Linux clustering

Morris Law, IT Coordinator,
Science Faculty, Hong Kong
Baptist University

PII 4-node clusters started in 1999

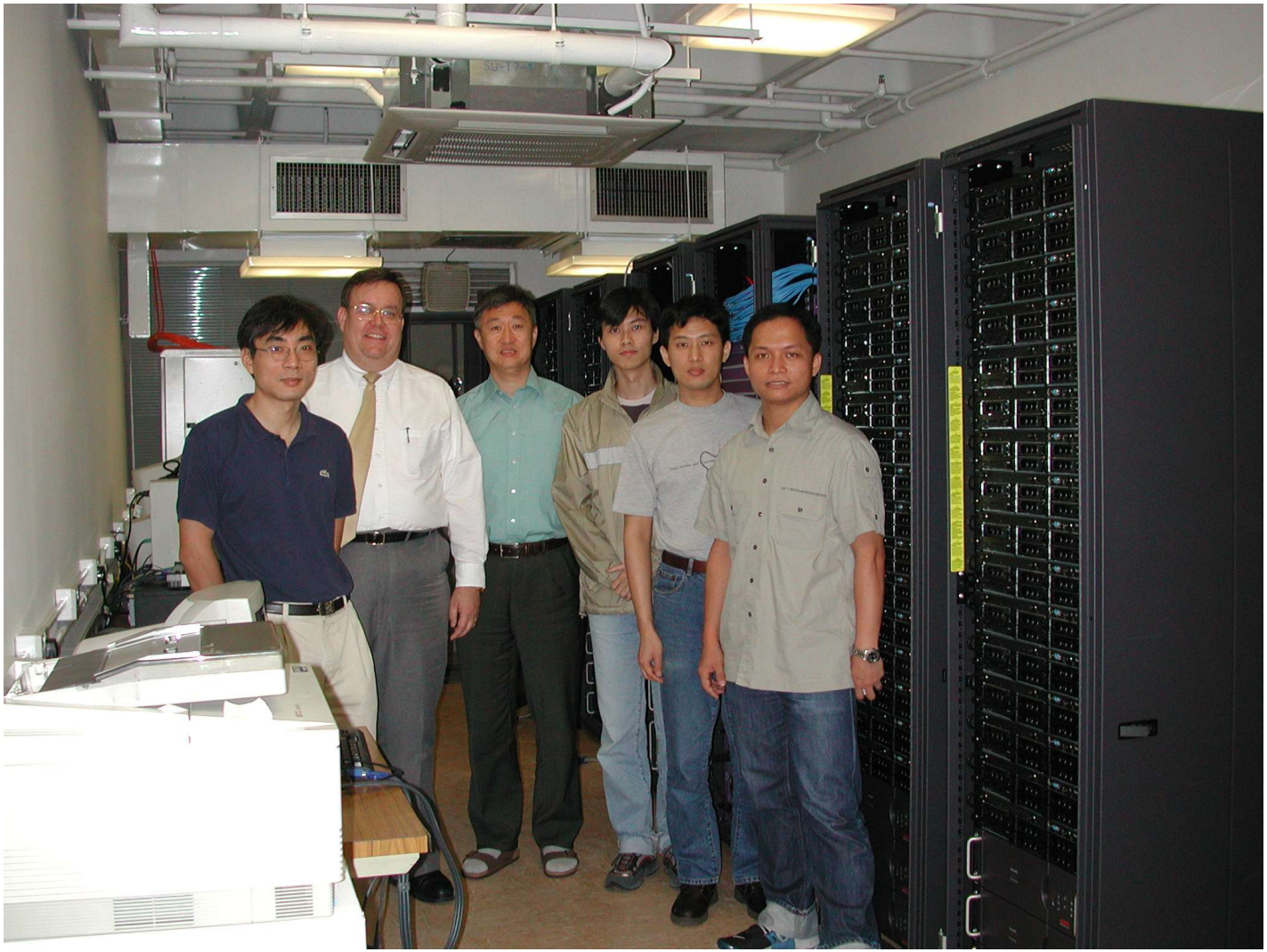
PIII 16 node cluster purchased in 2001.

Plan for grid

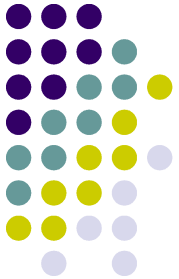
For test base

HKBU - 64-nodes P4-Xeon cluster at #300 of top500

The cluster management team

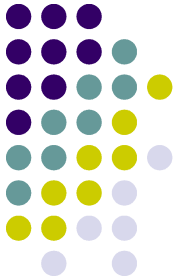


OUTLINE



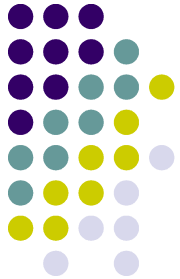
- What is PC cluster?
- Different kinds of PC cluster
  - Beowulf cluster example
  - SSI cluster example
- HPCC Cluster and parallel computing applications

What is a PC cluster?



