The GRAPE project: Current Status and Future Outlook

Jun Makino
University of Tokyo

Talk overview

- GRAPE Project
- Status of GRAPE-6
- Some new scientific results with GRAPE-6
- GRAPE-8?
- Summary

GRAPE project

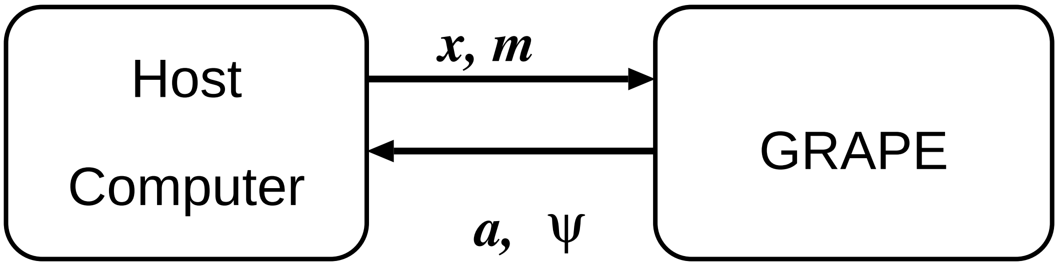
GOAL:
Design and build specialized hardware for simulation of dense stellar systems.

Rational:
You can do larger simulations (better resolution) for same amount of money.

GRAPE-6	(2001, ∼100 TF)	5M$
ASCI White	(2001, 12 TF)	100M$
Earth Simulator	(2001, 40 TF)	400M$

Basic idea of GRAPE

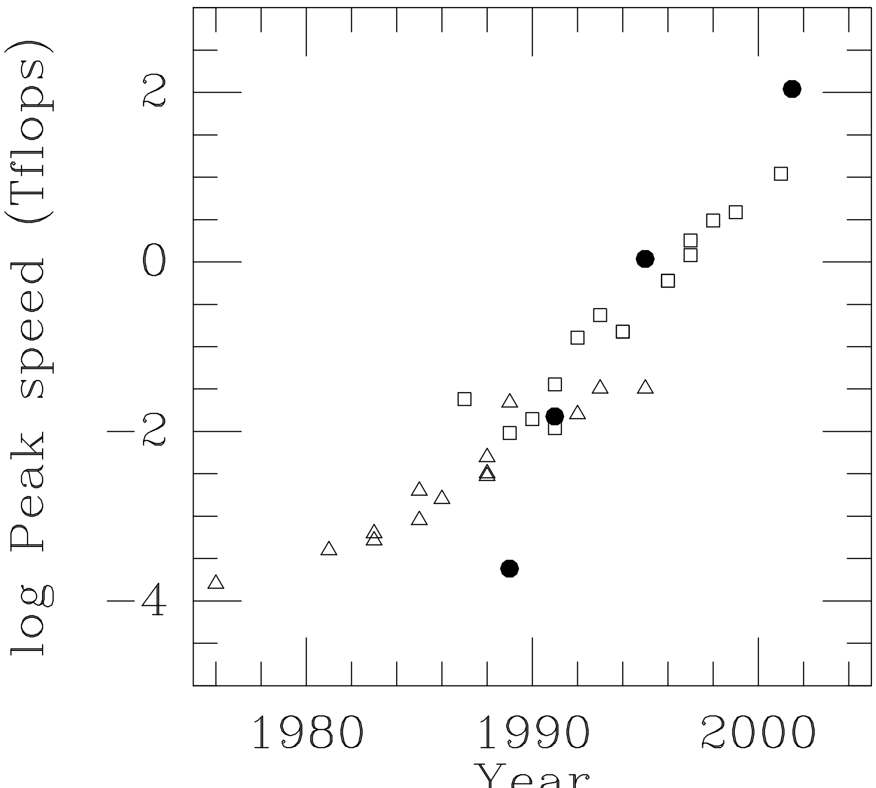
Special-purpose hardware for force calculation
General-purpose host for all other calculation



