1. Barooah, B. C. and Goswami, I. D., *J. Min. Met. Fuel*, 1972, 20, 368–373.
2. Golani, P. R., *J. Geol. Soc. India*, 1991, **37**, 31–38.

ACKNOWLEDGEMENTS. We are grateful to Dr. K. K. Dwivedy, Director and Shri Jagmer Singh, Associate Director, Atomic Minerals Division, Hyderabad for their guidance and encouragement during the course of investigation. We are also thankful to Dr A. V. Subrahmanyam for petrographic studies and S/Shri Raghav Saran and C. R. Khorge for analytical support.

R. MAMALLAN*
R. C. JAIN**
A. B. AWATI*
S. N. KAK**
B. M. SWARNKAR†

*Atomic Minerals Division,
Begumpet, Hyderabad 500 016, India
**RCER, Atomic Minerals Division,
Shillong 793 011, India
†RCER, Atomic Minerals Division,
Bangalore 560 072, India

COMMENTARY

Usability of parallel processing computers in numerical weather prediction

U. N. Sinha and Ravi S. Nanjundiah

On 23 November 1992, the Department of Science and Technology (DST) convened a meeting to discuss 'Future Supercomputing Strategies for Medium Range Weather Forecasting'. Subsequently it was decided to invite developers of indigenous parallel processing systems (PPS) to evolve suitable strategies of implementation of weather forecasting codes on their respective parallel machines. The aim of this project, as correctly stated by Basu in a recent report in this journal', was to demonstrate amongst the scientific community whether the PPS developed in India are capable of handling large applications with reasonable ease and also to benchmark the different PPS machines by running the same application code (namely the spectral model at T80 resolution) with identical initial and boundary conditions provided by a common agency (the NCMRWF). DST realized that India might have a headstart in the field of parallel computing, and its attempt to enhance and augment the indigenous technological base in this (then) emerging field for a well-defined national task was indeed commendable.

Basu was the co-odinator of this exercise and his paper summarizes his findings and views. In the present note, we present certain aspects which appear to have been overlooked by the author and therefore makes his assessment misleading, and offer a different perspective on the project and its international counterparts based on personal experience of one of us (RSN) in India and the US.

Are Indian PPS not good enough?

The title and abstract suggest that the paper is generally about the usability of parallel computing to weather forecasting, while the tone of the paper and its conclusion suggest that Indian PPS are not suitable to meet the requirements of NCMRWF. Basu tries to support this view with the following comments on the Indian exercise:

Poor sustained-to-peak ratio

Basu writes, 'The experience of parallelizing the global spectral forecast model operational at NCMRWF showed that the PPS computers designed and fabricated in India during 1994 could attain a sustained-to-peak performance close to 6%. Since this value is significantly less than the internationally accepted figure, it is possible that the basic design of processor boards used in the machines was not suitable for spectral forecast model.' During the same period as the Indian exercise, Drake *et al.*[2] have published sustained-to-peak ratios for the i860 processor (the processor used in India also by NAL, CDAC and BARC), and we reproduce their tables here. Table 1 displays the performance of the parallel computers in empirical studies, and Table 2 shows the processor's actual performance on meteorological codes.

Considering that the peak speed of i860 is 75 Mflop/s (according to Drake *et al.*[2]), peak of 6% achieved by the Indian PPS was on par with systems elsewhere. Drake *et al.*[2] admit that their experience with the i860 (one of the few processors that have been extensively used in parallel computing applications for meteorology) in regard to its sustained-to-peak speed ratio was less than satisfactory. Therefore it is wrong to conclude that the relatively Iow value of sustained-to-peak ratio is unique to the Indian PPS (as suggested by Basu). We are not aware on what basis Basu drew his conclusion about 'internationally accepted figures' in 1994.

