earth system models or predictions. Stochastic forward modelling is concerned with how uncertainties in parameters propagate through a physical model and affect its prediction in probabilistic terms. Probabilistic inversion is the only means to explore globally modeled parameter spaces and infer the whole set of inferred Earth models via Monte-Carlo like and maximum likelihood algorithms. In terms of computational models, quantifying uncertainties, stochastic ensemble simulations and parameter space exploration are tremendously expensive undertakings that raise entirely new sets of challenges for both mathematical and parallel algorithm development. Key issues here are asynchronous tasks and communications lowering global barriers, scalable coordination and synchronization, data movement across node's hierarchy memory, efficient and scalable parallel I/O and data transfer. Typical applications concern seismic tomography and imaging, seismic sources characterization, Earth magnetic and gravimetric fields…

Data assimilation is an emerging field of applications in solid Earth science. For computation fluid dynamics in the Earth's core, these methods need the construction of probability distributions of the assimilated states from deterministic dynamical simulation, including a noise term. For other stochastic systems, like seismicity models, probability distributions must be constructed from state variables. These applications are extremely demanding in computational and data capabilities (memory fabrics, scalable parallel I/O) due to the dimension of the state variables, the construction of the state probability distributions and the simulation of their evolution.

Other weak scaling issues are related to temporal scaling. Temporal scaling refers to the need for simulating the evolution of a system - with a given spatial and time resolution - over very long simulation times. Algorithmic and implementation issues are related to the control of the dispersion and dissipation errors and to the classical serial structure of time marching algorithms. Higher order time schemes - or new geometric integrators (symplectic time schemes) that Hamiltonian structure for elastodynamic wave simulation - are extremely important for long time simulation. This increases the ratio between computational work and memory. New parallel methods in time allow achieving higher order time accuracy, together with an increased level of parallelism in the vertical direction. Typical applications concern seismology (seismic wave propagation at global and regional scales)…

The above petascale and exascale weak scaling considerations will allow new transformative scaling of physically based front-end simulations of solid Earth systems. However, critical issues in a more operational mode or day-to-day research environment are more related to strong scaling. Strong scaling refers to the concept of applying more resources to the same problem size to get results faster. Very few applications will be actually amenable to the required strong scaling level, even in the horizon of petascale and exascale capabilities, without algorithmic innovations increasing parallelism in the vertical dimension (e.g., new parallel explicit time schemes, new parallel implicit solvers), increased asynchrony levels in parallel tasks and communications lowering global barriers, and efficient hybrid programming models and instruction set architectures (GPCPUs and array processors). This is foreseen as a critical issue for hazard related applications.

Cross-Cutting Considerations

The *software crisis* that many fields of natural sciences face within the context of scalable simulation methodologies also provides an opportunity for an new level of interaction between the various disciplines. Even though the science goals and underlying methodologies might be very different, the requirements are very similar. With very few exceptions, the core applications and codes have been developed and written by solid Earth scientists with an emphasis on the physics, but with a much smaller emphasis on using the latest technologies available from the computing science communities. The codes and tools are written in different computer languages and are dependent on different libraries and computational platforms. Many related problems are common to other domains:

- *refactoring:* identify potential re-usable data, and computation-oriented components, which can be extracted by refactoring existing methods and software implementations; and then improve their interfaces;
- *re-engineering:* indentify in these re-usable data and computation components those that need re-engineering, improvements to algorithms – or data and computational strategies modifications – to improve their performance and to better exploit the exascale capabilities.

Workflow development: analyze and identify the granularity of the different computation process elements and data exchange components of the applications; build efficient orchestration among those different components.

4.1.2.3 Multidisciplinary data assimilation

Technology and Science Drivers

The purpose of data assimilation is to obtain an optimal estimate of the state of the system at a given time, determined from a set of observations related to the system state and prior information. This state is normally referred to within the geophysical community as the analysis. The prior information normally consists of a short-range integration of the model from a previous analysis together with an estimate of the covariance structure of the error of this forecast. It may also contain other constraints imposed by the governing equations of the model. The main scientific drivers towards Exascale computing are a desire for using higher resolution, more complex models within the system and the increasing availability of observations, mainly from satellite based instruments.

Alternative R&D Strategies

In the last two decades we have seen dramatic developments in the area of data assimilation within the meteorological community and to a lesser degree within the oceanographic community. Firstly, the number of observations available for data assimilation has increased by several orders of magnitude with the introduction and proliferation of data from satellite based instruments. Secondly, there has been strong progress in the methods used for data assimilation. Virtually all methods for data assimilation in state-of-the-art systems are today based on different approximations of the Kalman Filter. The full Kalman filter can probably never be used within the geo-physical community, as the size of the state vector is too large. The approximation to the Kalman filter that has emerged as the algorithm of choice within the meteorological community is the so-called 4D-Var algorithm, which requires the development of a linearized version of the model and its adjoint. A further simplified version of 4D-Var (3D-Var) is also used within the community, mainly where a very fast solver is required for operational reasons. The last decade has also seen the emergence of the Ensemble Kalman Filter (ENKF) as an alternative to 4D-Var. One advantage of this algorithm is that it does not require a linearized model.

In general it is clear that for data assimilation systems it is more challenging to make efficient use of large processor counts than for the models themselves. There are several reasons for this, the main reason being that the 4D-Var algorithm as currently implemented is essentially sequential in nature. The 4D-var algorithm is a minimization algorithm and the method currently employed in the minimization (conjugate gradient) is an iterative algorithm. Within each iteration the integration of the tangent linear model and it adjoint is also sequential in the time-stepping. The parallel implementation is found only at the bottom, where the state variable is geographically distributed. This leads to a very fine granularity with frequent communications. Similar issues also make 3D-Var scale badly.

The emerging ENKF data assimilation system does not suffer from these problems. The mean state and a low-rank approximation of the covariance matrix are propagated with an ensemble of models. The actual analysis is performed independently for each element of the state vector, leading to a large number of degrees of freedom.

Another reason why data assimilation schemes have problems efficiently utilizing high processor counts is the heterogeneous nature of the observations. The different types of observations take radically different times to process, and the computation of the model equivalents that is at the core of all data assimilation systems is difficult to load balance as the nonhomogenous geographical distribution of the observations means that processors responsible for different parts of the state vector get unequal amounts of work.

Recommended Research Agenda

The main R&D strategy for enabling the use of Exascale computing within data assimilation has focused on the algorithmic side. Within the context of 4D-Var there exist possibilities for increased scalability by exploring parallelism within the time dimension (this requires a weak constraint formulation of 4D-Var) and also for exploring the use of parallel minimization algorithms. The ENKF approach shows promise from the scalability point of view, but it is unclear at this point if this approach can produce analysis of a similar quality as that of 4D-Var. The model always form an integral part of any data assimilation system, so any improvement in, e.g., the dynamical core also brings the data assimilation system closer to being able to utilize Exascale computing.

Cross-Cutting Consideration

As all modern Data Assimilation systems use the forecast model as an integral part of the scheme, the path for data assimilation systems to achieve the ability to us Exascale computing follows that of the models, especially of the dynamical cores.