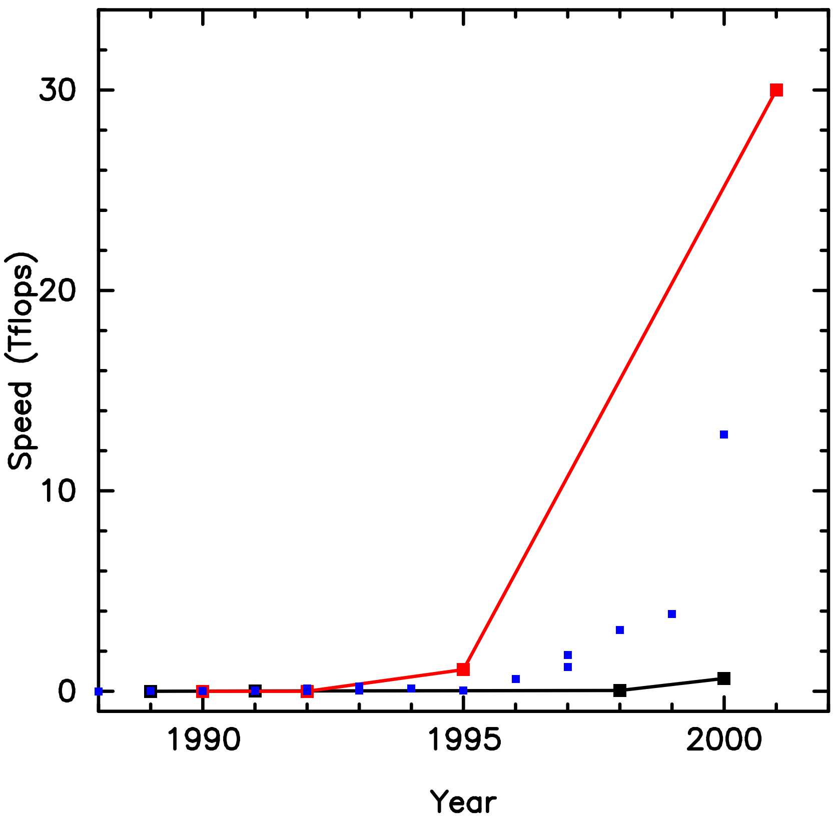
Advantages of GRAPE

- Special-purpose hardware
  - Hardwired logic → No software needed
  - Pipelined processor → High performance
- General-purpose host
  - Programmable in high-level languages
  - existing programs can be used with “small” modifications
  - can be used with Barnes-Hut tree/Aarseth scheme

Evolution of GRAPE systems



A different way to show the same data



GRAPE-4

Features:

- **Predictor pipeline in hardware (for individual timesteps)**
- **Massively parallel system with 1692 pipeline chips**
- **36 processor boards each with 47 processor chips**
- **calculates force on 94 particles**
- **Peak speed: 1.08 Tflops**

Project goals for GRAPE-6

- Peak speed $\sim$ 100Tflops
- Improved communication speed ($>$ 2GB/s)
- Minor fixes for some "features" of GRAPE-4...

Overall structure

- 8 clusters
- 8 processor boards per cluster
- 8 processor modules per board
- 4 chips per modules
- 6 pipelines per chip
- 60 arithmetic units per pipeline

In total,7.5×10^5 Arithmetic units operating at 90 MHz = 68 Tflops

Host Parallelization

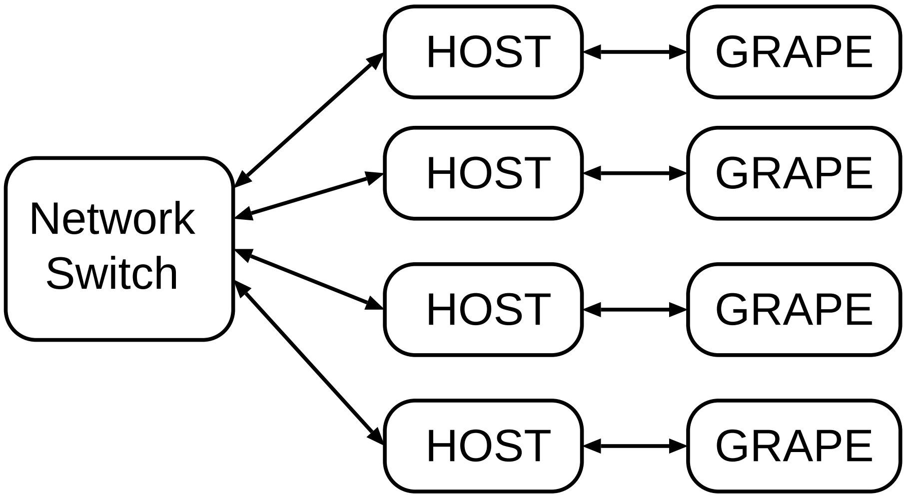
Major challenge: We need parallel host to serve GRAPE-6

- **GRAPE-6 is 100 times faster than GRAPE-4**
- **single CPU would become only 10 times faster in 5 years**

We have to use 10 hosts in parallel for GRAPE-6
How?

A Naive approach

Connect multiple GRAPEs and multiple hosts



This does not work

Two methods for parallelization

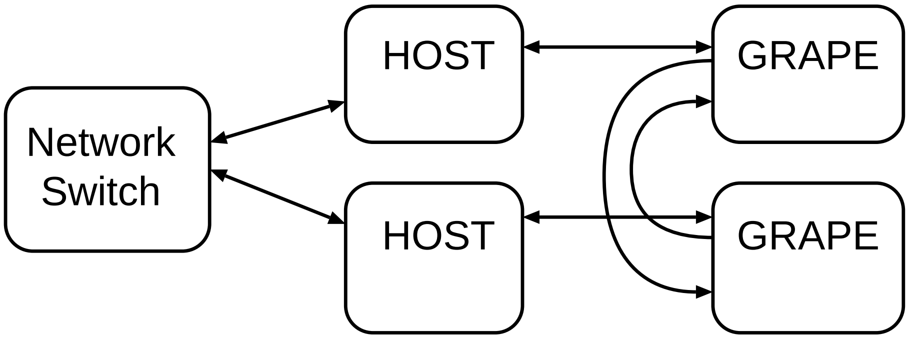
- **Each host has complete copy of the system. Time integration is distributed.**
- **Each host has its own share of particles.**

In either case, each GRAPE (and therefor each host) need the information of all particles to calculate the forces on its particles.

The communication bandwidth of a single host limits the scalability

Solution in GRAPE-6

Each GRAPE box has multiple input ports from which it can receive the data from all hosts in parallel.



For n hosts, GRAPE has n input ports.

