


Title goes here


**CS 594 Spring 2006**

Cluster Computing  
G E Fagg

Class 2: A very basic intro to making a cluster




**INNOVATIVE COMPUTING LABORATORY**  
COMPUTER SCIENCE DEPARTMENT  
UNIVERSITY OF TENNESSEE



**Overview**

- Cluster configuration
  - Headless, exposed or enclosed with a head
- Networking
- OS installing
  - Large scale OS installing
- File systems
- Batch systems
- Monitoring


1/19/2006 6:56 PM G.Fagg Cluster Computing 2



**Overview**

- Cluster configuration**
  - Headless, exposed or enclosed with a head
- Networking
- OS installing
  - Large scale OS installing
- File systems
- Batch systems
- Monitoring

1/19/2006 6:56 PM G.Fagg Cluster Computing 3



**Cluster configuration**

Assume we are running a standard operating system

- We need to make sure we can customize each install to match the task at hand
  - 3 target workloads

Assume we have a basic network, and a fast network


- The basic network needs to use a well known protocol
  - TCP/IP UDP/IP etc
- To support installing, logging in and maintenance

We have many machines can vary the hardware

- We can price match the systems to the work load (hence your h/w)

We are somehow going to be connected to the global internet somehow


1/19/2006 6:56 PM G.Fagg Cluster Computing 4



**Overview**

- Cluster configuration
  - Headless, exposed or enclosed with a head**
- Networking
- OS installing
  - Large scale OS installing
- File systems
- Batch systems
- Monitoring

1/19/2006 6:56 PM G.Fagg Cluster Computing 5



**Headless, exposed or enclosed with a head**

2 main decisions

- One: can all the nodes be seen from outside the cluster?
- Two: is one of the NODES special and less restricted?

1/19/2006 6:56 PM G.Fagg Cluster Computing 6

Title goes here

**Headless, exposed or enclosed with a head**

2 main decisions

One: can all the nodes be seen from outside the cluster?

This is a decision based on usage, h/w and s/w configuration, security etc

Can be answered by the users need to access the nodes directly

Workload

- Interactive: yes need to access the nodes from the outside
- HPC: maybe, depending on how debuggers and visualization tools work
- HPC production: no

1/19/2006 6:56 PM G.Fagg Cluster Computing 7

**Headless, exposed or enclosed with a head**

2 main decisions

One: can all the nodes be seen from outside the cluster?

1/19/2006 6:56 PM G.Fagg Cluster Computing 8

**Headless, exposed or enclosed with a head**

Advantages of an enclosed cluster

Security

- You can limit who logs into the system at only one point  
Firewall/gateway/head etc etc

Networking

- Naming  
You do not need to handle external addressing for all your hosts
- You can control traffic more effectively at a single point (as above)

Access management

- You can control how users log into the system  
Switch which single user has access today  
Speed up some h/w by not needing to check uids and ACL
- Helps stop them bypassing queuing systems

1/19/2006 6:56 PM G.Fagg Cluster Computing 9

**Headless, exposed or enclosed with a head**

Disadvantages of an enclosed cluster

Networking

- Naming  
You do not need to handle external addressing for all your hosts  
But you will need a gateway or NAT of some kind
- From outside most things will require two steps to access a node

1/19/2006 6:56 PM G.Fagg Cluster Computing 10

**Headless, exposed or enclosed with a head**

Advantages of an exposed cluster

Networking

- You can access any node directly when ever you want
- S/w does not have to handle problems with NAT systems

Access management

- User accounts can be same everywhere, no special configurations
- Can still be under batch queue systems  
But you might have to disable some system services

Configuration

- basically the same as any NOW but you can use short cuts to install all machine together quicker

1/19/2006 6:56 PM G.Fagg Cluster Computing 11

**Headless, exposed or enclosed with a head**

Disadvantages of an exposed cluster

Networking

- You need to configure an address for each node  
Large cluster might need a Class B subnet or many class Cs  
Make sure this cluster does not flood the rest of the normal network

Security

- Each node must be made secure  
Same as NOW

Access management

- Can still be under batch queue systems  
But you might have to disable some system services to control this

Configuration

- basically the same as any NOW  
But many machines equals a big administration overhead

1/19/2006 6:56 PM G.Fagg Cluster Computing 12

Title goes here

### Headless, exposed or enclosed with a head

2 main decisions

Two: is one of the NODES special and less restricted

Can make one of the nodes special:

Head node

This node is then configured like a server machine and the other are clients

Depending on network this can be done on an enclosed or exposed cluster

1/19/2006 6:56 PM G.Fagg Cluster Computing 13

### Headless, exposed or enclosed with a head

2 main decisions

Two: is one of the NODES special and less restricted

Fire-wall etc

Head

enclosed

Outside user

1/19/2006 6:56 PM G.Fagg Cluster Computing 14

### Headless, exposed or enclosed with a head

2 main decisions

Two: is one of the NODES special and less restricted

Log-in

Head

Job start

exposed

Debug directly

Outside user

1/19/2006 6:56 PM G.Fagg Cluster Computing 15

### Headless, exposed or enclosed with a head

Important point about a Head node is that it is a server of some kind for the cluster

