Applications of MN-Core Deep-Learning processor to large-scale astronomical simulations

Jun Makino
Department of Planetology, Kobe University

September 15, 2019, Challenges and Innovations in Computational Astrophysics, Saint Petersburg

Summary

- **MN-Core/GRAPE-PFN2 is accelerator processor designed mainly for deep learning.**
- **512 Tops/132TF/32TF for half, short and long word**
- **MN-Core/GRAPE-PFN2 will be ready "soon"**
- **We need to make use of its matrix multiplication unit to achieve high performance.**
- **We can use it for gravity (FMM), particle-based hydro (MLS) and possibly for other applications as well.**

Talk structure

- **What is MN-Core/GRAPE-PFN2?**
- **Past history, current status, and future plan.**
- **Overview of GRAPE-PFN2 (GPFN2)**
- **Some architecture details**
- **How MN-Core can be used for Astrophysical computing**
- **Summary**

What is MN-Core/GRAPE-PFN2?

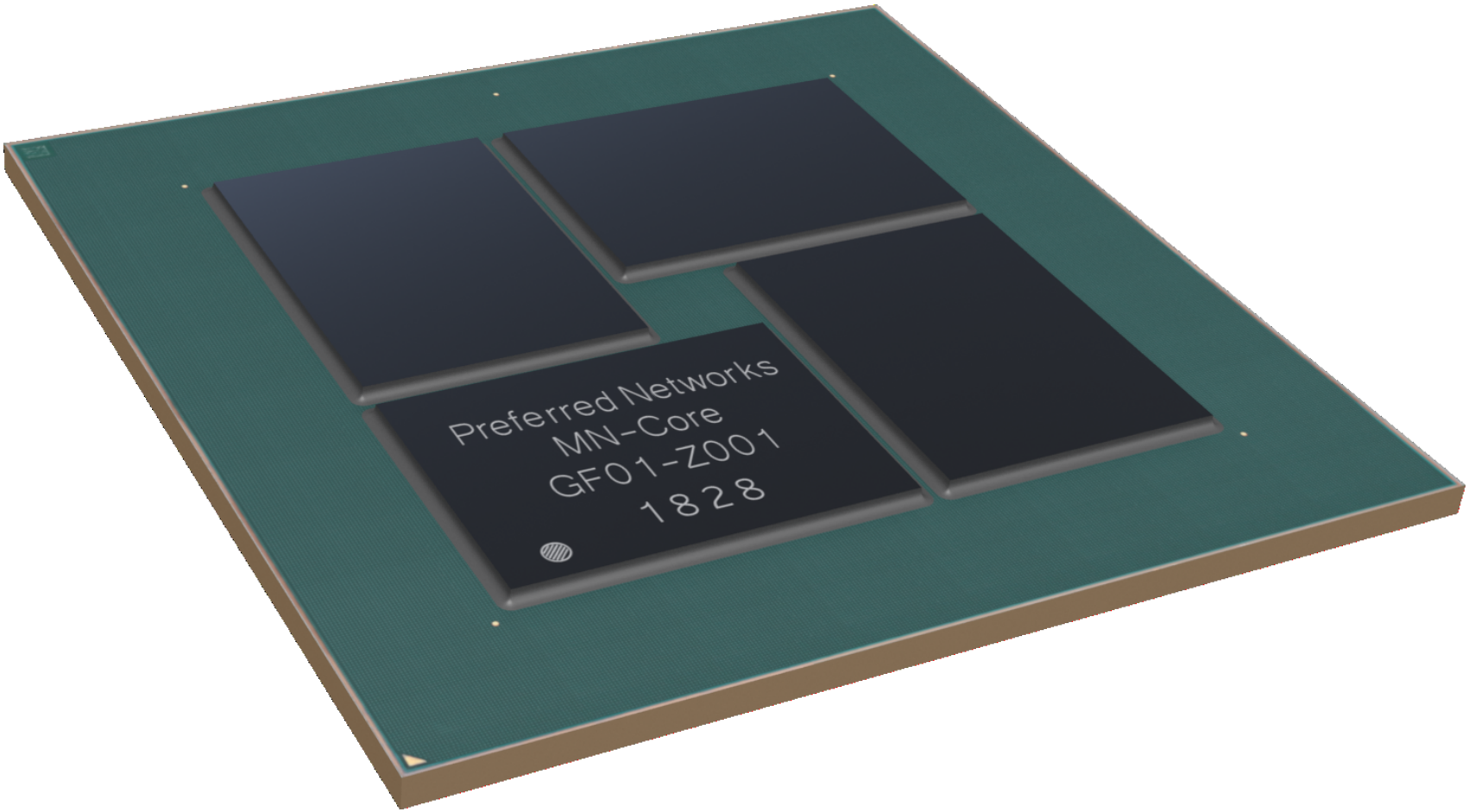
- The processor chip PFN (Preferred Networks) has been developing in collaboration with some of our group in RIKEN R-CCS/Kobe University.
- Goal: Highest performance and highest performance-per-watt for training DNNs (CNNs).
- Planned peak FP16(-equivalent) performance of single card: 524 Tops
- 131 and 33 TF for SP and DP operations

movie...

