We analyzed the performance of the PCG solver on a 512-node Paragon in further detail (row 6 in Table 1). The solver achieves a sustained speed of 18.3 GFLOPS for an 85000 observation problem. (On 512-PE T3D and T3E, the sustained speed in the PCG solver is 12.6 and 33.8 GFLOPS respectively.) This represents 36% of the theoretical total peak speed of 51.2 GFLOPS. The solver spends 27.5% of the time on communications for this problem on 512-nodes. During each iteration, most of the communication time is spent on sending or receiving (on average) 536 messages per processor with (on average) 166 floating point numbers in a message. These numbers indicate a tightly coupled system like T3E is necessary for this problem.

6. Conclusions

We have designed and implemented a set of efficient and scalable algorithms for the PSAS data assimilation package, and achieved a 740-fold solution time reduction on a 512-PE T3E parallel platform over a single head of a Cray C90. This clearly demonstrates that data assimilation problems are well suited for distributed-memory massively parallel computer architectures. In particular, this work demonstrates that irregular and unstructured problems such as the data assimilation problem can be efficiently implemented on this type of architecture, with good understanding of the problem, careful (re)design of all necessary algorithms, and effective use of explicit message passing. By focusing on the algorithms, the application achieves a high sustained performance. Since no specialized optimizations targeted for the particular computer are done, these performance numbers are representative of what could be achieved by average users.

Acknowledgment. This work was funded by a NASA HPCC ESS Grand-Challenge Application, adminstered through University of Maryland (BG97-06100). This research also uses resources of the National Energy Research Scientific Computing Center, which is supported by the Office of Energy Research of the U.S. Department of Energy through contract number DE-AC03-76SF00098.

References

[1] R. Daley, “Atmospheric Data Analysis”, Cambridge University Press, New York, 1991.

[2] J. Pfaendtner, S. Bloom, D. Lamich, M. Seablom. M. Sienkiewicz, J.Stobie, A. da Silva, Documentation of the Goddard Earth Observing System Data Assimilation System - Version 1, NASA Tech Memo 104604, v.4, Goddard Space Flight Center, January 1995.

[3] S.E. Cohn, A. da Silva, J. Guo, M. Sienkiewicz, and D. Lamich, Assessing the Effects of Data Selection with DAO’s Physical-space Statistical Analysis System, Tech Report 97-08, Data Assimilation Office, Goddard Space Flight Center, April 1997.

[4] C.H.Q.Ding and R.D.Ferraro, “A Parallel Climate Data Assimilation Package”, SIAM News, November 1996, pp.1-11.

[5] C.H.Q. Ding and R.D. Ferraro, “Slices: A Scalable Concurrent Partitioner for Unstructured Finite Element Meshes”, Proceedings of 7th SIAM Conference for Parallel Processing, pp.492-493, 1995.

[6] C.H.Q. Ding and R.D. Ferraro, “A General Purpose Parallel Sparse Matrix Solvers Package”, Proceedings of 9th International Parallel Processing Symposium, pp.70-76, April 1995.