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**
- **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.**

Talk structure

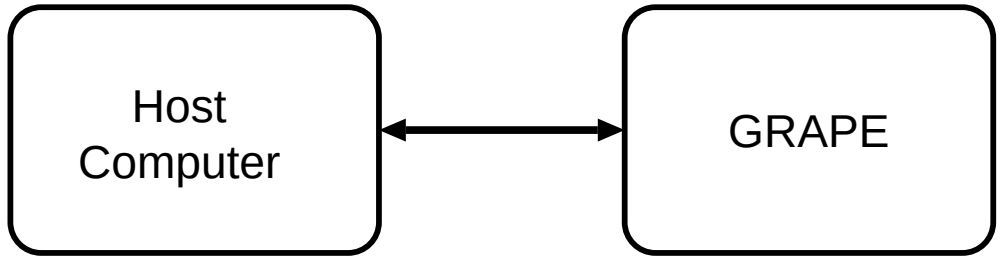
- Brief history of GRAPE project
- GRAPE-DR
- How many bits you need?
- Design of GRAPE-8 chip
- Measured performance and power consumption
- Summary

Brief history of GRAPE project

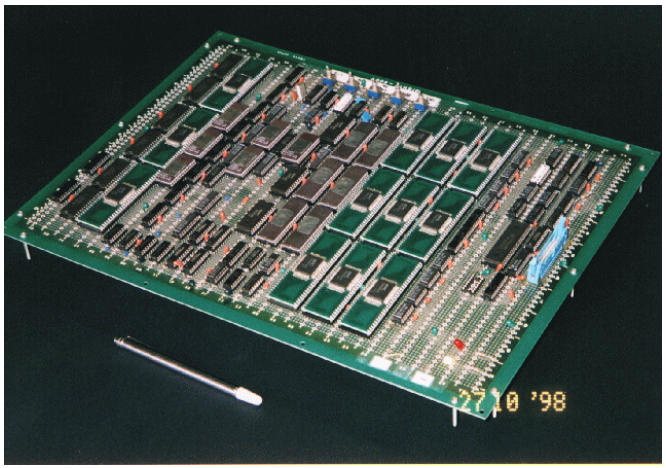
- **Basic Concept**
- **GRAPE-1 to 6**
- **GRAPE-DR**

Basic concept (As of 1988)

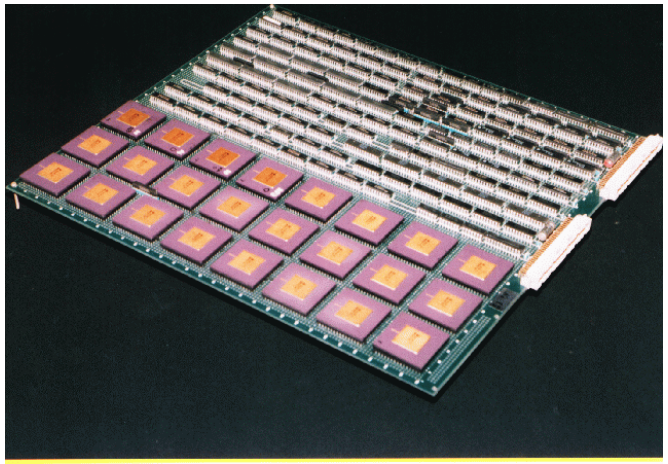
- **With astrophysical N-body simulation, almost all calculation goes to the calculation of particle-particle interaction.**
- **This is true even for fast $O(N \log N)$ or $O(N)$ schemes**
- **A pipelined hardware which calculates the particle-particle interaction can accelerate overall calculation.**
- **Original Idea: Chikada (1988)**



