

Report June 2001

NSF #ACI-0081791

TITLE: Integrated Design of High Performance Clustered Computing Systems and Algorithms for Large Scale FE Simulation

What People have worked on Project:

Bob McLay

Role: Research Associate

Extent of time: 160 hours – yes (part time on project)

What the person has done on the project: Dr. McLay is involved in the software development for the benchmark testing of the kernel.

Demographic data -- Gender: Male, Ethnicity: American, Race: Caucasian, Disability: no-disability, Citizenship: America

Ben Kirk

Role: Graduate Student

Extent of time: 160 hours – yes (part time on project)

What the person has done on the project: Mr. Kirk is involved in testing the Intel Pentium processors and in the performance studies.

Demographic data -- Gender: Male, Ethnicity: American, Race: Caucasian, Disability: no-disability, Citizenship: America

Bill Barth

Role: Graduate Student

Extent of time: 160 hours – yes (part time on project)

What the person has done on the project: Mr. Barth has assisted Mr. Kirk in the above studies.

Demographic data -- Gender: Male, Ethnicity: American, Race: Caucasian, Disability: no-disability, Citizenship: America

Saeed Iqbal

Role: Graduate Student

Extent of time: 160 hours – yes (part time on project)

What the person has done on the project: Mr. Iqbal has been looking at other related issues concerning architectures and algorithms

Demographic data -- Gender: Male, Ethnicity: Pakistan, Race: Asian, Disability: no-disability, Citizenship: Pakistan

Anand Ramachandran

Role: graduate student

Extent of time: > 160 hours - yes (part time for 2 semesters – supported by grant)

What the person has done on the project: ongoing research - memory management for high performance multiprocessor chips specialized to execute data dominated, computation intensive applications amenable to static analysis (such as finite element methods). Currently working on novel algorithms and architectural support for compiler direct (software based) cache management as well as compiler directed management of on-chip main memory partitions. Study on suitable level of abstraction and granularity of decomposition to be used for specifying the target application. Preliminary specification language for describing data streaming requirements of the application has been developed. Preliminary compiler directed cache management algorithms were developed and are currently being experimentally validated using a preliminary set of micro-benchmarks.

Demographic data -- Gender: Male, Ethnicity: Indian, Race: Asian, Disability: no-disability, Citizenship: India

Cagdas Akturan –

Role: graduate student;

Extent of time: > 160 hours - NO (only minor support from grant)

What the person has done on the project: ongoing research - advanced software pipelining (loop transformation) algorithms to be incorporated in optimizing compilers for high ILP clustered VLIW/EPIC machines. Currently extending/augmenting her previously developed software pipelining algorithm (which can effectively explore trade-offs between performance, code size, and memory/register file requirements) so as to handle target Clustered VLIW/EPIC machines.

Demographic data -- Gender: Female, Ethnicity: Turkish, Race: Caucasian, Disability: no-disability, Citizenship: Turkey

Viktor Lapinskii –

Role: graduate student;

Extent of time: > 160 hours - NO (only minor support from grant)

What the person has done on the project: ongoing research - Novel binding algorithms to be incorporated in optimizing compilers for high ILP clustered VLIW/EPIC machines. Novel design space exploration algorithms to identify specialized cluster configurations tuned to the needs of computation intensive applications amenable to static analysis (such as finite element methods). Algorithms have been developed and thoroughly validated with representative benchmarks.

Demographic data -- Gender: Male, Ethnicity: Russian, Race: Caucasian, Disability: no-disability, Citizenship: Russia

Satish Pilai –

Role: graduate student;

Extent of time: > 160 hours - NO (only minor support from grant)

What the person has done on the project: ongoing research – performance enhancing compiler transformations (advanced predication and static speculation techniques) for Clustered VLIW/EPIC machines executing computation intensive applications (such as finite element methods). Working on a novel SSA-PS (static single assignment predicated switching) compiler transformation and defining the required micro-architectural support. Preliminary algorithms were developed and experimentally contrasted (using a set of micro-benchmarks) with state of the art techniques, with promising results. The student is currently working on incorporating more aggressive resource constraints on the SSA-PS compiler transformation.

