Climate*prediction*.net: A System Designed to Predict the Global Climate in your home

Yang Song

Abstract: Climate*prediction*.net (CPDN) is a distributed computing project based on Berkeley Open Infrastructure for Network Computing (BOINC). Climate*prediction*.net aims to establish a grid upon the PCs of thousands of volunteers over the Internet for research purpose, just like the well-known project SETI@home. However, as it should be, Climate*prediction*.net also has its markedly typical characteristics. This paper is to provide a brief introduction to the system, explain its operating mechanism and bring the author's idea up upon its further developing tendency. It will also make a tentative discussion about the value of this form of public research and it is expected that the approach will be significant to raise similar projects.

Keywords: BOINC, Public resource, distributed computing, simulation

Introduction

The idea of establishing a large global Climate ensembles was firstly raised by Mayes Allen in 1997. However, this grandiose idea stayed unrealistic until the introducing of the first BOINC based program SETI@home in 1999. In September 12th, 2003, Climate*prediction*.net was ultimately launched by Oxford University, and was afterwards supported by the UK e-Science Programme. That enables it to become by far the largest full-resolution climate modelling experiment over the world.

Climate*prediction*.net aims to investigate and reduce uncertainties in the complex climate modelling. Based on BOINC system, CPDN established a grid over thousands of PCs. Upon the client on their PC, the volunteer participants run a complex climate model, each one receiving a different set of climate data for parameters in which there is uncertainty, and different initial conditions, which are also subject to uncertainty due to the chaotic nature of the weather. In that way, to a great extent, the system can utilize the spare CPU resources of thousands of personal computers of individuals. "Many a little make a mickle", this construction allows the system to analyze a great amount of data and run a massive ensemble of simulation using a foremost climate model, the new versions of the Hadley Centre Climate Model (HadCM3, HadSM3), the global set-up of the UK Met Office Unified Model (UM).

Climate*prediction*.net has many similarities with other public-resource, especially BOINC based programs. But it is still distinctive in a number of ways. The most important aspect is that is the primary work of CPDN is not one of search, but of simulation. That leads to a series problems including the complexity, time duration, demands for participant, and data volume. The following sections will give a description of BOINC platform, elaborate how Climate*prediction*.net is built up and how it works upon a model. What's more I'll talk about how these peculiarities lead to the success of Climate*prediction*.net, the challenges it faces and a forecast of its future development.

Climate*prediction*.net based on BOINC

The Berkeley Open Infrastructure for Network Computing (BOINC) is one of the most mainstream open source distributed computing system for volunteer on network. It was originally developed to support the SETI@home project, and soon applied by other programs spreading from mathematics, physics, and molecular biology to meteorology, catastrophology and so on.

Over the last 20 years, personal computer became more and more popular, and computation resources is no longer concentrated in the supercomputer centers and machine rooms. Under these circumstances, collect the computation power which distributed in thousands of individuals is not only an economical and efficient way to conduct massive project, but also a significant action to draw common people's attention, interest and enthusiasm towards scientific research/program and gives the public a measure of control over the directions of scientific progress.

BOINC was designed for SETI@home in 1999. SETI@home has attracted millions of participants worldwide, now runs on about 1 million computers, providing a sustained processing rate of over 70 TeraFLOPs (in contrast, the largest conventional supercomputer, the NEC Earth Simulator, provides about 35 TeraFLOPs)[1].

A typical BOINC based project corresponds to a research group or institute that does public-resource computing. One project is identified by a single master URL, which is also a webpage on web server in central server. Besides that, the core, or a central server, also including a database recording participants, the runs allocated and the data collected and client interaction server be in charge of participant registration and upload control. There are also a number of upload servers, the system uses it to upload the computed data generated by the participants' client on PC. The load on these servers may be high, and the communication overhead great. Thus, Climate*prediction*.net has set a series of duplicated servers around the world, to distribute the overall loads.

To get to work with Climate*prediction*.net and join in the program, Individuals or institutes can visit the home page at *www.climateprediction.net/* to download a software package and install on their PCs. The software includes a state-of-the-art model of the atmosphere / cryosphere / land surface / ocean system – commonly referred to as an Atmosphere Ocean Global Circulation Model (AOGCM).[2] After installing, the software would connect to the server and get a customized simulation data package. After that, the software would start to simulate the model a background progress which preserve a operation speed that won't hurt user's regular work. The model will cover a climate simulation period of 45-50 years. The process can last several weeks and after completed it again connects to the server and upload the simulation result and download another task again.

On the one hand, Climate*prediction*.net is going to provide data for the decision-makers, society institutions, and even schools/universities as teaching resources. Thus it gains some approving from government, social and academic forces. On the other hand, at present, climate change is no longer just an academic issue but also a social issue. Just like people's high enthusiasm in finding aliens at home, public are also very concern about their fortune in next decades. The project exactly provide us a opportunity to "predict" our future and join in the work of important social affair. In that way Climate*prediction*.net has been attracting more and more people to join in.