GRAPE-1 to GRAPE-6

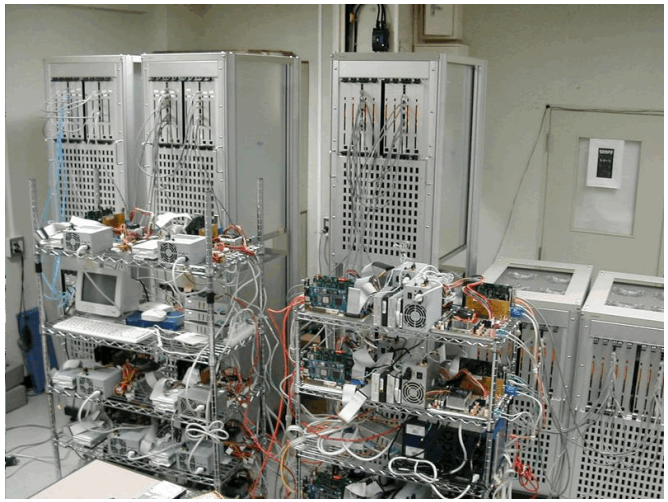


GRAPE-1: 1989, 308Mflops



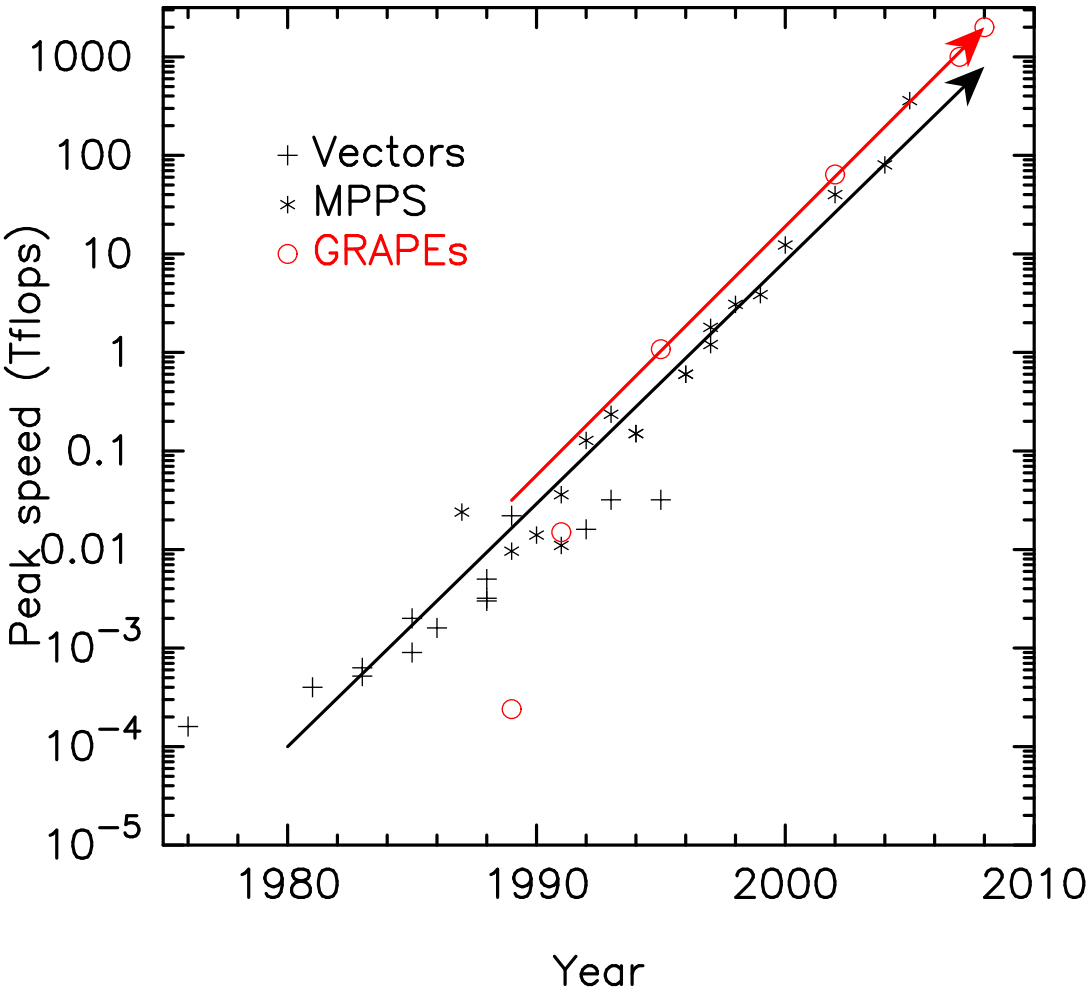
GRAPE-3: 1991, 14Gflops

GRAPE-4: 1995, 1.08Tflops



GRAPE-6: 2002, 64Tflops

Performance history



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?