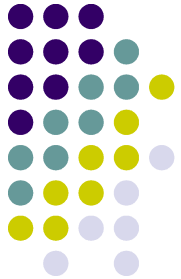
An ensemble of networked, stand-alone common-off-the-shelf computers used together to solve a given problem.

Different kinds of PC cluster



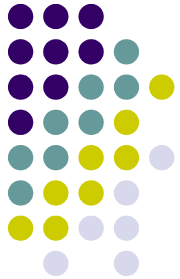
- High Performance Computing Cluster
- Load Balancing
- High Availability

High Performance Computing Cluster (Beowulf)



- Start from 1994
- Donald Becker of NASA assemble the world's first cluster with 16 sets of DX4 PCs and 10 Mb/s ethernet
- Also called Beowulf cluster
- Built from commodity off-the-shelf hardware
- Applications like data mining, simulations, parallel processing, weather modelling, computer graphical rendering, etc.

Examples of Beowulf cluster



Scyld Cluster O.S. by Donald Becker

http://www.scyld.com

ROCKS from NPACI

http://www.rocksclusters.org

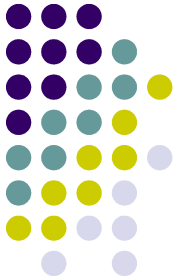
OSCAR from open cluster group

http://oscar.sourceforge.net

OpenSCE from Thailand

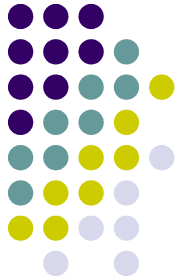
http://www.opensce.org

Load Balancing Cluster



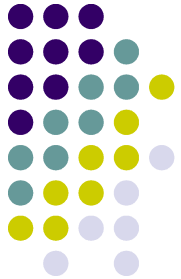
- PC cluster deliver load balancing performance
- Commonly used with busy ftp and web servers with large client base
- Large number of nodes to share load

High Availability Cluster



- Avoid downtime of services
- Avoid single point of failure
- Always with redundancy
- Almost all load balancing cluster are with HA capability