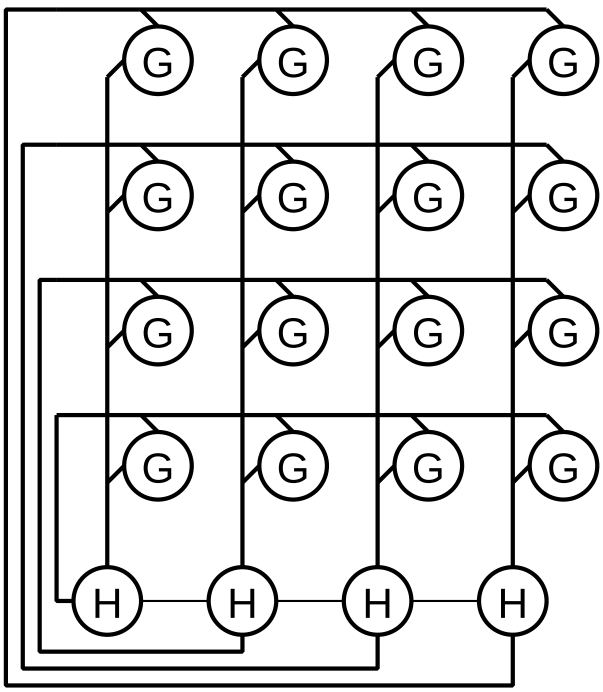
How to add multiple input ports?

Since a single cluster consists of multiple boards anyay, it has multiple port. If a cluster has n boards, it has n ports.

If we connect n^2 boards to n hosts, it should be okay.

GRAPE-6 : 100 Tflops, around 100 boards, around 10 hosts.

Network structure

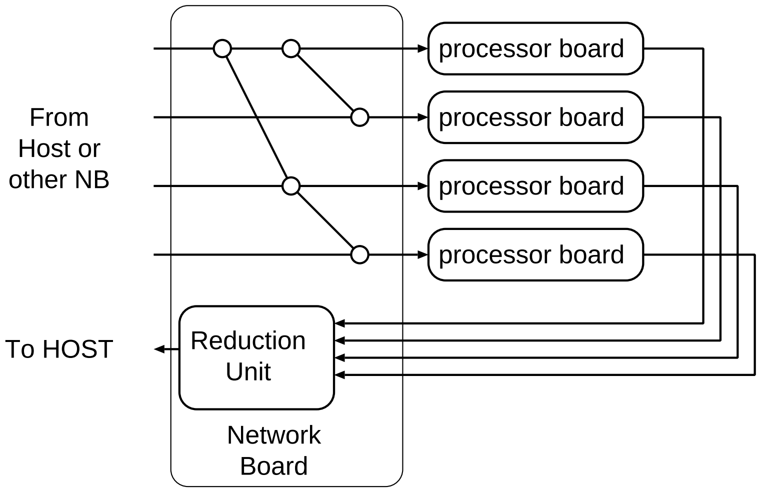


No host-host communication required (except for synchronization)

Low-latency message-passing interface is desirable

SCore? GAMMA?

A more flexible system

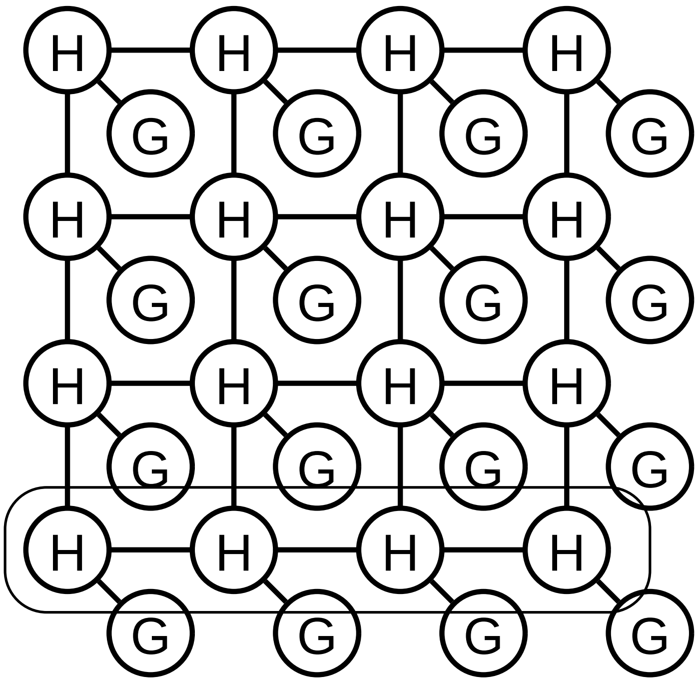


Can be configured for the broadcast from a single host, direct connections from n hosts, and multicasts from $n/2$ hosts etc.

