Accelerated Computing

Jun Makino
Interactive Research Center of Science
Tokyo Institute of Technology

IEEE Cluster 2011 Sept 28, 2011

Satoshi's three questions

1. **What does accelerated computing solve best?**
   — (Physics) simulation and visualization.

2. **What do you see as the current biggest technical challenge?**
   — Initial design cost of large LSIs, which effectively terminated academic efforts to make competitive chips.

   Research in academia is essential for the advance of the field.

3. **Do you believe that accelerated computing will become mainstream?**
   — In HPC, and in graphics, yes.

Talk structure

- **Brief history of GRAPE/GRAPE-DR**
- **Future of accelerators**

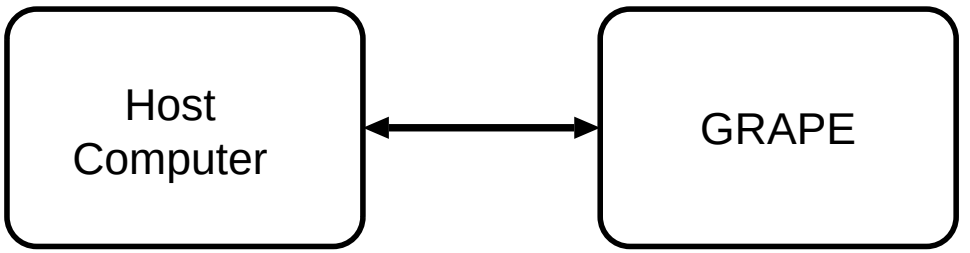
BTW, 6 times, not 7 (95, 96, 99, 00, 01, 03).

Brief history of GRAPE(-DR)

- Basic concept
- GRAPE-1 through 6
- GRAPE-DR

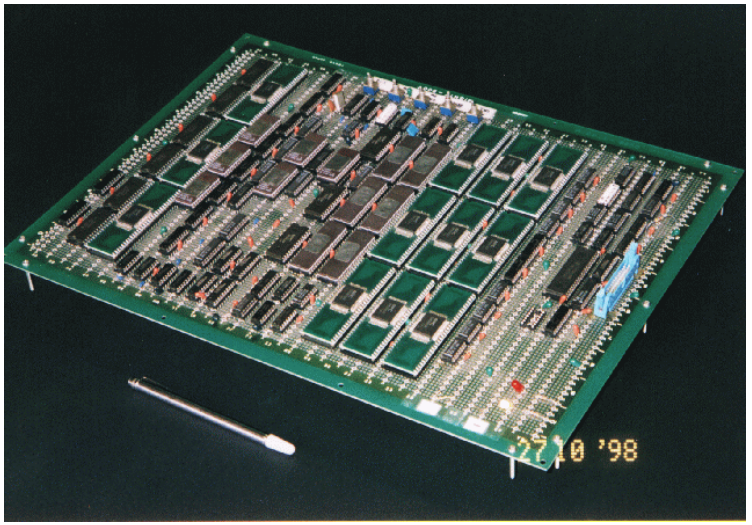
Basic concept (As of 1988)

