

Optimization of the Selected Quantum Codes on the Cray X1

**L. Bolikowski, F. Rakowski,
K. Wawruch, M. Politowski,
A. Kindziuk, W. Rudnicki,
P. Bala, M. Niezgodka**
ICM UW, Poland

CUG 2004, Knoxville



What is ICM?

Software ported and optimized

DFTB, VASP

Software under development

Gromos, Charmm, Gamess, Siesta

Cray X1 and Bioinformatics

Smith-Watermann

BMM filter

Summary

Questions...



Interdisciplinary Centre for Mathematical and Computational Modelling

High Performance Scientific Computing

state licenses for scientific software

Interdisciplinary research in computational sciences

modelling in biomolecular, material, atmospheric, environmental sciences
visual modelling and data processing

Network services for the Polish scientific community and for general public

Virtual Library, weather forecast, science festivals, academic television

Software ported and optimized

Density Functional – Tight Binding Method (DFTB)

quantum potential generator

do not calculate the Hamilton and overlap matrix elements at each step of the SCF cycle, but the convergence is reached on the level of Mullikan charges.

Vienna Ab-initio Simulation Package (VASP)

ab-initio quantum-mechanical molecular dynamics (MD)
using pseudopotentials and a plane wave basis set

DFTB

developed by prof Th. Frauenchaim at University of Paderborn
Fortran Code

Initial performance

Cray X1 MSP: **336 Mflops** (PC: 255 Mflops)

Changes

main loop vectorization (85% of time)
general loops' redesign to longer vector length
better data alignment for memory access

Results?

Performance is almost 10x higher!

3 213 Mflops for 1 MSP

active site of PKA kinase, consisted of about 400 atoms

single point energy calculations

convergence was reached after 20 SCF cycles

currently 65% of computation time is Linear Algebra

VASP

developed by prof Juergen Hafner and Juergen Furthmueller
at Vienna University

Fortran code with MPI

initially does not compile

we used port of previous VASP release to Cray X1 as help

initial SSP performance: **100 – 300 Mflops** (PC: 270 Mflops)

Changes

optimization of 3x main loops (90% of computation time)

better data alignment for memory access

playing with BLAS levels

Results

Performance rose to 800 – 1 555 Mflops per SSP

depends on kind of test

What else?

replacing MPI with CAF, due to poor parallel scaling

MSP version for large memory jobs

Software ported, but still in development

Gromos

Fortran code with CAF parallelization
vector version for Cray Y-MP of evaluation of nonbonded interactions between atom pairs

Performance

Cray X1 MSP: 640 Mflops
Cray X1 SSP: 540 Mflops
reasonable scalability (2.5x on 4 CPUs)

Future work

MSP directives
domain decomposition instead of force decomposition

Software ported, but still in development

Chemistry at HARvard Molecular Mechanics (CHARMM)

macromolecular simulations, including energy minimization,
molecular dynamics and Monte Carlo simulations.

huge and complicated Fortran 77 code with MPI

Current status

compiles! (and links...)

poor performance (about 100 Mflops on SSP)

poor scalability

problems with code structure

problems with vector code

Software ported, but still in development

General Atomic and Molecular Electronic Structure System (GAMESS)

Gordon research group at Iowa State University

ab-initio quantum chemistry package

Fortran code

Current status

poor real life performance: **160 – 366 Mflops** per SSP

(some parts of the code have over 1.5 Gflops)

problems with code structure

Software ported, but still in development

Spanish Initiative for Electronic Simulations with Thousands of Atoms (Siesta)

Fortran 90 code with MPI

ab-initio package

Current status

poor performance: **200 – 430 Mflops** per SSP

Bioinformatics

Smith-Waterman algorithm (SW)

one of most accurate algorithms for finding similarities of sequences
25M cells per second on SSP (PC: up to 7Mcps)

Filter for SW

using Bit Matrix Multiply (BMM) unit for some bit operations
80–160M cells per second on SSP (PC: up to 20Mcps)
decreases the number of sequences for SW 10x

Purpose?

fast search of whole sequence database (2M+) in 1h on 1 SSP
clustering of whole sequence database to speed up search
and find biologically significant similarities

Summary

64bits are still a problem

Cray to IEEE transition is a problem

**3 levels of parallelism
(inside MSP, inter-MSP, inter-node)**

**Ability to achieve sustained performance
25% – 50% of peak**



Where do you want to go *tomorrow*?