An alternative solution

- Configure n^2 hosts as a $n \times n$ grid.
- All hosts other than a single row of n hosts are used just as the network nodes described before.
- Cheap PCs woth GbE may be fine
- Final GRAPE-6 will partly use this approach

The host network solution



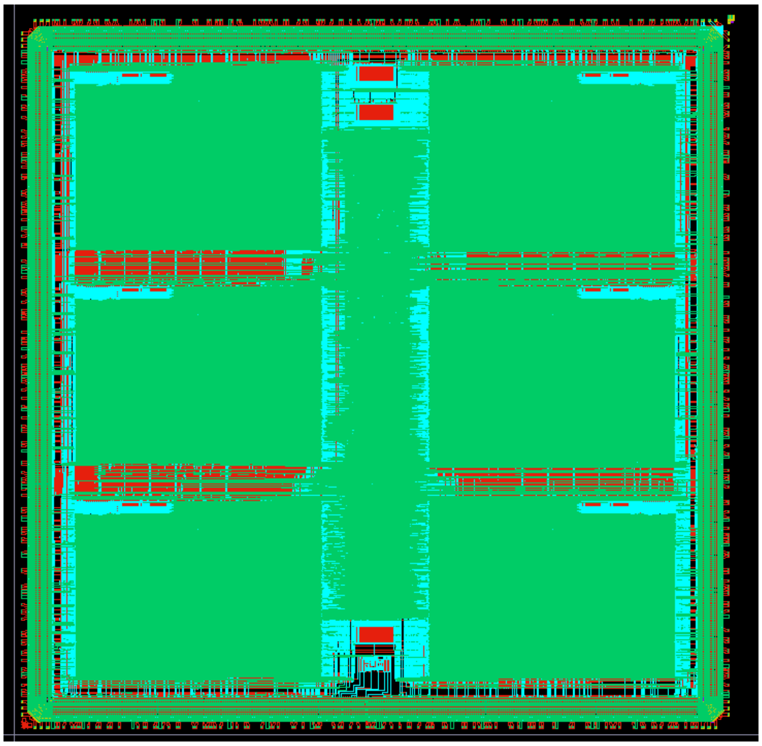
This approach can be used for parallelization on MPPs without GRAPE.

Much simpler than the Hyper-Systolic stuff...

GRAPE-6 Processor Chip

- **0.25 μm custom design (Toshiba TC240E, 1.8M gates)**
- **2.5 V power supply, 100 MHz operation (design goal)**
- **6 pipelines in a chip**
- **$>$ 30 Gflops (equivalent) per chip**
- **system clock: 25 MHz**
- **Single chip equivalent to single GRAPE-4 board**

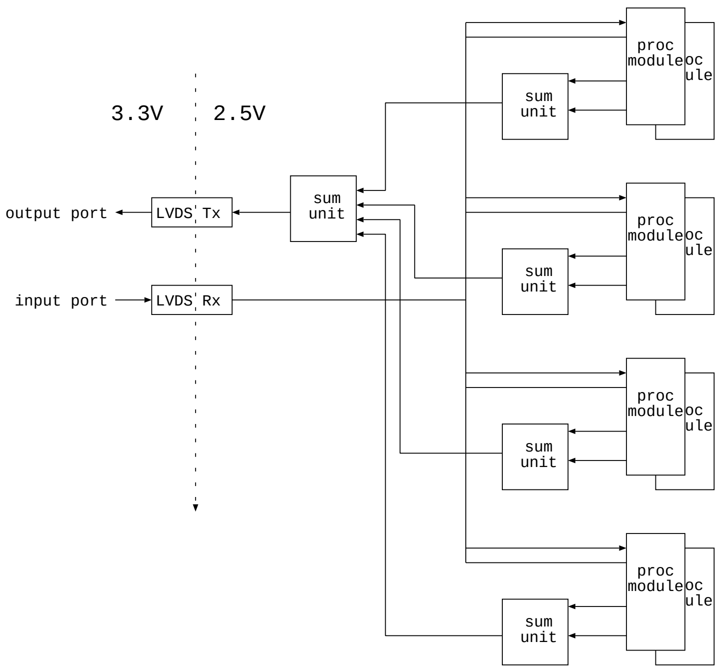
GRAPE-6 Processor Chip



GRAPE-6 processor module

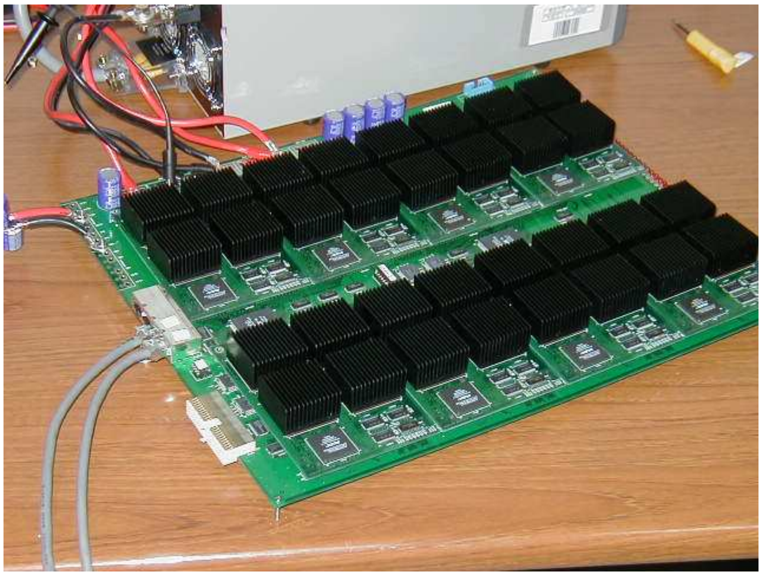
- 4 processor chips + Memory chips.
- Integrates all functions of GRAPE-4 processor board.

