PARALLEL PROCESSING AND METEOROLOGICAL MODELS

A DICKINSON

Rapporteur: S Caughey

PARALLEL PROCESSING AND METEOROLOGICAL MODELS

A. Dickinson
Meteorological Office
Bracknell, UK

Introduction

Meteorologists are dependent on the most powerful computers to simulate the motion of the Earth's atmosphere. Over the past 30 years two important application areas have evolved, both relying on similar computational fluid dynamics techniques. The first of these is numerical weather prediction (NWP), the real-time production of weather forecasts for a few days ahead. The second is the modelling of the general circulation of the atmosphere and oceans (AOGCM) with the aim of studying climate change. In the case of NWP, the time critical nature of the problem and the large amounts of data processing involved have established the case for using a supercomputer. Indeed, the accuracy of these computer models has advanced in step with the amount of computer power available. In AOGCM the objective is to investigate future climate change senarios. A large investment in computer time is necessary if results of acceptable quality are to be obtained in a reasonable amount of time.

Since the available technology limits the speed of any single computer, parallel computations have become necessary to achieve major increases in processing power. The current computer configuration at the Met Office comprises two 8 processor Cray Y-MPs. A unified climate-forecast model has recently been developed covering a range of ocean and atmosphere applications. It is written in a modular form using standard Fortran and makes full use of the automatic multitasking facilities available on Cray machines. Sustained performance levels of around one gigaflop are achieved for the global NWP model [1]. Work is now underway to investigate the implementation of the Unified Model on future massively parallel computers.

Unified Model

The Unified Model [2] consists of a numerical description of the Earth's atmosphere and ocean. The model can be run in three modes, atmosphere only, ocean only or as a coupled atmosphere-ocean model. In each mode a run consists of an optional period of data assimilation followed by a forecast. The model may be global or limited in either horizontal or vertical extent. In the latter cases boundary values of the prognostic variables must be given for the period of the integration. Table 1 gives examples of the model resolutions being used at the present time.

	Climate	**Forecast**
Number of levels	20	20
Number of points E-W	96	288
Number of points N-S	73	217
Time step (seconds)	1200	600
Grid box (degrees)	2.5x3.75	0.833x1.25

Table 1: Global atmosphere model dimensions used on the Cray Y-MP system for operational forecasting and climate studies.

To a first approximation, forecast models and the atmosphere part of climate models may be viewed as just different versions of the same program. In forecast mode the highest resolution that can be reasonably run in the available time is used. In climate studies there is a trade off between the accuracy in the representation of features and the length of time for which the model can be run, so that coarser resolutions are typically used. In the case of NWP the initial conditions determine the outcome, parameters such as the chemical composition of the atmosphere being held constant. In climate studies the objective is to determine what the average weather and its variability will be if those parameters are changed but the initial conditions no longer have influence. In climate configuration, more detailed physical parametrizations are used, the impact of which are unimportant on the short timescales used in NWP.

Both the ocean and the atmosphere component of the Unified Model consist of two distinct sets of processes referred to as dynamics and physics. In the current generation of models these contribute about equally to the cost of an integration. In the dynamics the governing equations, consistent with the motions of a thin layer of compressible fluid on the surface of a rotating sphere, are solved by explicit finite difference techniques. In simple terms, the integration of these equations requires the forecast domain to be covered by a regular grid of points at which the current values of the prognostic variables are stored. A set of linear equations, derived from finite difference approximations to the governing equations, are then used to step the integration forwards one time step at a time by modifying the values stored at each grid point and at each level of the forecast domain. The important feature of the dynamics is that the same calculations are generally applied at every point and this makes their solution fairly straightforward to parallelize.

The physics take place on spatial scales which are considerably smaller than the grid length of a meteorological model. They include the effects of clouds, solar and terrestrial radiation, precipitation, convection and turbulence. These sub-gridscale processes cannot be dealt with explicitly. Instead their statistical effect is represented in terms of the grid point variables. The schemes for modelling these processes are characterised by sets of calculations which generally apply only at non-contiguous subsets of the integration domain. For example, when calculating the transfer of moisture from the earth's surface into the atmosphere it is necessary to distinguish between land, sea and surface-ice grid points. More usually these subsets vary in position and size from timestep to timestep as in the case of developing areas of rainfall. The intermittent and conditional nature of these codes makes the physics far more difficult to parallelize because of the difficulties in achieving a balanced workload.