Examples of Load Balancing and High Availability Cluster



RedHat HA cluster

http://ha.redhat.com

Turbolinux Cluster Server

http://www.turbolinux.com/products/tcs

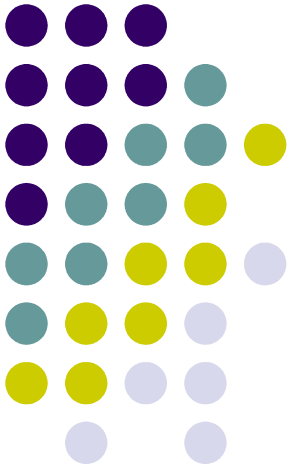
Linux Virtual Server Project

http://www.linuxvirtualserver.org/

Snapshots 1

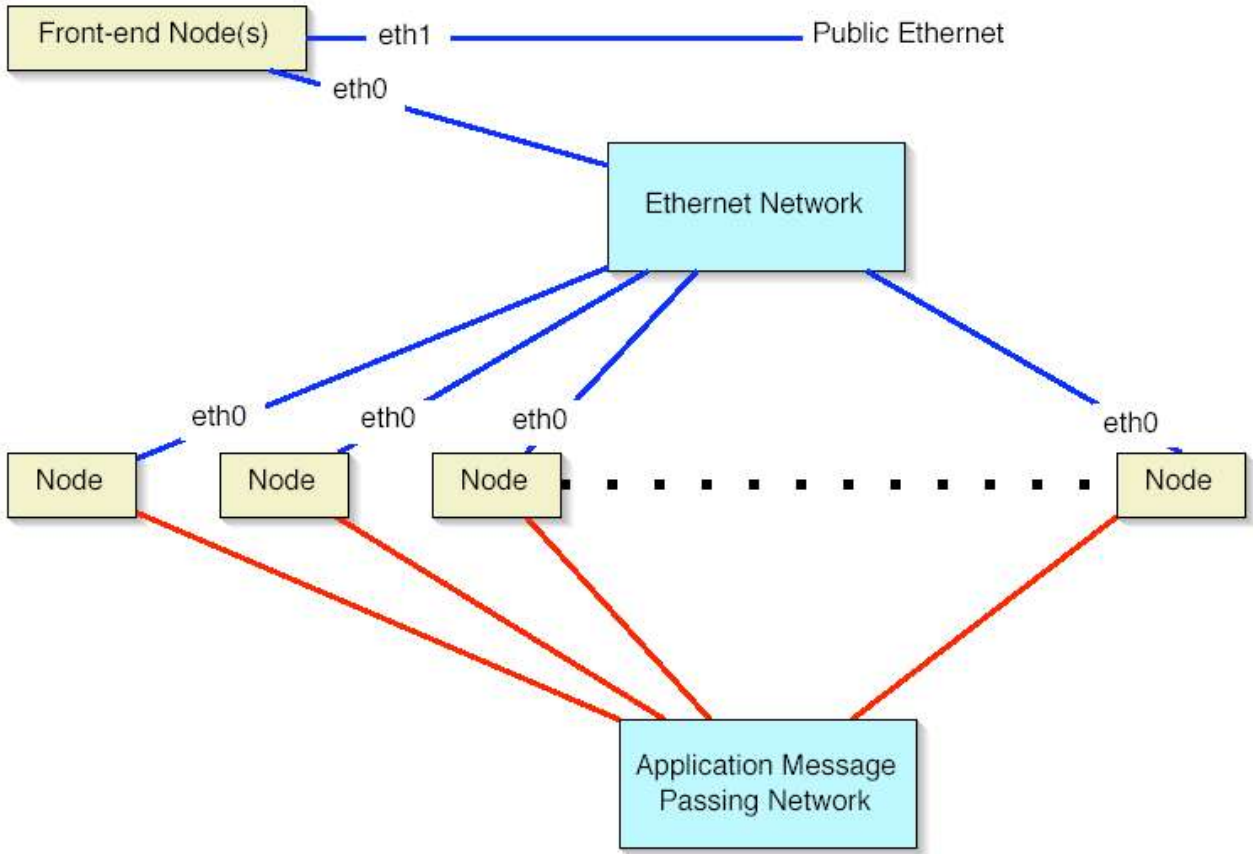
An example of Beowulf Cluster: ROCKS

(http://www.rocksclusters.org)



