Proposal n°2012060992

General information

Type of proposal: **Multiyear access**

Is this proposal a :

Resubmission : No **Continuation : No**

ID of the original proposal : – **ID of the previous proposal : –**

Note: *For proposals requesting access as continuation to previous access, it is mandatory to present the final report or a progress report at the time of the closure of the Call. This report should be sent to the Peer Review Team (**peer-review@prace-ri.eu**), and analysed by the Access Committee to evaluate the status of on-going access. The template document for this report is available on the PRACE website ("**Information for PRACE Awardees**"), and it must be carefully respected.*

Start date: 1st November 2012

Project name: HiResClim : High Resolution Climate Modelling

Research fields:

Please specify here after at least one research field along with a percentage of how much it matches your proposal. You may also provide a few additional keywords (at most six, each consisting in at most three words). This information will further be used during scientific assessment for choosing reviewers.

Research field :	How much fitted (%) ?
PE10_2 – Meteorology, atmospheric physics and dynamics	25
PE10_3 – Climatology and climate change	50
PE10_8 – Oceanography (physical, chemical, biological, geological)	25

Additional keywords :

High Resolution Coupled Global Climate Models Ensemble climate modelling Climate Change

Contact person for all correspondence

Name: Colin Jones

E-mail: colin.jones@smhi.se

Project leader (personal data and contact)

For this call, proposals from academia are eligible, as long as the project leader is a senior researcher employed in a research organisation homed in a European Union country or a PRACE Association member country.

☑ **The employment contract of the project leader with the research organisation must be valid at least 3 months after the end of the allocation period.**

See other eligibility criteria on the **PRACE website**.

Gender: Male

Title: Dr

First name:	Colin
Last name:	Jones
Initials:	G
Date of birth:	13101967
Nationality:	British
e-mail :	colin.jones@smhi.se
Phone number:	+46114958032
Fax number:	+46114958001

Project leader (organisation and job title)

Job title:	Head of Climate Research
Organisation name:	Swedish Meterological and Hydrological Institute (SMHI)

Organisation with an R&D mandate: **Yes**

For commercial companies,

- **Is the head office of the organisation in Europe? No**
- **% of R&D activity in Europe as compared to total R&D activity : –**

Department:	Research and Development
Group:	Rossby Centre
Address:	Folkborgsvagen 1
Postal code:	60176
City:	Norrkoping
Country:	SWEDEN

Collaborators

Organisation name:	CERFACS	**Address:**	42 Avenue G Coriolis
Department:	Global Change	**Postal code:**	31057
Group:		**City:**	Toulouse
		Country:	FRANCE

Dr Laurent Terray

Job title:	Research Director
Gender:	Male
Title:	Dr
First name:	Laurent
Last name:	Terray
Initials:	
Date of birth:	06061960
Nationality:	French
e-mail :	terray@cerfacs.fr
Phone number:	+33561193015
Fax number:	+33561193000

Dr Sophie Valcke

Job title:	Research Engineer
Gender:	Female
Title:	Dr
First name:	Sophie

Last name:	Valcke
Initials:	
Date of birth:	24021966
Nationality:	French
e-mail :	valcke@cerfacs.fr
Phone number:	+33561193074
Fax number:	+33561193000

Mr Eric Maisonnave

Job title:	Research Engineer
Gender:	Male
Title:	Mr
First name:	Eric
Last name:	Maisonnave
Initials:	
Date of birth:	24011971
Nationality:	French
e-mail :	eric.maisonnave@cerfacs.fr
Phone number:	+33561193012
Fax number:	+33561193000

Dr Christophe Cassou

Job title:	Researcher
Gender:	Male
Title:	Dr
First name:	Christophe
Last name:	Cassou
Initials:	
Date of birth:	24011971
Nationality:	French
e-mail :	cassou@cerfacs.fr
Phone number:	+33561193074
Fax number:	+33561193000

Organisation name:	Swedish Meterological and Hydrological Institute (SMHI)	**Address:**	Folkborgsvagen 1
Department:	Research and Development	**Postal code:**	60176
Group:	Rossby Centre	**City:**	Norrkoping
		Country:	SWEDEN

Dr Klaus Wyser

Job title:	Researcher
Gender:	Male
Title:	Dr
First name:	Klaus
Last name:	Wyser
Initials:	
Date of birth:	03121965
Nationality:	Swiss
e-mail :	klaus.wyser@smhi.se