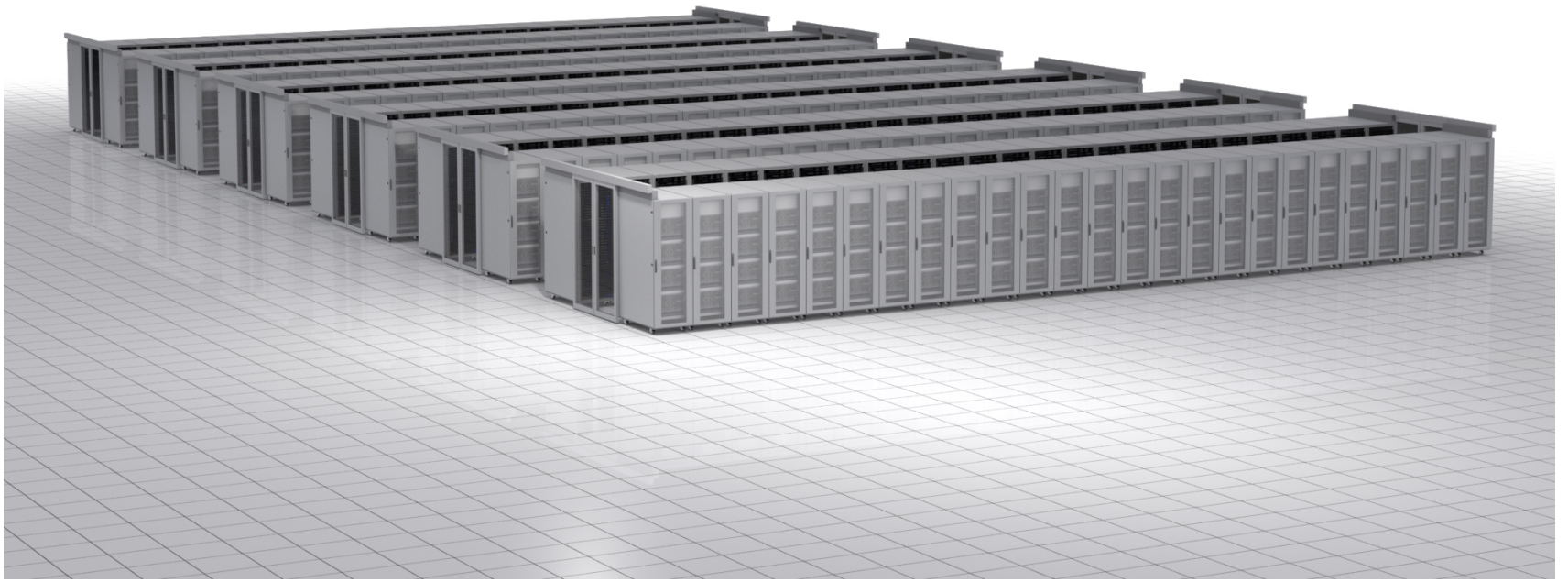
Past history, current status, and future plan.

- **Feb 2016: JM visited PFN at Hongo-3choume**
- **June 2016: Joint application to NEDO ("small" grant, 40MJYE/year $\times 2$) (PFN moved to Ote-machi)**
- **July 2016: PFN chip project started. Plan for two chips: GPFN1 by NEDO money (40nm, small chip), GPFN2 (12FFC, full-blown) by PFN internal money.**
- **2019 Evaluation of ES chips will be ...**
- **2020 "2EF" system (MN-3) will be ready.**
- **We have started the development of MN-Core2/GPFN3.**