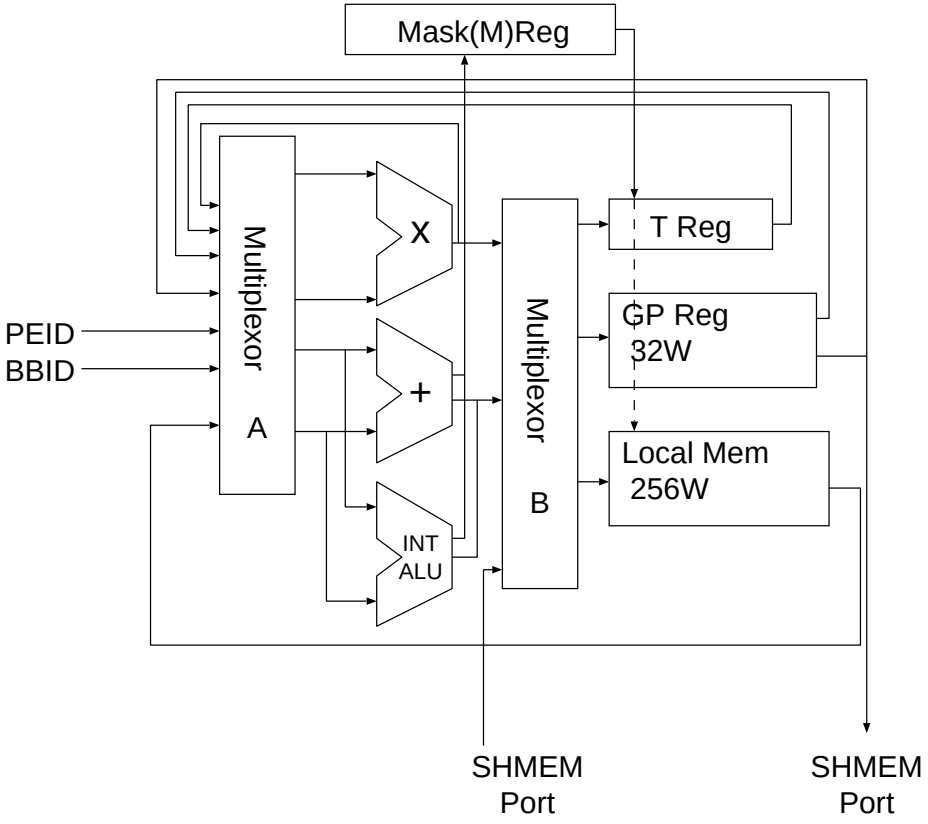
- Chip development cost becomes too high.

Year	Machine	Chip initial cost	process
1992	GRAPE-4	200K$	1μm
1997	GRAPE-6	1M$	250nm
2004	GRAPE-DR	4M$	90nm
2012?	GDR2?	$>$ 10M$	28nm?

How we can continue?