The GRAPE-6 Processor board

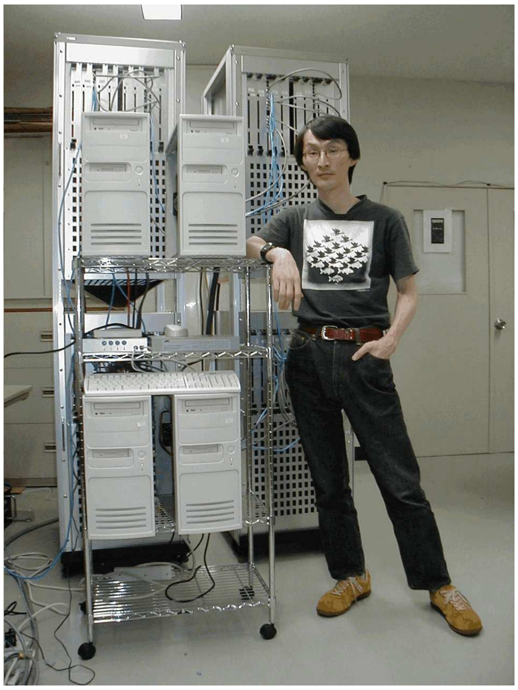


- **Single board with 32 chips**
- **Semi-serial (LVDS) inter-face (350MHz clock, 4 wires)**
- **Two input ports and one output port**
- **Result from 32 chips are summed by a reduction tree**

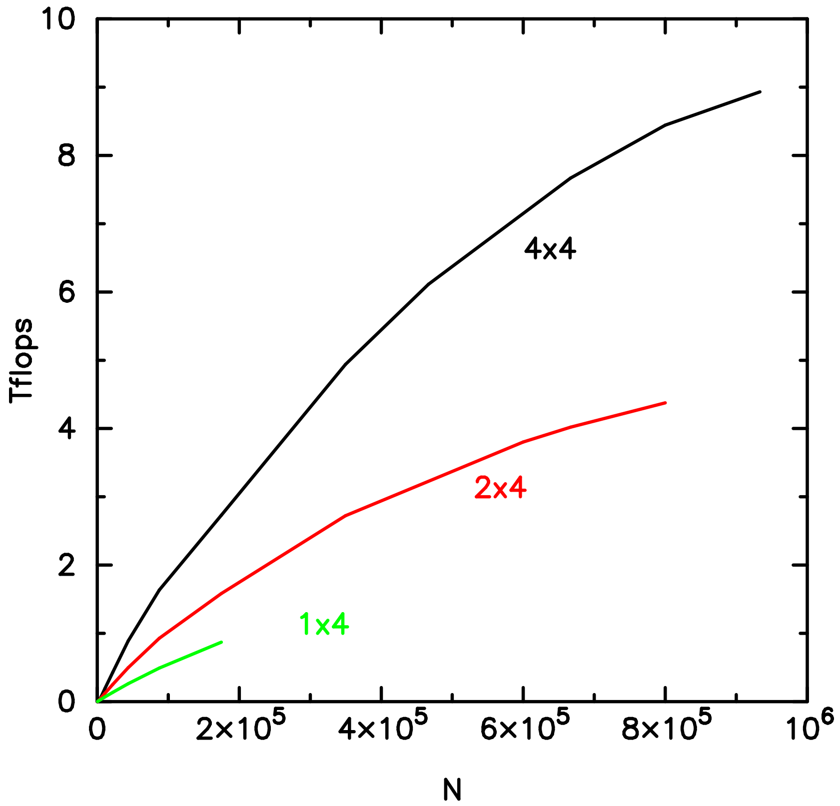
The GRAPE-6 processor board



The 32-Tflops GRAPE-6 system



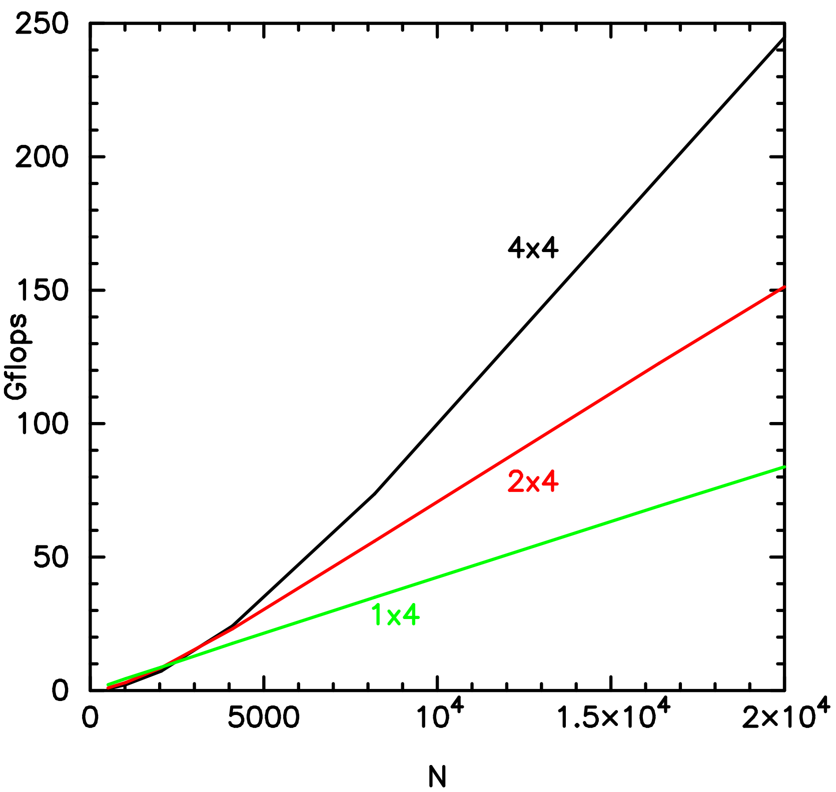
Performance of GRAPE-6



16-board,
8-board and
4-board system
with 4, 2 and 1
hosts,
respectively

Pentium 4 hosts
connected by
100BT Ethernet
(MPICH/p4)

A closer look



Even with slow MPICH, parallel overhead is small if $N > 10^4$.

