REPORT ON THE VISIT TO THE EARTH SIMULATOR FEBRUARY – MARCH 2003

1. Aim of the visit

Stimulated by the visit of the UK Science Minister, Lord Sainsbury, to Japan in April 2001 to discuss scientific collaboration with the Japanese Chief Scientist Professor Imura, a delegation of climate change scientists visited Japan in December 2001 to explore the potential for collaboration. The delegation was lead by Professor Alan Thorpe (NCAS, NERC Centres for Atmospheric Science) and included Dr David Griggs (Hadley Centre), Mr Trevor Guymer (Southampton Oceanography Centre) and Dr Jonathon Köhler from the Tyndall Centre. As a result of this visit a group of researchers from the Hadley Centre and the NCAS group from CGAM (Centre for Global Atmospheric Modelling) were invited by Professor Sumi, to attend a workshop held at CCSR, University of Tokyo in September 2002 and to visit the Earth Simulator to explore using it for collaborative experiments. The first stage of the UK/Japan collaborative project is to build the climate modelling framework that can be used by UK climate scientists to carry out their experiments on the Earth Simulator. The team designated to perform this task were Dr Malcolm Roberts (Ocean Scientist from the Hadley Centre), Dr Lois Steenman-Clark (computational scientist from CGAM) and Dr Jeff Cole (computational scientist from CGAM). In the UK the team were assisted by Dr Richard Woods (Team Leader), Mr Andy Heaps (CGAM system manager) and Dr Katherine Bouton (Data Specialist). The team visited the Earth Simulator from 11th February to 7th March 2003 and their goals were:

- to build the UK climate modelling infrastructure on the Earth Simulator
- to assess the performance of current UK climate models on the Earth Simulator
- to explore the issues of running UK climate experiments on the Earth Simulator.

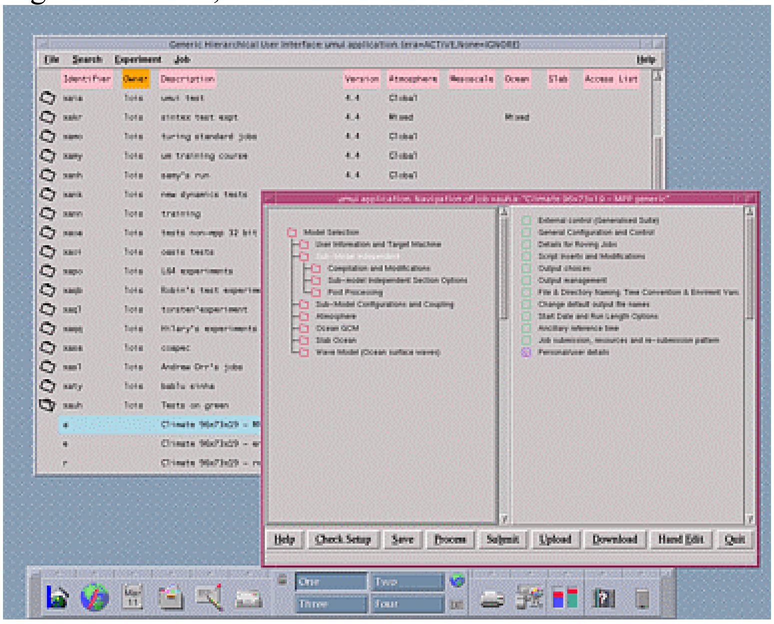
This report describes how these goals were successfully achieved.

2. Introduction to the Unified Model

The Unified Model (UM) is the modelling code that is used by both the UK Met Office for operational weather forecasts and the Hadley Centre for climate research. The UM has been developed and maintained by the UK Met Office over many years and it is also widely used by the UK academic research community. The UM is used for climate research within CGAM.

The computational infrastructure surrounding the actual UM model code is complex but the UM user normally never sees this, as UM experiments are controlled through a user interface. This Unified Model User Interface (UMUI) is a graphical user interface, which sets user controlled model switches producing the UNIX scripts which run the UM experiment. These UMUI scripts call other scripts and use pre-compiled executables to control all stages of the UM experiment. These scripts and executables are built during the initial installation of the UM on the computer and they have to be tailored to the particular computer on which the UM will be running.