4.1.2.4 Uncertainty and predictability

Technology and Science Drivers

It is now well established that the atmosphere is a non-linear, chaotic system which is highly sensitive to small uncertainties in initial condition specification. Two initial states, differing only by an infinitesimally small amount, can rapidly diverge in their solution at some point into the (simulated) future. Advanced atmospheric models, including data assimilation, suggest a purely atmospheric predictability limit may be around ~10-15 days, depending on the location, season and weather situation under consideration. We define a predictability limit as some point into a forecast where the simulated state of the system is no more accurate than a randomly selected state. Such a limit, arising from an incomplete specification of the initial conditions, is one factor contributing to the spread (uncertainty) in weather forecasts. The potential predictability of a system can be defined in terms of its mathematical properties and initial condition uncertainty. The actual predictability of the atmosphere is further limited by the imperfect nature of prediction models. An incomplete description of some atmospheric process (e.g. entrainment of air into a cumulus updraft) will result in an inaccurate forecast of a given variable (e.g. temperature), that may rapidly increase with forecast lead time. This results in a second type of practical predictability limit, referred to as model uncertainty and results in the deviation of two forecasts, started from identical initial conditions that are made with different weather/climate models, or by the same model with uncertain physical parameterizations perturbed within reasonable bounds for each separate forecast.

Such limits are well appreciated in weather and climate prediction and motivate the use of an ensemble approach to derive measures of forecast reliability and provide probabilistic estimates of event occurrence. Predictions of extreme weather in particular require an ensemble approach, so the risk for potentially low probability, but high impact, weather events can be assessed. The need for an ensemble approach leads to a compromise between high model resolution, known to improve the accuracy of a given forecast, and lower resolution combined with a large ensemble. As an example, the ensemble prediction system at ECMWF is run twice per day and consists of 51 modified initial states, each run at T639 (~32km) resolution for the first 10 days of a forecast, degraded to T319 (~60km) for forecast days 10 to 15. The ECMWF seasonal prediction system, run 4 times per year out to a 7 months lead-time, consists of 41 members at T159 (~120km) resolution. In comparison, a single deterministic forecast, run 4 times per day, employs a resolution of T1279 (~16km).

Ensemble forecasting is a perfectly parallel computational problem, whereby each ensemble member is independent and can therefore be run in parallel, simultaneously with all other members. To increase forecast accuracy it would be of great benefit to run the ECMWF ensemble system at T1279 resolution, with an ensemble size of ~100 perturbed initial states, combined with a range of perturbed or stochastic physics options to sample model uncertainty. As an example, we propose 10 perturbed/stochastic physics versions of the ECMWF model. This would lead to an increased compute cost of ~150-200 times that presently required. Due to perfect parallelization of the ensemble approach, if a machine with a sufficient number of compute cores and rapid inter-node communication was available, all forecasts could be made in parallel and this increase could be realized with no increase in wall-clock time for the forecast set. Similar arguments hold for seasonal prediction.

With respect to regional climate change due to increased greenhouse gas emissions (GHGs), initial condition and model uncertainty remain important contributors to total uncertainty. Uncertainty in future socio-economic conditions, expressed as future GHG emissions and land-use, also now contribute to overall uncertainty. Furthermore, a range of highly-uncertain Earth System processes must be accurately represented in models to simulate the temporal response of the climate system to increasing GHG forcing. Primary processes contributing to this model uncertainty include cloud-radiation feedbacks and the response of biogeochemical cycles to a changing climate. The sensitivity of the overall response of the climate system to such uncertain processes, under a perturbed GHG forcing, increases the absolute uncertainty of regional climate projections compared to that in weather and seasonal prediction. As a result 3 axes of uncertainty must be sampled to provide reliable regional climate change projections; (i) Initial condition uncertainty, manifested as uncertainty in the timing and location of natural climate variability with respect to slower global change, (ii) Uncertainty in future GHG emissions and land use. (iii) Imperfections in model representation of key climate phenomena determining global climate sensitivity and/or forced changes in regional climate variability.

Extensive collaboration means European regional climate change assessment in the near future will be routinely based on ~6 Global Climate Model (GCM) projections. Each GCM will sample ~3 plausible future GHG/land-use scenarios and make 3-5 simulations per GHG scenario, from slightly different initial conditions all representing pre-industrial conditions. This constitutes an ensemble of ~54-90 projections spanning the period 1860-2100. Due to its non-linearity, sensitive threshold (or tipping) points exist in the climate system that if exceeded, may result in rapid, potentially irreversible, climate change in response to increasing GHG forcing. To assess the risk for such changes an ensemble approach is necessary to sample a wide range of possible future climate states. Furthermore, the societal impact of possible forced changes in extreme weather events, such as tropical cyclones, necessitates a full exploration of forecast uncertainty space, preferably at model resolutions sufficient to simulate such phenomena. The need for an ensemble approach presently limits GCM resolution to ~200km, necessitating some form of downscaling to bring GCM projections to a spatial scale suitable for impact assessment and adaptation planning.

Decadal climate prediction is a new, rapidly growing field of research. It aims to fill the gap between seasonal forecasting and centennial climate projection, providing information on regional climate change and variability on ~1 to 25 year timescales. Decadal prediction builds on improvements in both GCM quality and ocean observations over the past few decades. The extended predictability range is based on observational initialization, and subsequent simulation, of slowly varying modes of ocean circulation, such as the Atlantic Meridional Overturning Circulation which, through modulation of ocean surface temperatures, imparts a potentially predictable signal on European climate. Due to the importance of accurate ocean initial conditions, the long forecast lead-times involved and the need to simulate detailed ocean circulation processes, initial condition uncertainty and model uncertainty both contribute significantly to total forecast uncertainty and require extensive sampling to provide useful forecasts.

From short-range weather prediction, through seasonal to decadal climate prediction and centennial climate change projection, an ensemble approach is an absolute necessity. None of these forecast areas are, nor will ever be, fully deterministic. Irrespective of the model resolution attained, an ensemble approach will always be required to provide predictions of practical use to society. The advantage of the ensemble approach is that it is a perfectly parallel problem and therefore ideally suited to massively parallel computer systems. Rarely does one find the confluence of societal need for information (weather and climate prediction), with the maturation of a scientifically based approach (ensemble forecasting) and the rapid development of a perfectly suited technology (massively parallel computer systems). Ensemble prediction, on all timescales, is both perfectly suited to future parallel computer systems, while also allowing the inherent uncertainty in weather and climate forecasting to be sampled and thereby useful forecasts developed.

Alternative R&D Strategies

As explained above, there is no real alternative to an ensemble approach in weather and climate prediction. Even in short range forecasting (12-24 hours), with high resolution (~1km) models, the systems targeted for prediction, namely deep convective cloud systems, have a stochastic nature to their onset and development. While a very high resolution model will improve the accuracy of a given simulated system, their onset, intensity-development and propagation will always have a non-deterministic component. Forecasts should, therefore, be expressed in a probabilistic sense and will benefit from an ensemble approach. As forecast lead times are increased the need for an ensemble approach to span the envelope of forecast uncertainty increases.