The data assimilation scheme adjusts the state of the model's ocean or atmosphere towards observations, providing initial fields for forecasts and analyses from which global climatologies can be generated. Observations of the state of the global atmosphere are provided through international agreement at set times of the day and come mainly from satellites, radiosondes and aeroplanes. Even so, much of the globe has only sparse coverage, particularly over the oceans and the Southern Hemisphere. In the scheme, model prognostic variables are updated by information from surrounding observations which are weighted according to their quality and distance from the current grid point. In data sparse regions the radius of influence of an observation can be as much as 1000Km. Data assimilation contributes something like 30% to the cost of producing a 6 day global forecast. Difficulties in parallelising the code arise from the inhomogeneous coverage and collocation of observations. A shared memory architecture is an advantage when coding this problem, otherwise the uneven geographical distribution of observations leads to increments being generated on one processor which are required to update grid points held in the memory of other processors.

Future Requirements

An indication of future computing requirements for NWP can be seen by considering the impact on computing power of improving model resolution by a factor of two.

Previous experience has shown that increases of this order will lead to a noticeable increase in the accuracy of weather forecasts. The factor of two involves doubling the number of points in the latitude and longitude directions, halving the timestep and increasing the number of vertical levels. The computing demand goes up by a factor of more than ten. Three such steps would thus require a machine delivering a sustained teraflop. This would reduce the horizontal gridlength of the global forecast model from 100Km down to approximately 10Km.

In practice, major revisions in formulation would be required in order to model the newly resolved spatial and temporal scales. A fully implicit treatment of the governing equations might be necessary as well as enhancements to the treatment of physical processes. In addition, the number and type of observations from satellites is expected to increase substantially over the next few years leading to increased costs in data assimilation. It therefore appears that teraflop performance will be required well before the 10Km resolution is reached.

The time constraints on climate modelling are not as severe as those on weather forecasting. However, with the recent political interest in the impact of man's activities on climate because of the potential adverse economic and social effects, the time constraints are becoming more severe. Work on studies of the impact of global warming has to be completed against staged deadlines and looks set to need sustained result rates of the order of one teraflop in the second half of this decade. This sort of target has been set by other climate modelling centres, most notably in the USA at the Department of Energy [3].

Computing trends at the Met Office over the past 40 years may be seen from Figure 1. The computers shown have been used by Met Office staff either for development, operational forecasting or research. There appears to be an increase in power of 34 times per decade. It is tempting to predict by extrapolation the approximate year of commercially usable, sustained teraflop performance. This appears to be early in the next decade.

Programming Considerations

The first proposal for a parallel numerical weather prediction system was made by L.F. Richardson in his 1922 book *Weather Prediction by Numerical Process* [4]. Writing well before the advent of the first electronic computer, Richardson realised the practical limits of producing such forecasts by hand. Nevertheless, in a flight of fancy, he imagined a forecast factory for global weather prediction made up of 64,000 people. In this factory, arranged somewhat like a large ampitheatre, separate teams calculated the state of the atmosphere by hand at each grid point using data provided by neighbours. A conductor at the centre of the theatre was responsible for synchronising the calculations at each timestep. This is a remarkable analogue of how one might implement a weather model on today's massively parallel computers. Indeed the degree of parallelisation envisaged by Richardson is still much greater than can be exploited in practice, since he proposed separate nodes (individuals) to calculate each term in each equation at each grid point.

In more recent times Tett [5] has examined the suitability of various integration schemes used in meteorological modelling for implementation on SIMD and MIMD architectures. He concluded that both types of machines are suitable. SIMD machines have an advantage because of their relative ease of programming, while MIMD machines gain because they are better suited to handling the inhomogeneities that occur in atmospheric modelling.

Ideally the Met Office would prefer to run the Unified Model with minimum modifications on any new system that might be acquired in the mid-90's. The Edinburgh Parallel Computing Centre has been asked to investigate a parallel implementation of the Unified Model. The preliminary conclusions [6] are that the overall structure of the Unified Model would transfer easily to a SIMD architecture but that the broadcast of instructions to processing nodes would limit performance. Efficiency could be low when dealing with the parametrization of sub-grid scale processes. Problems would arise on MIMD machines because of the need to store very large code segments in each node. These problems are such that a major restructuring of the Unified Model would probably be required.