MN-Core



MN-3

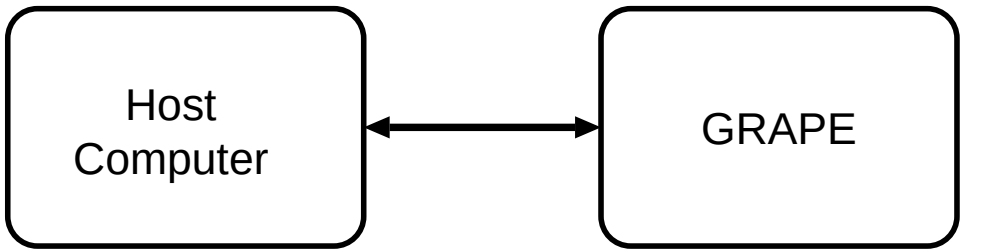


History before GRAPE-PFN

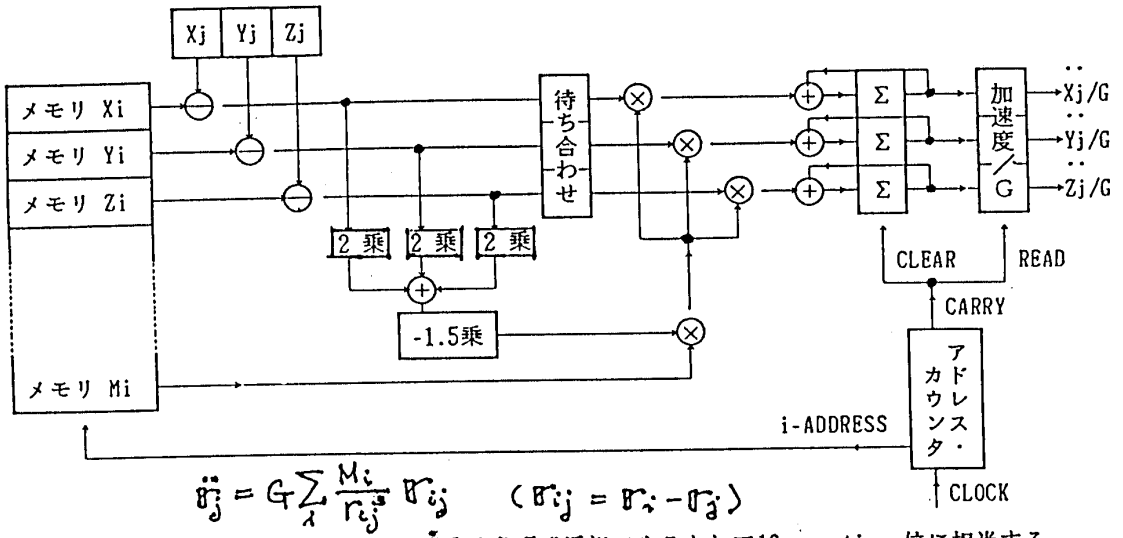
- **GRAPE (GRAvity PipE) processors**
- **GRAPE-DR**

Basic concept of GRAPE

- **With N-body simulation, almost all calculation goes to the calculation of particle-particle interaction.**
- **This is true even for schemes like Barnes-Hut treecode or FMM.**
- **A simple hardware which calculates the particle-particle interaction can accelerate overall calculation.**
- **Original Idea: Chikada (1988)**



Chikada's idea (1988)



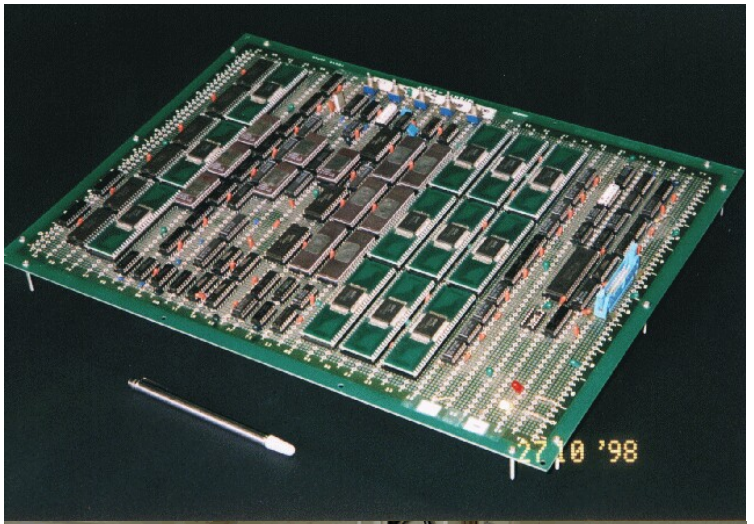
$$\ddot{\vec{r}}_j = G\sum_i \frac{M_i}{r_{ij}^3}\vec{r}_{ij} \qquad (\vec{r}_{ij} = \vec{r}_i - \vec{r}_j)$$

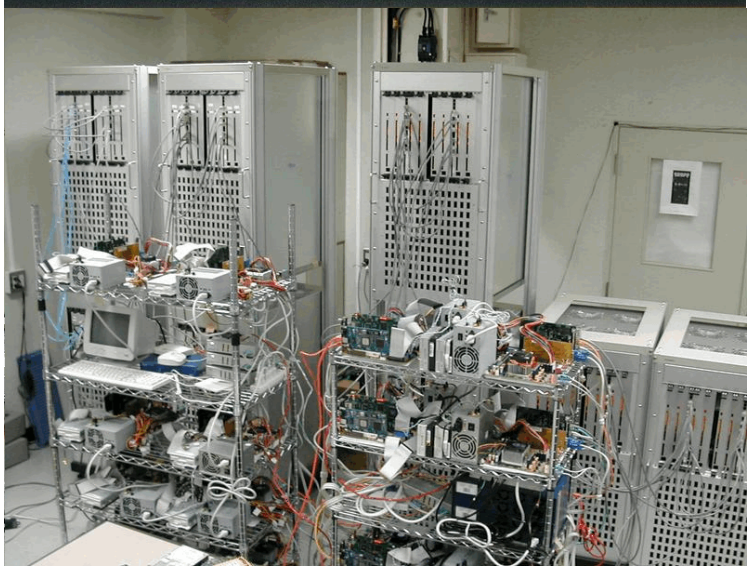
+, −, ×, 2乗は1 operation, -1.5乗は多項式近似でやるとして10operation 位に相当する.
総計24operation.
各operation の後にはレジスタがあって, 全体がpipelineになっているものとする.
「待ち合わせ」は2乗してMと掛け算する間の時間ズレを補正するためのFIFO(First-In First-Out memory).
「Σ」は足し込み用のレジスタ. N回足した後結果を右のレジスタに転送する.