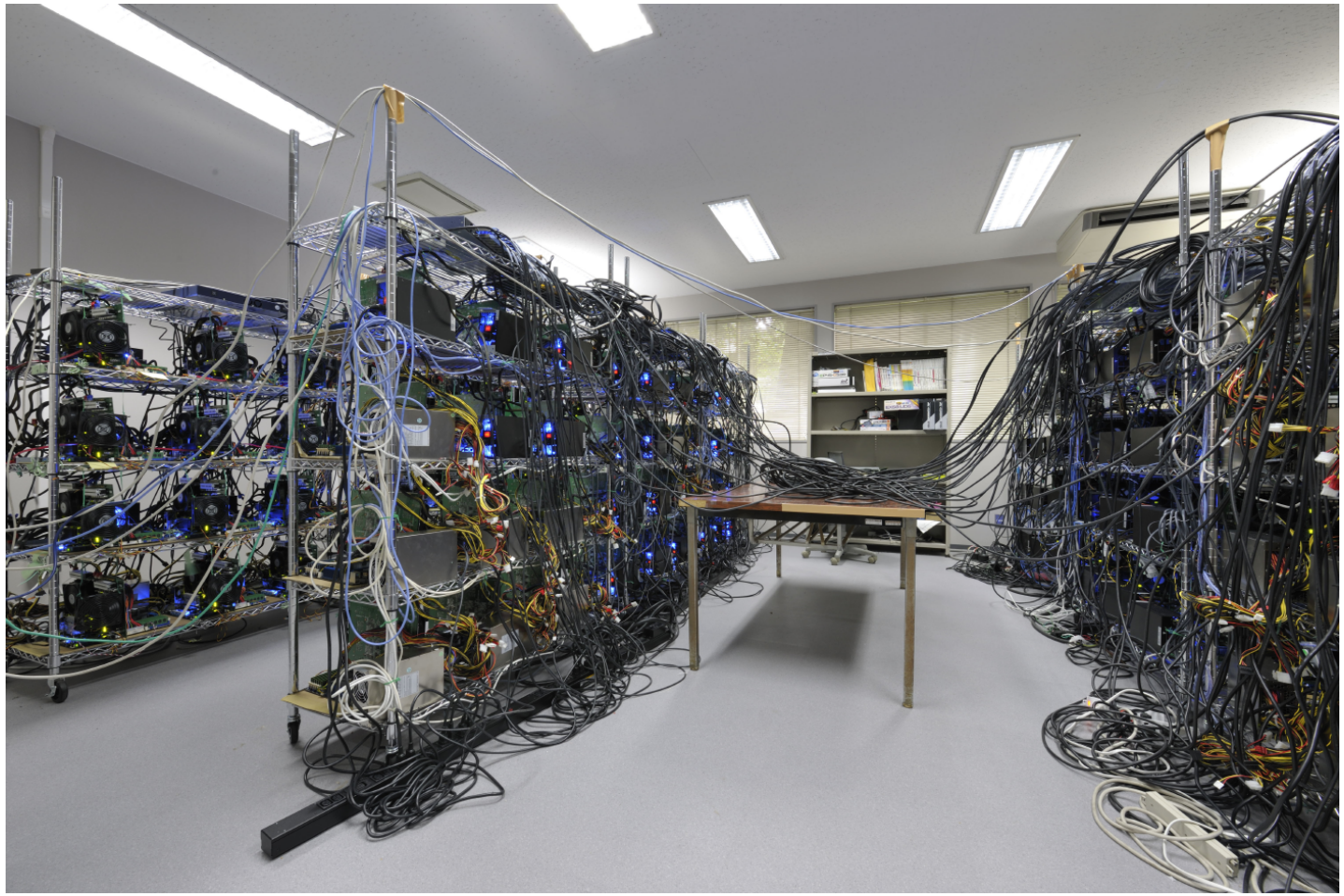
Widen application to justify the initial cost.

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

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

- 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.

	ASIC	FPGA	Str. A.
Initial cost (M USD)	10	0.01	0.2
Gate count(M gates)	200	10	10
Gate cost(USD/M gates)	3	100	20

(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 pairwize 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

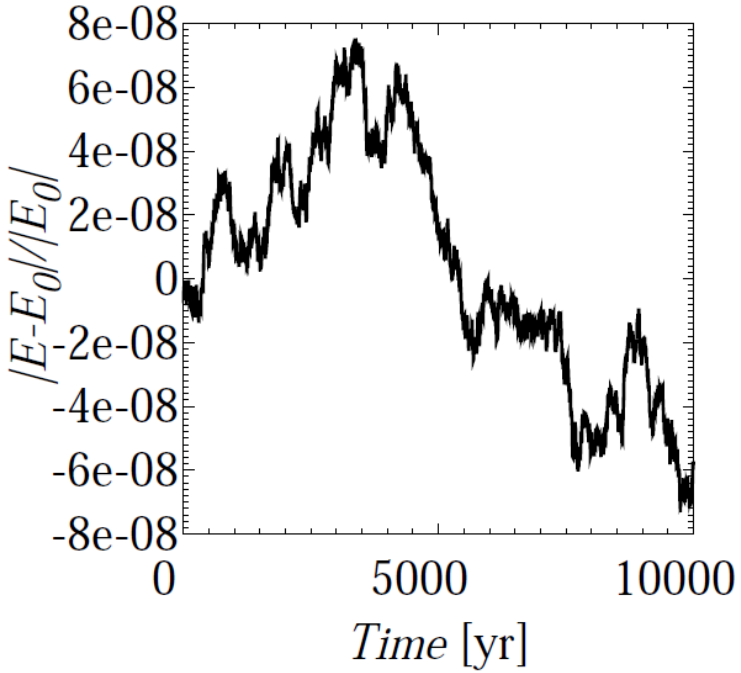
$$H = H_{Hard} + H_{Soft},$$

$$H_{Hard} = \sum_{i=1}^{N} \left(\frac{p_i^2}{2m_i} - \frac{G m_i m_\odot}{r_i} \right) - \sum_{i<j}^{N} \frac{G m_i m_j}{r_{ij}} W(r_{ij}),$$