In NWP and AOCGM the same programs are run many times over a period of years and so it is worth investing a great deal of effort into optimization to get the best out of a given machine. However, there is a limit to the effort that can be applied to tuning algorithms and programs, and scarce expertise is better devoted to scientific problems. On computers prior to the Y-MP, Met Office production code was always written in a low-level language. Although this led to fast efficient code, maintenance and modification were difficult. The Unified Model is written in Fortran and automatic features of the Cray compiling system are used to optimise the code. All of the special code used for multitasking and optimization is therefore hidden from the scientists using and developing the model.

Conclusions

Both NWP and AOGCM are suitable candidates for implementation on massively parallel computers, although the choice between SIMD and MIMD architectures for meteorological modelling is still unclear.

Although overall speed will continue to be a primary consideration when purchasing a new supercomputer, performance will need to be balanced against the difficulties in maintaining and porting large sections of non-standard code. This is being determined increasingly by consideration of costs and benefits.

Substantial efforts will be required by manufacturers of parallel computers and compiler writers to avoid the need for application programmers to have a detailed awareness of the hardware architecture and systems software. Parallel machines offering a virtual shared memory may have advantages in simplifying the coding of those parts of meteorological models which are non-local in character.

Acknowledgements

The author would like to thank Dr R. L. Wiley for his help in producing this paper. In particular, his insights into future trends in the supercomputer industry and their consequences for the Met Office have been freely quoted [7].

References

[1]Dickinson, A., 1990."Autotasking Forecast and Climate Models on a Cray Y- MP", Cray Channels, Vol 12, No 3.

[2]Bell, R.S. and Dickinson, A., 1991."The Meteorological Office Unified Model for Data Assimilation, Climate Modelling and NWP and its Implementation on a Cray Y-MP", to be published in the series "Topics in Atmospheric and Oceanic Sciences", Springer-Verlag.

[3]US Department of Energy, Office of Energy Research, Office of Health and Environmental Research, Atmospheric and Climate Research Division, December 1990."Building an Advanced Climate Model - Program Plan for the CHAMMP Climate Modelling Program".

[4]Richardson, L.F., 1922."Weather Prediction by Numerical Process", Cambridge University Press, Cambridge.

[5]Tett, S.F.B, 1991."Massively Parallel Computers and their Applications to Numerical Weather Prediction", Computer Modelling in the Environmental Sciences, Eds D.G. Farmer and M.J. Rycroft, Clarendon Press, Oxford.

[6]Hawick, K.A., 1991. "Unified Weather/Climate Model Parallel Implementation Study", unpublished preliminary report, 25 March 1991, details available from Meteorological Office.

[7]Wiley, R.L., 1991."Parallel Processing and Numerical Weather Prediction", Proc. 2nd International Seminar on "The Design and Application of Parallel Digital Processors", IEE, London, April 1991.

