

# An introduction to clusters and high performance computing

Andreas Engelbrecht Dalsgaard

April 1st 2011

# What is a Cluster Anyway

- It is NOT something that does any of the following:
  - Use magic
  - Automatically make your program run faster
  - Provide a single virtual OS image of all resources
  - Always makes your software faster
- Then what is it?

# Terminology

- Cluster
  - A set of closely connected computers
  - Usually homogeneous
  - Connectivity: GB Ethernet, Infiniband
  - They usually run some form of \*nix
- High Performance Computing (HPC)<sup>1</sup>
  - "above a teraflop . . . floating-point operations per second"
  - "near the currently highest operational rate for computers"
  - "use custom-made components"
  - Usually use Infiniband as connectivity

---

<sup>1</sup><http://searchenterpriselinux.techtarget.com/definition/high-performance-computing>

# Terminology

- Grid Computing
  - Virtual Supercomputer
    - Composed of several clusters
  - Distributed within organisation or globally
  - High-energy Physics E.g. CERN
- Cloud Computing
  - Software as a service
  - On demand resources
  - Cloud storage
  - Commodity hardware components
  - Usually virtual access to resources

# So what is a cluster used for and by whom

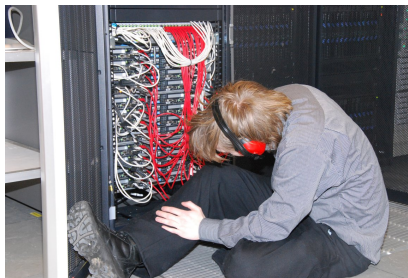
- Purpose

- To solve problems faster than on a single machine
- To solve problems that cannot be solved on a single machine
- Performance is everything
  - Users want the last 10%
  - Think about how the cache is used
  - Think about memory organisation
  - Consider communication overhead

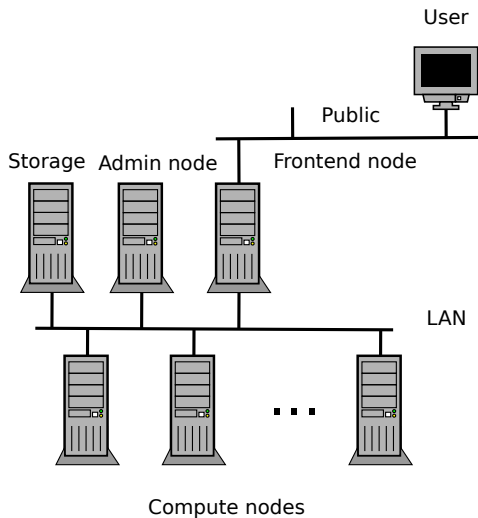
- Users

- Scientific researchers
- Engineers
- Academic institutions
- Government agencies E.g. military

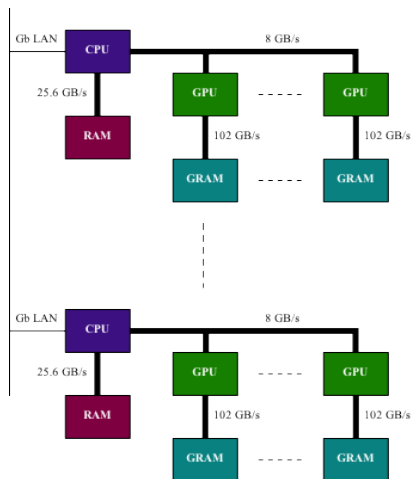
# What does it look like



# Cluster Overview



# Cluster Overview(2)





# How are clusters used

- A LRMS: Local Resource Management System
- Also called batch systems
  - Tasks are split into jobs
  - Jobs are executed by the batch system
  - Order of execution depend on scheduler
  - A job gets a set of nodes/CPU's (exclusively)
- How are most LRMS used
  - Make a job description
  - Submit it to the LRMS
  - The LRMS will decide when to run the job
- Torque/PBS, SLURM, LoadLeveler, Condor
  - Some are open source, some are proprietary

# What Runs a Cluster

- OpenSSH is used intensively
- Admin node
  - A LRMS server
  - OpenLdap server(user directory service)
  - Application file server (/pack)
    - Compilers, Gnu and Intel c,c++ and fortran
    - Libraries, E.g. openmpi
    - Firewall
  - Dhcp, tftp, email, mailinglist, monitoring, monitoring and monitoring
- Frontend node
  - LRMS client
  - Mount application filesystem
  - Firewall

# What Runs a Cluster(2)

- Compute nodes
  - LRMS client
  - Mount application filesystem
  - /scratch partition
  - Everything else is read only
- Storage
  - User data
  - Automounted on all nodes
  - Battery backed RAM, SSD, disks
  - No guaranties, use backup

