HPC in Japan — Past, Present, and Future.

Jun Makino
RIKEN Advanced Institute for Computational Science
Exascale Computing Project
Co-Design Team

Overview

- The past — the road to the Earth Simulator
- The present — K computer
- The future, or will the history repeat itself?
- Summary

Who am I?

Current position (as of 2014):

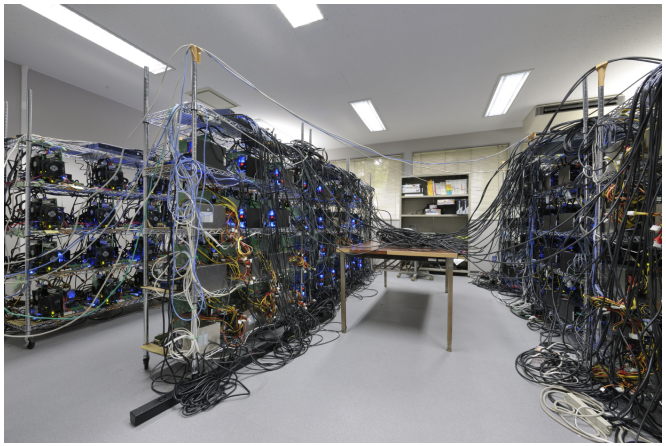
- **Team leader, Co-Design team, Exascale Computing Project, AICS, RIKEN**

What I have been doing for the last 20 years:

Develop GRAPE and similar hardware for astrophysical *N*-body simulations.

Use them for:

Planetary formation, star cluster dynamics, galactic dynamics, cosmology



The past — the road to the Earth Simulator

- First-generation vector machines
- Second generation machines
- From Numerical Wind Tunnel to Earth Simulator

First-generation vector machines

- **Hitachi S-810/20: October 1983, 630 Mflops**
- **Fujitsu VP-200: Jan 1984, 500 Mflops**
- **NEC SX-2: 1985, 1.3 Gflops**

Cray XMP/2: 400Mflops, as of 1982.

Characteristics of first-generation machines

- Single processor
- Multiple Pipelines
- Vector registers

In other words: Combination of VR architecture (Cray-1) and Multiple pipelines (Cyber 205), but with a single processor.

Advantage of Japanese Japanese companies then

- Vertical integration
- Wide range of computer products, supers and entire range of mainframes
- Development cost amortized in the large sales of mainframes.
- Advanced VLSI technology could be used.
  - Cray XMP: 16-gate MSI
  - First-generation Japanese machines: 1000-gate LSI

In the mid-1980s, the Japanese supercomputer industry looked very strong.

Second generation machines

- **Hitachi S-820/80, 1987, 3GF**
- **Fujitsu VP2600, 1988, 5GF**

Looked better than Cray YMP/8 (announced 1988, 2.667GF), but improvement from the first generation was small (factor of 5 in 5 years), still uniprocessor.

Japanese shared-memory vector parallel machines (2.5 Gen)

- NEC SX-3, 1989, 22GF
- Hitachi S-3800/480, 1992, 32GF
- 4-processor, multi-pipe machines.
- Cray C-90, 1991, 16-processor, 16 Gflops

Japanese machines still looked strong, compared to US vector-parallel machines.

From Numerical Wind Tunnel to Earth Simulator

Numerical Wind Tunnel: Joint development between National Aerospace Laboratory of Japan and Fujitsu

- **Delivered in 1993**
- **166-processor parallel-vector machine**
- **Distributed memory**
- **Crossbar connection**
- **around 10 times faster than Hitachi, NEC, or Cray vector machines**
- **Commercial Version: VPP500**

Characteristics of NWT/VPP500

- Bipolar/GaAs logic: Classic building blocks
- B/F=8: Classic vector parallel arch.
- Network B/F=0.5: fast enough to make implicit communication (in VPP-Fortran) usable.
- 1MW power consumption: 3-10 times better than shared-memory PVP machines, might be 3-5 times worse than scalar MPPs.