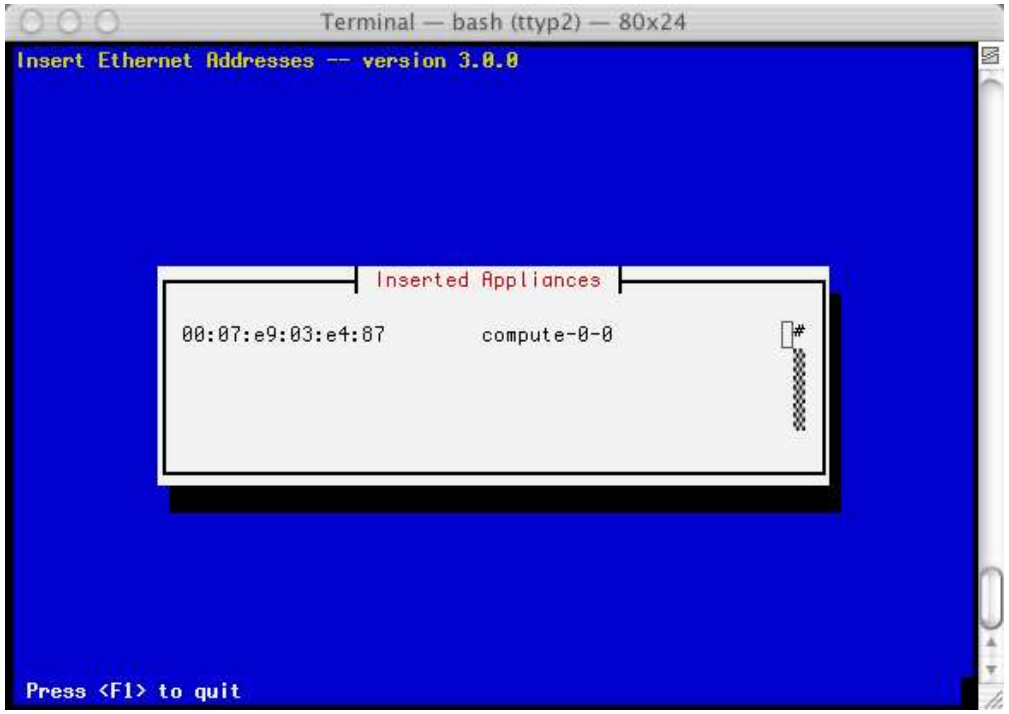
ROCKS SNAPSHOTS

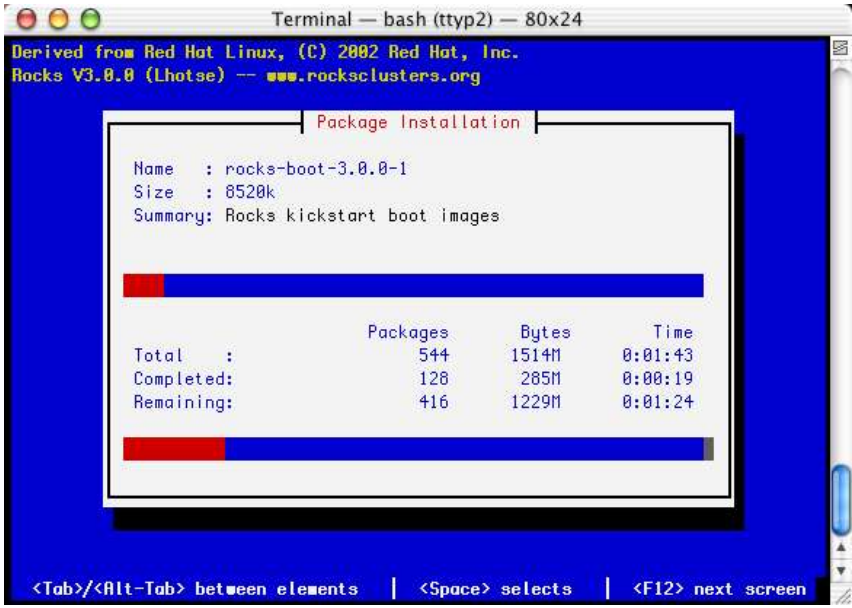
The schematic diagram of a rocks cluster