Demographic data -- Gender: Male, Ethnicity: Indian, Race: Asian, Disability: no-disability, Citizenship: India

Project Activities:

Task 1: Target Applications: Development of Detailed FE Models / Algorithms

- Study of the design of finite element methods and algorithms and their behavior on different chip and memory architectures.
- Theoretical studies on the methods to delineate classes of algorithms that are better suited to current generation commodity-off-the-shelf processors.
- Investigations on how the methodology and algorithms "suggest" the architecture and chip specifics might be significantly improved. Of particular interest are schemes that also lend themselves to parallel scalable HPC computations using Beowulf clusters or similar distributed and distributed-shared parallel computing systems.
- Preliminary numerical studies to investigate the improvements obtained with increased processor performance subject to memory bandwidth constraints and cache structure. Of particular note are the experiments on an SMP chip system recently obtained from Intel. In fall semester we obtained, via a grant from Intel, 32 Pentium III processors and supporting boards. Using support from other contracts we constructed a new Beowulf cluster for subsequent performance studies.
- Carried out some simple single-node studies to determine peak performance of the matrix-vector product kernel for our applications code on a number of different architectures. Following this, the performance of the 16-node Beowulf system was investigated and compared with that of a previous cluster of 16 Intel PII Processors, a 16 node Alpha cluster, and the CPlant clusters at Sandia, as well as our experience with SMP on SGI Origin systems. Despite the significant difference in floating-point performance, cache size, and cache speed available on these architectures, the single largest factor that influenced algorithm performance was found to be memory bandwidth. In fact, some of the results indicate that performance on newer Intel-based PCs is relatively independent of CPU clock rate. Dramatic increases in future performance are likely to come from an increase in memory bandwidth rather than a higher CPU clock speed.
- Scaled speed-up studies are being carried out on the 16 node, 32-processor Pentium III Xeon (550 Mhz) cluster in the CFD Lab. Linear scaling was observed when using one processor per machine (up to 16 processors) where the machine achieves a peak performance rate of 1.59 GigaFlops. The performance results were slightly improved by using the Myrinet interconnect instead of the Fast Ethernet. Scaling beyond 16 processors was observed at a much reduced rate due to memory bandwidth contention between the two processors. The performance of all 32 processors increased to only 1.92 GigaFlops, and the per-processor performance was nearly cut in half.

The element-by-element bi-conjugate gradient algorithm used in the code is limited by the memory bandwidth due to the low reuse of data in the processor cache. Therefore the addition of a second processor to the memory

bus was not seen to improve the scaling of the code. In fact, in scaled speed-up tests the performance was seen to change from linear in the number of processors to nearly constant above the 16 processor threshold, bolstering the idea that the memory bandwidth is the limiting factor for these types of solvers. This performance study demonstrates one of the key issues in our project concerning the unintentional limitations imposed by processor architectures on applications.

Task 2: *System Software*: Development of Algorithm Templates. Design of Specification Language.

- Study on suitable level of abstraction and granularity of decomposition to be used for specifying target data dominated, computation intensive applications (e.g., finite element methods). Development of preliminary specification language for describing data streaming requirements such applications.

Task 3: *System Software, System Design*: Study and Development of Retargetable Compilation Algorithms

- Development of preliminary novel techniques/algorithms for compiler direct (software based) cache management as well as compiler directed (software based) management of on-chip main memory partitions.
- Development of novel binding algorithms to be incorporated in optimizing compilers for high ILP clustered VLIW/EPIC machines.
- Development of advanced software pipelining (loop transformation) algorithms to be incorporated in optimizing compilers for high ILP clustered VLIW/EPIC machines (on-going).
- Development of preliminary compiler algorithms to implement on the novel SSA-PS code transformation (aggressive compiler directed speculative execution for high ILP clustered VLIW/EPIC machines)

Task 4: *System Design*: Definition of Parameterized S-CMP Architecture. Analysis of Cost-Performance Tradeoffs.