Host overhead is still large...

Some new results from GRAPE-6

- Triple massive BHs in galactic center
- The growth of a single BH in the center of a star cluster (J. B. Taylor + JM)

Triple BHs

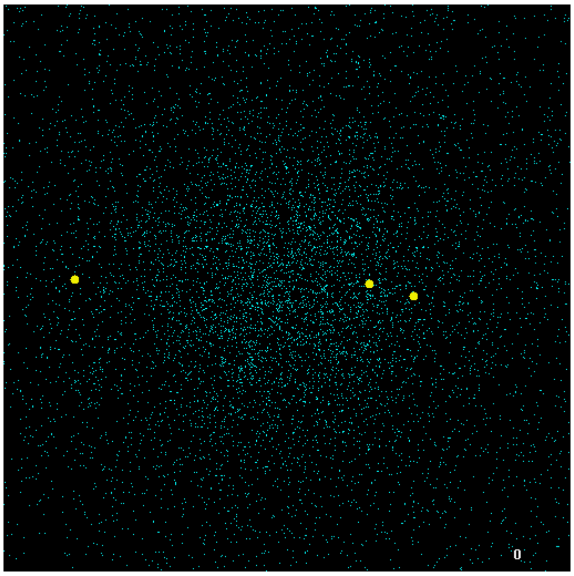
Numerical studies of binary BHs in the center of the galaxies seem to suggest rather long merging timestep (Makino 1997, Quinlan and Hernquist 1997)
If binary BHs have long lifetimes, systems of three (or more) massive BH must be common.

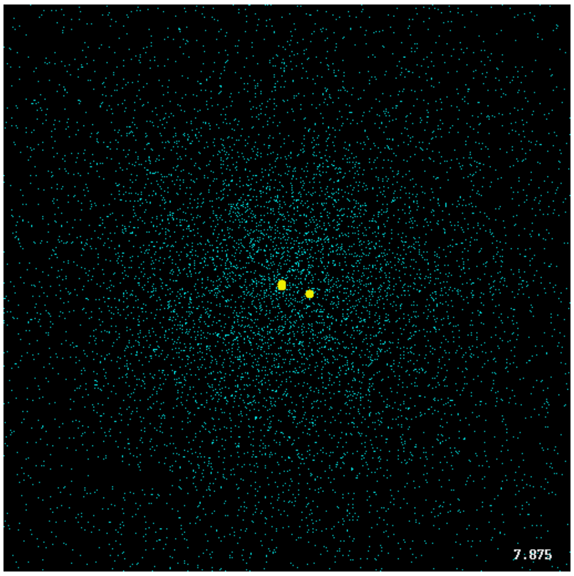
- Merging of galaxies is common — galaxies are formed "bottom-up", through merging of smaller objects.
- Merging of two galaxies, one with a central BH and the other with binary BH would form triple BHs in the center.

Triple BH simulation

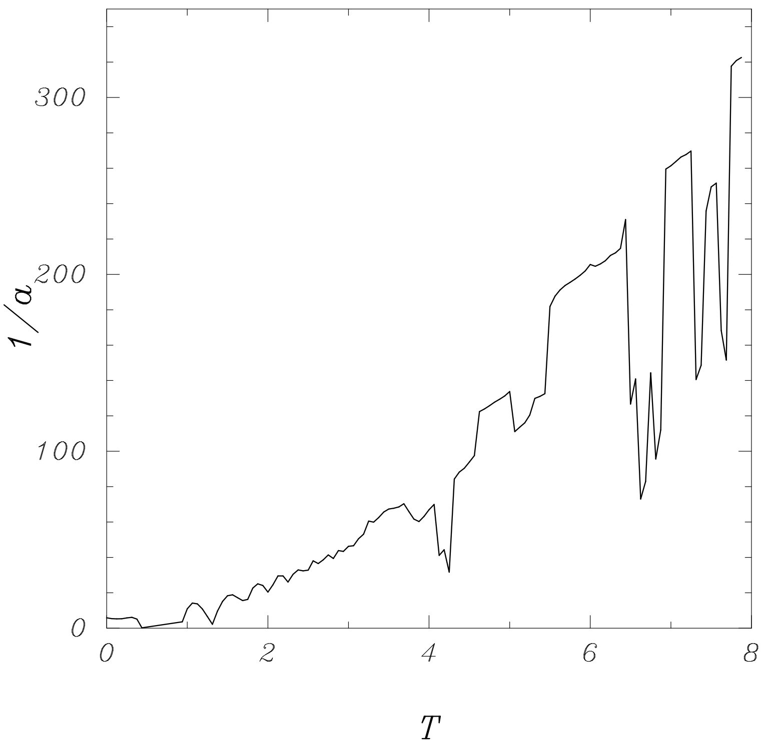
- Host Galaxy modeled by 786,432 equal-mass stars
- BHs have 1% of the mass of the host galaxy
- gravitational wave radiation ignored
- BHs placed “randomly” around the center