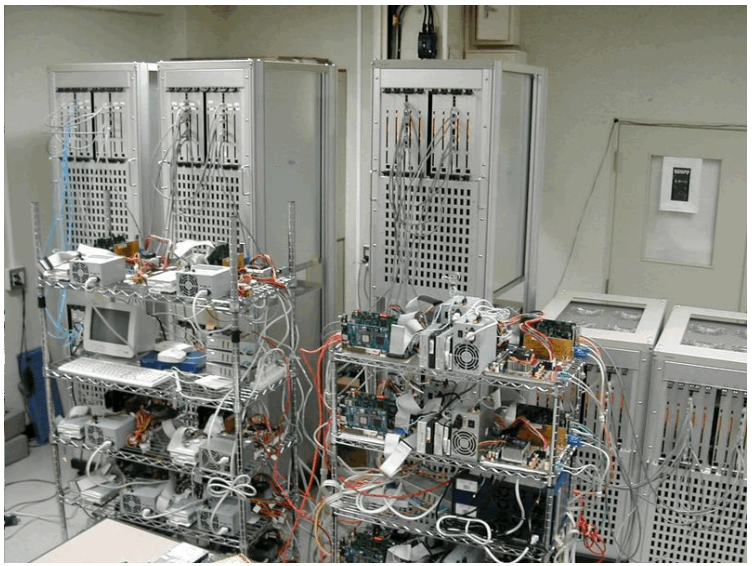
- **With astrophysical N-body simulation, almost all calculation goes to the calculation of particle-particle interaction.**
- **This is true even for fast $O(N \log N)$ or $O(N)$ schemes**
- **A pipelined hardware which calculates the particle-particle interaction can accelerate overall calculation.**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



Accelerated Computing two decades ago

GRAPE-1 to GRAPE-6



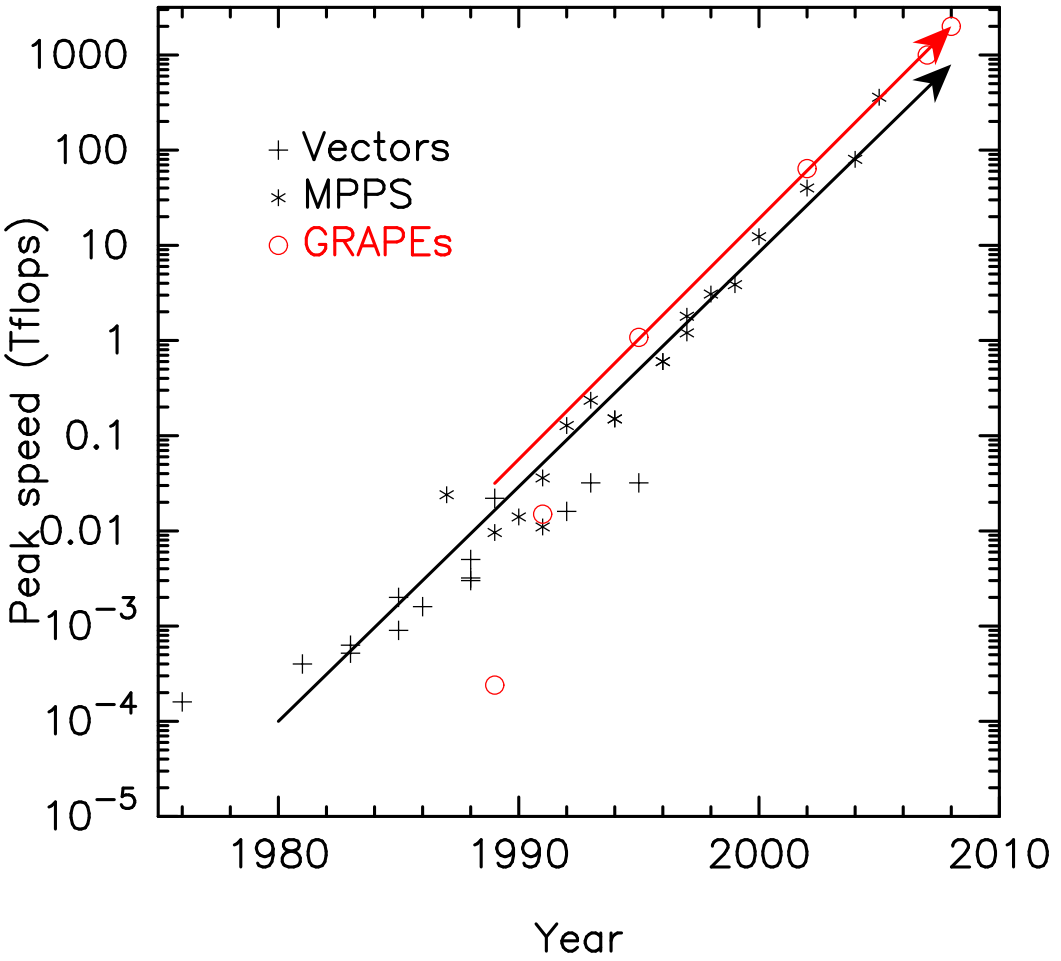


GRAPE-1: 1989, 308Mflops

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.