File server:	provides file services
Access point:	for enclosed clusters
Image server:	provides the OS images for the cluster nodes
Name service:	provides IP addresses etc for cluster nodes
Gateway/NAT/Networking:	provides access to the outside world for enclosed clusters
Login server:	log in location for course grain serial/parallel interactive clusters
Submission server:	for batch queuing system

1/19/2006 6:56 PM G.Fagg Cluster Computing 16

### Headless, exposed or enclosed with a head

The HEAD node may have differently configured operating system or even a different operating system completely

Cluster nodes may themselves only have an operating system that meets their needs

1/19/2006 6:56 PM G.Fagg Cluster Computing 17

### Overview

- Cluster configuration
  - Headless, exposed or enclosed with a head
  - Networking
  - OS installing
  - Large scale OS installing
  - File systems
  - Batch systems
  - Monitoring

1/19/2006 6:56 PM G.Fagg Cluster Computing 18

# Title goes here

**Networking**

Networking

Three parts:

- High speed connectivity (if applicable)
- Naming of nodes, local networking (if applicable)
- Head node itself (if applicable)

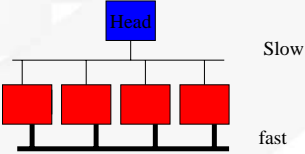
1/19/2006 6:56 PM G.Fagg Cluster Computing 19

**Networking**

Networking

High speed connectivity (if applicable)

Some systems will require a high speed network  
High IO throughput, HPC etc



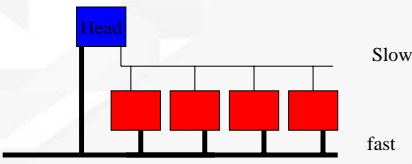
1/19/2006 6:56 PM G.Fagg Cluster Computing 20

**Networking**

Networking

High speed connectivity (if applicable)

Question, should the head node be on the fast network as well?



1/19/2006 6:56 PM G.Fagg Cluster Computing 21

**Networking**

Networking

High speed connectivity (if applicable)

Question, should the head node be on the fast network as well?

Advantage

- The head can then join in HPC operations
- Can install and move data faster

Disadvantages

- do you really want NFS running over you high speed interconnect?
- What happens with 256 nodes + head node with a 256 port switch ?

1/19/2006 6:56 PM G.Fagg Cluster Computing 22

**Networking**

Networking

Three parts:

- Naming of nodes, local networking
- If you have many nodes you can install each node and make sure they all know what their individual name is  
/etc/hosts and hostname
- Else  
You can configure some server (i.e. the head node) to do this for you via DHCP  
Head node is a server  
Cluster nodes are clients (with different MAC addresses)

