CRAY T3E and SGI Origin2000: Merging Architectures from the User's Point of View

Stephan Seidl and Wolfgang E. Nagel

Center for High Performance Computing (ZHR)
Dresden University of Technology
D-01062 Dresden, Germany
E-mail: {seidl | nagel}@zhr.tu-dresden.de

Abstract

While the T3E is very well established as a highly parallel machine in many compute intensive environments, large Origin2000 sites still have to optimize their usage profile to get effective cycles for parallel codes even for moderate numbers of processors. The paper compares T3E and Origin2000 systems, highlighting some details with respect to parallel programming and runtime behavior of appropriate applications. The goal is not to favor one system over the other, but to give recommendations how to design applications which are able to run efficiently on both architectures.

Users are mainly faced with two differences between both systems. First, on a T3E a parallel application is statically parallel from the beginning. In case of the Origin2000, an application gets parallel during execution time when the user has control. Second, once started on a T3E, a parallel application is always running as fast as possible. On an Origin2000, this does only happen under certain circumstances. The paper will give some background about these facts and will demonstrate the strong dependence of the runtime behavior of parallel programs on different runtime situations. Comparative performance illustrations of both machines will color the overall picture of the merging worlds.

Originally published in Proceedings of the Fourth European SGI/Cray MPP Workshop (Sep 10-11, 1998, IPP, Garching, Germany), H. Lederer and F. Hertweck, eds., IPP R/46, Oct 1998, pp. 6-19.
Also available as internal report ZHR-IR-9803.

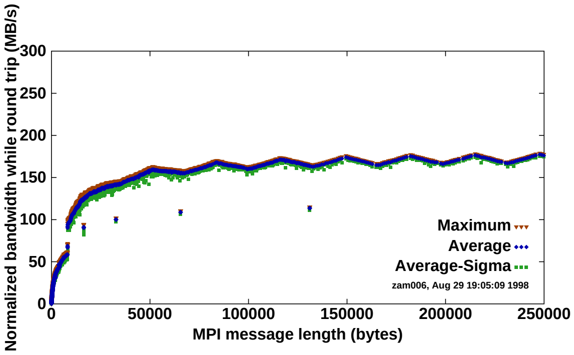

Figure 1: MPI point-to-point communication on T3E-600 with streams off, long messages with *MPI_Send()*

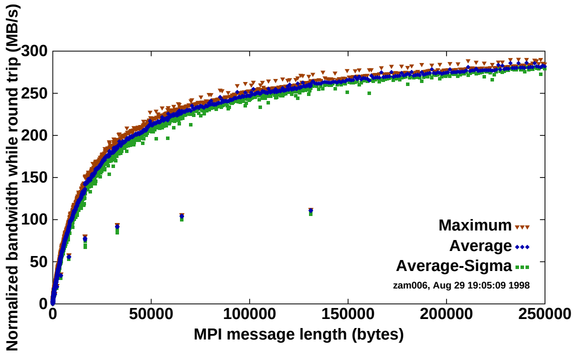

Figure 2: MPI point-to-point communication on T3E-600 with streams off, long messages with *MPI_Ssend()*

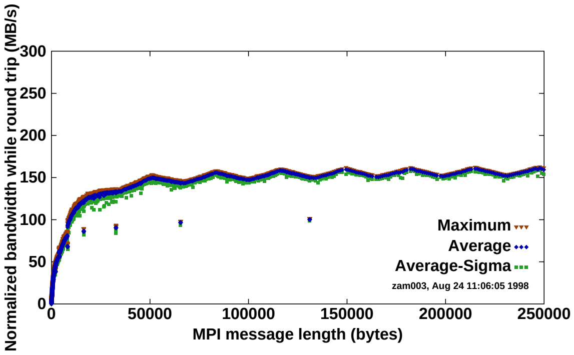

Figure 3: MPI point-to-point communication on T3E-900 with streams on, long messages with *MPI_Send()*

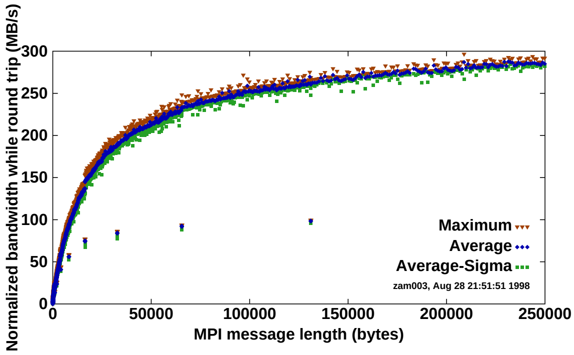

Figure 4: MPI point-to-point communication on T3E-900 with streams on, long messages with *MPI_Ssend()*