Recommended Research Agenda

Future efforts in ensemble prediction will always benefit from increased model resolution and therefore increased model efficiency. This strongly motivates continued efforts to improve the scaling efficiency of weather and climate models on parallel computer systems. Nevertheless, this effort should complement, not replace, the use of present-generation global weather and climate models on such systems, where the ensemble approach is both scientifically necessary and near-perfectly scalable. As an example, evaluation of the EC-Earth coupled GCM indicates reasonable scaling for a ~0.1° coupled system out to ~15,000 compute cores. Individual components of this coupled model would likely attain better scaling. The need to sample both initial condition and model uncertainty space in seasonal and decadal climate prediction, motivates the parallel integration of O(400) such simulations, indicating the potential for productive use of ~6 million compute cores.

Future research should concentrate on fully investigating the geographical and temporal distribution of initial condition uncertainty in terms of its impact on the accuracy of seasonal to decadal predictions, as has occurred in medium-range weather forecasting over the past decades. Systematic model biases, arising from a poor understanding and model representation of key climate processes, remains a major factor in forecast inaccuracy and requires continued improvement. More effort should also be directed towards the development of efficient methodologies for encompassing structural model uncertainties, such as non-definable (or non-unique) constants in physical parameterization schemes and their contribution to forecast uncertainty. In particular, the new technique of introducing a stochastic component to model parameterization schemes should be furthered developed and fully evaluated against more standard multi-model and single model/perturbed physics approaches.

A key problem to address for the successful application of massively parallellized, ensemble weather and climate prediction is the large data volumes that will be rapidly produced. In terms of model read/write performance, continued efforts in efficient parallel I/O will contribute to improved scalability. A significant degree of data thinning can be achieved by on-the-fly ensemble post processing. Nevertheless, subsequent downscaling and the application of ensemble forecasts in numerous downstream activities, requires the archival and, particularly for weather forecasting, rapid distribution of ensemble output to a wide set of users.

Cross-Cutting Considerations

Cross-cutting areas include the need for stable compute node performance during model simulation and the requirement for reliable, preferably automated, job restartability in relation to node failure. Efficient job-control and node allocation across numerous, large parallel jobs will be an important factor determining ensemble performance. Finally, treatment and archiving of large data volumes will be a cross-cutting activity for all aspects of weather and climate modeling on massively-parallel systems.

4.2 Data Intensive applications in Earth Science

4.2.1 Introduction

This section focuses on data intensive applications in Earth Science. In particular “Scientific Data Management” and “Data analysis and Visualization” have been recognized as the two key aspects of this section. It is important to note that general data management aspects like “Input/Output” and “Data Discovery” have been discussed in this roadmap as well, but they are presented into Section 4.3 Cross-cutting dimensions, since they have been recognized as cross-working group topics.

4.2.2 Scientific and technical vision

4.2.2.1 Scientific Data Management

The management of data faces many challenges as rapidly increasing computational power and similarly fast progress in satellite sensor technology create floods of data. Science is not alone in facing an explosion of data, with the consequence that a range of technological solutions are emerging as industry responds to sector wide demand for storage devices at lower cost and lower power use. Scientific data centres need to be able to exploit these new technologies at the same time as software deployed at the archive can be thought of in layers – some terminology is introduced here to focus the discussion below:

L0: software that extracts data from the storage media and delivers it through a single API which is independent of the media. If a data centre uses a single media for its user-accessible data, the L0 layer can be trivial, or at least very generic. E.g., an NFS file system for a spinning disk archive.

L1: software which acts on the data locally and presents results through a local API. This layer will contain any computationally intensive tasks which are integrated into the data management stack. For instance, visualisation and data analysis software described in section 4.2.2.2 could be located in L1, or could be external to the management stack, accessing data through the external APIs provided by L2 below.

L2: software which presents an external API, such as Open Geo-spatial Consortium (OGC) web services, Data Access Protocol (DAP).

L3: software which presents a user orientated API, such as a browser interface to OGC services.

The definition of these levels is not intended to be prescriptive: some archive configurations will have a single software package spanning multiple levels (e.g., a database with a built in web API) and others will use multiple packages within these levels. The levels are only intended to give a well defined terminology facilitate a concise description of the role of software packages and bundles in the archive.

Technology and Science Drivers

Climate Science Drivers

The CMIP5 archive is pushing the boundaries of data management within the climate modelling community, with an expected volume of around 10Pb. Rapidly increasing HPC performance will be reflected in increased data volumes, pushing towards and 1Eb archive within a decade. Dealing with such volumes of data will require fundamental shifts in data management and analysis methodologies.

Consider a query, of the kind that might be posed to an archive of the exa-scale future, requesting the projected frequency of intense tropical cyclones in some region of the globe for input into an impacts model. The projections of cyclone activity will need to be sampled across a probability distribution of outcomes from multiple climate realizations from different models, stored at different locations. The execution of the query will involve several steps that, in an exa-scale environment, will need to execute automatically:

- evaluation of provenance and quality control meta-data to determine which datasets to include;
- despatch of queries to data nodes, negotiating authentication and access control layers;
- collection of results from the data nodes, evaluation of return codes for fault detection;
- further calculations to combine collected results;
- archive results for re-use;
- delivery of processed results to the end-user, perhaps in deferred fashion if the associated computation needs to be scheduled on a "cloud".

The provision of quality control information poses both informatics and climate science challenges: techniques for assessing uncertainty are evolving rapidly as data volumes grow. The informatics challenge is to develop a structured vocabulary to describe characterisations of uncertainty.

	Tera-world	Peta-world	Exa-world
Step one	Copy data to local network	Copy selected fields to local network	Evaluate a test function on the data to determine potential “regions of interest”
Step two	Plot some fields	Evaluate a test function on data to determine “regions of interest”	Copy region of interest for selected fields
Step three	Detailed analysis	Copy additional data relevant to region of interest	More detailed analysis to determine region of interest
Step four		Detailed analysis	Download data
Step five			Detailed analysis

In order to reduce the data transfer demand, additional capability will be required to evaluate the data prior to download. In the climate science context, evaluation should include simple statistical measures (means, distributions) evaluated on arbitrary domains. Domain specification should be flexible enough: e.g. “variance of temperature in clear sky, in the tropics” – the “clear sky” constraint here restricts the calculation to regions of the domain with no clouds.

Technology drivers

Increasing rate of data supply

The rapid increase of HPC centre productivity and parallel increases in sensor technologies can be expected to increase data flow by a factor of a 1000 in the coming decade, taking us from Peta- to Exa-scale archives.

Power supply constraints

The rapidly falling cost of computation is expected to lead to a correspondingly rapid increase in data generation. Data volumes can be expected to increase by a factor of 100-1000 over the next 10 years. Over the same time period, the fall in energy usage per byte of stored data held on disk may only be a factor 10, compared to a projected fall in procurement costs by a factor of 200. We are likely to move from a regime in which procurement costs dominate to one in which power costs dominate. This will greatly increase the cost of holding data on disk, demanding more power aware data management strategies. Major commercial data centres are responding by placing major data archives near sources of cheap and sustainable cooling and power.

Heterogeneity of storage media

The growing complexity of storage systems will challenge management strategies. A storage centre might contain a selection of storage technologies: traditional disk, micro-servers, tape, solid state, WORM1. In addition, the disks will have multiple modes of operation: full speed, slow, idle, rest. A multi-state cache algorithm will be needed to optimise usage where more than two technologies are deployed and disks are in multiple states.

