

Cluster Evolution at the
GENOME SCIENCES CENTRE

Bernard Li
bli@bcgsc.ca



Genome Sciences Centre

- Canada's Michael Smith Genome Sciences Centre
- Part of BC Cancer Research Centre
- High-throughput DNA sequencing and mapping
- Other research includes bioinformatics, proteomics, gene expression, gene regulation, cancer genetics
- 100+ node HPC
 - Intel PIII, Xeon and AMD Opteron
- 40TB storage
 - IBM SAN, NetApp, other NFS storage
- Linux and Windows Environment
 - RedHat and SUSE
 - Windows



Introduction

- Started with 100 IBM x330s (dual PIII 1.4GHz) with commodity interconnect
- Named *xofy* to facilitate displaying of monitoring information in a matrix
- No batch system, users `rsh/rlogin` to execute tasks
- Nodes with multiple jobs running -> Users need to waste time looking for free nodes to run their jobs -> Need better system!
- Cluster expansion – 64-bit Opteron
- LinuxNetworx's turnkey solution?
- Decided to buy more nodes and look for open source solution

The Beginning

- Perl script that uses PVM to execute batch jobs on least loaded nodes
- Evaluation of PBSPro -> still need other cluster software
- ROCKS? OSCAR?
- Mostly a RedHat shop (have some SUSE boxes running 64-bit)
- OSCAR can be installed on top of selected RedHat and Mandrake versions
- OSCAR is modular in the form of packages



HPC Infrastructure

- Hardware

- 70 Sun v20z dual Opteron 1.8GHz with 2.0GB RAM and 36GB SCSI HD
- 40 Appro AMD dual Opteron 2.0GHz with 2.0GB RAM and 80GB IDE HD
- 14 IBM BladeCenter dual Xeon 2.4GHz with 1.5GB RAM and 40GB IDE HD
- 80 IBM x330 dual PIII 1.4GHz with 1.0GB RAM and 18GB SCSI HD

- Software

- OSCAR v3.0
- SUN Grid Engine v5.3p5
- Ganglia v2.56
- BLAST (NCBI, Paracel)



HPC Infrastructure (cont'd)

- Interconnect
 - 100MBps and Gigabit Ethernet
 - 96-port HP Procurve switch
 - 24-port Cisco Catalyst 3750 switches
 - Cisco Catalyst 4500 Core routers
- Storage strategy
 - NetApp NAS960 Filer for NFS storage
 - /home and /scratch directories mounted on each client



Cluster jobs specification

- BLAST, gene prediction, physical map assembly, sequence assembly, regulatory elements prediction
- With new OSCAR cluster, users are encouraged to run intensive jobs on nodes and also migrate existing pipeline analyses to the HPC infrastructure
- Current system much more efficient than before
- SGE's Share-Tree Policy makes job scheduling fair
- Many short-running jobs
 - Home grown script: mqsub
 - Makes use of SGE's 'array job'



Creating packages for OSCAR

- Packages are RPM-ized programs with XML meta-file descriptor and corresponding install scripts
- OSCAR Package HOWTO and mailing-lists are good resources
- OSCAR Packages
 - Ganglia
 - Sun Grid Engine
 - OpenLDAP
 - Bioinformatics



Creating packages for OSCAR (cont'd)

- Ganglia Package
 - Version 2.5.6
 - Need to patch official Ganglia RPM (gmond)
 - Works with RedHat 9.0 and Mandrake 9.2
 - Available now from OSCAR Package Downloader for OSCAR 3.0
- Sun Grid Engine Package
 - Version 5.3p6
 - No official RPM available from Sun – uses third party RPM as basis
 - Work in progress – should be finished by end of June



Creating packages for OSCAR (cont'd)

- OpenLDAP Package
 - User authentication with OpenLDAP Server
 - Alternative to OPIUM which syncs user password information in OSCAR
- Bioinformatics Packages
 - Some work in integrating applications with Sun Grid Engine's Parallel Environment
 - FPC and Paracel BLAST
 - May package individual programs if interest arises



OSCAR-specific issues

- Upgrading OSCAR
 - No upgrade path
 - Overhead might be too great to bring down cluster from production just to upgrade
 - Needs to be addressed for wider adaptation
- New Distribution Support
 - RedHat's free workstation distribution is EOL
 - Mandrake is most supported distribution after RedHat
 - Debian IA-64 port in development
 - Fedora port also in development



Future Work

- Use mpiJava to bridge our Java applications with HPC resource
- Vast majority of jobs will still be embarrassingly parallel – nature of bioinformatics analyses (run many instances of same program with different inputs)
- Linking Bioinformatics grids (OSCAR-enabled?)



Conclusion

- OSCAR has made life easier for users and administrators of the cluster
- Quick and simple to install – powerful deployment, administration tools and programming libraries
- Encourage more users to use OSCAR and to contribute in:
 - Creating more packages
 - Core development
 - Testing



Acknowledgements

- Mark Mayo, Asim Siddiqui and Steven Jones
- HPC developers and users at the GSC
- OSCAR developers and users