GRAPE-DR: Why and what?

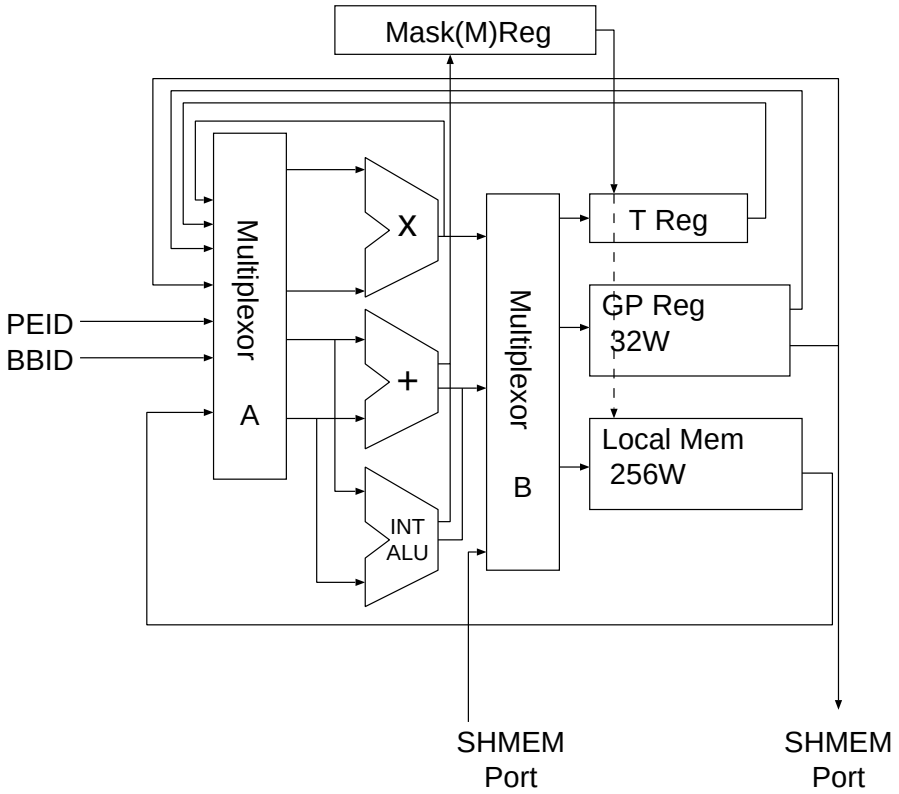
- Chip development cost becomes too high.

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?