$$H_{Soft} = \sum_{i<j}^{N} \frac{G m_i m_j}{r_{ij}} (1 - W(r_{ij})),$$

- **Divide interaction to near and far terms, as in P^3M**
- **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



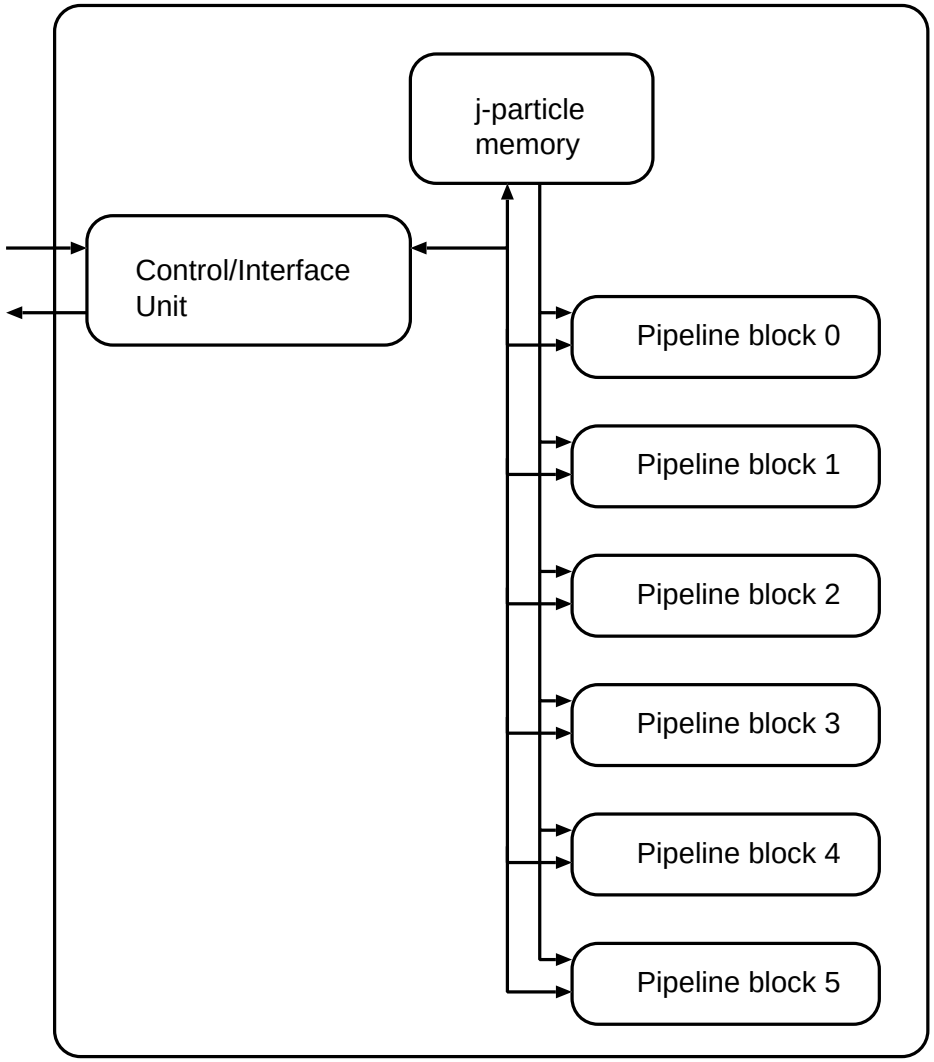
10^4 planetecimals in Earth region

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 pairwize 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