# A Little About Hardware

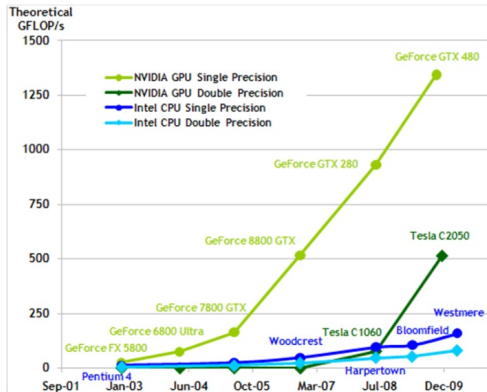
- For common jobs
- 84 IBM blade computers in 6 bladecenters.(killing1-84) 672 cores, 1,3 Tb
  - With two Xeon E5345 quad core 2.33 GHz CPUs, 16 GB RAM, a 53 GB scratch partition, Gbit ethernet and Infiniband interconnect.
- 24 HP computers (lion1-24) 288 cores, 3,5 Tb RAM
  - With two Xeon X5650 six core 2.66 GHz CPUs, 145 Gb RAM, a 53GB scratch partition, Gbit ethernet and Infiniband interconnect.
- Multi-core CPUs is the norm
- This is why this course is important

# A Little About Hardware(2)

- Test jobs
- 5 SUN computers (sister1-5)
  - With a Xeon X3220 quad core 2.40GHz CPU, 8GB RAM, a 66 GB scratch partition, Gbit ethernet interconnect
- Interactive jobs
- 4 SUN computers (tiger1-4)
  - With two Xeon X5570 quad core 2.93GHz CPU with hyperthreading(HT) enabled, 68GB RAM, Gbit ethernet interconnect.
- GPU jobs
- 10 Colfax computers (cub1-10)
  - With two Xeon X5570 quad core 2.93GHz CPU with HT enabled, 48GB RAM, Gbit ethernet and Infiniband interconnect, 3xTesla C2070 or Tesla C1050(soon GTX 580)

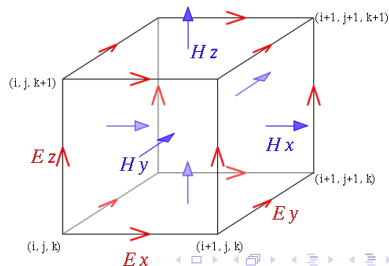
# A Little About Hardware(3)

- Trends
  - GPU used for calculations
  - Massive parallelism
  - 10 gbit ethernet
  - This is why this course is important!



# What people use a cluster for

- Scout
  - Analysis of integrated circuits
  - Continues systems
  - E.g. Noise Analysis in integrated circuits
  - Mobile communication
- Fyrkat
  - Continues systems
  - Simulate an antennas behaviour
    - Antenna near field simulation using Finite-difference time-domain method
  - Acoustic simulations
  - Grid computing



# PART 2



# What I do

- Development
  - Resources trading project
  - Irmsurgen
  - Scripting
  - MultiAdm
  - Torque Power Management
- User management
  - More than 100 users
  - Many new user request
    - Verifying need for resources
    - Automate as much as possible

The screenshot shows a web browser window titled "NDCG A&U - Fyrkat - Mozilla Firefox". The address bar shows the URL "https://fyrkat.grid.aau.dk/ausersaccount/". The page content includes a navigation bar with links like "Home", "Getting Started", and "Latest Headlines". Below this is a form titled "To apply for an account on Fyrkat please fill out the form below. All fields are required for student, master students and Ph.D. students." The form contains several input fields: "Name", "Username", "Email", "Phone", "Nationality" (with a hint "e.g. dk"), and "Organization" (with a hint "e.g. aau"). There is a large text area for "Computational problem description". At the bottom, there is a dropdown for "Account type" (currently set to "Student"), a field for "Supervisor email" (with a note "Required for all forms of student account types."), and a "Submit" button.

# What I do(2)

- Mails back and forth
- Run adduser script
- Add user to groups
- Send sms with password
- Add user to scheduling groups
- Add user to `/etc/aliases`
- Add user to mailinglist
- Add user to shared google doc document
- Send mail with link to wiki

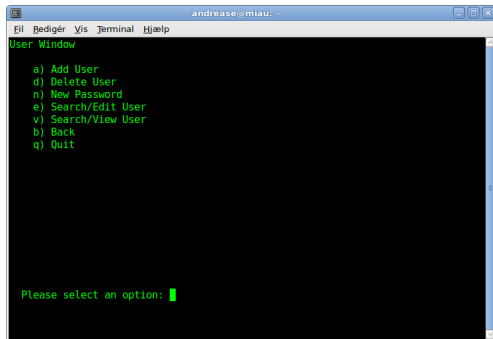
# What I do(2)

- Mails back and forth
- Run adduser script
- Add user to groups
- Send sms with password
- Add user to scheduling groups
- Add user to `/etc/aliases`
- Add user to mailinglist
- Add user to shared google doc document
- Send mail with link to wiki



# What I do(3)

- Automate as much as possible
- Standardise as much as possible
- Make tools to assist in common tasks



```
andrease@miau: ~  
File Redigér Vis Terminal Hjælp  
User Window  
a) Add User  
d) Delete User  
n) New Password  
e) Search/Edit User  
v) Search/View User  
b) Back  
q) Quit  
  
Please select an option: █
```