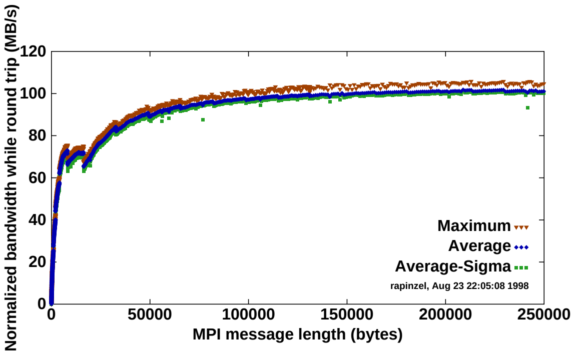

Figure 5: MPI point-to-point communication on Origin2000, long messages with *MPI_Send()*

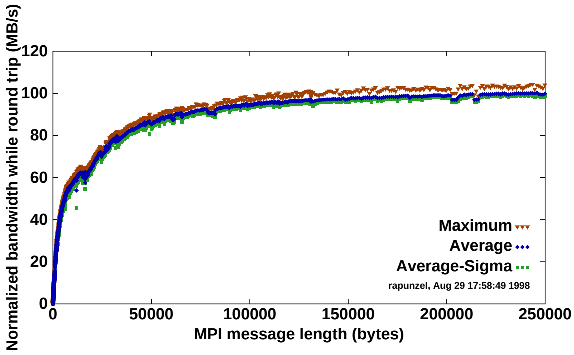

Figure 6: MPI point-to-point communication on Origin2000, long messages with *MPI_Ssend()*

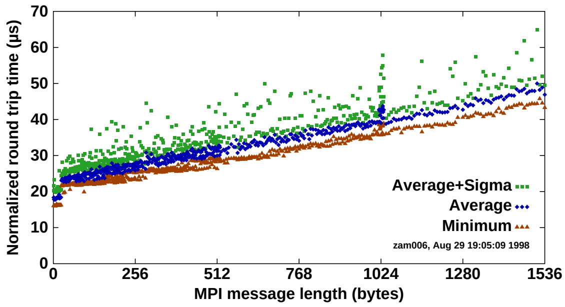

Figure 7: MPI point-to-point communication on T3E-600 with streams off, short messages with *MPI_Send()*

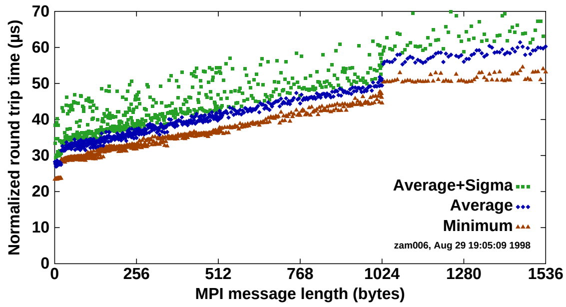

Figure 8: MPI point-to-point communication on T3E-600 with streams off, short messages with *MPI_Ssend()*

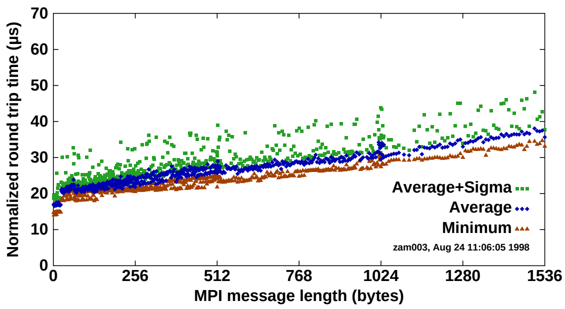

Figure 9: MPI point-to-point communication on T3E-900 with streams on, short messages with *MPI_Send()*

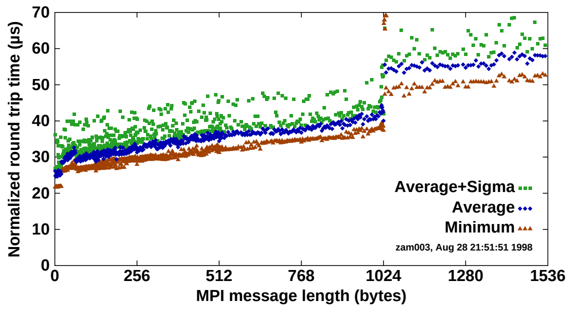

Figure 10: MPI point-to-point communication on T3E-900 with streams on, short messages with *MPI_Ssend()*