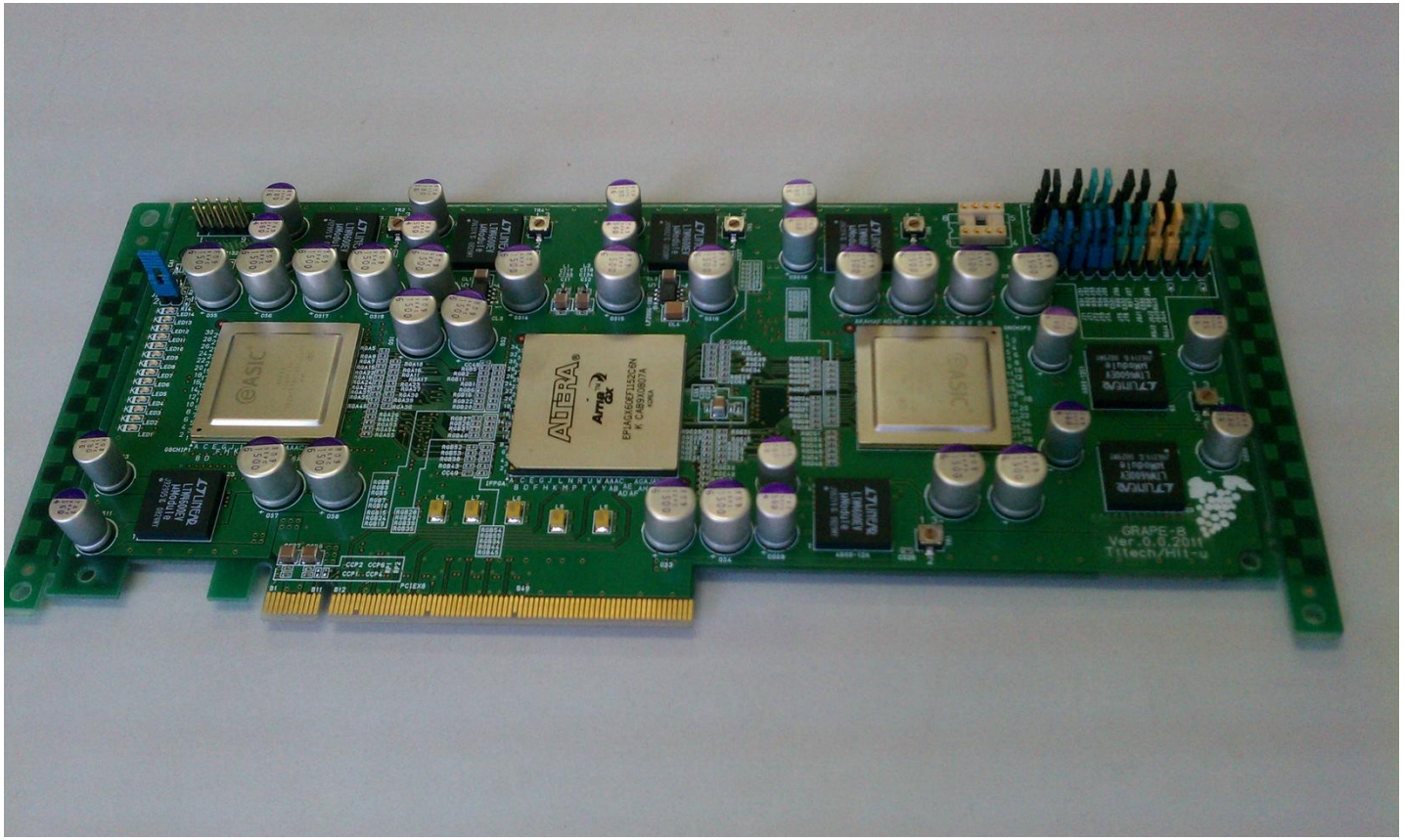
Calculates:

$$\mathbf{a}_i = \sum_j G m_j g(r_{ij}/r_0) \frac{\mathbf{r}_{ij}}{(r_{ij}^2 + \epsilon_i^2 + \epsilon_j^2)^{3/2}}$$

$$\phi_i = \sum_j G m_j g(r_{ij}/r_0) \frac{1}{(r_{ij}^2 + \epsilon_i^2 + \epsilon_j^2)^{1/2}}$$

