



## Demonstration of Mgrid cluster installation

13.8. Olli-Pekka Lehto

8/19/2004



## Overview of the Mgrid topology

- Front node (FN)
    - User logins, job submission
    - Interactive testing
  - Compute node (CN)
    - Batch job processing
    - No user logins
  - Management node (MN)
    - Console access to other nodes
    - Admin access from CSC
    - Possible backups
    - Restricted access
- CSC:
- Test cluster
    - Testing of software before deployment
  - Distribution server
    - ISO images, sources
    - RPM updates
    - Possibly also
      - CVS
      - network install

8/19/2004



## Overview of Rocks Linux

- Design philosophy
  - Consists of a base distribution and add-on Rolls (i.e. SGE)
  - Relies on RPM and a set of XML files for configuration
  - Nodes can be cleaned up and new ones added quickly
- Based on Red Hat Enterprise Linux 3.0
  - Well-supported among different vendors
- Issues:
  - 32-bit support is still ugly (Better system is coming soon)
  - Network install of front nodes is under construction
  - RPM system can be problematic at times
  - Making changes to the distribution can be complicated

8/19/2004



## Installation of the distribution to the Frontnode

1. Insert the CD marked Rocks base and boot
2. Type `frontend` in the bootloader
3. Insert the roll CDs when prompted
4. Fill in the cluster information
5. Choose manual partitioning and use the
6. Configure the private network (10.1.1.1, 255.255.255.0)
7. Configure the public network
8. Configure the NFS network (10.2.1.1, 255.255.255.0)
9. Setup the Gateway and DNS to point to the **public network's** GW and DNS
10. Insert the complete (FQDN) name of the (i.e. csc-mgrid.csc.fi) FN
11. Type the root password as agreed with CSC
12. Installation will commence and the installer will ask for the roll CDs
13. Machine will reboot and installation is done

8/19/2004



## Installation of compute nodes

1. Run `insert-ethers` in the front node
2. Select 'Compute' from the list of options
3. Power up the compute nodes one by one in starting from compute-0-0
4. The `insert-ethers` will detect each one and add it to the 'Inserted appliances list in the frontnode

8/19/2004



## Base and HPC

- RHEL distribution packages
- Essential Rocks packages
  - Rocks utilities
    - 411
    - Ganglia
  - Kickstart system
  - Cluster monitoring
  - Etc.
- MPI libraries
- ATLAS
- Benchmarks

8/19/2004



## Other rolls

- SGE
  - Sun Grid Engine batch queuing system
  - Sets up a default mpich –based environment
- AMD
  - A set 32-bit libraries obtained from a default i386 install
  - A temporary solution. Rocks maintainers are working on a better way to do this
- Grid
  - NSF Middleware (NMI) Globus, Condor, NWS, MDS etc.
  - Separate NorduGrid roll will probably be used
    - SGE integration of NorduGrid still under development

8/19/2004



## CSC Roll

- Support for multiple compiler/architecture combinations
- LAM and MPICH
- Portland Group Compilers
- ACML (AMD Core Math Library)
- SCALAPACK and BLACS
- Cscsu
- Autoupdate, Cfengine
- Additional 32-bit libraries
- Scripts
- Overrides for default Rocks-configuration
- Management node configuration

8/19/2004



## Adding users

- Simplest way is to use the `adduser` -command
  - This automatically initiates the 411 NIS system to distribute this change to all the compute nodes
- You can also include your own user management system
  - After modifying the `passwd`, `shadow` and `group` files you should run `/etc/rc.d/init.d/411 commit`
  - `Cscadm -user` should be kept in the files

8/19/2004



## Adding software

- Adding a package to all nodes (locally)
  - RPMs are the preferred way to install software to every node
    - Frontnode
      - `rpm -ivh`
    - Compute nodes
      - RPM: `/home/install/contrib/enterprise/3/public/AMD64/RPMS`
      - Conf: `/home/install/site-profiles/3.2.0/nodes/skeleton.xml`
      - Parse the distribution using `rocks-dist dist`
    - Additional software should reside under the `/opt` tree
- A simpler way is to install software under the `/home` directory
- The distribution server will contain a more complete set of RPMs and source files than the basic distribution.
- Make backups and contribute your work!

8/19/2004



## Paths

- Rocks distribution is in `/home/install`
  - `rocks-dist` parsed version of RPMs for the compute node
    - Installer
    - RHEL base
    - Additional rocks packages
    - Rolls
    - XML rules
- Directories under `/home` are automounted
- `/opt/` is the preferred path for local installation of software

8/19/2004



## Some important Rocks utilities

- `rocks-dist`
  - Parse the final distribution
  - Add rolls to the installation
- `411`
  - NIS system
  - Used to distribute configuration files (i.e. `/etc/passwd`) to compute nodes
- Cluster database
  - MySQL database containing information on the cluster
  - `Dbreport` generates configuration files (i.e. `/etc/hosts`)

8/19/2004