Phone number:	+46114958000
Fax number:	+46114958001

Dr Uwe Fladrich

Job title:	Computer Scientist
Gender:	Male
Title:	Dr
First name:	Uwe
Last name:	Fladrich
Initials:	
Date of birth:	03051968
Nationality:	German
e-mail :	uwe.fladrich@smhi.se
Phone number:	+46114958000
Fax number:	+46114958001

Organisation name:	Catalan Institute of Climate Sciences	**Address:**	Doctor Trueta 203
Department:	Climate Forecasting Unit	**Postal code:**	08005
Group:		**City:**	Barcelona
		Country:	SPAIN

Mr Muhammad Asif

Job title:	Research Engineer
Gender:	Male
Title:	Mr
First name:	Muhammad
Last name:	Asif
Initials:	
Date of birth:	01011982
Nationality:	Pakistani
e-mail :	masif@ic3.cat
Phone number:	+34935679976
Fax number:	

Mr Domingo Manubens

Job title:	software engineer
Gender:	Male
Title:	Mr
First name:	Domingo
Last name:	Manubens
Initials:	
Date of birth:	25091982
Nationality:	Spanish
e-mail :	dmanubens@ic3.cat
Phone number:	+34935679976
Fax number:	

Prof Francisco Doblas-Reyes

Job title:	Head of Unit
Gender:	Male

Title:	Prof
First name:	Francisco
Last name:	Doblas-Reyes
Initials:	
Date of birth:	07061968
Nationality:	Spanish
e-mail :	f.doblas-reyes@ic3.cat
Phone number:	+34935679976
Fax number:	

Organisation name:	Linkoping University	**Address:**	Linkoping University
Department:	National Supercomputing Center	**Postal code:**	58183
Group:		**City:**	Linkoping
		Country:	SWEDEN

Dr Chandan Basu

Job title:	Application Expert
Gender:	Male
Title:	Dr
First name:	Chandan
Last name:	Basu
Initials:	
Date of birth:	21101964
Nationality:	Indian
e-mail :	cbasu@nsc.liu.se
Phone number:	+4613282579
Fax number:	

Dr Torgny Faxen

Job title:	Computer scientist
Gender:	Male
Title:	Dr
First name:	Torgny
Last name:	Faxen
Initials:	
Date of birth:	04101950
Nationality:	Swedish
e-mail :	faxen@nsc.liu.se
Phone number:	+46709525798
Fax number:	

Organisation name:	Royal Netherlands Meteorological Institute (KNMI)	**Address:**	P.O.Box 201
Department:	Global Climate	**Postal code:**	3730
Group:		**City:**	De Bilt
		Country:	NETHERLANDS

Prof Wilco Hazeleger

Job title:	Head of Division
Gender:	Male
Title:	Prof

First name:	Wilco
Last name:	Hazeleger
Initials:	
Date of birth:	08101967
Nationality:	Dutch
e-mail :	wilco.hazeleger@knmi.nl
Phone number:	+31302206718
Fax number:	

Dr Richard Bintanja

Job title:	Researcher
Gender:	Male
Title:	Dr
First name:	Richard
Last name:	Bintanja
Initials:	
Date of birth:	07081965
Nationality:	Dutch
e-mail :	richard.bintanja@knmi.nl
Phone number:	+31302206499
Fax number:	

Dr Camiel Severijns

Job title:	Computer Scientist
Gender:	Male
Title:	Dr
First name:	Camiel
Last name:	Severijns
Initials:	
Date of birth:	10111964
Nationality:	Dutch
e-mail :	camiel.severijns@knmi.nl
Phone number:	+31302206215
Fax number:	

Single machine request

- Computing time specified in core-hours for the 12 months period.
- The maximum value for each machine is specified in the **Call announcement**.
- See also the **Summarized Technical Requirements for each machine**

- **Curie Thin Nodes (TN)**: 38 000 000

If less than 5Mio is required, the use of the Tier-0 system has to be justified (Maximum 500 words) :

Not Applicable

If the request of a single machine exceeds 30% of the resources available in the machine for the specific call,
please justify the large request of resources (Maximum 500 words) :