- **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



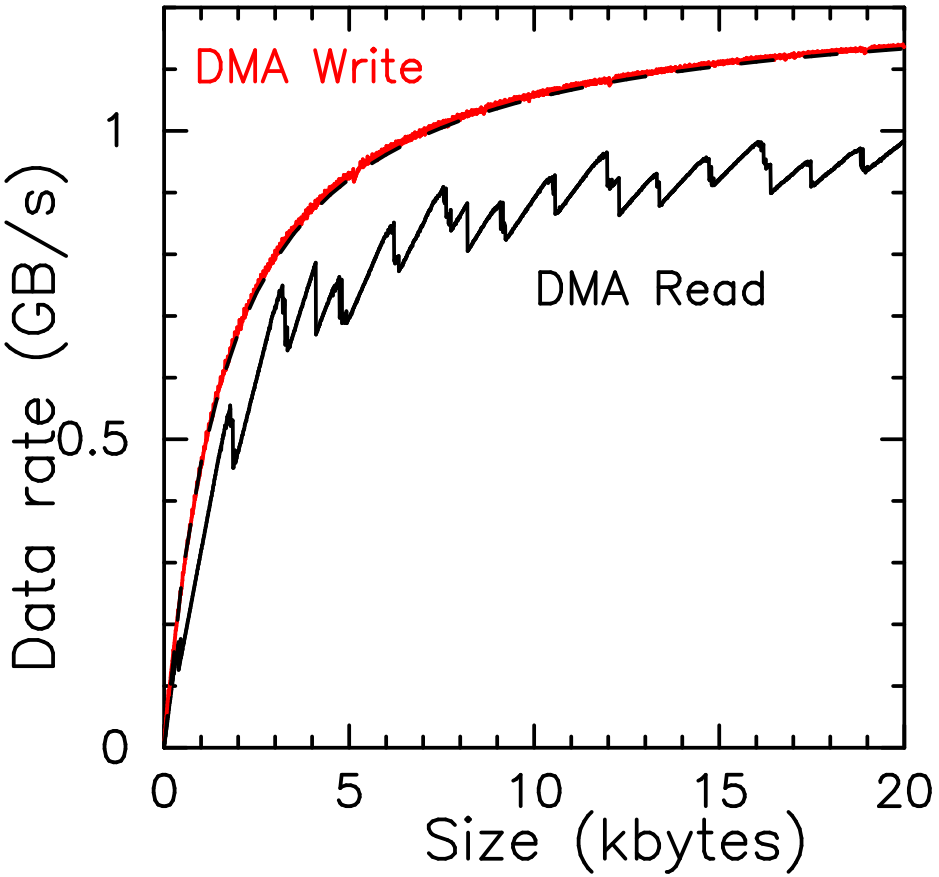
Two GRAPE-8 chips and one FPGA for PCIe interface
960Gflops peak, 46W (G8 chips:26W, FPGA: 20W)