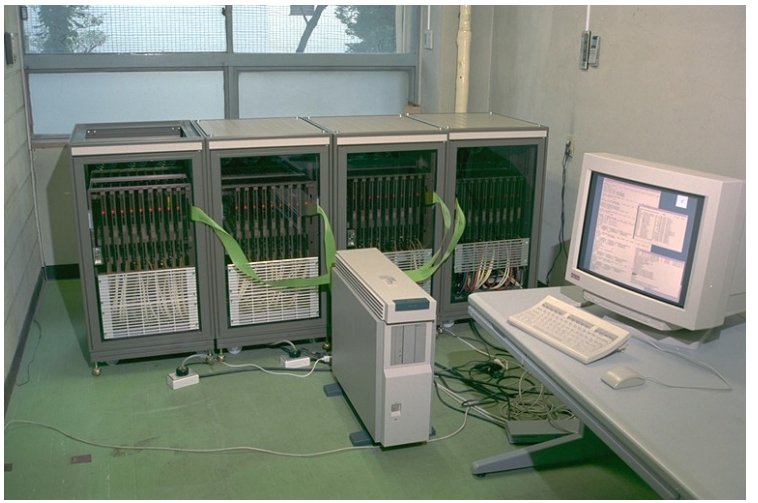
図2. N体問題のj-体に働く重力加速度を計算する回路の概念図.

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

GRAPE-1 to GRAPE-6







GRAPE-1: 1989, 308Mflops

GRAPE-4: 1995, 1.08Tflops

GRAPE-6: 2002, 64Tflops

From GRAPE-6 to GRAPE-DR

Chip development cost has become too high.

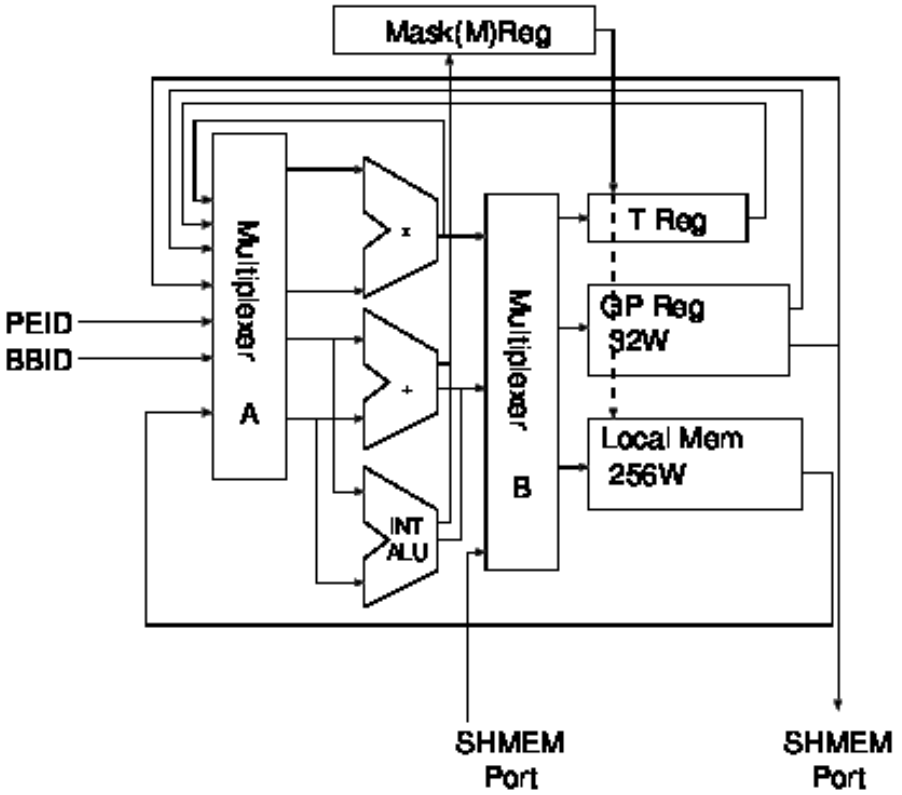
Year	Machine	Chip initial cost	process
1992	GRAPE-4	200K$	1μm
1997	GRAPE-6	1M$	250nm
2004	GRAPE-DR	2M$	90nm
(2019	?	$>$ 20M$	"7nm")

How to deal with high initial cost?