Figure 1. UMUI, The Unified Model User Interface



The UK Met Office has ported the UM to other computer systems, in addition to their current operational / research computer which is a Cray T3E. A "Portable Unified Model" has been developed which has an install procedure that can tailor the UM infrastructure to a particular computer.

It was decided that it was important to install the whole of the UM infrastructure on the Earth Simulator as

- it would be compatible with the environment used by the Hadley Centre and the UK academic community so that testing and exchange of ideas is easier
- it would provide a familiar modelling framework to the team of researchers planning to run UM experiments on the Earth Simulator
- the installation of the central model code and computational infrastructure would be easy to maintain
- modelling or scripting changes needed for optimisation for the Earth Simulator of the modelling code would be used by all users

3. The Earth Simulator

The Earth Simulator is currently the largest and fastest computer in the world and it has therefore has a very high level of security to protect it. The computer is isolated from the external network and is only accessible from terminal buildings in the Earth Simulator complex from the researchers own computers, after they have been registered with the Earth Simulator administration. The security arrangements for the team's visit were put in place by Dr Syota Emori (FRSGC) before the team arrived and so no delays in accessing the Earth Simulator were experienced. We are grateful to Dr Emori for his careful preparation.

The Earth Simulator has 640 nodes with 8 processors per node. Each processor is a vector processor with a peak speed of 8 Gflops. The Earth Simulator is accessed via a front-end machine; in this case a 12 processor Hewlett-Packard. The front-end machine is used for housekeeping, managing file space etc and for compiling using a cross compiler, as these operations are best done on scalar processors rather than on vector processors. The Earth Simulator runs the batch jobs submitted from the front-end machine. The Earth Simulator is divided into 40 clusters where cluster 0 is for the S-system and the other clusters form the L-system. There are two classes of batch jobs, those for the S-system for small test jobs for development and testing and those for the L-system for production runs.

4. The UM on the Earth Simulator.

4.1 Initial Set-up

There are a number of initial tests that needed to be made before the installation of the UM infrastructure could be started.

The code and the scripts for the UM are managed using NUPDATE, a code maintenance tool originally from Cray but no longer maintained and now in the public domain. CGAM has ported this code maintenance tool to many other platforms and so the original Cray version could be installed on the HP front-end machine. Because it was known from problems encountered on other NEC computers that the UKMO portable version, a stripped down version of NUPDATE, would have to be used on the Earth Simulator itself.

The UM is written mostly in Fortran with some C routines for portable I/O called from the Fortran. The interface of the C to Fortran needs to be determined before the initial installation for both the front-end and the Earth Simulator itself.

The complex set of scripts used to install and run the UM use a variety of low level UNIX commands in a manner that is not necessarily reproducible on all UNIX platforms. In fact the sed utility, a stream editor that reads text files, making editing changes according to a script of editing commands, had to be replaced on the HP front-end by the GNU sed, which is known to have the behaviour expected by the UM scripts.

The UMUI is written in Tcl/Tk. It was hoped that one central UMUI could be installed on the Earth Simulator front-end but problems were encountered which had not been expected as the UMUI runs on HP machines elsewhere. The Earth Simulator HP front-end runs a version of the operating system to which we don't have access to in the UK so a solution to this problem has not yet been found. The fall back position for the UMUI was to install the UMUI on all the laptops connected to the Earth Simulator network. Although not an ideal situation experiments, which are stored in the UMUI database, could still be shared by creating a centrally accessible directory of UMUI experiment files on the HP front-end. The only issue to having the UMUI on individual laptops is that the management of the UMUI itself has to be undertaken by the individual researcher.