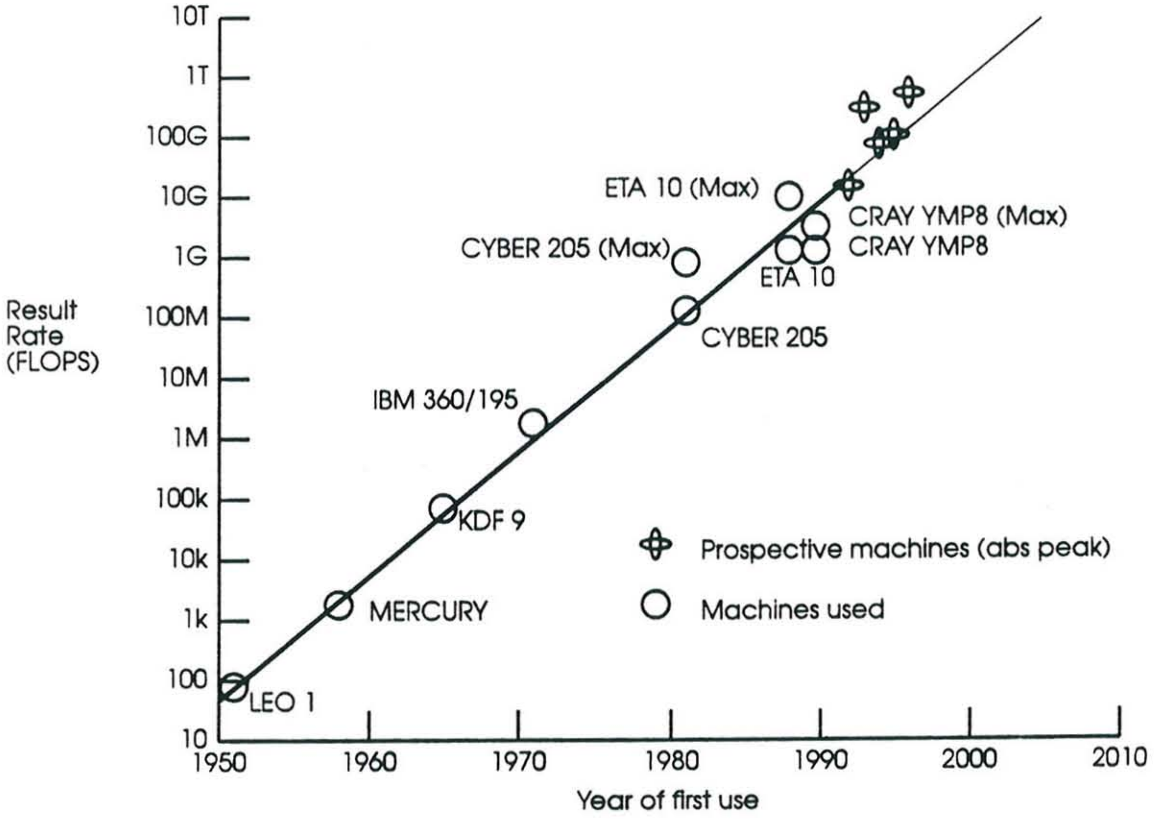

Figure 1: Supercomputer trends at the Met Office. Unlabelled plots, shown as double ellipses, are based on some of the designs known to the Met Office but are not as yet in the public domain.

DISCUSSION

Rapporteur: S. Caughey

During his lecture, Dr. Dickinson showed a slide indicating the Met. Office Cray is only used for 1 hour in a 12 hour cycle for global forecasting. Professor B. Randell asked if that was the only use of the Cray and was told that it was also used for other forecasts and for research. Later in the lecture, Prof. Randell asked if each Met. centre uses the same program, and when Dr. Dickinson replied that they did not, Dr. C. Holt asked if the different programs tend to agree. Dr. Dickinson replied that forecasting is an inexact science, so the various results do not always agree, however the results of other forecasts are taken into account.

In response to a graph showing the improved quality of forecasting vs. the costs, Professor A. Tanenbaum pointed out that the improvements did not match the costs. In response Dr. Dickinson reminded the audience of the "inevitable law of diminishing returns".

Mr. N. P. Holt enquired whether every element of the grid into which the global surface was divided for the purpose of global forecasting needed to be of the same size, and was informed that certain regions such as polar areas do not require the same resolution.

After the lecture Prof. Randell asked if there is much cooperation between various Met. Offices. Dr. Dickinson said that, yes, there was considerable discussion and, indeed recently a number of Met. Offices had submitted a joint Esprit proposal.

Professor Dr. D. Swierstra commented on the difficulties of porting the Met. Office program onto new machines and asked how they hope to resolve this. Recently, he was told, the Met. Office had taken a step in to improve this situation by standardising on Fortran.

Professor M. H. Rogers asked how much code was involved and Dr. Dickinson answered that there was something in the area of 1/2 million lines.

Professor Dr. Swierstra commented that it was surprising that such a large program was required "for a few pages of equations", and was informed that the problem was rather more complex than that.

Mr. I. Barrow suggested that the existing code was probably highly optimised, and when it was transferred to a massively parallel machine this optimisation would have to be first undone. Dr. Dickinson said there was a certain amount of optimisation but he felt it was well understood.

Professor P. Hall asked how much code was concerned with IO and was told it was very little, perhaps 5%.