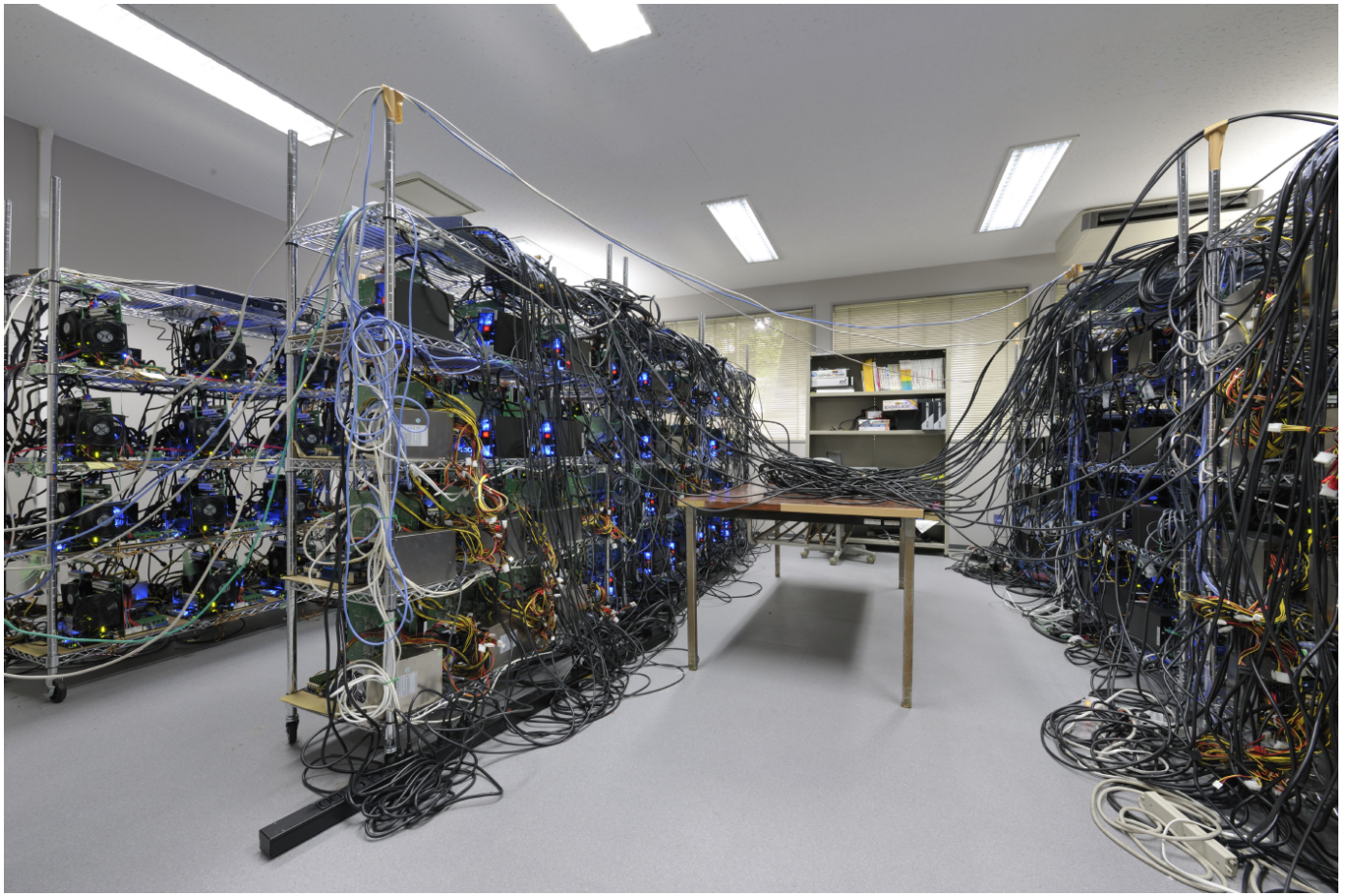
Widen application to justify the initial cost.

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.

Little Green 500, June 2010

Green500 Rank	MFLOPS/W	Site*	Computer*	Total Power (kW)
1	815.43	National Astronomical Observatory of Japan	GRAPE-DR accelerator Cluster, Infiniband	28.67
2	773.38	Forschungszentrum Juelich (FZJ)	QPACE SFB TR Cluster, PowerXCell 8i, 3.2 GHz, 3D-Torus	57.54
2	773.38	Universitaet Regensburg	QPACE SFB TR Cluster, PowerXCell 8i, 3.2 GHz, 3D-Torus	57.54
2	773.38	Universitaet Wuppertal	QPACE SFB TR Cluster, PowerXCell 8i, 3.2 GHz, 3D-Torus	57.54
5	536.24	Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw	BladeCenter QS22 Cluster, PowerXCell 8i 4.0 Ghz, Infiniband	34.63

