

LA-UR- 10-08134

Approved for public release;
distribution is unlimited.

Title: Optimizing the Inner Loop of the Gravitational Force
Interaction on Modern Processors

Author(s): Michael S. Warren, T-2, LANL

Intended for: The Future of AstroComputing Conference



Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by the Los Alamos National Security, LLC for the National Nuclear Security Administration of the U.S. Department of Energy under contract DE-AC52-06NA25396. By acceptance of this article, the publisher recognizes that the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes. Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy. Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee its technical correctness.

Optimizing the Inner Loop of the Gravitational Force Interaction on Modern Processors

Michael S. Warren
Los Alamos National Laboratory
`msw@lanl.gov`

We have achieved superior performance on multiple generations of the fastest supercomputers in the world with our hashed oct-tree N-body code (HOT), spanning almost two decades and garnering multiple Gordon Bell Prizes for significant achievement in parallel processing. Execution time for our N-body code is largely influenced by the force calculation in the inner loop. Improvements to the inner loop using SSE3 instructions has enabled the calculation of over 200 million gravitational interactions per second per processor on a 2.6 GHz Opteron, for a computational rate of over 7 Gflops in single precision (70% of peak).

We obtain optimal performance some processors (including the Cell) by decomposing the reciprocal square root function required for a gravitational interaction into a table lookup, Chebychev polynomial interpolation, and Newton-Raphson iteration, using the algorithm of Karp. By unrolling the loop by a factor of six, and using SPU intrinsics to compute on vectors, we obtain performance of over 16 Gflops on a single Cell SPE. Aggregated over the 8 SPEs on a Cell processor, the overall performance is roughly 130 Gflops. In comparison, the ordinary C version of our inner loop only obtains 1.6 Gflops per SPE with the `spux1c` compiler.

Inner Loop Performance

Processor	libm	Karp
533-MHz Alpha EV56	76.2	242.2
933-MHz Transmeta TM5800	189.5	373.2
375-MHz IBM Power3	298.5	514.4
1133-MHz Intel P3	292.2	594.9
1200-MHz AMD Athlon MP	350.7	614.0
1800-MHz AMD Athlon XP	609.9	951.9
1250-MHz Alpha 21264C	935.2	1141.0
2530-MHz Intel P4 (icc)	1170.0	1357.0
2530-MHz Intel P4 (SSE)	6514.0	
2600-MHz AMD Opteron (SSE3)	7380.0	
PowerXCell 8i (single SPE)	16356.0	

Table 1: Mflop/s obtained on our gravitational micro-kernel benchmark.

Historical Performance

Year	Site	Machine	Procs	Gflop/s	Mflops/proc
2006	LANL	Coyote (SSE3)	448	1880	4200.0
2004	LANL	Space Simulator (SSE)	288	1166	4050.0
2003	LANL	ASCI QB	3600	2793	775.8
2002	NERSC	IBM SP-3(375/W)	256	57.70	225.0
2000	LANL	SGI Origin 2000	64	13.10	205.0
1996	Sandia	ASCI Red	6800	464.9	68.4
1995	JPL	Cray T3D	256	7.94	31.0
1995	LANL	TMC CM-5	512	14.06	27.5
1993	Caltech	Intel Delta	512	10.02	19.6

Table 2: Performance of HOT on a variety of parallel supercomputers.

Cell Code for Inner Loop

```
for (i = 0; i <= n/16; i++) {  
    mass = *(vector float *)p;  
    dx = *(vector float *)(p+4);  
    dy = *(vector float *)(p+8);  
    dz = *(vector float *)(p+12);  
  
    dx = spu_sub(dx, pposx);  
    dy = spu_sub(dy, pposy);  
    dz = spu_sub(dz, pposz);  
  
    dr2 = spu_madd(dx, dx, eps2);  
    dr2 = spu_madd(dy, dy, dr2);  
    dr2 = spu_madd(dz, dz, dr2);  
  
    phii = spu_rsqtrte(dr2);  
    mor3 = spu_mul(phii, phii);  
    phii = spu_mul(phii, mass);  
    total_mass = spu_add(total_mass, mass);  
    p += 16;  
    mor3 = spu_mul(mor3, phii);  
  
    phi = spu_sub(phi, phii);  
    ax = spu_madd(mor3, dx, ax);  
    ay = spu_madd(mor3, dy, ay);  
    az = spu_madd(mor3, dz, az);  
}
```