Heterogeneity of computing clusters

In addition, there are options for deploying software at various locations in the archive. If data is held on RAID arrays, each array might have 4 CPUs attached. The archive might have another 100 or so CPUs in the same rack as the data. The institution hosting the archive may also have a computational resource of several 1000 CPUs in the same building. In the Exa-scale future these numbers will change, but the range of options is likely to increase. The falling cost of CPUs will increasingly mean that rather than keeping CPUs busy, the resource optimisation will be targeted at using those CPUs which require least data movement.

Specialised processors

New opportunities are offered by the emergence of affordable and flexible accelerators (such as graphical processing units - GPUs) and field-programmable gate arrays (FPGAs). Data centres offering processing services will want to focus on the provision of resources for data intensive calculations. The potential of specialised processors to deliver such services should be investigated.

Data-aware grid scheduling

There is a need to develop and deploy scheduling algorithms which can take into account data transfer costs: e.g., a job requires 1000CPU hours and 2Pb of data; it can be run near (on a 20Gbps link) the archive on 8 cores, or at a remote centre (on a 400Mbps link) on 500 cores – where should it be scheduled?

Web services

Users need more flexible access to data in order to deal with increasing data volumes and complexity. A plethora of portals have emerged providing a vast range of services targeted at many different user groups. The archives should seek to provide a core set of standardised services on which more generic, community-specific services can be build. The challenge will be to select these standardised services in such a way as to maintain flexibility and efficiency.

Alternative R&D Strategies

Taking the computation to the data:

- A) providing access to a set of tools configured to run at very high efficiency on accelerator enabled cluster at the archive;
- B) provide access to generic computational cluster with high band width link to archive;
- C) ensure that key national supercomputing resources have dedicated link to archive;

In solution (A), the instruction set will have to be limited to ensure that the limited resources located next to the archive are used efficiently. Solution (B) gives the user greater flexibility, which in turn requires that a greater computational resource be made available to support this flexibility. Solution (C) links into existing computational resources and focuses on making those resources close to the data in network space. There may be limits on how efficient such a link can be.

Grid vs. cloud

GRID and cloud computing can be seen as alternatives or used together. The key difference from the user perspective is whether they have a dedicated machine with complete freedom to specify installed libraries, or share access to a pre-configured machine. If the latter option ischosen, the implementation cost will be reduced, and the utility to users increased if there is some standardisation of the resource configuration. A cloud solution could allow users to have cloned virtual machines running at a range of data centres.

Scientific Workflow Tools

In an exa-scale environment, users will not want to inspect the result of each operation, but rather chain multiple operations together. Accurate and robust provenance meta-data and its propagation through the workflow will be central to the creation of a robust and resilient system. Reliable quality control will also be required at each step of the workflow. Users should be able to test and build workflows on their local machines and submit them for execution on resources near archives. Development of systems to support workflows with the appropriate level of resilience and user support needs to be investigated. A decision needs to be made on the choice of supported workflow language (e.g. BPEL or YAWL).

Scalable data formats

Standards for the description of uncertainty sampling strategies need to be adopted, in order that analysis tools can treat ensembles of simulations appropriately. The current NetCDF standard limits to one "unbounded" dimension, which limits flexibility of data generation and analysis software: we need to relax this constraint. NetCDF 4 allows a greater flexibility through user-defined data structures: we need to develop conventions for exploiting this flexibility in a way which is accessible to analysis to scalable analysis tools. Developing conventions for user-defined data structures is unlikely to be a smooth process. An initial target will be to draw up a requirements document. A range of options for describing complex data features exists (E.g. CSML, GML, KML).

Search and Query tools

We need to integrate function evaluation in the search interface, to allow for searches which are contingent numerical calculations rather than simply on matching character strings. Some portals already have this capability implemented for spatial coordinates (i.e., the ability to search for data within a user-specified latitude and longitude range). A design study will need to be conducted to establish a useful and realistic scope for this activity, as there are numerous design choices to be made. Can such a search be implemented for numerical calculations on the data (e.g., "*datasets with maximum 10m wind speed greater than 30ms-1*"), or should it be limited to calculations on the meta-data?

Storage media, backup and curation

Industry is offering a growing range of storage solutions, and there is an even larger range of tools to support efficient use of the media. Some storage media options are listed in the "technology drivers" section. Any combination of such media will require a software layer which can present a clear interface to the users. It will be necessary to choose between commercial and open source/community developed solutions. The importance of long term stability and guaranteed access to the source favours open source, but commercial solutions may be better able to respond to rapid changes in hardware capabilities.

Hardware deployment within the archive

Efficient use of data archives will depend on coordinated deployment of resources, and flexible management strategies. The storage facilities at current HPC centres generally constitute a severe restriction for the climate science community, but there is no technical barrier to increasing the size of the storage facility. However, it may be more cost-effective to enhance the data transfer between HPC centres a data archives by purchasing dedicated links.

Archive locations

As energy usage becomes more important consideration should be given to the optimal location for data centres. Decisions by major commercial data centres to locate close to cheap electricity and natural cooling resources demonstrate the importance of power costs in determining optimal locations. Are there economies of scale to be gained by, for instance, creating an exa-scale archive, at an optimal location, for scientific data, with domain specific L1-L3 software maintained by domain specific data centres running on top of generic L0 software?

Robust meta-data collection

There are two primary options for collecting robust meta-data:

- A) Collect meta-data automatically at run-time;
- B) Use meta-data to drive configuration.

The collection of meta-data at runtime (option A) requires that the Earth System modelling codes be capable of defining their configuration in a standardised and extensible syntax. This will place a heavy requirement on the code developers, but there will be significant co-benefits in the form of self-documenting code: with the current generation of models, the modelling centres must dedicate a significant staff resource to entering into the archive meta-data catalogue information about the model configuration which has already been specified in the configuration files used to run the experiments. This approach does not scale.

Option (B) achieves the same, but requires that models be designed to extract information from the meta-data. On the basis that translation into one's own language is easier than translation into someone else's, this approach will be easier than asking modelling centres to provide automated translation of their configuration files. On the other hand, many of their experiments will be exploratory and not destined for archive. The overhead of using a generic meta-data system for all experiments may be considered too large. Consultation with the modelling groups will be required to resolve this.

The same problems do not apply to satellite datasets, since there the space agencies have established quality control procedures. An upward harmonisation of meta-data quality control would be beneficial.

Adherence to REST principles

The REST2 principles establish a set of design rules for Web Services which are intended to facilitate parallel development of different components of the system and enhance interoperability. In the short term the constraint imposed on the developers of such systems may appear detrimental, such that many systems depart from REST principles. Should adherence to REST be advanced as a desirable or essential feature of components of the exa-scale service stack?