#1: GRAPE-DR, #2: QPACE: German QCD machine
#9: NVIDIA Fermi

Transistor count

Chip	Total Trs	SP operations	Trs/op	Ratio
GRAPE-6	8M	400	20k	1
GRAPE-DR	200M	1024	200k	10
Intel LRB	1.7B	1024?	1.7M?	85?
Intel Westmere	1.2B	48	25M	1250

Not all accelerators are created equal

Several important dimensions:

- **Application-specific/General-purpose**
- **SIMD/MIMD**
- **Local Memory/Cache**
- **Narrow/Wide external memory**

My view

One should go for: Application-specific, SIMD, Local Memory, and Narrow external memory.

Reason: Gives you the best performance

- **for the same number of transistors**
- **for the same power consumption**

It can beat other approaches by two orders of magnitudes.
But: initial cost has become too high. Need some compromisation.

Solution?

- **FPGA? — Always "the future of HPC" for the last quarter century**
- **Structured ASIC? — Maybe, if manufacturers survive.**
- **Programmable SIMD processor? — Yes, if you can raise fund.**

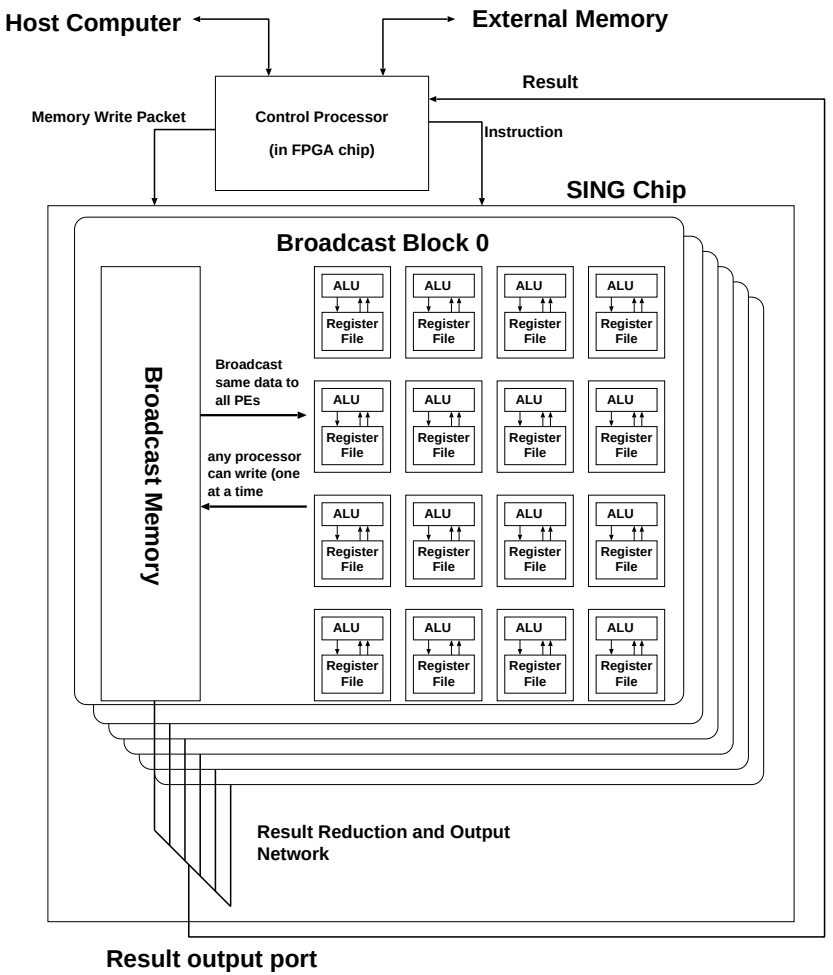
GRAPE-DR is the only processor designed in academia which appeared in Top 500 list in the last 10 years (Tsukuba CP-PACS in 1996)