Several options:

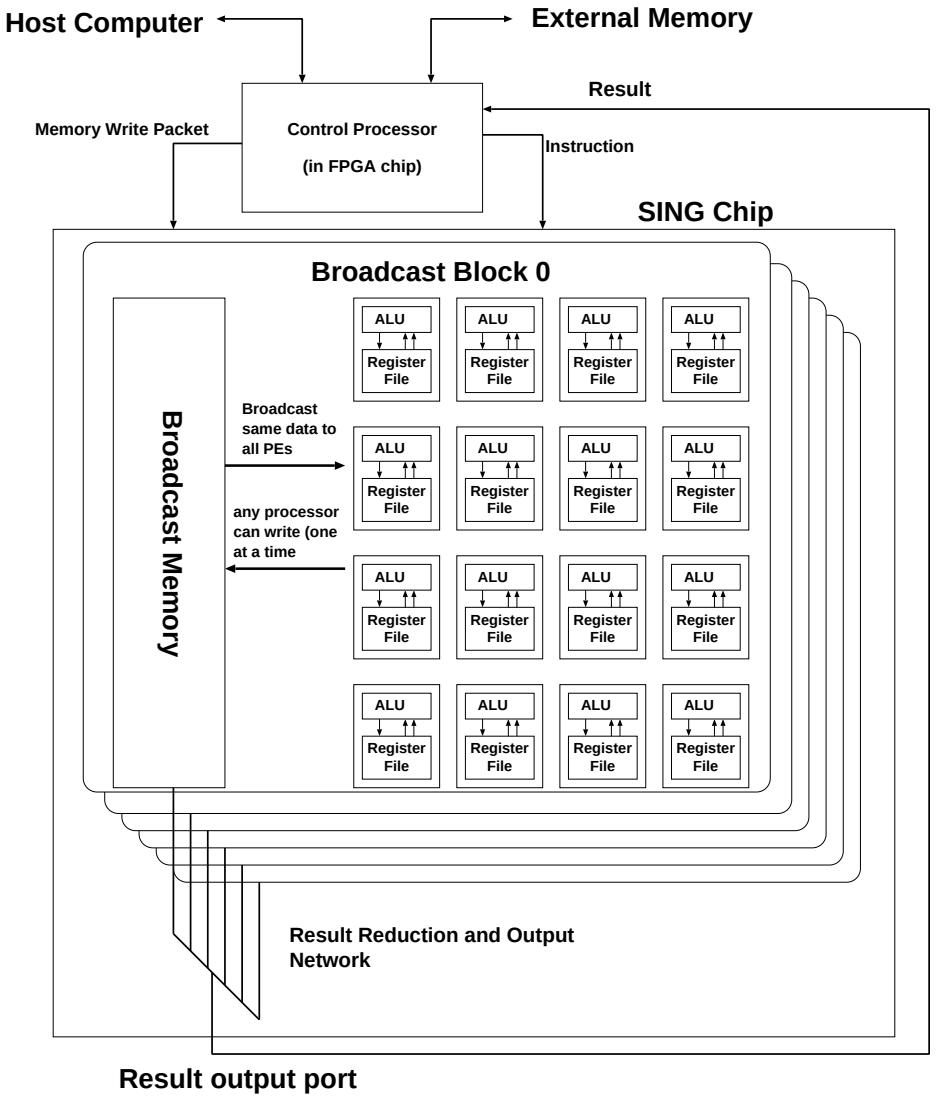
- Forget about making hardware, use x86 or GPU
- Use FPGA
- Develop hardware with wider range of application
  — our decision for GRAPE-DR
  - an SIMD processor chip with very large number of processing cores (512)
  - simple on-chip network (broadcast/reduction tree)
  - particle-particle interaction, dense matrix operation, and other computationally expensive applications

GRAPE-DR Processor architecture



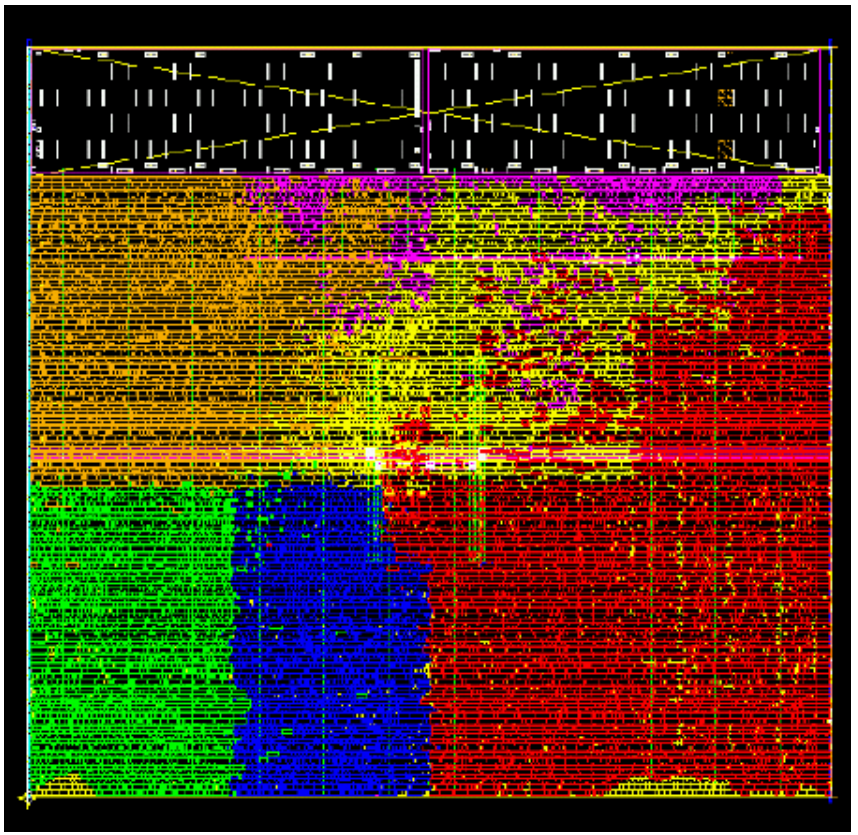
- DP Float Mult
- DP Float add/sub
- Integer ALU
- 32-word registers
- 256-word memory
- communication port