Recommended Research Agenda

Timeframe	Targets and Milestones – Scientific Data Management
2010-11	• Adopt conventions for in-file metadata describing ensembles and grid structures • Implement WPS with chainable security infrastructure • Standardised quality control of datasets • White paper on implications of community vs. commercial L0 software layer
2012-13	• Specify service requirements • Enhanced automated meta-data collection from experiment configuration • Integrate data archive hardware infrastructure with data producing HPC centres • White paper on potential for use of specialised processors in L1 software layer
2014-15	• Programmatic processing enabled on storage clusters, with extensive range of functions and operations • Libraries supporting run-time meta-data generation • Programmatic search query specification document
2016-17	• Caching of analytics enabled • Workflow for cataloguing and publication of derived products established
2018-19	• Work-flows spanning data ingestion to impacts analysis • Coupling of processing services and visualisation services

Cross-Cutting Considerations

With Earth System Modelling

Collection of meta-data is closely linked to Earth System Model workflow development.

With "Data Analysis and Visualization"

Here, as in the IESP roadmap, data analysis issues have been included in the scope of "Scientific Data Management".

4.2.2.2 Data analysis and Visualization

Technology and Science Drivers

In all areas of Earth and environmental sciences, the amount of data is increasing at an exponential rate. Due to technological advances, more and more observational data acquired by satellites as well as results of simulations with numerical models of the Earth system need to be stored, analyzed and visualized at a continuously increasing spatial and temporal resolution. In Geosciences, as in most scientific fields, international standards for data and corresponding metadata exist and data are openly accessible through dedicated data centers. Many of the geo-scientific problems with pronounced societal relevance (e.g. climate change, extreme weather, natural hazards) have a strong interdisciplinary character and will require the integration and analysis of gigantic heterogeneous data volumes with tools that do not exist today.

The growing data size, complexity and the overall data volume within single research projects poses considerable challenges to the scientific workflow. Problems are induced by:

- limited bandwidth of data transfer via networks
- limitations in storage capacity and performance
- lack of scalability of analysis and visualization tools
- limited scalability of metadata

As the network bandwidth increases at a slower rate as the data size grows, analysis and visualization will need to get more closely integrated with the data producing applications, or, more specifically, large scale international data centers will have to get collocated with supercomputing facilities.

Alternative R&D Strategies

Data analysis and visualization tools have to go far beyond the current state of the art. Today, post-processing, analysis and visualization is done offline in separate steps after the acquisition or generation and storage of the data. Numerous different non-standardized tools, most of them not even parallelized, are used by the community for these steps. In many cases, some of the steps are done on separate specialized systems such as graphics workstations.

In the Exascale world, this strategy might no longer be feasible. Due to the high spatial and temporal sampling of the original data, it cannot be stored completely at feasible costs or within reasonable time. The data volume to be stored has to be reduced during runtime of the data producing application. This imposes a drastic change of the typical workflow applied today. Consequently, the scientific questions posed to the data need to be precisely defined before the simulation starts. A large fraction of the data is only available during runtime of the simulation or during data acquisition. Questions arising after filtering and storing the data might pose the need to repeat a simulation to restore the original highly resolved information.

When scientific questions require analysis and visualization of results at full available spatial and temporal resolution, but the full resolution cannot be stored, this part of the workflow will have to be integrated with the application itself. In case of model simulations, this would require to filter and visualize the data directly on the supercomputer, while it is still in memory (In-situ analysis). Analysis and visualization parameters will need to be defined prior to simulation.

With respect to this data challenge, research and development strategies should include:

- the development of scalable post-processing tools,
- the development of cross-disciplinary parallel data analysis and visualization tools
- integration of post-processing, analysis and visualization with the data generation workflow.

Recommended Research Agenda

Due to the scale of the necessary software developments, one of the most important aspects is the coordination within and across the related disciplines at an early stage. Some of the communities are only now starting to unify their e-infrastructure with respect to data formats, metadata conventions, standard products and workflows. A standardization of these general aspects across disciplines needs to be established as early as possible.

Furthermore, the focus needs to be set on scalability issues. The workflows have to be analyzed with respect to optimal data reduction strategies, parallelization and potential for knowledge based data analysis and visualization.

Specific research areas should include:

- definition of domain specific requirements across fields
- establishment of standard workflows, data and metadata standards and data exchange methodologies
- investigation of new strategies with respect to the scalability of data analysis, filtering, post-processing, and (interactive) visualization
- knowledge based data analysis and visualization

Cross-Cutting Considerations

In Earth system modeling, ensemble simulations are regularly used in order to reduce the uncertainty in the simulation results. Visualization of statistical measures such as uncertainty, probability, variability and its temporal development is still an active research issue in computer science. Results from this research will have to be integrated in the analysis and visualization solutions in order to satisfy the scientific needs.

Key issues in Earth and Environmental Sciences are the integration of large-scale multi-disciplinary observational and synthetic data into workflows that take scientist from the problem position to the testing of specific hypotheses. This will require not only the standardization of model and data formats, the handling of derived data, but also long-term planning of storage and access in the vicinity of exascale computing facilities.

With increasing data size, analysis and visualization services need to be tightly coupled with the supercomputing facilities and data repositories. Additionally, in-situ analysis and visualization might benefit from hybrid supercomputing systems: the use of embedded GPGPU type accelerators should be evaluated with respect to visualization tasks.

4.3 Cross-cutting dimensions

4.3.1 Introduction

This section presents from a scientific point of view the most relevant cross-cutting dimensions. In particular the working group experts have focus on:

- performance
- programmability
- workflow
- data management

For the sake of completeness, it is worth mentioning here, that the data management part deals with two key aspects like: Input/Output and Data discovery.

4.3.2 Scientific and technical vision

4.3.2.1 Performance

Linpack benchmarking represents a good starting point to evaluate the performances of one exascale computer. [3]

However, a performance measure adapted to our applications is still needed and we propose here to use a unit adapted to climate or weather applications like number of simulated years (days) per day[4]. A similar measure unit is relevant for Earth Science applications.

This measure should include all the processes from initialisation until outputs are produced and ready for analyses. This measure should help to understand and to improve load balancing between different concurrent processes.

An important measure concerns the scientific performance of the model itself i.e. the realism of the simulation. A set of metrics exists, but the most appropriate depends on the scientific question considered. This field currently implies a lot of active research. [5]

Exascale computers will give us the opportunity to increase the resolution in horizontal and vertical mesh. But the scientific performance of the simulation will not necessary be improved, considering the metrics mentioned above. Even with an increase of resolution (both vertical and horizontal), we will still be required to run a large number of simulations to improve the realism of the physical processes so to improve the different metrics.

In terms of resolutions (both vertical and horizontal) some thresholds exist and require important developments to model the different processes adequately (e.g. at 10 km non hydrostatic models are mandatory).

4.3.2.2 Programmability

Programmability is an abstract property that expresses the property of a computing system – from processor to user interface – to be easily used for a specific purpose, in our case to build, run and maintain software serving the climate science community's needs. It indirectly includes the capability of such software to survive technology changes without a complete redesign and without losing the core features of the original application. As a cross-cutting property it involves not only the ready

[3] Top500.org

[4] ECMWF, Fourteenth Workshop on Use of High Performance Computing in Meteorology, 1 - 5 November 2010

[5] IPCC Expert Meeting on Assessing and Combining Multi Model Climate Projections National Center for Atmospheric Research, Boulder, Colorado, USA, 25-27 January 2010, Stocker, T., Q. Dahe, G. Plattner, M. Tignor, P. Midgley (Eds.) Meeting Report.

availability of extensive information about the computer systems, including computer architecture, compilers, programming languages, software stacks and libraries, but also a clear view on the architectural specifics of future systems, and a thorough understanding of the special needs of a specific scientific area.