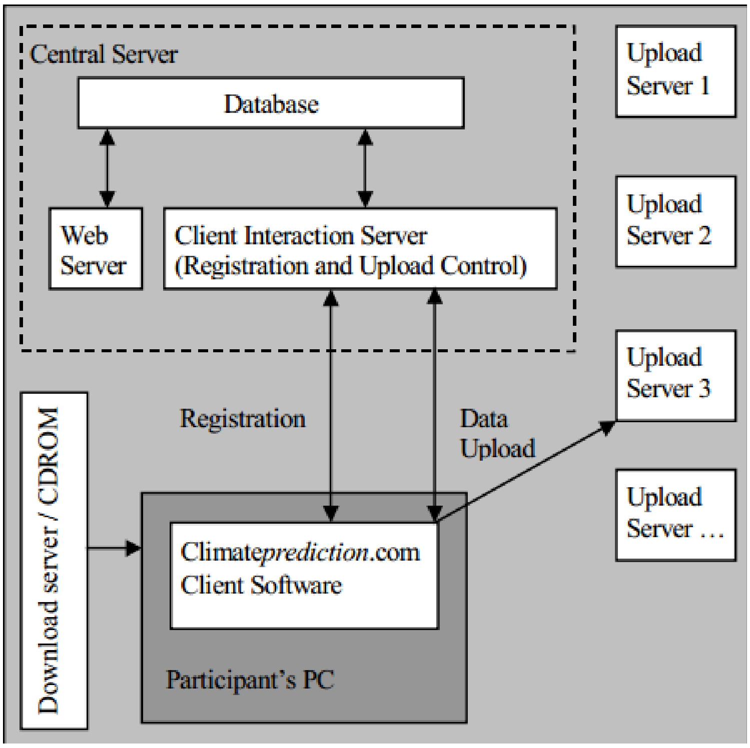

Figure1. Overall system design of Climate*prediction*.net[3]

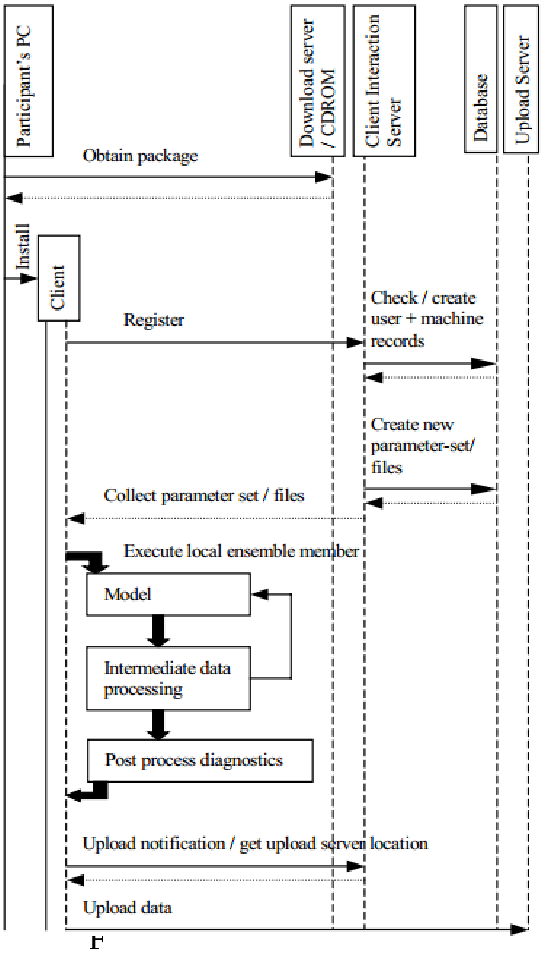

Figure2. Sequence of event[4]

Climate model

Global climate models is now widely used in climate analysis and forecast. It denotes a worldwide hierarchy system that reflect the temperature change, ocean flow as well as atmosphere fluxion; in that way summed up a regularity mathematical model that helps we estimate the future climate change. It contains imprecisely defined parameters that account, approximately, for sub-grid-scale physical processes. This uncertainty makes the model very unstable and needs a great mount of computation and testing. And the parameters would also varies from year to year. The simulation should be conducted continuously and substantially. Thus, public-resource computing is highly demanded and also a very efficient way to approach that.

The climate model which is used by Climate*prediction*.net is Hadley Centre Global Environmental Model. It is a coupled climate model developed at the Met Office's Hadley Centre.

HadCM3 (Pope et al. 2000; Gordon et al. 2000) became operational late in 1997. Projecting a normal rate of increase in computer power up to 2003/04 (the target for operational delivery of HadGEM1) implied that a model having approximately 15–20 times the computational cost of HadCM3 would be affordable by that time.[5] It is a grid point model and has a horizontal resolution of 3.75 $\times$ 2.5 degrees in longitude $\times$ latitude. This gives 96 $\times$ 73 grid points on the scalar (pressure, temperature and moisture) grid; the vector (wind velocity) grid is offset by 1/2 a grid box (see Arakawa B-grid).[6]

For the model, computer is going to solve a dynamic equation for each point at one step. By solving the equation, a series of parameters is calculated which refers to the effects of processes on a corresponding tiny scale. The effects includes atmosphere status, cloud motion, ocean current, temperature function as well as presentative wave equation. These parameters in a overall scale would be analyzed and refer the climate change of the study area.