Chip architecture



- 32 PEs organized to "broadcast block" (BB)
- BB has shared memory.
- Input data is broadcasted to all BBs.
- Outputs from BBs go through reduction network (sum etc)

PE Layout

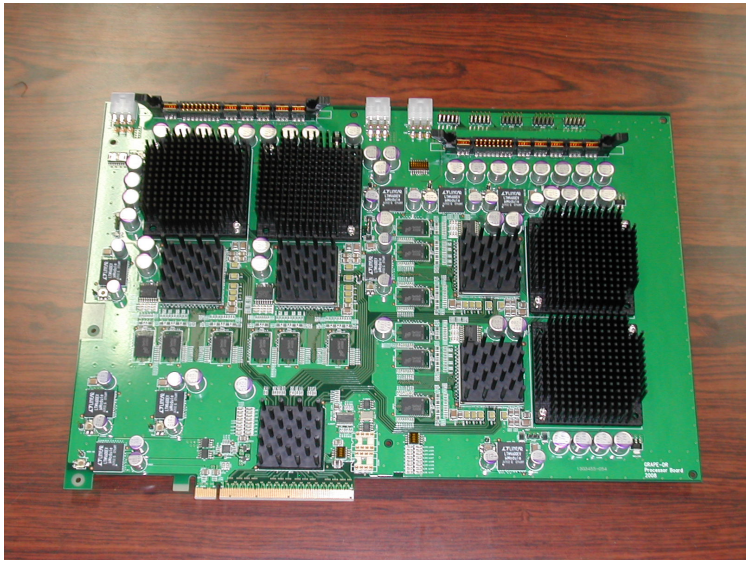


Black: Local Memory
Red: Reg. File
Orange: FMUL
Green: FADD
Blue: IALU

0.7mm by 0.7mm
800K transistors

0.1W@400MHz
800Mflops/400Mflops peak
(SP/DP)