In this sense, the modelling community is inherently dependent on the computer system community to provide tools, methodologies and expertise to exploit the upcoming exascale computing platforms. More intelligent tools, such as smart compilers and easier-to-use run-time environments, should unburden researchers and scientific programmers from architectural and programming details such as task granularity and accelerators, tolerance to system faults, architecture optimizations or data locality. At the same time, the modelling community should provide directions on future algorithmic and numerical requirments of their applications, and be prepared to incorporate newly developed tools as soon as they become available.

The science of climate modeling continues to advance, but it would be of particular importance to this community to reduce uncertainties in climate sensitivity and the different feedbacks. Climate models will also be useful to investigate mitigation proposals as well their limitations and uncertainties; and this requires many long years' simulations with high resolution. In consequence, strategies of a new programming able to use the potential of the new machines are a key issue.

Also, to infer climate change impacts on air quality, water, ecosystem, health, forest and last but not least societal behaviour from model simulations depends upon simulations performed by global climate models complemented by regional models. This will require stronger interaction with a large interdisciplinary community of users and the development of models coupling a large range of impact models to different climate models.

Next generation climate models will often include chemical gases and aerosols that increase the computation time by a factor of 5 to 10. We need to quantify the uncertainty through ensemble runs, so more systematic ensembles of 10 or more runs should be performed to better account for the internal variability. We need long time simulations to know possible climate thresholds and gain paleo-climate understanding.

Technology and Science Drivers

The majority of models currently in use are evolutions of older, sequential software. They now incorporate MPI layers to take advantage of commodity clusters or OpenMP layers to exploit the more recent multi-core rush. Models should continue to scale up, both in the direction of multi-node and multi-core systems, and by adding modelling processes and variables and increasing resolutions. But the outcome of exascale computing system will strongly depend on the ability of the computer science community to provide especially adequate software, intelligent tools and common guidelines to exploit the exascale intrinsic power.

Still, according to the latest IESP roadmap, it is unclear if current ESFM, CC, NWP or AQM models will be able to exploit the power of exascale systems, or if models which are programmed in a new – (revolutionary?) way will overcome them. Until then, the modeling community needs to constantly rethink their current models, preserving the knowledge and experience inside of them, while redesigning them with modularity, reusability, portability and flexibility of the I/O as major aims.

Alternative R&D Strategies for Programmability

Exascale computations will allow a dramatic increase in model resolution, process, phenomena and variables. This will in turn generate massive amounts of data which need to be stored, searched, analyzed and transformed according to users' needs. This challenges the input/output methodologies as well as the storage systems. The exponential increase in data volume poses a serious contest to climate science, and new approaches are needed to minimize potentially crippling inefficiencies.

Thus a common framework for data storage, search, and retrieval would be the next requirement, and data locality and access will be its major performance and usability drivers. The potential to search, retrieve and merge existing data from different sources calls for a widespread usage of common data representation and visualization formats, which can provide high-quality contents through consistent interfaces among the entire community

Additionally, the current models mesh allows a thorough evaluation of model-coupling opportunities and their ability to increase the global research value by subdividing the expertise among different groups.

Recommended Research Agenda

Modularization: Modularization allows isolating core model components from common routines. As a consequence, the likely need for code rewriting as new programming models emerge would involve much less effort both in term of programming and debugging. Common interfaces permit rapidly changing only those components or sub-components which mostly affect the performance of a given application.

Code reusability: Code reusability involves the use of routines to perform common operations. Their use should rely as much as possible on standardized or open source libraries. This is especially critical for numerical routines. which seem to be the most architecture-dependent portion of the software stack, and most likely the one with greater impact on the portability of future models.

Portability: Can be defined as a set of procedures that premits the use of the same source code on different architectures and run-time environments. It involves knowledge of programming languages and compiler features among different architectures, and it often crosses paths with code optimization, thus influencing the performance across different architectures up to the point of rendering the application useless. Portability implies access and testing on a pool of diverse and heterogeneous computing resources, from mini-clusters inside smaller research facilities to larger governmental or cross-country super-computation centres. Portability might cost some architectural benefits, such as specific architectural optimizations, which should be pushed directly into libraries, but comes with the advantage of a broader penetration for portable models inside the entire scientific community.

Flexible I/O: New models in the exascale range will allow for higher resolutions and a larger number of processes and variables, resulting in massive amounts of data, for which I/O will be a major performance driver. Thus a flexible, scalable, yet transparent high-performance I/O layer will be the strongest need. Without this layer, exascale applications might be unable to meet the near-time requirements of many scenarios, rendering the use of exascale machines for the climate modelling community an unrealistic approach.

Cross-Cutting Considerations

A greater effort should be directed in parallel to formal testing, which is a less common practice inside the modelling community; a standard development cycle could shorten software iterations and guarantee an adequate level of software quality.

Finally, code openness and access would help, together with portability, the dissemination of high quality software through the scientific community. Further advances in the new era of data-and computer-intensive science depend on how well researchers collaborate with one another and with technologists in the areas of database architectures, programmability, workflow management, visualization, etc.

The “forecast performance” rate (years simulated per real time day) is the most important metric to assess the performance of a climate model code on any given machine. This measure depends on the machine performance and the software employed. The present performances are in the range of 1-10 years/day, which would be completely insufficient for the increase of resolution and complexity, and for paleo-studies.

The enhancements in programming are basic to improve this rate, accompanied by measures of quality control and testing.

4.3.2.3 Workflows

The IESP roadmap defines scientific workflow as a series of structured activities, computation, data analysis, and knowledge discovery that arise in scientific problem-solving. Typically these workflows today employ tens to thousands of processors, cope with tens to thousands of experiments, produce GB to TB of data, and run for days to months, but normally employ only one numerical model instance, though with different boundary conditions, parameter sets, etc.

Technology and Science Drivers

So far workflows - as discernible software entities – play only a minor role in numerical Earth-System-Science, and if so, mainly in operational environments, and not very much in research environments. Major reasons seem to be the need for the researchers to go through a complete software development circle – development, testing, running, experimentation, evaluation – with their software tools – the models – over and over again. Albeit the fact that this may look like an easily automatable process, it obviously employs many different steps, tools and scientific issues in a – single or even a few - complex environment(s). Additionally, a typical numerical climate experiment lasts month to years and is not likely to repeat itself without major changes to the experimental setup. Thus, the readiness to add the complexity of development (like eclipse) and execution (like Kepler) workflow toolboxes is minimal, and the benefit employing such tools is not readily seen.

But these toolboxes, as we know from other research fields, can help hide parts of the complexity in general, and avoid repetitious work patterns. In principal the scripting technologies employed oftentimes today can be dubbed workflow systems, though without a GUI. But they are hardly scalable, not transferable across sites, and oftentimes very difficult to comprehend since they are developed in an ad-hoc, not very well planned and/or documented manor on an individual basis. The question arises as to how much this is a sustainable approach when it comes to carrying out these workflows on exascale machinery.

