GRAPE-8 — An Accelerator for Gravitational \( N \)-body Simulation with 20.5Gflops/W Performance

Jun Makino, TiTech,
Hiroshi Daisaka, Hitotsubashi Univ.

SC12 Nov 15, 2012
Summary

- GRAPE-8 is a specialized hardwired processor for low-precision calculation of gravitational interaction between particles.
- eASIC N2X740 device is used. Around 300K USD initial cost.
- Board level: 20.5 Gflops/W.
- Expect to achieve 2-3 times better than best GPU treecode, or 10 times better than best CPU treecode, per single card.
- Structured ASIC is a very promising way to achieve extremely high performance-per-watt.
Talk structure

- Brief history of GRAPE project
- GRAPE-DR
- How many bits you need?
- Design of GRAPE-8 chip
- Measured performance and power consumption
- Summary
Brief history of GRAPE project

- Basic Concept
- GRAPE-1 to 6
- GRAPE-DR
Basic concept (As of 1988)

- With astrophysical $N$-body simulation, almost all calculation goes to the calculation of particle-particle interaction.
- This is true even for fast $O(N \log N)$ or $O(N)$ schemes.
- A pipelined hardware which calculates the particle-particle interaction can accelerate overall calculation.

![Diagram showing Host Computer and GRAPE connected for time integration and interaction calculation.](image-url)
GRAPE-1 to GRAPE-6

GRAPE-1: 1989, 308Mflops

GRAPE-3: 1991, 14Gflops

GRAPE-4: 1995, 1.08Tflops

GRAPE-6: 2002, 64Tflops
Since 1995 (GRAPE-4), GRAPE has been faster than general-purpose computers.

Development cost was around 1/100. Performance per Watt was around 100 times better.
GRAPE-DR: Why and what?

- Chip development cost becomes too high.

<table>
<thead>
<tr>
<th>Year</th>
<th>Machine</th>
<th>Chip initial cost</th>
<th>process</th>
</tr>
</thead>
<tbody>
<tr>
<td>1992</td>
<td>GRAPE-4</td>
<td>200K$</td>
<td>1µm</td>
</tr>
<tr>
<td>1997</td>
<td>GRAPE-6</td>
<td>1M$</td>
<td>250nm</td>
</tr>
<tr>
<td>2004</td>
<td>GRAPE-DR</td>
<td>4M$</td>
<td>90nm</td>
</tr>
<tr>
<td>2012?</td>
<td>GDR2?</td>
<td>&gt; 10M$</td>
<td>28nm?</td>
</tr>
</tbody>
</table>

How we can continue?

Widen application to justify the initial cost.
Processor architecture

- Float Mult
- Float add/sub
- Integer ALU
- 32-word registers
- 256-word memory
- communication port
GRAPE-DR cluster system

(As far as I know) Only processor designed in academia listed in Top500 in the last 10 years.
Transistor count

<table>
<thead>
<tr>
<th>Chip</th>
<th>Total Trs</th>
<th>SP operations</th>
<th>Trs/op</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>GRAPE-3</td>
<td>100K</td>
<td>30</td>
<td>3k</td>
<td>0.15</td>
</tr>
<tr>
<td>GRAPE-6</td>
<td>8M</td>
<td>400</td>
<td>20k</td>
<td>1</td>
</tr>
<tr>
<td>GRAPE-DR</td>
<td>200M</td>
<td>1024</td>
<td>200k</td>
<td>10</td>
</tr>
<tr>
<td>NVIDIA GK110</td>
<td>7.1B</td>
<td>4992</td>
<td>1.4M</td>
<td>70</td>
</tr>
<tr>
<td>Intel Ivy Bridge</td>
<td>1.4B</td>
<td>64</td>
<td>22M</td>
<td>1100</td>
</tr>
</tbody>
</table>

- GRAPE DR Much better than GPU/CPU
- Even so 10 times more trs/flop compared to G6
- High initial cost = long development cycle ...
Alternative approach

Basic idea: Design highly optimized hardware for low-precision calculation

- GRAPE-3 used **five-bit mantissa** for pairwize force, which was sufficient many applications
- As a result, it requires 3-4 orders of magnitude less transistors compared to GPUs or CPUs.
- For such low-accuracy pipelined hardware, we can use somewhat inefficient technologies, if their initial costs are small.
- FPGAs, Structured ASIC, Older technologies (e.g. 130 or 90 nm)

For GRAPE-8, we used Structured ASIC
Structured ASIC

Something like single-metal-programmable FPGA.