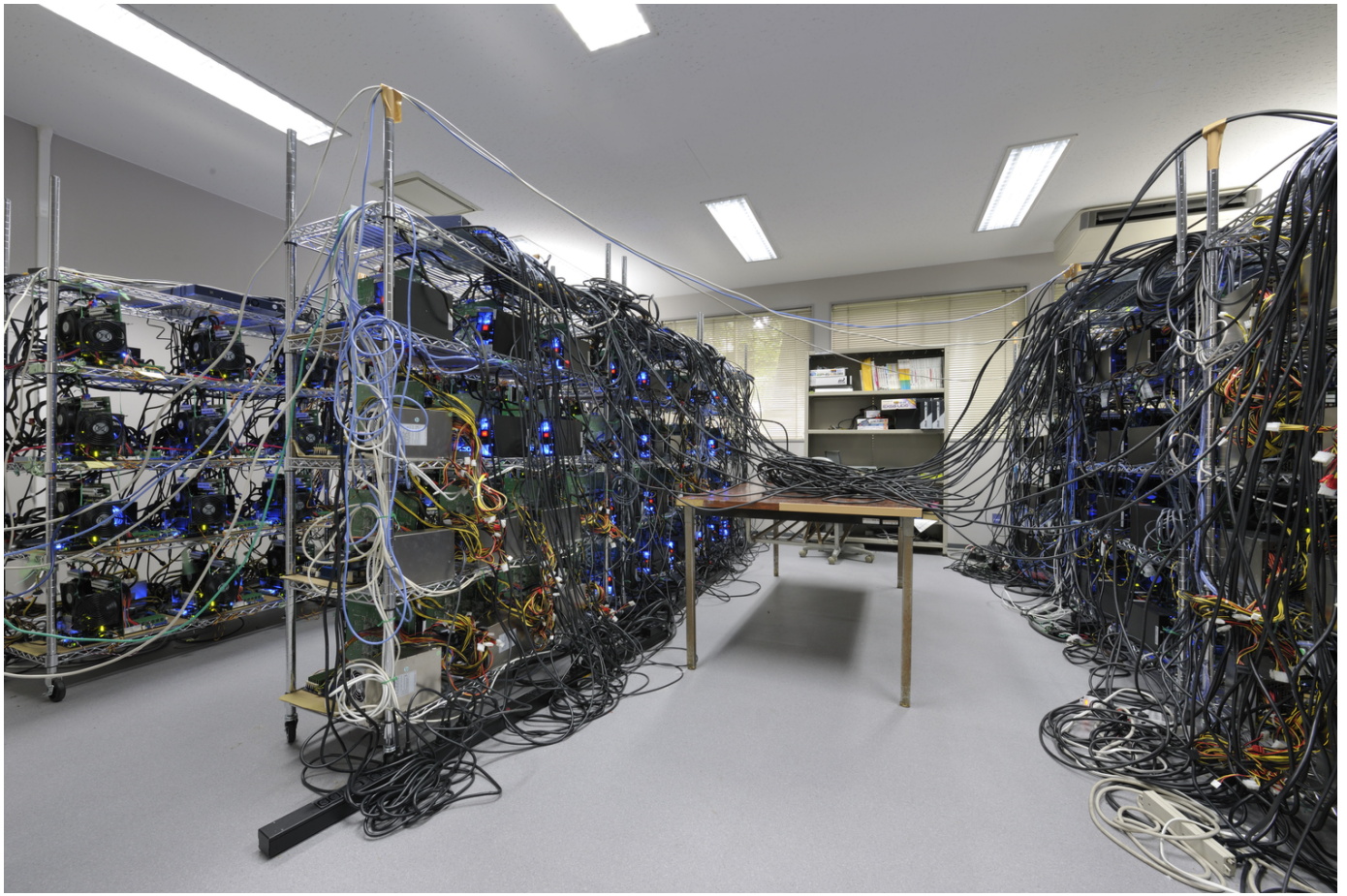
Processor board



PCIe x16 (Gen 1) interface
Altera Arria GX as DRAM controller/communication interface

- Around 200-250W power consumption
- 819Gflops DP peak (400MHz clock)
- Available from K&F Computing Research (www.kfcr.jp)

GRAPE-DR cluster system



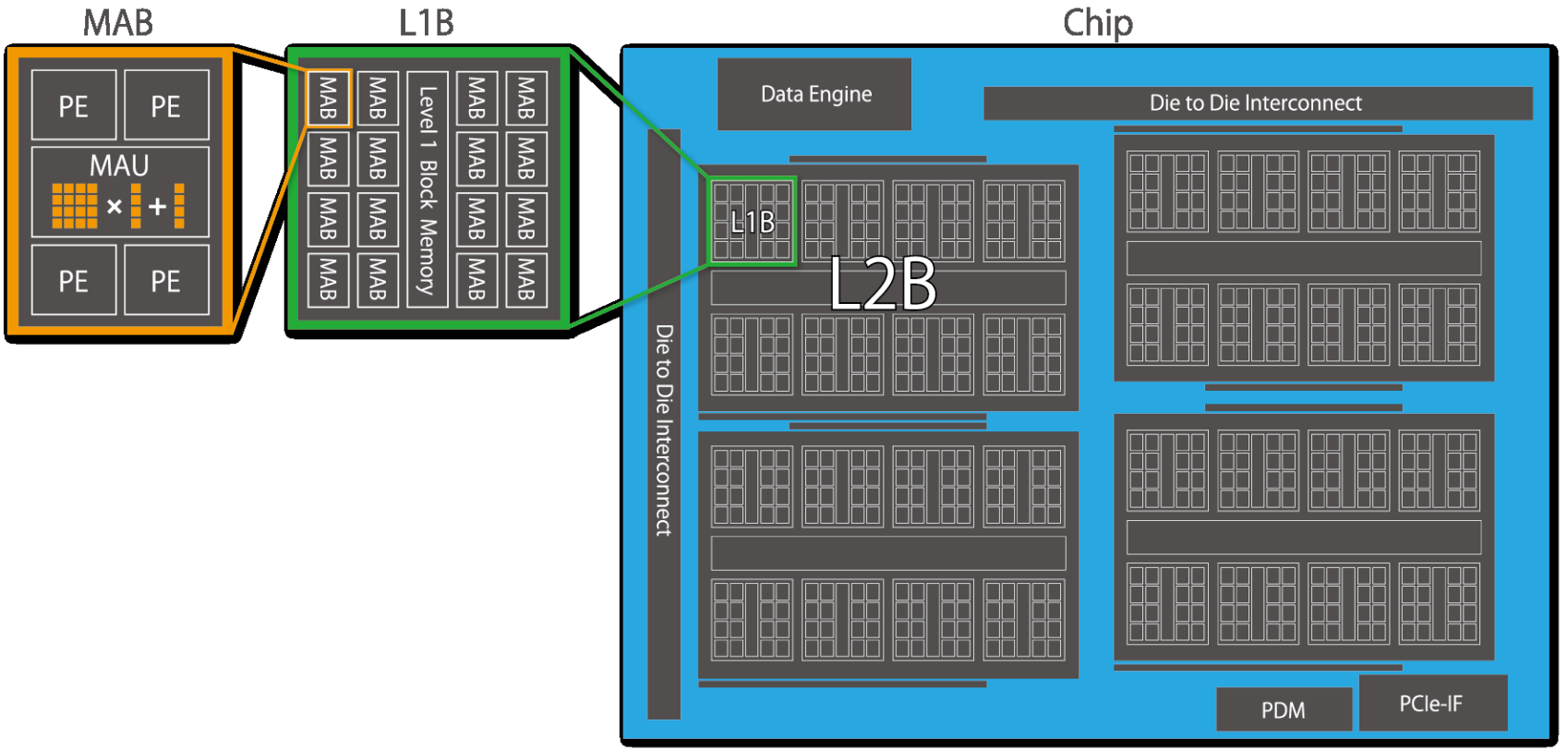
(As far as I know) Only processor designed in academia listed in Top500 in the last 20 years.