Another consideration in conjunction with exascale architectures and workflows is the potential demand to deploy models not only together with their forcing (i.e., input) data, but also with their compilation and runtime environment, down to the operating system. This obviously depends very much on the strategies the providers of exascale services follow, but could offer the possibility for the model "operators" to fine-tune and adjust all the necessary ingredients for a model experiment "at home", on a smaller scale machine with the same architecture as the tier 0 capability machine, and not waste precious cycles on that tier0 machine for testing and adjustment purposes, as it can be observed in many instances today.

It is not clear to the authors if multi-model-ensembles on single machines will play a larger role on future exascale machines, although this would obviously also influence the workflow development.

Turning the view away from large-scale modelling more into the evaluation part of our science, it can be estimated that web access and workflow tools will play a significantly larger role in the times of federated data centres to come, like in the German c3 Grid or the ESGF.

Alternative R&D Strategies

Given the potential of modern workflow tools and the recent advances in developing metadata standards like in the METAFOR project, as well as the complexity of modern architectures, it is hard to imagine how the field can continue to progress without employing these more modern tools. The alternative – continue to develop overly complex, architecture and site-dependent, script-based environments with short life times – shows a high potential for failure and will turn into a competitive dis-advantage for those institution stuck with it.

Recommended Research Agenda

So harvesting the community for best practices in terms of workflows, and translating script-based environments into their workflow-tool equivalents will be the method of choice. This approach demands to set up projects on the following issues: Adaptation and/or development of a workflow description language, expression of typical workflows in this language, generalisation and optimization of these workflows, development and teaching of best practices in this area. It should be stressed that a single experiment or set of related experiments (as for example ensemble simulations) may last many run for many calendar months and produces a continues and voluminous data stream during that time. Therefore bookkeeping and fault tolerance in earth system workflows is probably of much greater importance than in other disciplines.

To become widely accepted by climate modellers, workflow description should be easy to use and flexible and should include braches and loops, features that are not well supported in current tools.

Cross-Cutting Considerations

Issues of major concern for climate modellers are I/O, memory bandwidth, power efficiency, and – scientific and technical - stability of the models. How much better workflow representations can help in these issues needs to be seen. Better tools to monitor I/O, memory bandwidth and power efficiency behaviour of the models on the new architectures will be increasingly important. The stability of models will become more and more of an issue when MTBFs on the order of seconds will become commonplace. It is hard to imagine that the models can become more error-resilient than they are today.

4.3.2.4 Data Management

As stated in the introduction of Section 4.3, this paragraph presents both Input/Output and Data Discovery topics.

4.3.2.4.1 Input/Output

Technology and Science Drivers for Input/Output

The cost of computation is falling much faster than the cost of data movement, and this trend is expected to continue with the introduction of CPU accelerators into the HPC arena.

Alternative R&D Strategies for Input/Output

Co-development of I/O libraries

There is a clear need to work with the hardware providers to develop efficient I/O libraries which can cope with the demands of the user communities without requiring them to be familiar with all the details of the rapidly evolving computer architecture.

Early quality control

Quality control on the output of numerical simulations is often performed after output of results. Where I/O is a major cost factor, it will make sense to carry out quality control at an earlier stage to reduce the burden of unnecessary I/O traffic.

Data compression and information harvesting

Data compression may save significantly on I/O costs. Attention should also be given to the information content in the data stream.

Recommended Research Agenda for Input/Output

Timeframe	Targets and Milestones – Scientific Data: Input/Ouptut
2010-11	White paper on I/O requirements and costs Establish a co-development strategy for I/O libraries
2012-13	Co-development roadmap and MoUs
2014-15	I/O library, beta release
2016-17	Evaluation of library performance
2018-19	I/O library, operational release

Cross-Cutting Considerations

I/O library

The I/O library is likely to be of wider use: the co-development strategy should identify the appropriate scope and engage with the relevant communities.

4.3.2.4.2 Data discovery

Technology and Science Drivers for Data discovery

The growing importance of data intensive science and associated needs for reliable (i.e. robust and repeatable) automated discovery services. Repeatability is a major requirement in scientific research – current discovery services place the onus on the searcher and provide no explicit support for finding the results of a given query executed in the past.

Alternative R&D Strategies for Data discovery

Search and Query tools

Need to integrate function evaluation in the search interface, to allow for searches which are contingent numerical calculations rather than simply on matching character strings. Some portals already have this capability implemented for spatial coordinates (i.e. the ability to search for data within a user-specified latitude and longitude range). A design study will need to be conducted to establish a useful and realistic scope for this activity, as there are numerous design choices to be made. Can such a search be implemented for numerical calculations on the data (e.g. “datasets with maximum 10m wind speed greater than 30ms-1”), or should it be limited to calculations on the meta-data?

Generic, domain and community standards

The use of standards promotes interoperability, but the standards must be fit for purpose. Standards which are too generic encourage the development of multiple profiles within the standard while standards which are too specific encourage the development of extensions. The widespread use of profiles within standards or extensions to standards will damage interoperability.

Vocabularies and ontologies within standards

There are many approaches to maintaining vocabularies and ontologies. There is a need to establish/adopt standard procedures which will support co-development. Software developers, whether commercial or in the academic sector, cannot support standards if they are subject to arbitrary or unpredictable changes. This is particularly true of changes in ontologies.

Recommended Research Agenda for Data discovery

Timeframe	Targets and Milestones – Data Discovery
2010-11	Adopt conventions for in-file metadata describing ensembles and grid structures White paper on standards needed to support data discovery at generic, domain and community levels.
2012-13	Specify discovery service requirements (excluding programmatic search)
2014-15	Programmatic search query specification document
2016-17	Repeatable search service implemented
2018-19	Production level services for data discovery

4.4 WCES impacts and findings at European Level

4.4.1 Societal, environmental and economical impact

The European climate community has made significant contributions to the Intergovernmental Panel on Climate Change (IPCC) assessment reports, and has long been working within several European Framework Programs via projects devoted to improve the European knowledge of climate change and related issues. The scientific results of the Weather, Climate and Earth Sciences studies are short, medium and long term; they have a very strong impact on economy, soil, coast, agriculture, and health.

Climate scientists work also on the development of tools for scenario production of climate risk associated with several processes such as forest fires, desertification and land degradation, biodiversity changes and primary productivity. In addition, specific activities refer to changes in quality and quantity of agriculture food production as well as water use and consumption.

This will become more and more computing power demanding in the next decade and even more challenging from societal, economical and environmental points of view.

4.4.2 Strengths and weaknesses

A strong point is that the European climate modelling community forms a very active network of scientists and engineers. Indeed, several European partners agreed to create the European Network for Earth System Modelling -ENES- with the purpose of working together and cooperating towards the development of a European network for Earth system modelling. These institutions include university departments, research centres, meteorological services, computer centres and industrial partners.

Another strength is the diversity of the European community in terms of their models and approaches to scientific issues (*climate science relies on such a diversity*): Quite regularly there are several (sometimes competing) solutions developed for the same problem by different groups across the European countries. This allows climate scientists and other researchersto take into account several relevant aspects and issues from different angels. It can also be considered a weakness since such an approach requires a lot of resources. There is the need to manage the trade-off in coordinating these efforts keeping the (scientific) diversity as main asset.

From a software point of view, a good point is that most of the EU groups share common coupling software. For instance, most of the groups within this network use the OASIS coupler to assemble the Earth System model components developed independently by different research groups to form global coupled climate models.