4.2 Compilation of the UM code

A UM experiment has three stages: compilation, reconfiguration to set up the base initial data file for the particular experiment and running the model. The first two stages are serial operations and are therefore best suited to running on the Earth Simulator front-end. The third stage is parallel and therefore should run on the Earth Simulator itself. Stage two, the reconfiguration, uses some of the model code so the UM had to be built on the front-end using the HP compiler for the reconfiguration and on the Earth Simulator using the NEC cross compiler on the HP. Suitable compiler options needed to be selected for both compilers. The most important options were for the NEC Earth Simulator compiler where the model performance was critical. Because of previous experience on the NEC SX5 in Paris it was decided that a conservative 'middle way' option should be used. Of the three levels of optimisation

vsafe -> vopt -> hopt

the middle, vopt, option would be the best option for the initial tests. Even using vopt one subroutine, st_diag1.f, failed to compile and so a compile overrides file, which are used in the UM scripts to alter the makefiles for the UM compilation, was written to reduce the level of optimisation for this routine. Run time errors were encountered that were only resolved by rebuilding the whole of the UM using the compiler option

-Wf "-pvctl noassume loopcnt=10000"

For the initial tests we compiled the UM code with the ftrace option so that we could obtain basic performance information for the UM. There is a certain overhead to using ftrace so when production experiments are run it would be better not to include this compiler option. So two versions of the UM were built, one compiled with ftrace and one without.

The UM has a number of subroutines defined in the EXTERNAL lists in the UM code but not actually called. On some computers this is not a problem but on others it is necessary to supply dummy routines. We selected to use the compiler option that only give warnings if there were unresolved external routines but if one of missing routines is actually called it then the model crashes with a segmentation fault, which is not very helpful or informative. So in future it may be sensible not to use this compiler option and to provide a dummy routine generator, which would print a message and abort properly if one of the missing routines were actually called.

The compilation of the pre-compiled UM code sections took about 1.5 hours to compile using the NEC cross-compiler on the HP front-end. The compilation of the UM experiment code typically took of the order of 0.5 hours on the HP front-end.

4.3 Job submission

The UM only has two options for job submission via the UMUI, using the Cray version of qsub and 'at' for workstations. The Earth Simulator front end, which is used for compilation and reconfiguration, seems only to allow interactive jobs and the version of qsub used for batch jobs is the generic not the Cray version. A job scheduler controls the Earth Simulator batch jobs. The job scheduler controller commands need to be included in the scripts before the qsub commands. The syntax of the Earth Simulator job scheduler commands is relatively simple for jobs for the S-system but they are more complex for the L-system, as file pathnames need to be specified. Only the S-system was used during the teams visit. Rather than change the UMUI to include these Earth Simulator options, it was easier to use post UMUI scripts to edit the submission scripts, which were more flexible and easier to adapt quickly.

The normal method of running the UM is to run the reconfiguration on one processor before running of the experiment on whatever number of processors specified. The reconfiguration is an essentially serial job and so it would very inefficient on a vector Earth Simulator processor. So the UM scripts were adapted so that both compilation and reconfiguration are run interactively on the front-end and only the parallel running of the model would be submitted to the Earth Simulator batch job system using the generic qsub commands.

4.4 Performance

The version of the UM (4.5) was written at a time when the ratio of processor CPU speed to inter processor communications bandwidth was very different to today's machines. This version of the UM, ported to the Earth Simulator during this visit, has an old method of communications with a very large number of short messages so it was not anticipated that the scaling would be good, especially at lower resolutions. The new versions of the UM (>5.3), which will be ported to the UK Met Office NEC SX-6 later this year and which will be used for the FAR IPCC report will have a very different communication pattern, better matched to today's computers.

The UM was only run on the Earth Simulator S-system, where jobs share the processors and the memory, so timings were variable according to the job load. A range of timings is therefore shown in the table below.

The aim of the testing of the UM modelling infrastructure was to ensure that different model resolutions could be run and to assess the performance of the model at different resolutions. Therefore four experiments were run:

- HadAM3 (SRES), the current standard climate resolution atmosphere only model
- HadAM3 (MRES), the medium climate resolution atmosphere only model
- HadCM3, the coupled ocean-atmosphere model at the standard climate resolution
- HadCEM, the coupled ocean-atmosphere model with an eddy permitting ocean model

In the time available only limited optimisation was performed. As a result of work done on the UM atmosphere code for the Paris NEC SX-5, a number of optimisations were easily included for the Earth Simulator. These optimisations included:

- the use of NEC compiler directives
- re-ordering loops in key routines to improve vectorisation
- pre-calculating some array indices to improve vectorisation in key subroutines
- changing the calling mechanism of the long wave radiation to make it more suitable for vector machines.