<table>
<thead>
<tr>
<th></th>
<th>ASIC</th>
<th>FPGA</th>
<th>Str. A.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial cost (M USD)</td>
<td>10</td>
<td>0.01</td>
<td>0.2</td>
</tr>
<tr>
<td>Gate count (M gates)</td>
<td>200</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Gate cost (USD/M gates)</td>
<td>3</td>
<td>100</td>
<td>20</td>
</tr>
</tbody>
</table>

(Estimate as of 2009)

Lower initial cost than full-custom ASIC, lower per-gate cost than FPGA

(Higher initial cost than FPGA, higher per-gate cost than ASIC...)

For the total budget in between 0.3 and 10M USD, it is “Best solution”
How many bits you need?

- In cosmological calculations or simulations of galaxies, the accuracy of the result is determined by the number of particles.
- There is $O(1/\sqrt{N})$ error (particle noise), independent of the error in the pairwise force.
- To model thin disks, we need a few more bits.
- We handle close collisions separately (Particle-Particle Particle-Tree method).
Particle-Particle Particle-Tree

\[ H = H_{\text{Hard}} + H_{\text{Soft}}, \]

\[ H_{\text{Hard}} = \sum_{i=1}^{N} \left( \frac{p_i^2}{2m_i} - \frac{Gm_im_\odot}{r_i} \right) - \sum_{i<j}^{N} \frac{Gm_im_j}{r_{ij}} W(r_{ij}), \]

\[ H_{\text{Soft}} = \sum_{i<j}^{N} \frac{Gm_im_j}{r_{ij}} (1 - W(r_{ij})), \]

- Divide interaction to near and far terms, as in P$^3$M
- Far term is handled with Barnes-Hut tree (and low-accuracy hardware), and near term with high-accuracy direct integrator
P$^3$T Test calculation

10$^4$ planetecimals in Earth region

Error is due to the approximation in the tree algorithm
Very high accuracy is achieved
We can use this scheme also for star clusters and other systems
Design of GRAPE-8 chip

- Use eASIC’s N2X740 device (740K logic elements, approx 7M gates)
- Integrate everything except for PCIe interface (not available on N2X)
- 9-bit mantissa for pairwise force, 32-bit fixed point for coordinates, and 64-bit fixed point for accumulation of force
- Designed so that calculation and sending to/from the chip can be fully overlapped.
Chip top level structure

- Interface unit
- one “j-particle” memory (memory to store particles which exert forces)
- six pipeline blocks (each with 8 pipelines)
The Pipeline

Calculates:

\[ a_i = \sum_j Gm_j g \left( \frac{r_{ij}}{r_0} \right) \frac{r_{ij}}{(r_{ij}^2 + \epsilon_i^2 + \epsilon_j^2)^{3/2}} \]

\[ \phi_i = \sum_j Gm_j g \left( \frac{r_{ij}}{r_0} \right) \frac{1}{(r_{ij}^2 + \epsilon_i^2 + \epsilon_j^2)^{1/2}} \]

- Arbitrary cutoff function \( g \)
- 9-bit mantissa for pairwize force, 32-bit fixed point for coordinates, and 64-bit fixed point for accumulation of force
GRAPE-8 board

Two GRAPE-8 chips and one FPGA for PCIe interface
960Gflops peak, 46W (G8 chips: 26W, FPGA: 20W)

20.5Gflops/W
Performance

- basic communication
- actual computation
basic communication

- 50-60% of the theoretical peak of PCIe (x8, Gen 1)
- enough to achieve > 50% of peak for 1k-2k particles (Important for treecode)
Calculation performance

More than 50% of peak for $N = 2048$

Expect to achieve 5M particle steps/sec (pss) with treecode (We are still working on this...)

Bonsai GPU code on Fermi: 2M pss/card
GreeM on K: 0.5M pss/node
Breakdown of calculation time

force: calculation + sending back
IP: sending particles which receive forces
JP: sending particles which exert forces

Communication speed better than 60% of the raw speed
Comparison with other approaches

<table>
<thead>
<tr>
<th></th>
<th>FPGA</th>
<th>GPU</th>
<th>GRAPE-8</th>
</tr>
</thead>
<tbody>
<tr>
<td>hardware cost</td>
<td>high</td>
<td>lowest</td>
<td>low</td>
</tr>
<tr>
<td>Power consumption</td>
<td>moderate</td>
<td>high</td>
<td>low</td>
</tr>
<tr>
<td>Flexibility</td>
<td>moderate</td>
<td>high</td>
<td>low</td>
</tr>
</tbody>
</table>

If you know exactly what you want, literally to every bits, Structured ASIC is an attractive solution.
Summary

- GRAPE-8 is a specialized hardwired processor for low-precision calculation of gravitational interaction between particles
- eASIC N2X740 device is used. around 300K USD initial cost
- Chip performance: 480Gflops, 13W (board 960GF, 46W)
- Board level: 20.5 Gflops/W
- Expect to achieve 2-3 times better than best GPU treecode, or 10 times better than best CPU treecode, per single card.
- Structured ASIC is a very promising way to achieve extremely high performance-per-watt.