Not Applicable

Summary of the project

If the project is successful this will be published on the PRACE website unless you mark it as confidential below. Please make this summary understandable to a general audience. (500 words)

HiResClim aims to make major advances in the science of climate change modelling . This will be achieved by addressing the dual requirements of; increased climate model resolution and increased number of ensemble realizations of future climate conditions for a range of plausible socio-economic development pathways. Increased model resolution aims to deliver a significant improvement in our ability to simulate key modes of climate and weather variability and thereby provide reliable estimates of future changes in this variability. A large ensemble approach acknowledges the inherent uncertainty in estimating long-term changes in climate, particularly in phenomena that are highly variable and, of which, changes in the occurrence of the rare but intense events are those impacting society and nature most strongly. To provide credible risk assessment statistics on future change in phenomena such as; extra-tropical and tropical cyclones, heatwaves, droughts and flood events, the combination of high climate model resolution and a large ensemble approach is unavoidable. In HiResClim we attack both of these requirements in a balanced approach, which, as well as being the most efficient way to utilise the most advanced HPC systems of today, is also the only path to providing more robust and actionable estimates of future climate change.

Recent bibliographic references that are relevant to the project

Hazeleger, Wilco, and Coauthors, 2010: EC-Earth: A Seamless Earth-System Prediction Approach in Action. Bull. Amer. Meteor. Soc., 91, 1357?1363. doi:http://dx.doi.org/10.1175/2010BAMS2877.1

Donners, J., Basu, C., McKinstry, A., Asif, M., Porter, A., Maisonnave, E., Valcke, S., Fladrich, U., 2012: Performance analysis of EC-EARTH 3.1, PRACE White Paper, to be published

Maisonnave, E., Coquart, L. and Valcke, S., 2011: PRACE preparatory access for high resolution ARPEGE-NEMO porting on Bullx TGCC platform ,Working Note, WN/CMGC/11/77, SUC au CERFACS, URA CERFACS/CNRS No1875, France

Meurdesoif, Y., Denvil, S., Masson, S., 2012: Production et dissémination de données à l'IPSL, 29th ORAP Forum, 2012, Paris

Redler, R., Valcke, S. and Ritzdorf, H., 2010: OASIS4 - A Coupling Software for Next Generation Earth System Modelling, Geoscience Model Development, 3, 87 - 104, DOI:10.5194/gmd-3-87-2010.

Taylor, Karl E., Ronald J. Stouffer, Gerald A. Meehl, 2012: An Overview of CMIP5 and the Experiment Design. Bull. Amer. Meteor. Soc., 93, 485?498.
doi: http://dx.doi.org/10.1175/BAMS-D-11-00094.1

For multi-year access proposals

Please explain why a multiyear access is needed and provide an estimation of the resource requested for the second year (please note that this is to be considered only as an estimation , as the awarding of hours and resources for the second year requires a review and the availability of the system)

(Maximum 500 words)

A detailed explanation of the need for multi-year access is provided in the accompanying mandatory questions: HiResClim plans to make an ensemble of high-resolution global climate simulations of the past and coming century, building on the specifications of the 5th Coupled Model Intercomparison Project (CMIP5). To achieve this, a first step is to run the 2 project GCMs using CMIP5 pre-industrial forcings for sufficiently long simulated time so that the coupled model, particularly the deep ocean,attains a stable equilibrium with the forcing. From the last 100-200 years of this pre-industrial run a set of model states are saved for subsequent use as initial conditions for the, so-called, historical and future projection runs. Hence in the first year of the project, the 2 GCMs will perform the long pre-industrial integrations. Year 2 of the project concentrates on producing an ensemble of historical and future climate simulations. Hence year 2 is where the bulk of the core-hour request lies. For year 2 we estimate a need of 112 million core hours for the core experiments and a further 13.5 million core hours for the optional experiments.

Computer resources requested

- For Multi-year access proposals: **1st year only.**
- For all fields below, you can find information on range limits in the **Technical Guidelines for Applicants**.