ROCKS SNAPSHOTS

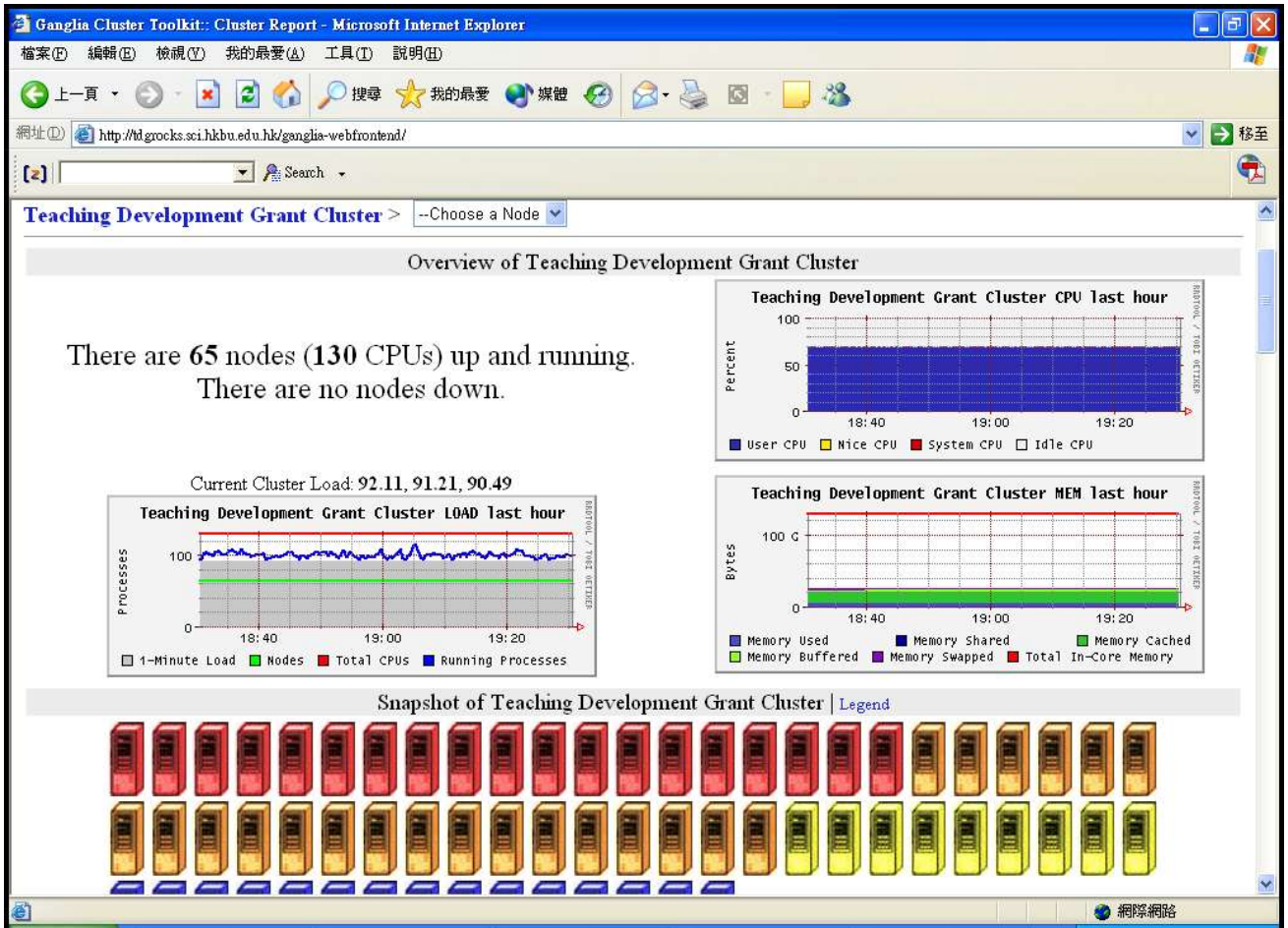
Installation of a compute node



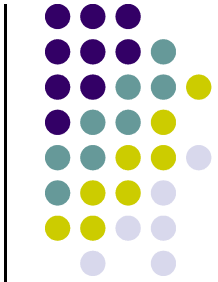


ROCKS SNAPSHOTS

Ganglia Monitoring tools

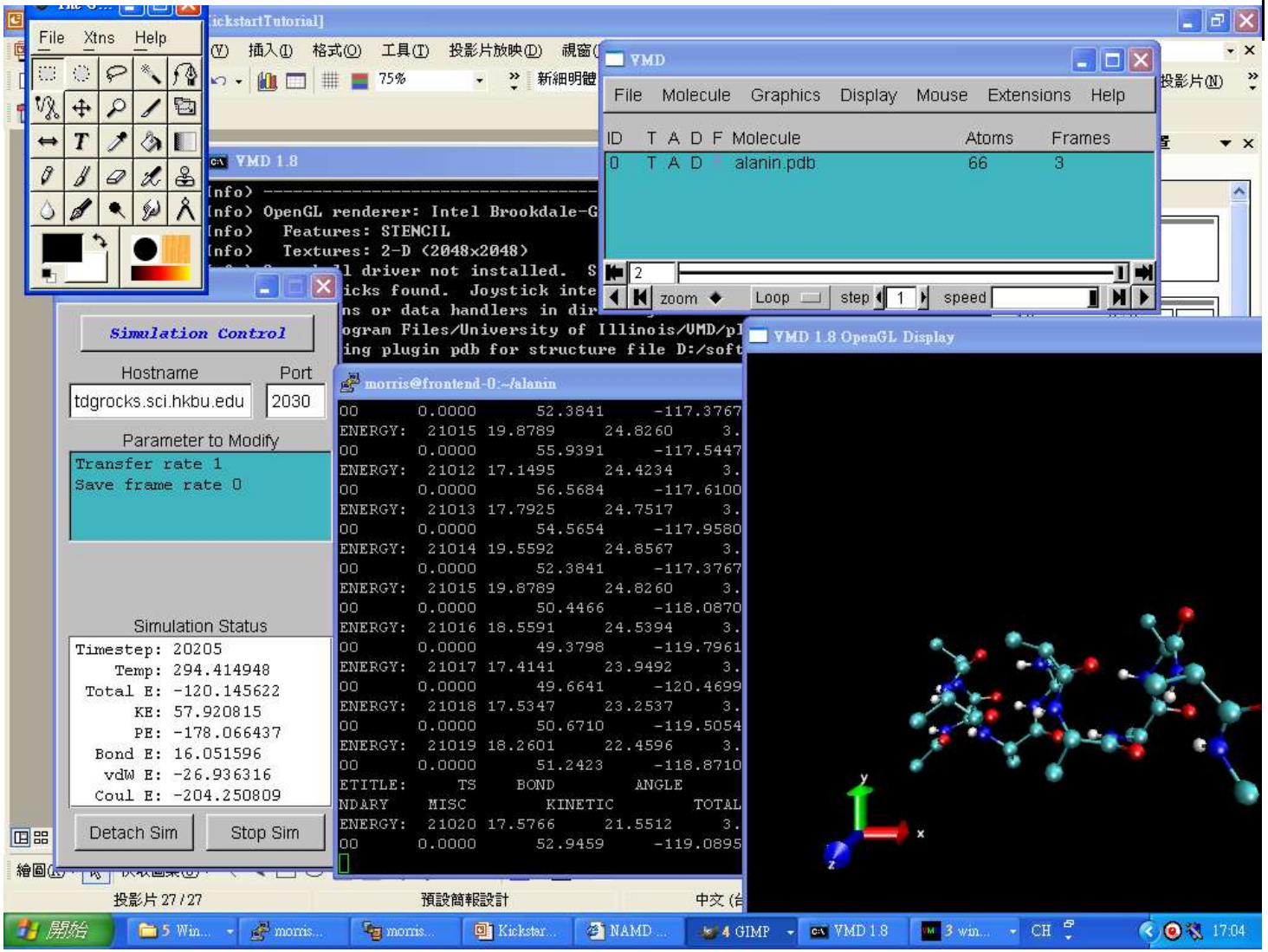


HPCC Cluster and parallel computing applications



- Message Passing Interface
  - MPICH (http://www-unix.mcs.anl.gov/mpi/mpich/)
  - LAM/MPI (http://lam-mpi.org)
- Mathematical
  - fftw (fast fourier transform)
  - pblas (parallel basic linear algebra software)
  - atlas (a collections of mathematical library)
  - sprng (scalable parallel random number generator)
  - MPITB -- MPI toolbox for MATLAB
- Quantum Chemistry software
  - gaussian, qchem
- Molecular Dynamic solver
  - NAMD, gromacs, gamess
- Weather modelling
  - MM5 (http://www.mmm.ucar.edu/mm5/mm5-home.html)

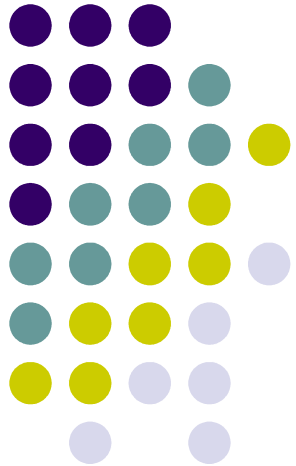
NAMD2 – Software for Quantum Chemistry



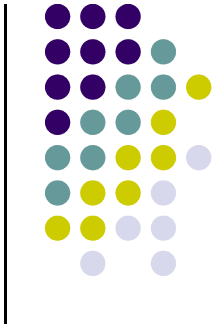
Single System Image (SSI) Cluster

MOSIX

openMosix