The UM ocean code has not really been optimised for machines other than the Cray T3E and so more work needs to be done to improve the vectorisation of the ocean code on the Earth Simulator so results for the coupled ocean-atmosphere cases are preliminary.

HadCEM, the UM experiment with the high resolution ocean code, has only ever been run on the UK Met Office Cray T3E so there were a number of porting issues which needed to be addressed before the experiment would run on the Earth Simulator. These porting issues were difficult to resolve and so no effort was made to optimise HadCEM for the Earth Simulator. The HadCEM timings shown in the table below should therefore be considered as preliminary results.

Model	Resolution	Elapse time (seconds) /model day	Processors
HadAM3 (SRES)	3.75° x 2.5° L19	24.9 – 57.3	1 x 2
HadAM3 (MRES)	1.875° x 1.25° L30	128.8 – 132.2	1 x 2
		101.8	1 x 4
		68.5 – 92.0	2 x 2
		54.7 – 90.3	2 x 3
		63.7	2 x 4
HadCM3	3.75° x 2.5° L19 1.25°x 1.25° L20	49.7	1 x 2
HadCEM	3.75° x 2.5° L19 0.33°x 0.33° L40	3,321	1 x 1
			1 x 2
			2 x 2
			2 x 3
			2 x 4

5 Earth Simulator experiment data

The work by CGAM and the Hadley Centre on the Earth Simulator has generated considerable interest within the UK climate modelling community. We want to build on this interest to strengthen collaboration with Professor Sumi's group at the Centre for Climate Science Research, University of Tokyo, as well as the Frontier Research System for Global Change and the National Institute for Environmental Studies groups. We would also like to use this opportunity to strengthen collaboration within diverse UK climate modelling groups by sharing information and data from Earth Simulator experiments, experiments that could not be performed in the UK.

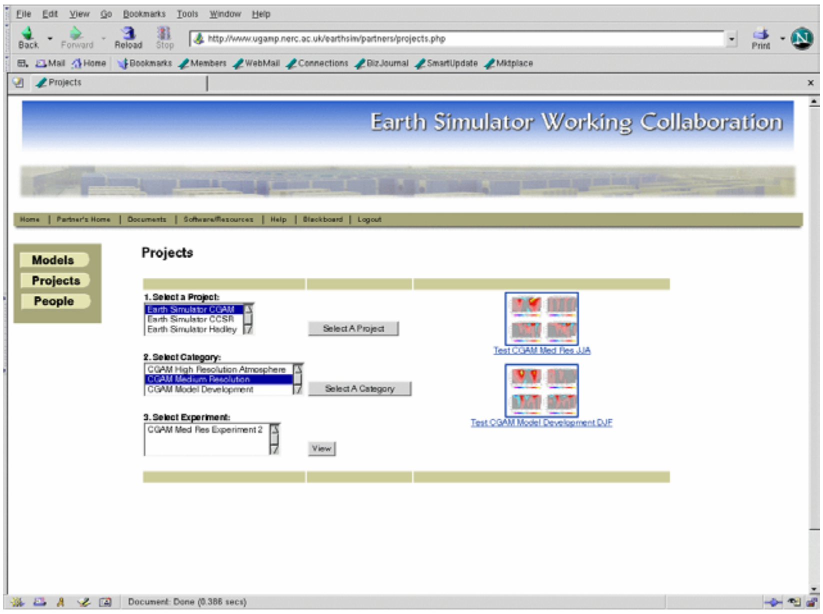
We have considered different ways of disseminating data and results from Earth Simulator experiments. As a first step we have designed and implemented a web site with a 'visualisation gallery', a catalogued repository of the visualisations from experiments. The experiments could therefore be analysed and visualised at the Earth Simulator centre and only the visualisations would be uploaded to this web site and made accessible to the climate community. A second step is to consider access to the Earth Simulator experiment raw data and how to store, catalogue and make it available to a wide community of climate scientists. We describe here the on-going work to make Earth simulator experiment data and results available to climate researchers.

5.1 The Earth Simulator Working Collaboration WWW site and database.

The primary requirement for the Earth Simulator Working Collaboration WWW site and database was to be able to present information about the projects and people involved in the Earth Simulator Collaboration through a web browser in an intuitive and easy-to-use manner.