Comparison with other works

(From Nov 10 Top 500 list)

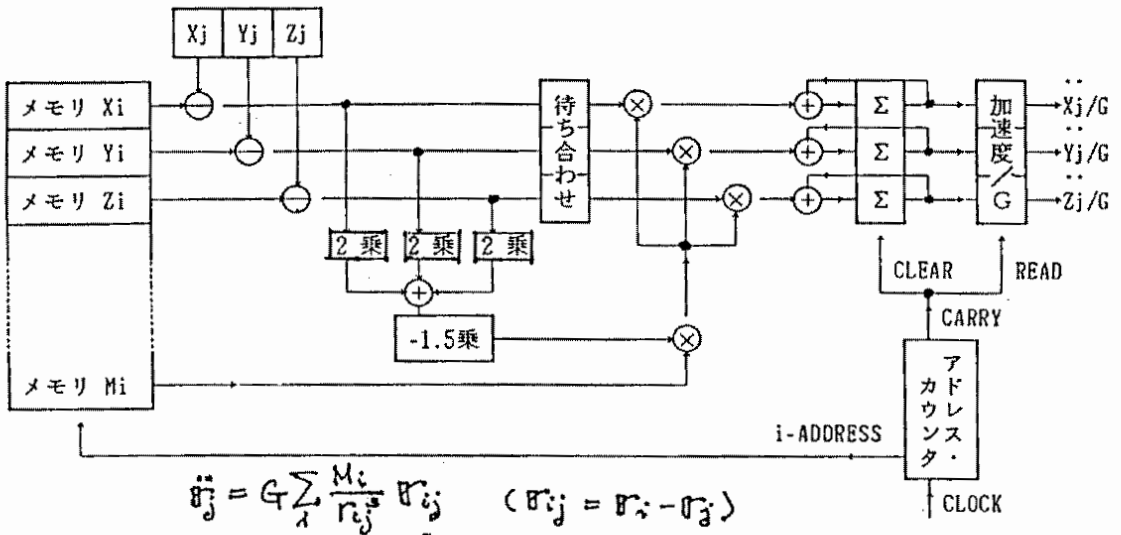
Accelerator /System	CPU /Clock	Performance /Efficiency	Acceleration over host
Fermi Tianhe-1A	Xeon 6c 2.93GHz	2.566PF 54.4%	2.83
Fermi Tsubame 2.0	Xeon 6c 2.93(3.19?) GHz	1.192PF 53.5%	6.13
GRAPE-DR	Core i7 4c 3GHz	37.4TF 53.2%	10.6

Chip architecture



- 32 PEs organized to "broadcast block" (BB)
- BB has shared memory. Various reduction operation can be applied to the output from BBs using reduction tree.
- Input data is broadcasted to all BBs.
- "Solved" data movement problem: Very small number of long wires and off-chip IO.

Chikada's idea (1988)



$$\ddot{\vec{r}}_j = G\sum_i \frac{M_i}{r_{ij}^3}\vec{r}_{ij} \qquad (\vec{r}_{ij} = \vec{r}_i - \vec{r}_j)$$

+, −, ×, 2乗は1 operation, -1.5乗は多項式近似でやるとして10operation 位に相当する.
総計24operation.
各operation の後にはレジスタがあって, 全体がpipelineになっているものとする.
「待ち合わせ」は2乗してMと掛け算する間の時間ズレを補正するためのFIFO(First-In First-Out memory).
「Σ」は足し込み用のレジスタ. N回足した後結果を右のレジスタに転送する.

図2. N体問題のj-体に働く重力加速度を計算する回路の概念図.

- Hardwired pipeline for force calculation (similar to Delft DMDP)
- Hybrid Architecture (things other than force calculation done elsewhere)