From VPP500 to the Earth Simulator, and after that

- Fujitsu continued this line to VPP5000 (1999, 9.6GF node performance)
- NEC switched to hybrid of shared-memory and distributed-memory machine with CMOS-based SX-4.
- The Earth Simulator was NEC design, 640-nodes, 8-processor per nodes, each processor is single CMOS chip.
  - B/F=4, performance per watt three times worse compared to Intel P4
  - Not much worse compared to US ASCI machines
- NEC is still making SX-series of machines. (B/F=1 for SX-ACE)

Distributed-memory scalar parallel supercomputers in Japan

- As in the US, many research projects used micro-processors to construct parallel supercomputers, in 1970s and 1980s in Japan
- The most successful and important is PAX project

Machine	Year	Number of processors	CPU	Peak speed (PE/total)
PACS-9	1977	9	6800	1k/10k
PACS-32	1980	32	6800/Am9511A	16k/0.5M
PAX-128	1983	128	68B00/9511A-4	30K/4M
PAX-64J	1984	64	LSI-11	0.1M/6M
QCDPAX	1989	432	68020/L64133	28M/12G
CP-PACS	1996	2048	extended HP-PA	300M/614G

CP-PACS/Hitachi SR2201

- Joint development between Tsukuba University and Hitachi
- HP-PA architecture with some changes (sliding register window)
- B/F=1.
- High-bandwidth 3D hyper-crossbar network.

SR-8000: IBM power based. After that, Hitachi has been selling IBM power series under the Hitachi brand name.

Special-purpose systems

- GRAPE systems, specialized for gravitational interaction between particles
  - started in 1989
  - GRAPE-3, 1991, 14.4GF, 48-chip, two wire-wrapped cards.
  - GRAPE-4, 1995, 1.08TF, 1792-chip, 36 PCBs, 20KW
  - GRAPE-6, 2002, 64TF, 2048-chip, 36 PCBs, 40KW
- Machines for molecular dynamics
  - MDGRAPE-2, 2002, 100+ TF
  - MDGRAPE-3, 2006, 1PF
  - (ANTON, 2009, 500TF?, by DESRES: Integrate microprocessors and networks as well as pipeline processor for particle interaction calculation)

GRAPE-DR

- Developed by JM and Prof. Hiraki
- SIMD parallel processor with 512 cores
- Machine completed in 2008
- Achieved #1 in Little Green500 list in June 2010
- #2 in Green500 in Nov 2010 (#1: BG/Q)

Initial cost of custom LSI skyrocketed:

- 1992: USD 200K
- 1997: 1M
- 2004: 2-3M
- 2014: 20M?

Too large for special-purpose machines with narrow application range.

Summary of the past history

- Quick rise with the first generation machines
- Gradual fall, with second-generation machines
- Some recovery with NWT
- Gradual fall with vector-parallel systems, shift from vector to scalar architecture

Representative HPC machines through 1983-2003 in Japan and US

Year	Japan	US
1983	Fujitsu VP-200	Cray X-MP
1988	Fujitsu VP-2600	Cray Y-MP
1993	NAL NWT	Cray T3D
1998	CP-PACS	ASCI Red
2003	ES	ASCI Q

Difference between US and Japan

- Switch from vector-parallel to Scalar-parallel architecture was more or less finished in 1993 in the US, but yet to happen in Japan at 2003
- Vertically integrated structure of Japanese companies made this adherence to vector architecture possible, with high B/F numbers
- The advance of CMOS semiconductor technology made the high B/F number very costly (the "Memory Wall"). As a result, Japanese vector machines have become costly, and two out of three companies have switched to scalar processors
- One of them stopped development of HPC systems

Change in the semiconductor industry

- As of 2014, there is NO Japanese company with fabs for LSIs below 40nm.
- Fujitsu made the Venus processor (used in K computer) with their 45nm line
- NEC made 65nm SX-9 processor with their line
- These are last processors they made with their fabs
- In 1980s, vertical integration helped
- It delayed the transition from vector to scalar, and in the meantime Japanese vector-parallel machines had lost the market share.