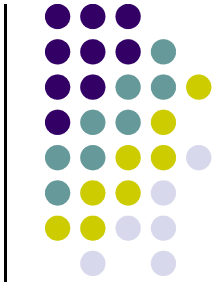
MOSIX and openMosix



MOSIX: MOSIX is a software package that enhances the Linux kernel with cluster capabilities. The enhanced kernel supports any size cluster of X86/Pentium based boxes. MOSIX allows for the automatic and transparent migration of processes to other nodes in the cluster, while standard Linux process control utilities, such as 'ps' will show all processes as if they are running on the node the process originated from.

openMosix: openMosix is a spin off of the original Mosix. The first version of openMosix is fully compatible with the last version of Mosix, but is going to go in its own direction.

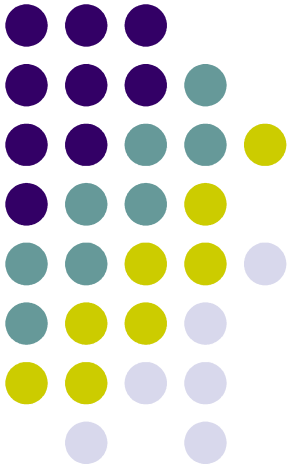
OpenMosix installation



- Install Linux in each nodes
- Download and install
  - openmosix-kernel-2.4.22-openmosix1.i686.rpm
  - openmosix-tools-0.3.4-1-RH80.i386.rpm
  - and related packages like thoses in
  - www.openmosixview.com
- Reboot with openmosix kernel