Initial and final snapshots





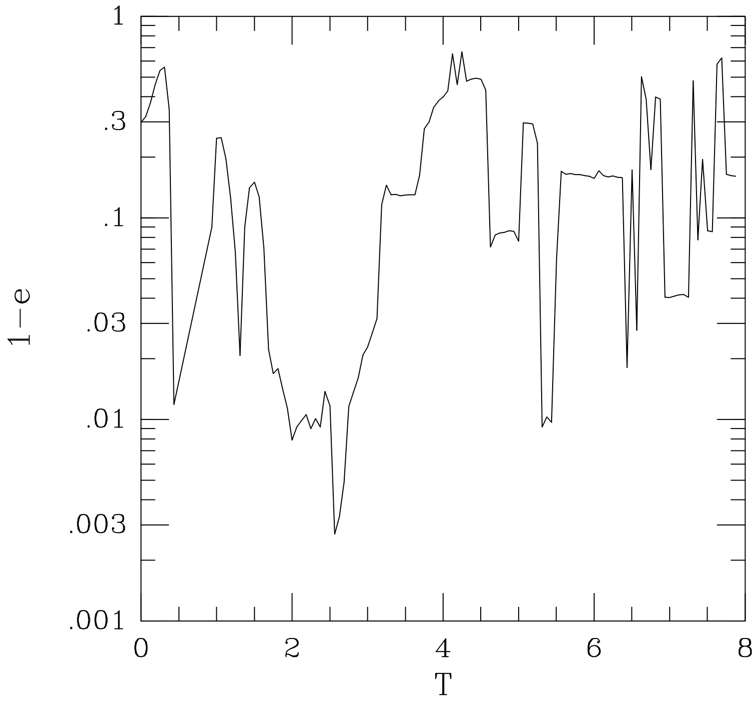
Result – separation



Rather quick hardening

Not surprising since the BH binary and the 3rd BH both sink to the center

Result – eccentricity



Eccentricity can go very close to one

— gravitational wave would quickly let the bound pair merge.

Growth of a single BH in the core of a star cluster

Young compact clusters are not very rare.

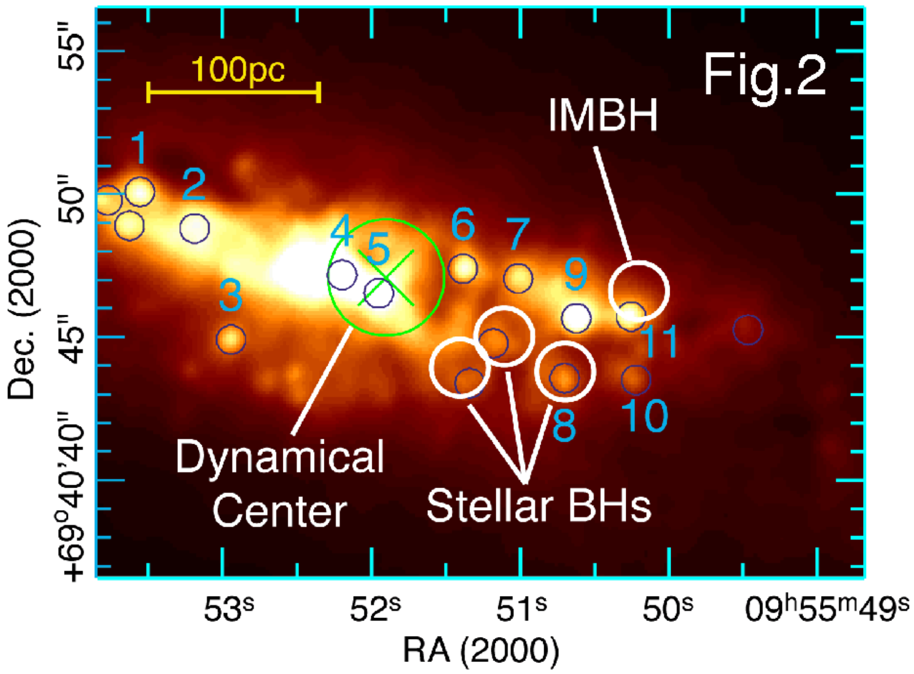
The center of our galaxy: Two known (Arches, Quintuplet)

- Very young (several Myrs)
- Fairly massive ($10^4 \sim 10^5 M_\odot$)
- Very compact ($r_h < 1$pc)
- Close to GC (30-50pc)

At least one in LMC: R136
Quite similar to Arches except for the distance to GC

Similar object in M82?