- Defining architectural support (novel memory management subsystem) required by our compiler direct (software based) cache management technique, as well as our compiler directed management of on-chip main memory partitions (see task 3).
- Novel design space exploration algorithms to identify specialized cluster configurations tuned to the needs
- Defining micro-architectural support to predicated switching code (minor extension to support provided in EPIC machines to standard predication and speculative execution).

Task 5: *Target Applications*: Advanced FE Studies

- The applications problem is coupled transport simulation for fluid flow and heat transfer - a central problem in computational fluid dynamics. We have been enhancing the methodology and solver strategies of an existing grand challenge HPC code we developed under prior NASA support. This has put us in a position to accelerate the work on this NSF project. The solver work has been the center of our recent effort.
- We have implemented a class of Krylov subspace iterative solvers with due attention to distributed parallel solution. The computational kernel requires repeated matrix-vector products and this is a key focal point of our "architecture" studies.

Task 6: *Validation*

- Development of a "representative" applications code for benchmarking and testing ideas and for performance studies. In this context we have also acquired the performance measurement software PAPI from Jack Dongarra's group at the University of Tennessee to expedite our work and pinpoint performance bottlenecks in the algorithms. PAPI is a portable library that provides a uniform interface to the hardware performance counters that are difficult to get access to otherwise. It allows one to measure performance behavior, including floating point instruction counts, with very little overhead.
- Preliminary algorithms for compiler direct (software based) cache management as well as compiler directed management of on-chip main memory partitions are currently being experimentally validated using a set of micro-benchmarks.
- Advanced binding algorithms have been thoroughly validated with representative benchmarks.
- Design space exploration algorithms have been thoroughly validated with representative benchmarks.
- Preliminary algorithms for the SSA-PS code transformation were experimentally validated and contrasted with state of the art techniques for compiler directed speculative execution using a set of micro-benchmarks.

Project Findings:

No definitive findings as yet (in core research).

Training and Development

The project/grant has directly funded one research associate and four graduate students, and provided research topics for three additional graduate students. Viktor Lapinskii has completed his Ph.D. in Fall 2001.

Journal Publications

Plaza, A., M. A. Padron and G. F. Carey, "Simulation of 3D evolution problems", *AMSE Journal*, Submitted May 2001.

Cagdas Akturan and Margarida F. Jacome, "RS-DFRA: A Register Sensitive Software Pipelining Algorithm for Embedded VLIW Processors," submitted to the IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, Jan 2001.

Viktor Lapinskii and Margarida F. Jacome, "Application Specific Clustered VLIW Datapaths: Early Exploration on a Parameterized Design Space," submitted to the IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, May 2001.

Books or other non-periodical (e.g., conferences)

Margarida F. Jacome, Gustavo de Veciana, and Satish Pillai, "Clustered VLIW Architectures with Predicated Switching," to appear in Proceedings of IEEE/ACM Design Automation Conference (DAC'2001), Las Vegas, June 18-22, 2001.

Viktor Lapinskii, Margarida F. Jacome, and Gustavo de Veciana, "High Quality Operation Binding for Clustered VLIW Datapaths," to appear in Proceedings of IEEE/ACM Design Automation Conference (DAC'2001), Las Vegas, June 18-22, 2001.

Valli, Andrea M.P., A. L. G. A. Coutinho and G. F. Carey, "Control Strategies for Timestep Selection and Convergence Rate of Nonlinear Iterations in Simulation of Rayleigh-Benard-Marangoni Flows," Proceedings of CILAMCE 2000 in Rio de Janeiro, Brazil, December 6-8, 2000.

Web-site

A link has been established on the CFD Lab and this contains information related to this project.

Other specific Products:

Software implementations of:

Binding algorithms for Clustered VLIW/EPIC Processors – module to be incorporated in a retargetable compiler

Design space exploration algorithms for Clustered VLIW/EPIC Processors

Software pipelining algorithms for Clustered VLIW/EPIC Processors - module to be incorporated in a retargetable compiler

Contributions (within discipline:, other disciplines, etc...)

None as yet (too early)

Objectives and Scope:

- not changed from the original proposal