Scalability

Discussing this issue Basu says: 'To ensure scalability of an application code is not a trivial task even for multitasking, shared memory, vector processing computer. Distribution of data and optimization of inter-processor communication make it even more difficult for a distributed memory PPS.' He further contends, 'Indian machines, however, have not demonstrated scalability clearly and some more effort is

Table 1. Parallel computers used in empirical studies, characterized by operating system version, microprocessor, interconnection network, maximum machine size in experiments (N), message passing startup cost (t_s), per-byte transfer cost (t_b), and achieved per-processor Mflop/s at single and double precision (from ref. 2)

Name	OS	Processor	Network	N
Paragon	SUNMOS 1.6.5	i860XP	16 x 64 mesh	1024
SP2	AIX + MPL	Power 2	multistage crossbar	128

Name	t_s (µs)	t_b (µs)	MB/s (swap)	Mflop/s Single	Mflop/s Double
Paragon	72	0.007	282	11.60	8.5
SP2	70	0.044	45	44.86	53.8

Table 2. Elapsed time per model day and computational rate at T170 resolution on the Paragon and SP2 for double precision and single precision (ref. 2)

			Computational rate	
Name	Nodes	Time/model day (s)	Gflop/s	Mflop/s/node
***Double** precision*				
Paragon	512	1510	1.71	3.3
	1024	814	3.18	3.1
SP2	128	1092	2.27	18.5
***Single** precision*				
Paragon	1024	525.6	4.93	4.8
SP2	64	1606	1.61	25.2
SP2	128	1077	2.40	18.8

required'. Basu is well aware of the fact that a small sequential element in a program can significantly limit the effectiveness of the parallelizing exercise. But the fact that such a small element existed was neither apprehended by the experts at NCMRWF nor by the developers (who, it must be stated, did not have much earlier experience with the T80 code). The NCMRWF T80 global spectral model has its origins in the NCEP model, which has been largely shaped by Sela with Basu as one of the co-authors[3]. Sela's experience in parallelizing this model on a shared memory vector parallel machine (clearly the author's favourite), viz. C90, is very succinctly summarized in Figure 1 (reproduced from Sela[4]).

We would like the reader to note that the efficiency of the C90 with 4 processors was 77.5% and with 8 processors, it was 68.75%. Hence the Indian efforts in the DST project were comparable to efforts elsewhere at the same time (with the disadvantage of little support from the industry in contrast to the close interaction between industry and research groups in most efforts elsewhere).

Basu correctly states that, unlike the implementations of the ECMWF model' and the NCAR model[2] where considerable effort was devoted to developing codes that were scalable, Indian PPS developers did not make efforts in this respect. The following must, however, be stated:

1. This was the first implementation of the model, and the general experience is that such first implementations of any software are rarely optimal.
2. The project was closed in March 1996, just as these initial implementations were completed.
3. The efforts of PPS developers after March 1996 have not been considered in Basu's paper, on the pretext that model outputs have not been examined!

Out of scientific curiosity we have conducted further studies on the scalability of this model[6]. We have found that the initial parallel implementation of the NCMRWF code has a sequential component of 4.7% and its scalability on an ideal machine (i.e. with maximum theoretical efficiency, with infinite bandwidth for communication) is presented in Table 3.

Table 3. Comparison of maximum theoretical and actual achieved efficiencies on a 4 processor SGI power challenge (ref. 6)

No. of procs	Maximum theoretical efficiency		Efficiency achieved	
	Case A	Case B	Case A	Case B
1	100.0	100.0	100.0	100.0
2	95.5	99.5	93.3	96.1
4	87.6	98.6	81.4	88.9
8	75.1	96.9	–	–
16	58.5	93.5	–	–
32	40.5	87.5	–	–
64	25.2	82.3	–	–

The cause of poor efficiencies in Sela's or our earlier implementation can now be explained on the basis of this table. Sela's implementation uses the strategy of parallel implementation of grid space computations (Case **A** in Table 3). However, we have further refined the load decomposition strategy (Case B). This refinement now includes concurrent computing of the linear part of the model, in addition to decomposition of loads in physical space. The sequential part by this strategy reduces to 0.34%, and the scalability consequently improves dramatically. It must be pointed out that in the present version, the computation of the linear part is conducted on the summed coefficients, whereas this could be done on the modes themselves (as modes do not interact in this part of the model).

Had these modifications been performed as part of the Indian project on the PPS, the results would have been less misleading. However, we need to point out that we could arrive at these conclusions and alternate strategies only after the initial parallelizing exercise and after studying the results of this effort. It is disconcerting that Basu (one of the co-authors of the NCEP/NCMRWF model) missed this critical aspect of Sela's parallelization, i.e. the technique of parallelizing computations in physical grid space alone would not