The present — the K-computer

- Planning
- Concept design and its evaluation
- The machine built
- Lessons (not quite) learned

Planning

- Discussion started in the computational science and technology working group, under the information science and technology committee under MEXT, in 2004.
- 2006 report:
  - heterogeneous system with vector-parallel, scalar-parallel, and "special processing" (accelerator), with 1, 3, and 20PF
  - From today's viewpoint, looks rather reasonable

Concept design and its evaluation

- Jan 2006: the project headquarter for the development of the next generation supercomputer as a part of RIKEN.
- Summer 2006: Eight "joint research project for the construction of the concept of the next generation supercomputer" initiated by the HQ.
- each participant was requested to evaluate their proposed system against 21 benchmark codes for the peak speed of 10PF
- Price performance or performance per watt were not questioned (except the power limit of 30MW)
- Design decision was made and approved by yet another MEXT committee by Summer 2007.
- The meeting materials are now open to public, after the request made by the House of Representatives made in Dec 2011.

Presentation at March 27, 2007

- Compared six proposals, three from three companies and the other three from Universities and a National lab (Astronomical Observatory, NAOJ)
- NEC and Hitachi proposed vectors, and merged. U Tokyo and NAOJ proposed accelerators, and "merged"
- U Tsukuba proposal regarded to lack feasibility
- "Vector" proposal from NEC and scalar proposal from Fujitsu were very similar in B/F (both 0.5), peak performance, estimated efficiency, and power consumption (15-17MW)
- Power consumption of accelerators were reported to be $\sim$ 10MW (actual numbers in NAOJ and U Tokyo reports were 0.66 and 0.88MW).
- No mention on the performance of accelerators

Conclusions of the presentation

- **RIKEN is currently studying two possibilities: Either chose one from NEC-Hitachi or Fujitsu proposals, or let both two groups design a heterogeneous system.**
- **RIKEN decided that the accelerator is not necessary, since both NEC-Hitachi and Fujitsu proposals meet the project goals.**

The final decision:

Let both Fujitsu and NEC-Hitachi develop the machines based on their proposals, but it seems the B/F number for NEC-Hitachi proposal was increased to 1.0 from 0.5.

The machine built

- **Fujitsu showed off the processor chip they made with their 45nm line in spring 2009.**
- **NEC decided to drop out from the project in May 2009**
- **RIKEN decided to continue the project with Fujitsu machine only, and the machine was (mostly) completed by November 2011 (Linpack 8.6PF)**
- **Stayed at #1, till June 2012 (IBM BG/Q reached 16.3PF)**
- **Two Gordon-Bell prizes, 2011 and 2012 (for the latter JM was one of the awardees). Problems: Real-space DFT (Huge Matrix diagonalization) and Cosmological N-body simulation.**

Lessons (not quite) learned

We can summarize the problems with the K-Computer project as

- **adherence to old architectures, vector and traditional scalar "big" scalar cores**
- **Too much faith in the Japanese semiconductor industry, which was about to vanish**

The future, or will the history repeat itself? — The Exascale Project

- April 2011: a working group for the study of the R&D in the future HPC technology.
- July: Another three WGs for that FY, made report in March 2012
- July 2012: Three "feasibility study" projects, for vector, scalar parallel, and accelerator architectures, and one for scientific goals and application software.
- FS projects just finished (end of FY 2013)
- RIKEN proposed Scalar+Accelerator system by Summer 2013.
- Project (sort of) approved in Dec 2013

Why Scalar+Accelerators?

Two main reasons

- Accelerator gives significantly better price performance and performance per watt. With scalar-only architecture, an Exaflops machine in 2019 (10nm technology assumed) would consume 60-100MW.
- Accelerator with its own low-latency network can in principle provide much better scalability for "small" problems ("Small" means typically less than 10B grid points...)

Whether or not this heterogeneous architecture will be actually built remains to be seen

Conclusion

- National Projects played crucial roles in determining the direction of R&D of Japanese HPC companies.
- Except for the case of NWT, however, it seems rather clear that architectures selected in the National projects were too old, based on the faith in the superiority of Japanese semiconductor industry.
- It seems Japanese national projects have been tied with the illusion of the superiority of the Japanese industry, or have had too much inertia to change the direction.