Screenshots 2

OpenMosix cluster management

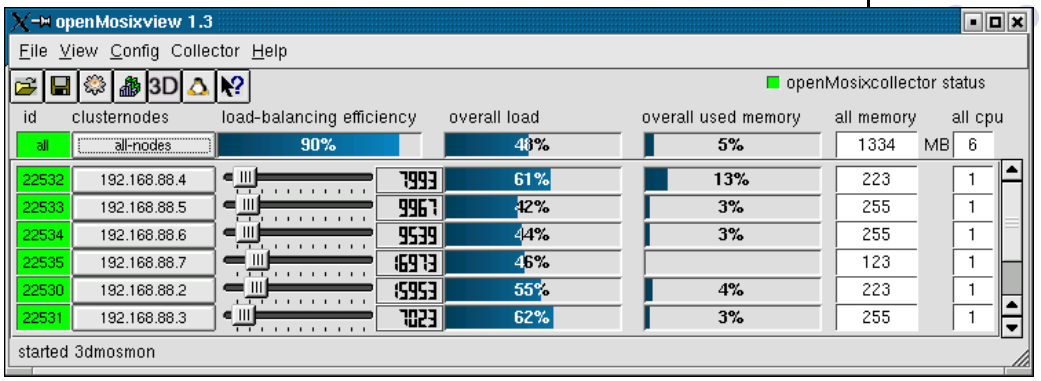


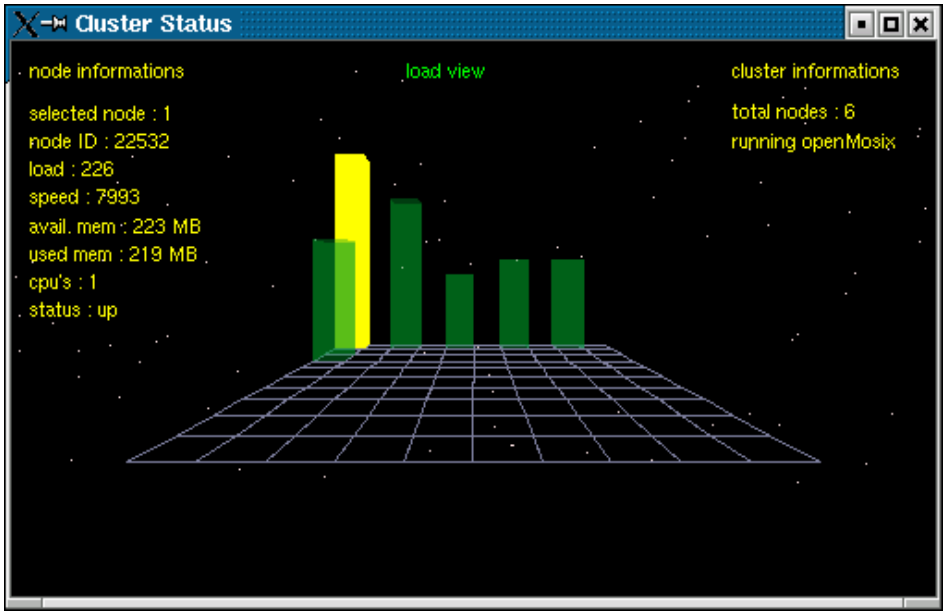
openMosix cluster management tools

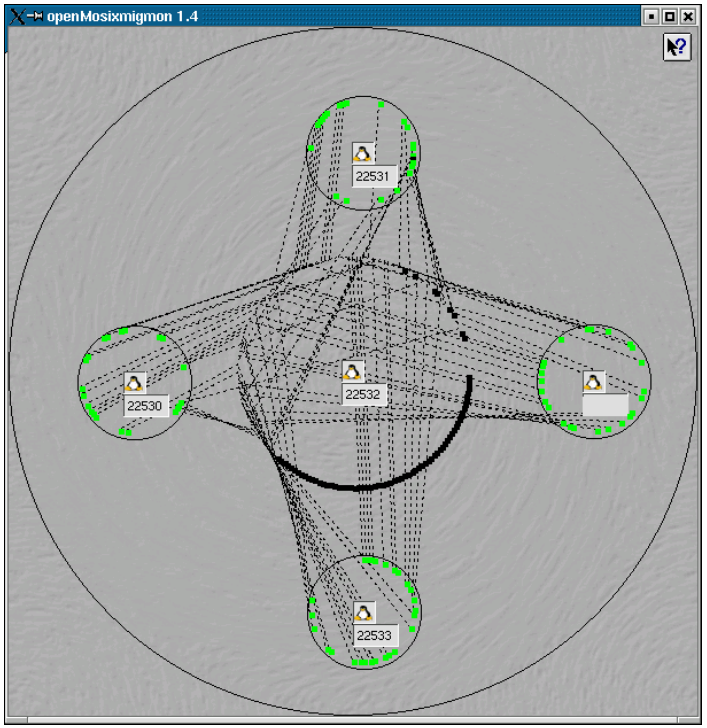
openMosixView

openMosixmigmon

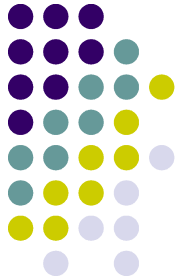
3dmosmon





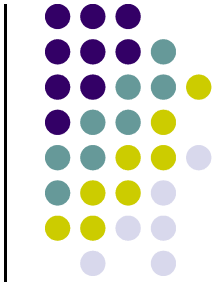


Advantage of SSI cluster



- Not need to parallelize code
- Automatic process migration, i.e. load balancing
- Add / delete nodes at any time
- Well aware of hardware and system resources

Reference URLs



Clustering and HA

Beowulf , parallel Linux cluster.

ROCKS from NPACI

OPENMOSIX , scalable cluster computing with process migration

HKBU Science Faculty Cluster

High Performance Cluster Computing Centre S

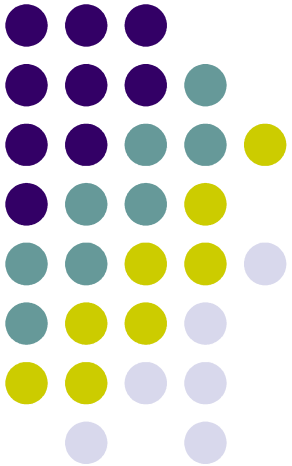
Linux Cluster Information Center

Thank you!