A weak point highlighted by the WCES group, in particular for the climate domain, is the weak scalability of climate models. This will probably make the rewriting of several climate codes for the new Exascale machines mandatory. Co-design represents the proper way to address this point.

The European Plate Observing System (EPOS) is another strength for the WCES community, in particular for the solid earth part. It represents a relevant research infrastructure and e-Science for data and observatories on earthquakes, volcanoes, surface dynamics and tectonics.

4.4.3 Needs for education and training

The need for education and training will strongly increase for the Exascale age, due to the complexity of the HPC environment and the relevance of the scientific challenge. Training programs will allow WCES scientists to improve their HPC background as well as to establish stronger links between the HPC community and their own domain. In this respect funding specific actions to support training activities, summer/winter schools, intra-European fellowships as well as international incoming and outgoing fellowships, will play a strategic role to prepare new scientists with a stronger and more interdisciplinary background. Given the expected increased complexity of the component models and of future exascale computing platforms, a lot of resources should be devoted to the technical aspects of coupled climate modelling; the coupler development teams should be reinforced, including experts in computing science remaining at the same time very close to the climate modelling scientists.

4.4.4 Potential collaborations outside Europe

Collaborations outside Europe are a key point for the WCES communities to globally share and face the Exascale challenges. Such collaborations are already taking place thanks to international intercomparison projects such as CMIP5, in particular in the development of common metadata (EU METAFOR and US CURATOR projects) and in the distribution of data (EU IS-ENES project and US ESG). However, these collaborations need to be reinforced. For example, in terms of development of coupling software, it is expected in the coming years that the different coupler developers interact more closely at the international level and share more infrastructure building blocks (e.g. regridding algorithms, decomposition description, etc.); in this area as in others, best practices in coupling should be discussed, identified, and promoted.

An important example for the fostering of such collaborations is the G8 call: Proposals were prepared from the beginning by internationally recognized experts in the fields in question. Another way to have a stronger level of collaboration outside Europe could be by means of specific actions funding series of workshops focusing on specific themes (coupling software, dynamical cores, scientific data management, etc.) to be held on a regular basis (each 9 or 12 months for instance). Such an activity could represent something like a “synchronization point” at a worldwide level for the community and could drive the community itself towards a more effective solution of problems and the establishment of new collaborations.

4.4.5 Agenda, need of HR, provisional costs

The research agenda and steps towards Exascale should include both hardware and software funding. It has been widely recognized by the working group experts as a key point that Exascale issues must be addressed at several levels from the hardware to the applications.

The establishment of co-design centres is crucial to allow people with different expertise and background to share the global context and mission, and at the same time focus on different aspects of the Exascale stack. This kind of centres will represent an aggregation point for scientists: Here future WCES scenarios and use cases could be addressed in a more effective way than ever.

Establishing co-design centres at the European level, represents the best scenario (plan A) from the WCES working group perspective. The WCES challenges are extremely complex and need more articulated and multifaceted answers. Weather, Climatology and Earth Sciences focus on global issues and for this reason without the proper support, the Exascale opportunity could result in loss of competitiveness for the European scientific community. A plan B could be to continue to apply for European calls on different topics of the same (Exascale) medal, with a lower level of interaction across the Exascale stack. Obviously the level of effectiveness in terms of achieved results could not be the same of plan A.

As we can argue, co-design centres need strong support in terms of resources and for this reason, the number of these centres should be carefully evaluated, along their geographic location, the number of years for these programs (at least 5), etc.

For example the CRAY XT6 serving as the premier supercomputer resource for NOAA’s National Climate-Computing Research Center (NCRC) has been funded within a five year program worth 215 million dollar program that includes funds for the machine as well as for manpower to operate the system.

In the following table, a provisional costs estimation related to the main topics discussed in this roadmap is presented.

Roadmap issue	**FTE/year à 2015**		**Integrated (4 years) Provisional Costs 2012 à 2015**		**FTE/year 2015 à 2020**		**Integrated (5 years) Provisional Costs 2015 à 2020**	
	Min	**Max**	**Min**	**Max**	**Min**	**Max**	**Min**	**Max**
Earth Science Models								
Earth System Models optimization	35	45	11200	14400	70	90	35000	45000
Dynamical cores	20	25	6400	8000	40	50	20000	25000
Flexible couplers and coupling frameworks	5	10	1600	3200	10	20	5000	10000
Multiscale Simulation	12	15	3840	4800	24	30	12000	15000
Data Assimilation	15	20	4800	6400	30	40	15000	20000
Uncertainty and predictability	12	15	3840	4800	24	30	12000	15000
Data Intensive appl. in Earth Science								
Scientific data management	25	30	8000	9600	50	60	25000	30000
Data analysis and visualization	20	25	6400	8000	40	50	20000	25000
Cross-cutting dimensions								
Performance	15	25	4800	8000	30	50	15000	25000
Programmability	10	15	3200	4800	20	30	10000	15000
Data Management - I/O	25	30	8000	9600	50	60	25000	30000
Data Management - Data Discovery	15	20	4800	6400	30	40	15000	20000
Others	10	15	3200	4800	20	30	10000	15000
Total	**219**	**290**	**70080**	**92800**	**438**	**580**	**219000**	**290000**
Cost for 1 FTE/year (KEuro) - 2012-2015	**80**							
Cost for 1 FTE/year (KEuro) - 2016-2020	**100**							

5. Conclusions

This document presented (at the application level) issues, requirements, expectations, agenda and impacts related to the Weather, Climate and Earth Sciences domains in the Exascale timeframe. It has been prepared by the EESI WCES Working group, that is a team of European scientists with complementary expertise and knowledge operating in these challenging application domains.
The complementarity of the team has represented a key point for this document to provide a clear and complete understanding about the future scenarios towards exascale computing in the WCES domains.

6. WCES Experts

EESI - WG 3.2	WCES (Weather, Climatology and Earth Sciences)		
Name (alphabetical order by country)	**Organization**	**Country**	**Area of Expertise**
Giovanni Aloisio (chair)	ENES-CMCC	IT	Exascale Computing
Massimo Cocco (co-chair)	INGV	IT	Seismology
Nadia Pinardi	CMCC	IT	Oceanography
Joachim Biercamp	DKRZ	DE	High Performance Computing
Reinhard Budich	MPI	DE	Earth System Modeling
Heiner Igel	LMU	DE	Solid Earth, Geophysics
Marie-Alice Foujols	IPSL	FR	Climate Modeling
Sylvie Joussaume	IPSL	FR	Earth-system Modeling
Sophie Valcke	CERFACS	FR	Coupled Climate Models
Jean-Pierre Vilotte	IPGP	FR	Solid Earth, Seismology
Mats Hamrud	ECMWF	UK	Data Assimilation
Bryan Lawrence/Martin Juckes	BADC	UK	Climate Data Management
Tim Palmer	ECMWF	UK	Weather/Climate Modeling
Graham Riley	Manchester Univ.	UK	High Performance Computing
Jose Baldasano	BSC	ES	Earth Sciences
Johan Silen	FMI	FI	Geophysics & Computing
Colin Jones	SMHI	SE	Climate Change Modeling

