Micro-Analysis of the Titan's Operation Pipe

John Sanguinetti
Ardent Computer Corporation

Abstract

Much of the performance analysis done in designing a computer is based on fundamental operation rates, like cycle time and number of pipe stages in an operation pipeline. This kind of analysis yields peak computation rates which, in fact, may never be realized. Resource contention between different units, each of which has a fundamental operation rate adequate to support a given overall peak, may cause the actual obtainable rate to be much less. In order to determine the effects of interactions between different requestors and common resources, micro-analysis, or detailed modelling of the part of the system in question, is necessary. In this paper, we report on the results of such an analysis on the operation pipe of the Titan graphics supercomputer. The Titan is a new class of machine with a supercomputer-style architecture implemented in a technology appropriate for a single-user machine. The micro-analysis reported here resulted in enhancing the actual obtainable computation rate from 10.8 Mflops to 14.6 Mflops for a particular real application, while the fundamental operation rate is 16 Mflops.

Description of the Titan

The Titan was conceived to be a "Visualization Tool" — a machine which would allow an engineer or scientist to model a physical entity and then visualize the results of the model. Consequently, the machine would have to first be able to do the computation required for physical modelling and second be able to render the resulting image, all in a reasonable amount of time. The goal of the Titan was to provide a significant fraction (about 25%) of the performance of a super-computer and provide superior graphics performance for a price less than $100,000.

Permission to copy without fee all or part of this material is granted provided that the copies are not made or distributed for direct commercial advantage, the ACM copyright notice and the title of the publication and its date appear, and notice is given that copying is by permission of the Association for Computing Machinery. To copy otherwise, or to republish, requires a fee and/or specific permission.

© 1988 ACM 0-89791-272-1/88/0007/0190 $1.50

Proceedings of Supercomputer Conference, St. Malo, June 1988.

Particular performance goals included executing a Linpack 100x100 benchmark at a rate of 6 Mflops (compiled) on a single processor. This was considered an important performance goal since many real problems have characteristics similar to the Linpack benchmark and typical supercomputer rates on this benchmark range from 20 to 40 Mflops — a Cray 2S (1 processor) is rated at 23 Mflops [Dongarra].

Aside from computational power for physical modelling problems, floating point power is also needed for doing 3-D graphics rendering. Drawing a picture on a 1024x1280 screen which occupies 1/2 of the screen area using a ray tracing algorithm with shading, 3 light sources, and environmental reflection would typically take about 5 billion floating point operations. Titan was intended to have enough computational power to draw such pictures in "reasonable" time. A rate of 2 Mflops for scalar computation would allow such rendering to be done in about 40 minutes on a 2 processor Titan.

In order to satisfy its performance goals, the Titan needs a fast processor and a high bandwidth processor-to-memory interconnection. In order to do double-precision vector triad operations, the processor-to-memory connection must deliver 16 bytes to the processor and 8 bytes to memory for each operation, making a requirement of 24MB per Mflop (this requirement can be reduced by chaining in the floating point unit for some operations).

The Titan is a symmetric multiprocessor, with up to four identical processors. The organization of the processor is fairly typical of a vector computer. There is an integer processor (IPU) and a floating point unit (FPU). The FPU contains two load pipes to memory, one store pipe, a single operation pipe which can accommodate a chained multiply and add, and a reasonably large set of vector registers (8192 cells) which can be configured in a variety of vector sizes. Figure 1 is a block diagram of the processor and memory of the Titan.

The fundamental rates of the various units are matched so that a peak of one result every 2 clock cycles can be