Challenges

At present, under years of Climate*prediction*.net has built up a large data that predicts significant changes to the Earth's climate in the coming century. But the predictions range greatly, signifying high levels of uncertainty. Governments make relevant policy according to the prediction data to a great extent, while there is an increased probability that the models may over-estimate or under-estimate the speed and scale of climate change. That means, we may invest too much to avoid a trouble which is predicted serious but proved to be not that grave. We may also ignore or do so little to a trouble that is predicted negligible but proved to be very serious. Also, if the time and location is falsely predicted, what we do will also become somewhat a vain. Thus, to cope with this problem, we need to quantify and control the uncertainty in these predictions.

There are two complementary approaches to assess and reduce uncertainty:

1. Improve the parameterisations while narrowing the range of uncertainty in the parameters. This continuous process requires:

i) Improving the models by using the latest supercomputers as they become available.

ii) Gathering more and more (mainly satellite) data on a wide range of atmospheric variables (such as wind speed, cloud cover, temperature…).

2. Carry out large numbers of model runs in which the parameters are varied within their current range of uncertainty. Reject those which fail to model past climate successfully and use the remainder to study future climate.[7]

Climate*prediction*.net now takes the second approach. By running the model thousands of times on thousands of computers, it is expected to find out how sensitive individual models are to small disturbances and parameter uncertainty. But their still existed two major problem as I am concerned:

Climate*prediction*.net is now also not a wide-spread program to common people. Comparing to the popular SETI@home program, climate forecast is absolutely not as attractive and interesting as searching extraterrestrial life. An extensive outreach and reward would encourage public's enthusiasm to join in more and more. As the prediction is associated with decision-makers' policy, national governments also have the obligation to promote this program.

The second problem come with the autologous imperfection of distributed system. Distributed programs should operate on different OS and platform. Professions are still holding different ideas upon how much the software should occupy the computational power, if the clients can be asked to do more. These problems are expected to be less as our deep put into practice but it still not neglectable. Besides, the system is highly depend on the network and data from clients. As individual is not fully guaranteed to have a stable contribution and the network is also not guaranteed to be always unimpeded and reliable. This demands the group to have its own supercomputer or computing center, which can ensure a steady contribution to data to some degree.

Conclusion and Future development

As of December 2010, there are more than 32,000 active participants from 147 countries with a total BOINC credit of more than 14 billion, reporting about 90 teraflops (90 trillion operations per second) of processing power.[8]

Of course, this program is still in progress and there will be continuing need of data regenerate. Both proposals and critical will also be continuously given by the participants as the program growing. Anyhow, support and views from all quarters are indispensable for the improvement of Climate*prediction*.net.

As the models getting more and more complicated and the data became larger and larger, this is not only a need of more participants and faster machine, but also new architecture of distributed system, software optimization and settlement of additional complication. Whether these problems are properly resolving will relate to the proceeding and development Climate*prediction*.net.

In particular, the program come up with the idea that for a scientific project, scientists can concentrates more on design-related work: Modeling, analyzing and statistics, and handing the work of day-by-day calculation, simulation by public involvement. It will promote efficiency on scientific work, as well as maintain public interest, raising public awareness of the scientific issues. And it is hoped that my discussion will be relevant to the establishment and popularization of more and more similar scientific project.

---

[1] Anderson D P. BOINC: A system for public-resource computing and storage[C]//Grid Computing, 2004. Proceedings. Fifth IEEE/ACM International Workshop on. IEEE, 2004: 4-10.

[2] Stainforth D A, Allen M R, Frame D, et al. Climateprediction. net: a global community for research in climate physics[M]//Environmental Online Communication. Springer London, 2004: 101-112.

[3] Stainforth D, Kettleborough J, Martin A, et al. Climate*prediction*.net: Design Principles for Public-Resource Modeling Research[C]//IASTED PDCS. 2002: 32-38.

[4] Stainforth D, Kettleborough J, Martin A, et al. Climate*prediction*.net: Design Principles for Public-Resource Modeling Research[C]//IASTED PDCS. 2002: 32-38.

[5] Johns T C, Durman C F, Banks H T, et al. The new Hadley Centre climate model (HadGEM1): Evaluation of coupled simulations[J]. Journal of Climate, 2006, 19(7): 1327-1353.

[6] http://cera-www.dkrz.de/IPCC_DDC/IS92a/HadleyCM3/Readme.hadcm3.grid

[7] http://www.climateprediction.net/about/

[8] "Detailed user, host, team and country statistics with graphs for BOINC". boincstats.com. Retrieved 2010-12-13.