** Storage: Maximum amount of data needed at a time.*

Total storage* (scratch) – scratch files during simulation, log files, checkpoints (GB/TB)	20 TB
Total storage (work) – result and large input files (GB/TB)	10 TB
Total storage (home) – source code and scripts (GB/TB)	2 TB
Total storage (archive) (GB/TB)	30 TB

** Maximum number of files to be stored.*

*** Mio: Millions*

Number of files* (scratch) (Mio)**	1.00
Number of files (work) (Mio)	1.00
Number of files (home) (Mio)	0.50
Number of files (archive) (Mio)	1.00

Total amount of data to be transferred to/from the production system (GB/TB)	25	TB

If the amount of data is larger than 20TB or larger than 1TB a day, describe your strategy concerning
the handling of data (pre/post processing, transfer of data to/from the production system, retrieving relevant data for long-term) (Maximum 500 words) :

Global Climate Modelling groups are familiar with the need for large data transfer and archival. This is built on over 20 years of experience in international coordinated projects under the CMIP protocols. In CMIP5 the distributed and federated Earth System Grid has been used for archival and distribution of simulation results. HiResClim simulations will all ultimately be archived in one or more ESG data centres, primarily in France (Meteo France) and the National Supercomputing Center (NSC) in Sweden. As data is generated by the 2 models, each has a direct set of post-processing routines that extract and calculate the range of diagnostics required for subsequent scientific analysis. The procedure significantly thins the amount of basic data for subsequent transfer. Once post-processed data is available in suitable temporal chunks it will transfered off the Curie system, with the CERFACS group being responsible for ARPEGE data and the 4 EC-Earth institutes sharing responsibility for retrievel and removal of EC-Earth data. More details on post-processing and data transfer are provided in the response to the mandatory questions.

Number of jobs that can run simultaneously, i.e. do not depend on each other 10

Wall clock time of a typical simulation 3.00 H

For wall-clock time over 10 months:

- **Are you able to write checkpoint? Yes**
- **Maximum time between 2 checkpoints (hours)** 4

Expected job size (number of cores) and job memory (total memory usage over all cores of jobs)

	Number of cores:	Memory (GB):
Minimum:	2 048	700.00
Average:	4 096	800.00
Maximum:	14 336	1 000.00

I/O intensive ? Yes

Please note that if you plan to transfer more than 20 TB of data, you application is de facto considered as I/O intensive.

**If the answer is Yes, please describe your strategy concerning I/O (for example usage of I/O libraries, MPI I/O, netcdf,
HDF5 or other approaches):**

Both models produce output in netcdf (NEMO) and GRIB (EC-Earth) formats. In the ARPEGE-NEMO model the new "XIOS" (XML-IO-Server) NEMO I/O server is planned to be used (Meurdesoif 2012). XIOS, written in C++, relies on a server-client paradigm with a XML-based I/O description. The servers are cores fully dedicated to I/O tasks while the clients are MPI processes of the computational codes. XIOS uses parallel file systems based on netdcf4/hdf5 and MPI_IO calls in the end, reducing drastically the post-processing step. Experience gained with the XIOS server in ARPEGE-NEMO will be extremely useful for the EC-Earth teams who plan to implement the same I/O server into EC-Earth during the is-ENES2 project.

In both models the number and frequency of output variable will be significantly reduced compared to the standard CMIP5 protocol. We will mainly save only monthly variables for both phases (control, historical and RCP simulations, with a few carefully selected daily fields added only to the second phase (historical and RCP simulations) of the project. Both models will run a post-processing "after-burner" which will continuously calculate averaged diagnostics from the raw output fields, allowing significant reduction in the amount of final data to be transferred from the compute system.

Parallel I/O strategy :

I/O data traffic (write and read) per hour for typical production job (GB) :	85
Number of files generated per hour for typical productions job :	10 000

Describe what work has already been done to develop the codes

This should include the following: describing the main algorithms, how they have been implemented and parallelized, and their main performance bottlenecks and the solutions to the performance issues you have considered. For each code that needs to be optimized, please provide the details listed below.

1. Name and version.
2. Webpage and other references.
3. Licensing model.
4. Contact information of the code developers.
5. Your relationship to the code (developer, collaborator to main developers, end user, etc.).

Both GCMs have already been run and tested on the Curie platform, with performance analysed out to ~2000+ cores per single job. For ARPEGE the CERFACS team in collaboration with Meteo-France have extensive experience running ARPEGE-NEMO on the Curie system. EC-Earth was the subject of an is-ENES-PRACE1IP collaboration to test and improve performance of a T799/ORCA025 version on PRACE systems. All of the HiResClim partners were involved in this effort, hence significant experience already exists in the team on the performance bottlenecks with both models, specific to the Curie system.