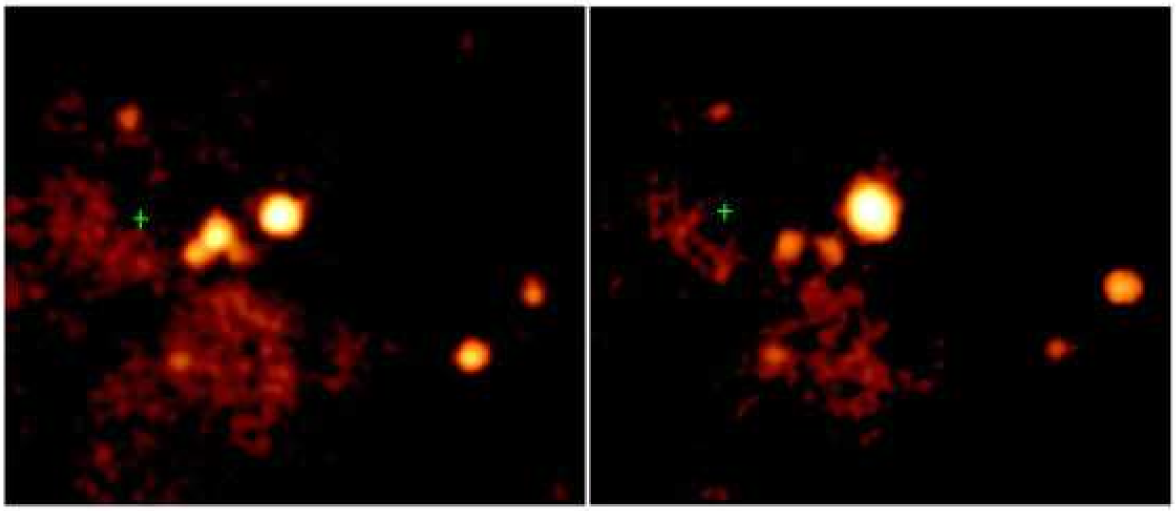
SUBARU IRCS view of the center of M82 (K' band)



> 10 compact cluster candidates found.

Link to Massive BH?

Two Chandra X-ray views of central 30 arcsecs of M82 (distance 17Mpc) with 2 months interval



N-body simulation of the growth of BH in a cluster center

Numerical work: Not much was done since Monte-Carlo studies by Shapiro and his collaborators in early 1980s.
We tried simple N-body simulations

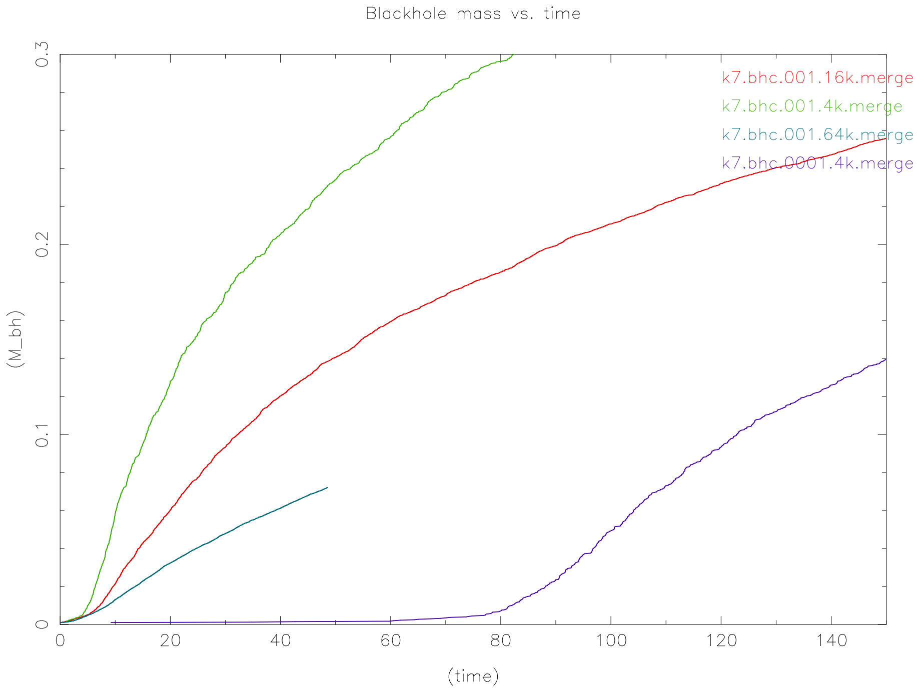
- field stars are all equal-mass particles.
- the seed BH is placed on the center.
- BH and a star are merged when their separation becomes smaller than the tidal disruption radius.

Initial conditions

- **Initial BH mass 10^{-4}**
- **King 7 model**
- **N =4K, 16K and 64K**
- **initial BH “radius” 10^{-4}, 10^{-3}**

BH growth

Green, Red and Blue: 4, 16 and 64K
Purple: $\times$10BH radius



BH growth

- **2 (3?) phases of growth.**
- **Growth timescale $\sim \sqrt{N}$???**
- **Dependence on BH radius: seemingly linear**

Not quite consistent with Monte-Carlo result...
With MC, Growth timescale $\sim N$

GRAPE-8?

What can we build in 5 years from now?

- **Technology: sub-0.1 μm: 6-10 times more transistors and 2-3 times faster clock**
- **Initial design cost for chip: currenly unknown...**

Assuming we can use the same amount of money for the mass production of the chip, 1–3 Petaflops is within reach.

Hardware Design challenge

Major issues

- What host computer should we use?
- How should we connect hosts and GRAPEs?
- How should we organize the memory architecture?

Software challenge

Major issues

- Fast algorithms with high accuracy?
  - High-order tree with individual timestep?
  - Ahmad-Cohen scheme?
- Softwares for MPI (or other message-passing model)

Summary

- **32 Tflops GRAPE-6 is up and running.**
- **Multiple-petaflops GRAPE-8 seems technically feasible, but we do need careful design studies.**