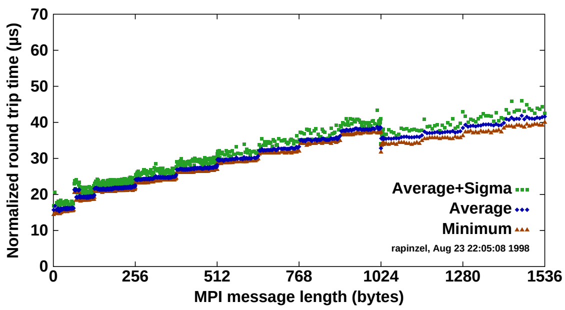

Figure 11: MPI point-to-point communication on Origin2000, short messages with *MPI_Send()*

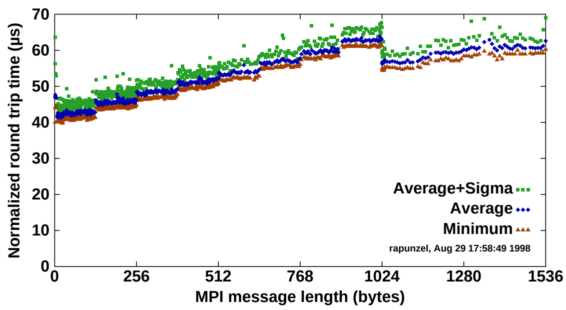

Figure 12: MPI point-to-point communication on Origin2000, short messages with *MPI_Ssend()*

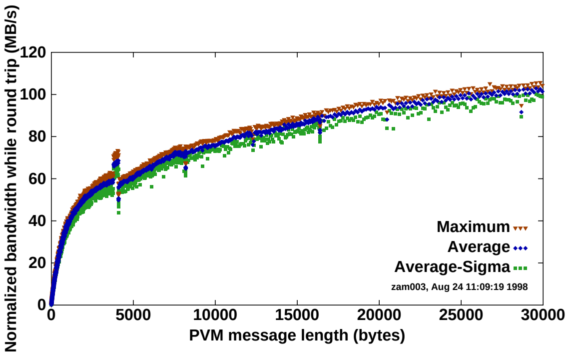

Figure 13: PVM point-to-point communication on T3E-900 with streams on, medium size messages

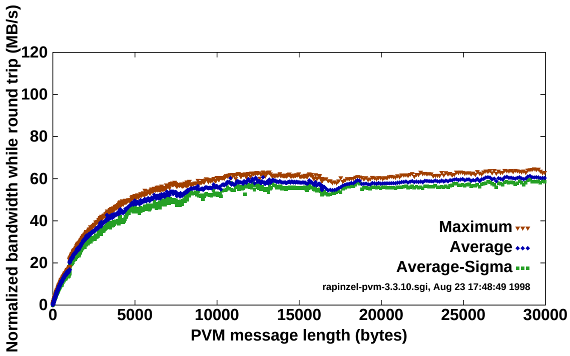

Figure 14: PVM point-to-point communication on Origin2000, medium size messages

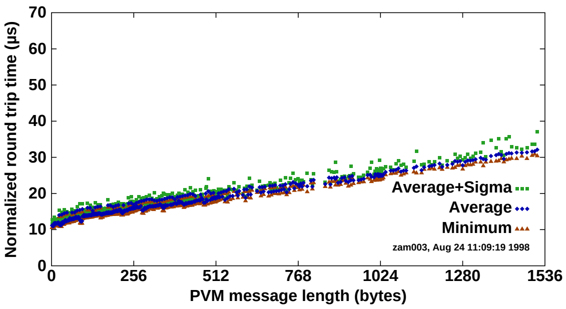

Figure 15: PVM point-to-point communication on T3E-900 with streams on, short messages

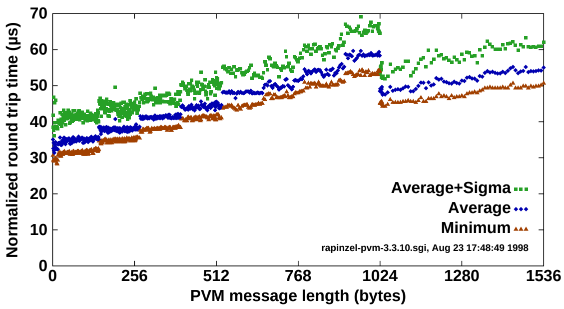

Figure 16: PVM point-to-point communication on Origin2000, short messages