A Global Perspective on Supercomputing Research and Education for Scientific Computing in Japan

- Historical overview of Japan and US HPC’s

-What is the difference between Japan and US trends?

-Education and Training for the Next Generation Supercomputer in Japan

Yoshio Oyanagi

Kogakuin University, Tokyo

Next-Generation Supercomputer

- To be available in 2012 in Kobe
- In collaboration with leading supercomputers in universities and laboratories ….
- Strategic committee (in Mext)
  - Define five strategic fields
  - Human resource development for supercomputers

Human Resource

- What is the problem?
  - Education of supercomputer users.
- Objectives
  - Let users harness the computing power.
  - Petaflops machines are not easy to use (performance depends critically on the tuning of programs).
  - So far no systematic education or training of supercomputers have been given to young scientists.
  - Computational scientists in Japan tend to be indifferent to computer science.

Education Initiatives

- To innovate computational sciences thru comprehensive use of supercomputing power
  - New modeling for new architecture
- To innovate computer science in collaboration with applications
  - New algorithm for new architecture
- To attract people in application fields to computational sciences

My Personal Retrospective

- I started using computers for particle physics in the late 1960's.
- My adviser said, "First-class physicists don't user computers."
- That was the common sense in Japan, not only in physics.
- Not true at all：von Neumann, Fermi, Alder, Feynman, …..

Is computing something inferior?

- Most scientists do not speak about computing, even if they use computers extensively (in Japan)
  - In interdisciplinary meetings of computational scientists, people tend to talk only on their scientific results, but not on modeling, algorithms, tuning, architectures etc.
  - "I wrote a program. Someone should tune it for me"
- Characteristics of computational sciences
  - Common algorithms or computational techniques in different fields
  - Exchange of experience is important, among different fields or with computer science.

Why no education so far?

- In the 1970's: Study the FORTRAN grammar and you will be able to write scientific programs.
  - As if speaking English just by knowing the grammar.
  - Self-taught brute-force modeling, algorithms, coding.
  - Programs transmitted from senior laboratory members without knowing what is written in them.
- Computers then were slow and had simple memory structure and relatively high bandwidth.
  - Tuning was not critical.

Next-Generation Machine

- One chip contains eight cores (CPU’s). Each core has two SIMD processors which can execute two mult’s and tow add’s in a clock.
  - Peak performance is available only when all these floating units are busy.
  - How to write programs.
- Relatively narrow bandwidth between CPU and memory
- Super-massive parallelism
  - Between cores, chips, nodes
  - OpenMP, MPI, PGAS, ….

Who will tune my program?

- User may hope "I bring my program to the center and supporters there will tune it for me. "
  - They may give you some advices.
  - They do not know how the program works.
  - Users themselves know it best. Their commitment is indispensable.
- Tuning is a collaborative work of users and supporting staffs.
  - Users have to have a certain level of technology.
  - Especially when they aim at high performance.

Different levels of users

- 松”Pine”: Most sophisticated users. They tune the program to the limit and attain the peak performance.
  - Most expected, but not many.
- 竹”Bamboo”: Middle-class users. They tune the program using various tools.
- 梅”Plum”: Beginner users. They depend on compilers. They will one day be elevated to higher levels.
- Balance is important.

Historical Perspective

- Vector and parallel supercomputers
- User formation at each time
- Observation

1970's (red for vector machines)

- USA Vendors: ASC(72), STAR-100(73), ILLIAC-IV(73), Cray-1(76), HEP (79)
  - Y. Muraoka, K. Miura and others learned at ILLIAC IV.
- UK: ICL DAP (79)
- Japan. Vendors: FACOM 230/75 APU(77), HITAC M180 IAP(78)
- Kyoto U (Electric Eng.): QA-1(74), QA-2 (VLIW)
  - Signal processing, Image processing
- Kyoto U (Nuclear Eng.): PACS-9(78) (→U. Tsukuba)
  - Reactor simulation

1980's (Vectors)

- USA Vendors:
  - Cyber-205 (81), XMP-4 (84), Cray-2 (85), IBM 3090 VF (85), ETA-10 (87), YMP (88)
  - Convex C1 (85), SCS-40 (86), Convex C2 (88), Supertek S1 (89)
- Japan. Vendors:
  - Hitac S810/20 (83), S820 (87)
  - FACOM VP200 (83), VP2600 (89)
  - NEC SX-2 (85), SX-3 (90)

Japanese Vectors in the 80's

- Vectorization compilers at that time was not good enough.
  - But they were better than Cray's.
- Various know-how was necessary to tune the program. You could get 20 to 30 times speed up sometimes.
- Japanese users came to believe that you can speed up the program on supercomputers only with a small modification.

1980's (US Parallel)

- Parallel Ventures in US:
  BBN Butterfly (81), Cosmic Cube (83),
  Elxsi 6400 (83), Pyramid 90x (83),
  Balance 8000 (84), nCUBE/1 (85),
  Alliant FX/8 (85), Encore Multimax (86),
  FPS T-series (86), Meiko CS-1 (86),
  CM-1 (86), CM-2 (87),
  Multiflow Trace/200 (87)

1980's (Japan Parallel)

- Japan. Activities (mainly for research):
  - U. Tsukuba: Pax-32 (80), Pax-128 (83), Pax-32J (84), qcdpax (89) for qcd
  - Fifth Generation (ICOT) of MITI 82-92 PIM machines for inference
  - Supercomputer Project of MITI 81-89 PHI, Sigma-1 (dataflow), CAP, VPP (GaAs)
  - Osaka U.: EVLIS (82) for LISP
  - Keio U.: SM^2 (83) for sparse matrix
  - U. Tokyo: Grape-1 (89)

1990's (USA)

- USA Vectors: C90 (91), Cray-3 (93), T90 (95), SV1 (98)
- USA Parallel (now mainstream):
  - CM-5 (92), KSR-2 (93), SPP (94)
  - SP1 (93), SP2 (94), ASCI Blue Pacific (97), Power 3 SP (99)
  - T3D (93), T3E (96)
  - ASCI Red (97)
  - Origin 2000 (96), ASCI Blue Mountain (98)

1990's (Japan)

- Japan. Vectors: S3800 (93), NWT (93), VPP500 (93), SX-4 (95), VPP300 (95), VPP5000 (99)
- Japan. Parallel:
  - cp-pacs (96), SR2201 (96), SR8000(98)
  - AP1000 (94), AP3000 (97)
  - Cenju-2 (93), Cenju-3 (94), Cenju-4(97)
  - Except SR's, they are sold as a testbed.
- RWCP project (MITI, 92-02): Cluster connected by Myrinet. Score middleware.

Observation (1/3)

- Until late 1990’s, Japanese vendors focused on vector machines.
- Users exploited the power of vectorization.
- Vendors thought parallel machines were for specialized purposes (eg. image processing). Most application users dared not try to harness parallel machines in the 80’s.
- Some computer scientists were interested in building parallel machines, but they were not used for practical scientific computing.

Observations (2/3)

- Practical parallel processing for scientific computing in Japan was started by application users: qcd-pax, NWT, cp-pacs, GRAPE's, ES.
- Softwares
  - Very good vectorizing compilers.
  - Users were spoiled by them.
  - Users found difficulties in using message passing.
  - HPF efforts for the Earth Simulator.
  - OpenMP
  - Score from RWCP

Observation (3/3)

- Japan is at least ten years late in parallel processing for scientific computing as compared to US.
- Education in parallel processing is a urgent issue for the Next Generation Supercomputer.
- More collaboration of computer scientists and application scientists is needed.