20.5Gflops/W

Performance

- **basic communication**
- **actual computation**

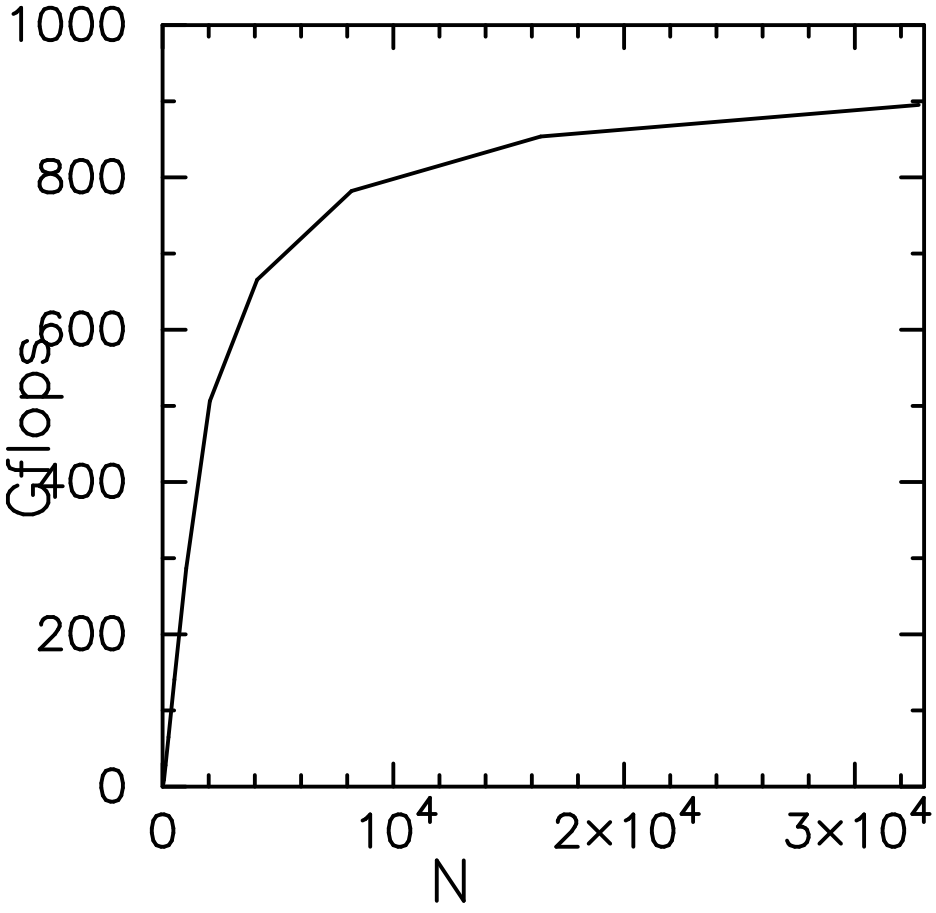
basic communication



- 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

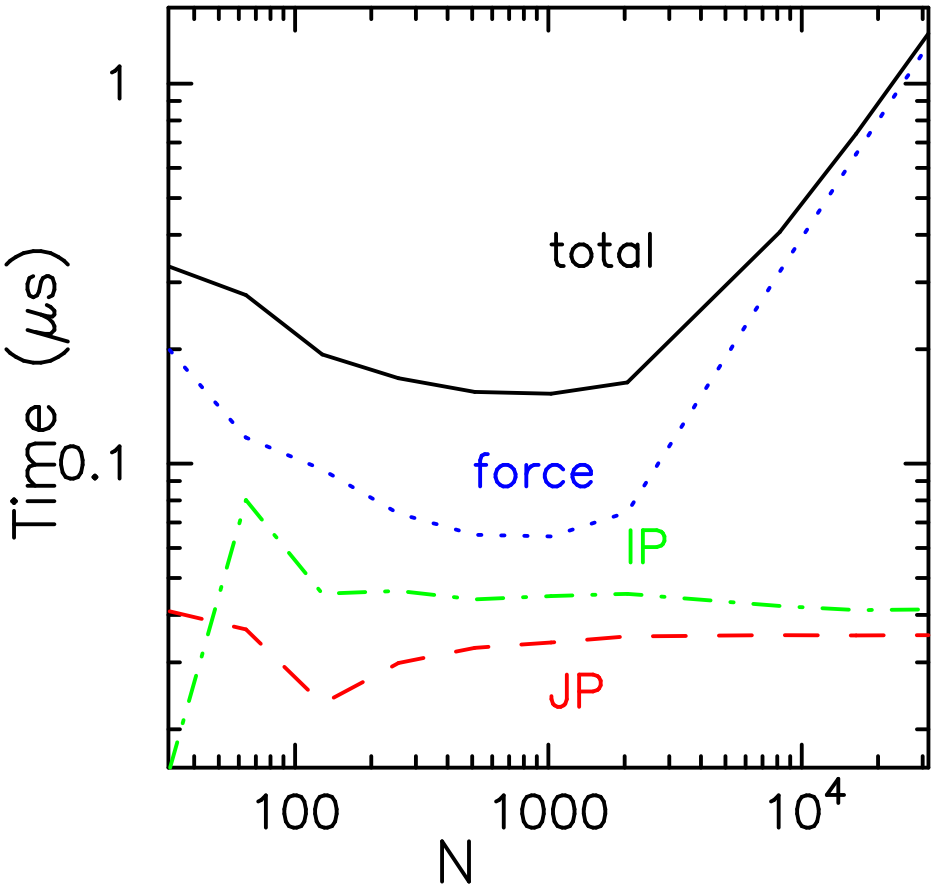
More than 50% of peak for $N = 2048$



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

Bonsai GPU code on Fermi:d 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

	FPGA	GPU	GRAPE-8
hardware cost	high	lowest	low
Power consumption	moderate	high	low
Flexibility	moderate	high	low

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.**