It is well established that the main bottleneck for massively parallel performance of coupled GCMs is I/O, due to the need for large amounts of raw output data, from which numerous post-processed diagnostic fields are required. As detailed in the response to mandatory questions file attached to this proposal, both groups have developed a number of automated post-processing procedures, that continually calculate required diagnostics to be saved from the raw model output. This data thinning and reduction procedure occurs continuously in parallel with the production runs. Furthermore, the CERFACS team will test a new

parallel IO server option within their model for further data thinning.

The second primary performance bottleneck for high-resolution, coupled GCMs is the need for frequent exchange of data and regridding through a coupler (OASIS) between component models (atmosphere, ocean, land, sea-ice). The performance of the OASIS coupling system has recently greatly improved for large jobs, through the development of the OASIS-MCT system. Both models have the OASIS-MCT option available and will utilize this in the planned simulations. The developers of the OASIS-MCT at CERFACS are partners in HiResClim (Valcke and Maisonnave)

EC-Earth is developed from the ECMWF IFS system, with strong communication bewteen EC-Earth and the IFS developers. SMHI has developed the latest version of EC-Earth (version 3), to be used in HiResClim, with the main developers (Wyser and Fladrich) partners in the proposal. Furthermore, KNMI developed version 2.3 of EC-Earth, the forerunner to version 3 and the code version used for CMIP5 simulations. The main developers of this code version (Severijns and Bintjana) are also project partners. IC3 partners are developing and testing the Autosubmit system for EC-Earth, a python software that wraps a number of independent model simulations into a single executable and manages all job-control of this bundled executable during integration. The IC3 developers are also part of the HiResClim team.

Finaly, both models use the NEMO community ocean model developed and supported by the Paris-based NEMO system team. Through a range of projects the HiResClim team have extensive knowledge of the NEMO code and work closely with the NEMO team in various aspects of NEMO performance.

Discuss the routes that you will use for dissemination of the project and for any appropriate knowledge transfer. This should include any resources that you will be using to support this. (500 words)

Dissemination of HiResClim will mainly follow 2 paths, (i) Dissemination of the actual results (data) of the project through archival of simulation output on the CMIP5 ESG nodes, thereby making the output of the project openly available to the research community. (ii) The main route for dissemination of science findings arising from the project will be peer reviewed articles. Further to these channels, European climate researchers will be updated of progress through the ENES network and directly within the is-ENES2 project and other relevant EU and national research collaborations. European policy-makers will be targeted for dissemination via suitable publications and conferences and, where feasible advances made in climate modelling will be communicated to the interested public.

Confidentiality

Is any part of the project covered by confidentiality ? **No**

If YES, specify which aspect is confidential and justify (Maximum 500 words) :

Do you have any other support for this application e.g. from your national funding council, the EC or international collaborations? Please give details of this below.

All partners have national and European Union research funding to support the science activities underpinning the proposed work. These include EU FP7 projecst such as : is-ENES, is-ENES2, EMBRACE, SPECS. No direct funds are available for the implied computer costs involved in this

application.

You can nominate up to three reviewers to assess your proposal

Please note that the reviewers you nominate should not be a member of your research group or a member of a group with whom you work on a regular basis. Preferably you should nominate at least one reviewer from outside your own country. PRACE will aim to use one of these nominees to review your proposal but there is no guarantee that any of the nominees you indicate will be used.
Please give the names, affiliations and e-mail addresses of the nominated reviewers.

	Name & email		Affiliation
First nominee	**Name:**	James Hack	Director, National Center for Computational Sciences, Oak Ridge National Laboratory, Tenessee, USA.
	e-mail:	jhack@ornl.gov	
Second nominee	**Name:**	James Murphy	Hadley Centre, UK. Meteorological Office, Exeter, UK.
	e-mail:	james.murphy@metoffice.gov.uk	
Third nominee	**Name:**	Tim Pugh	Ocean Forecasting, The Centre for Australian Weather and Climate Research (CAWCR),Melbourne, Australia.
	e-mail:	t.pugh@bom.gov.au	

☑ **I have read and I comply with the terms and conditions of PRACE 5th Call for proposal available at http://www.prace-ri.eu/Call-Announcements**