Welcome to visit HPCCC, HKBU

http://www.hkbu.edu.hk/hpccc/

http://www.hkbu.edu.hk/tdgc/

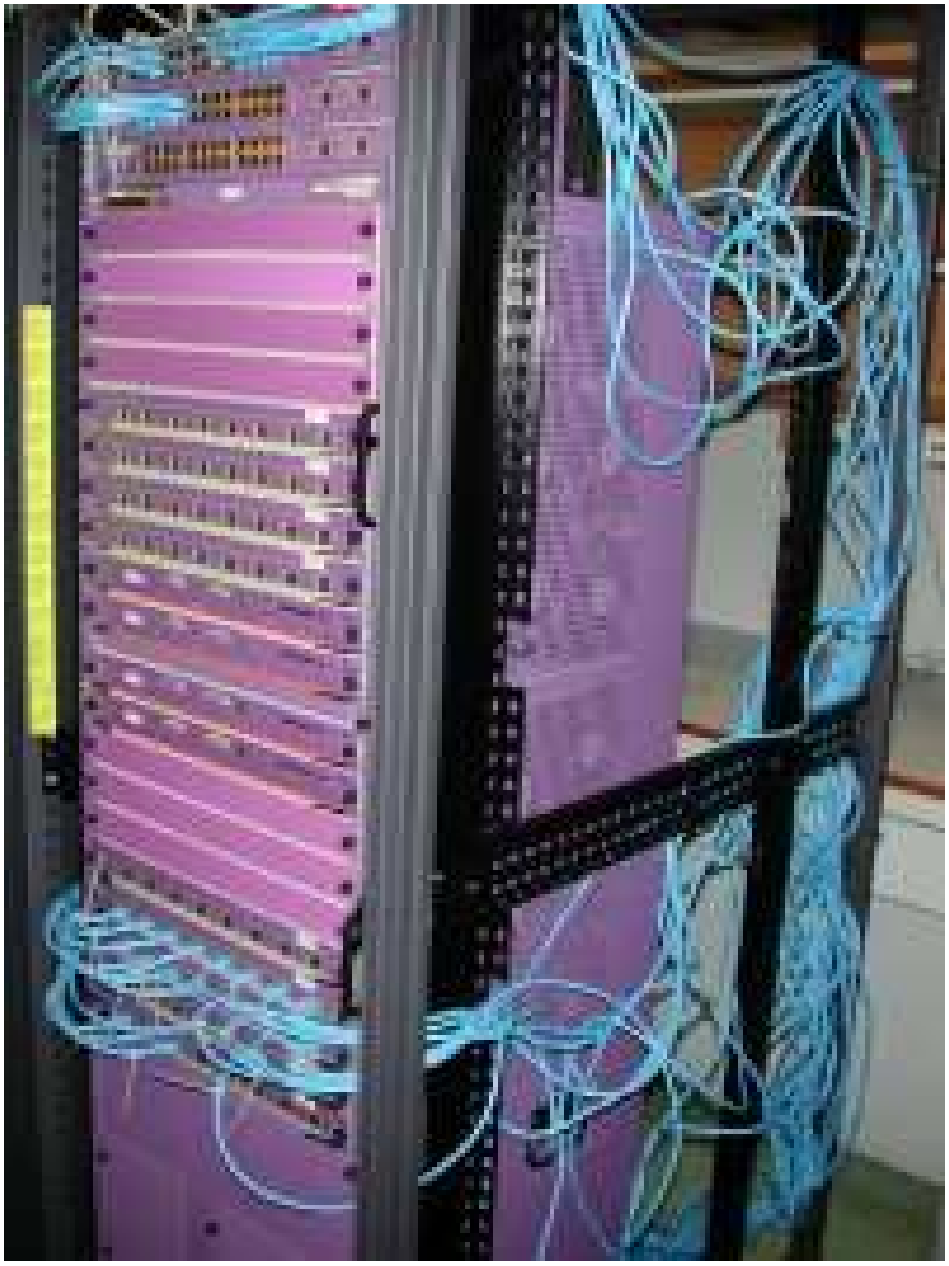


The Scientific Computing Lab.

Opening of the High Performance Cluster Computing Centre Supported by Dell and Intel

TDG cluster configuration

- Master node:
  - DELL PE2650 P4 Xeon 2.8GHz x 2
  - 4GB ECC DDR RAM
  - 36GB x 2 internal HD running RAID 1 (mirror)
  - 73GB x 10 HD array running RAID 5 with hot spare
- Compute nodes x 64 each with
  - DELL PE2650 P4 Xeon 2.8GHz x 2
  - 2GB ECC DDR RAM
  - 36GB internal HD



Interconnect configuration

Extreme BlackDiamond 6816 Gigabit ethernet switch