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

GPFN2 architecture



Overview of GPFN2

- One card: One "module"
- One module: four chips in one package
- One chip: one PCIe interface, DRAM interfaces, four "Level-2 broadcast blocks" (L2Bs)
- One L2B: eight L1Bs
- One L1B: 16 MABs (Matrix Arithmetic Blocks)
- One MAB: four Processor Elements combined to perform FP64, FP32, or FP16 matrix-vector multiplication.
- One PE can be also used as 64-bit scalar processor. We added many special instructions for DL.
- All PE/MAB/L1B/L2B operate on single clock and single instruction stream (card-level SIMD)

Changes made from GRAPE-DR

- Second layer of on-chip tree network
- Integration of PCIe and DRAM interface
- Addition of MAB
- Much larger memory
- Many other changes in on-chip network

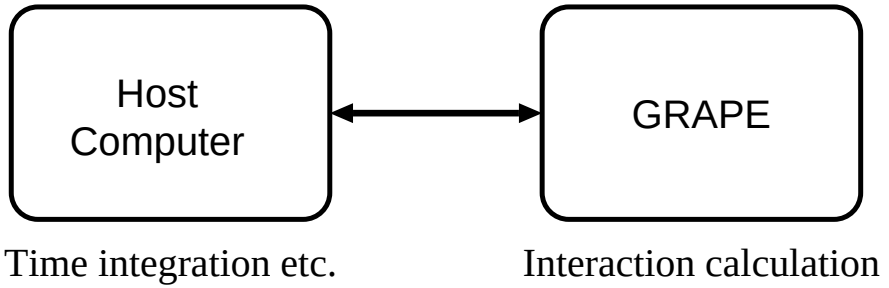
Some numbers

- **32.8TF (long), 132TF (short) and 524Tops (half)**
- **Memory bandwidth: (not yet open)**
- **Link to the host PC: PCIe**

How MN-Core can be used for Astrophysical computing

- The simplest way — as a new generation of GRAPE
- More “advanced” ways?

As a new generation of GRAPE



- **Performance would be similar to that of NVIDIA Volta (Iwasawa+ 2019)**
- **Matrix multiplication unit unused**

Can we accelerate N-body simulation with matrix multiplication hardware?

Use of FMM

- **FMM consists of the following steps: p2M, M2M, M2L, L2L, and L2p (and pp part)**
- **M2M, M2L, and L2L can be implemented using matrix multiplication**
- **We can probably use mixed-precision arithmetic. DP or SP for low-order terms and HP for high-order terms**
- **If matrix multiplication is c times faster, overall acceleration of FMM is $\sqrt{c}$.**

Other considerations

- PCIe communication bandwidth and the speed of host is too low.
- Speed ratio (30 DP TF-16GB/s) is similar to that of PEZY-SC2 system (2DP TF-1GB/s)
- So we can use the same algorithm as we used on PEZY-SC2 (Iwasawa+ 2019b)
- In this algorithm, we still construct the tree on host, but all other calculations are done on the accelerator side. Also, the tree is constructed only once per multiple steps.
- On PEZY-SC2, 40% of the theoretical peak has been achieved.

Particle methods other than gravity

SPH?

- It's not clear if we can use matrix multiplication to accelerate standard SPH (or its variations).
- Methods based on moving least squares use moderately large matrix-matrix multiplication to derive least squares, and thus can be accelerated.
- Example of such methods: CPHSF (Yamamoto and JM 2018)

How about other types of calculations

- grid hydro: can be implemented. Matrix unit?
- radiative transfer: If we need o solve linear system directly, we can use matrix units quite efficiently.

Programming environment

- Currently rather poor (PFN people are preparing DL frameworks but not much else)
- We are currently working on programming environment on which we can develop kernel functions (such as particle-particle interactions and matrix multiplications) efficiently
- More sophisticated programming environment will take some more time...

Summary

- MN-Core/GRAPE-PFN2 is accelerator processor designed mainly for deep learning.
- 512 Tops/132TF/32TF for half, short and long word
- MN-Core/GRAPE-PFN2 will be ready "soon"
- We need to make use of its matrix multiplication unit to achieve high performance.
- We can use it for gravity (FMM), particle-based hydro (MLS) and possibly for other applications as well.
- Actual hardware will be available "soon".