1/19/2006 6:56 PM G.Fagg Cluster Computing 23

**Networking**

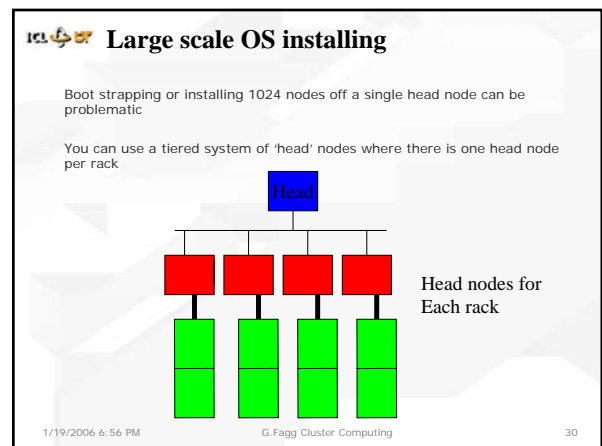
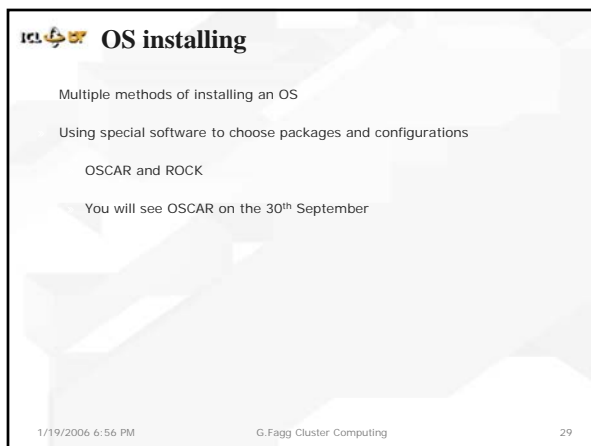
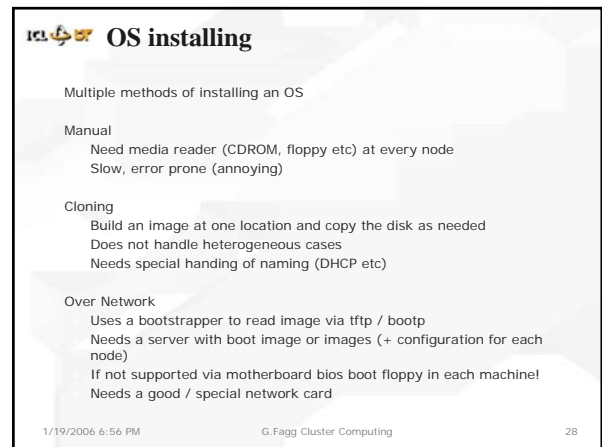
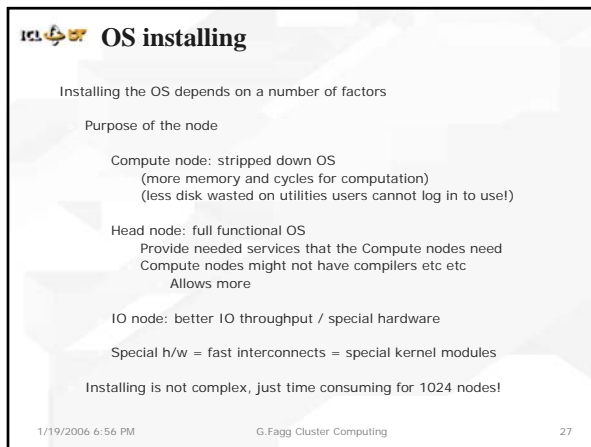
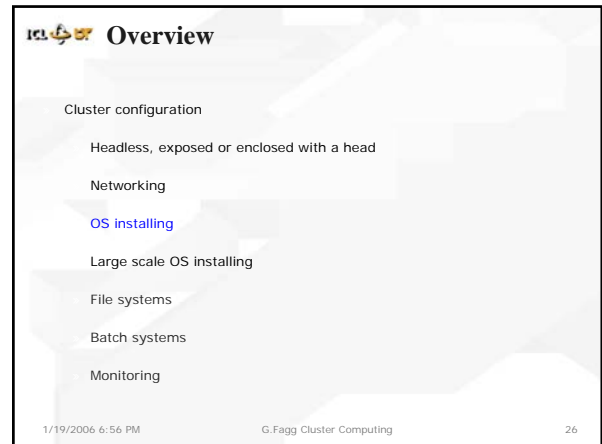
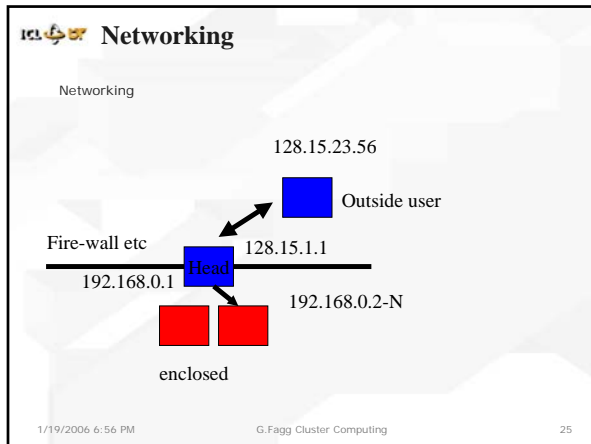
Networking

Three parts:

- Head node itself
- Will usually need to have two network connections  
One to the external world and one to the cluster
- If the cluster is enclosed and you want cluster nodes to talk BACK to the outside world (visualization) then the head node has to be configured as a gateway (with NAT)

1/19/2006 6:56 PM G.Fagg Cluster Computing 24

Title goes here



Title goes here

## Overview

- Cluster configuration
  - Headless, exposed or enclosed with a head
- Networking
- OS installing
- Large scale OS installing
- File systems**
- Batch systems
- Monitoring

1/19/2006 6:56 PM G.Fagg Cluster Computing 31

## File systems

There are a number of file systems available

Local:	per node	/ fast
RAID:	per node	/ faster
NFS:	distributed	/ slow cannot handle many simultaneous access
PVFS:	parallel	/ fast
PVFS2:	parallel	/ faster
Luster:	parallel	/ complex but very fast and scalable built on to of ext3

1/19/2006 6:56 PM G.Fagg Cluster Computing 32

## File systems

Big question

- How do users gain access to their files??

Individual file systems on each node

- Good for scratch temporary files
- Means lots of copying

NFS

- Good for users
- Bad for the network and NFS server when lots of concurrent users

1/19/2006 6:56 PM G.Fagg Cluster Computing 33

## Overview

- Cluster configuration
  - Headless, exposed or enclosed with a head
- Networking
- OS installing
- Large scale OS installing
- File systems
- Batch systems
- Monitoring**

1/19/2006 6:56 PM G.Fagg Cluster Computing 34

## Monitoring

Once I have the cluster running how do I check what it is doing?

- How many users?
- What is the load on the nodes?
- Has anything broken (yet)?

Can be done by monitoring systems

Ganglia: daemon on each node, sends output to one point but it makes very nice web pages and graphs

NWS: good between clusters and racks (also has cpu/memory probes)

PCP: network + node monitoring (open source) files and webpages output

1/19/2006 6:56 PM G.Fagg Cluster Computing 35