Information about the projects would allow a user to select a project/experiment of interest and view and/or download visualizations of the relevant Earth Simulator model results from that experiment. Information about the people involved would include contact information, listing of publications and relevant URLS such as home pages or model results pages and links to the results. This information needed to be organized and presented dynamically as new information became available and as each user selected different options. To do this we developed a system based on WWW access to a relational database.

Figure 2. The prototype Visualisation Gallery



We examined two open source (free) databases - MySQL and Postgres. Available in-house expertise in MySQL, and the lack of need for transaction analysis which is Postgres' main strength, lead us to choose MySQL (v3.23.55), which is recognized for its speed and reliability. Database administration was provided through phpMyAdmin (v2.4.0) - an online database administration system. The dynamic web pages, the online search and display of database information, security and user authentication was done through PHP v4.3.1. While there was some difficulty in installing the required software on the WWW server (due to non-standard placement of files) available documentation was good and the installation was relatively painless. The many platforms that these software products are available on, makes it a very portable system.

With redundancy and reliability in mind, development and production servers were kept on two separate servers. Development is done on a Microsoft Windows XP laptop using Macromedia DreamweaverMX, FireworksMX, and NetStudio and transferred via FTP to the production server for public viewing. The production database resides on a Unix workstation (Sun Ultra 10) and is served on the web via Apache (v1.3.27). The databases and WWW pages are backed up on a regular basis and copies are kept both on and off facilities.

No standards currently exist for the cataloguing and display of model visualizations. Metadata formats are currently being developed at several institutions. As a result our database structure was based on an examination of, and discussions with, the Visualization and Enabling Technologies Section (VETS) at UCAR (http://www.vets.ucar.edu/vg/), the NERC Data Grid (http://ndg.nerc.ac.uk), and British Atmospheric Data Centre (http://badc.nerc.ac.uk). The development of the Earth Simulator Working Collaboration database is intended to fit in with further developments of international visualization data formats.

5.2 Raw Data from Earth Simulator Experiments.

Setting up the website with access to visualisation galleries is a first step for collaborators to share results from Earth Simulator experiments. There will be however a requirement to retrieve raw data from Earth Simulator experiments to enable the community to perform their own analysis of the data rather than just relying on the analysis used for the visualisation gallery. No long experiments were performed during this visit to the Earth Simulator, as the primary purpose of the visit was to build the modelling infrastructure. However sufficient information has now been acquired to set up preliminary mechanism for handling with raw data from the Earth Simulator experiments. There are several points to be taken into account:

- the Earth Simulator has many TBytes of disk space and an initial allocation has been given to our group. The practical limits of a group's disk space that can be requested are not known.
- the Earth Simulator is isolated from the external network so all data has to be retrieved via the researchers laptop attached to the Earth Simulator network.
- the archive system is only accessible from the L-system. There are stringent performance conditions for both vectorisation and parallelisation that need to be passed before access to the L-system can be granted. The team did not try to use the L-system during their visit.
- for this visit we did not take an external data storage device (DVD or disk pack) to test rates of transfer from the earth Simulator to the storage device. Tests will be carried out at CGAM before long Earth Simulator experiments are carried out.

From the limited testing we were able to do on the Earth Simulator we could see that the UM slowed down considerably as the size of the model output data was increased. We need to do some further tests on the UM to assess the optimum output data size for the UM at various resolutions.

To devise practical methods of retrieving UM model data from the Earth Simulator we need further work to assess

- the UM performance with the resolution and output requested
- the rate of moving data from the Earth Simulator
- the rate of saving data to the chosen external data storage device
- the data accumulation that is acceptable on the Earth Simulator given that raw data retrieval will have to be a manual operation only possible during working hours

Work needs to continue on designing mechanisms for the retrieval, storage and dissemination of Earth Simulator data and results during the next phase of experiment design.

6. **Future Experiments**.

Following the team's visit to the Earth Simulator Centre there was a very successful second UK/Japan climate workshop held on the 27th - 28th February at the Hadley Centre. Discussions are now on-going for the second phase of this project, the design, testing and then running of the UK climate experiments on the Earth Simulator.