# What I do(3)

- Automate as much as possible
- Standardise as much as possible
- Make tools to assist in common tasks

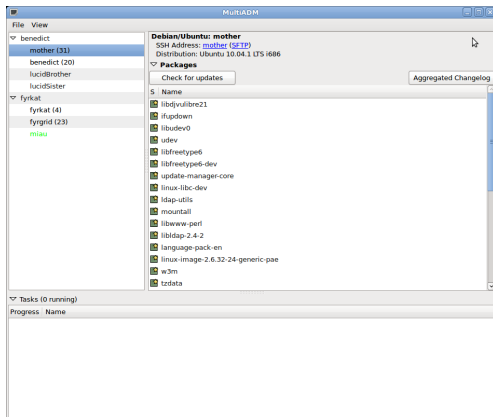
```

1  <?xml version="1.0" encoding="UTF-8"?><!DOCTYPE usercreationdata [
2  <ELEMENT usercreationdata (usercreationdata+)>
3  <ELEMENT usercreationdata (name,username,email,mobile,nationality,organization,expiredate,guid)>
4  <ELEMENT name (#PCDATA)>
5  <ELEMENT username (#PCDATA)>
6  <ELEMENT email (#PCDATA)>
7  <ELEMENT mobile (#PCDATA)>
8  <ELEMENT nationality (#PCDATA)>
9  <ELEMENT organization (#PCDATA)>
10 <ELEMENT expiredate (#PCDATA)>
11 <ELEMENT guid (#PCDATA)>
12 ]>
13 <usercreationdata>
14 <usercreationdata>
15   <name>Andreas Engelbrecht Dalsgaard</name>
16   <username>andrease</username>
17   <email>andreas.dalsgaard@gmail.com</email>
18   <mobile>60110868</mobile>
19   <nationality>dk</nationality>
20   <organization>aau</organization>
21   <expiredate>2015-01-01 10:12:12</expiredate>
22   <guid>/O=Grid/O=aau/OU=aau.dk/CN=Andreas</guid>
23 </usercreationdata>
24 </usercreationdata>
25

```

# What I do(3)

- Automate as much as possible
- Standardise as much as possible
- Make tools to assist in common tasks



# Why don't people use clusters

- Getting parallel program to perform well is immensely difficult
  - But challenges are fun!
  - The need for skilled multi-core programmers is raising
  - Apps using multi-core is becoming the norm
    - Not just in clusters
    - Also in desktops and smart phones

# How people use clusters

- Parameter sweeps
  - Serial code run in parallel
- Distributed jobs
  - Programs are made explicitly parallel
  - This is hard work
  - OpenMP, Pthreads, MPI
  - MPI is the hard part
    - Usually it is more of a rewrite than a port
  - Many MPI solutions for different interconnect types

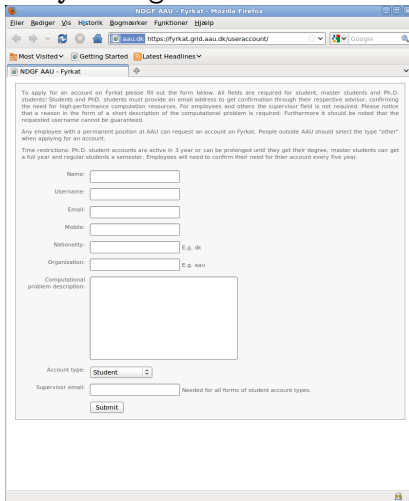


# Considerations when using a cluster

- Task generation
  - Static task generation
    - Matrix-multiplication
    - Parameter sweeps
    - State-space exploration E.g. 15-puzzle
  - Dynamic task generation
    - Ray tracing
    - State-space exploration E.g. 15-puzzle
- Task size
  - Uniform versus non-uniform
- Size of Data Associated with Tasks

# How to get an account

<https://fyrkat.grid.aau.dk/useraccount>



The screenshot shows a web browser window titled "NDGF AAU - Fyrkat - Mozilla Firefox". The address bar displays "https://fyrkat.grid.aau.dk/useraccount/". The page content includes a navigation menu with "Filer", "Bediger", "Vis", "Hjstori", "Bogmarker", "Funktioner", and "Hjlp". Below the menu, there are "Most Visited" and "Latest Headlines" sections. The main content area is titled "NDGF AAU - Fyrkat" and contains a form for creating a user account. The form includes a detailed instruction paragraph, a "Name:" field, a "Username:" field, an "Email:" field, a "Mobile:" field, a "Nationality:" field with a dropdown menu (showing "E.g. dk"), an "Organization:" field with a dropdown menu (showing "E.g. aaau"), a large text area for "Computational problem description:", an "Account type:" dropdown menu (showing "Student"), a "Supervisor email:" field with a note "Needed for all forms of student account types.", and a "Submit" button.

To apply for an account on Fyrkat please fill out the form below. All fields are required for student, master students and Ph.D. students! Students and Ph.D. students must provide an email address to get confirmation through their respective advisor, confirming the need for high-performance computation resources. For employees and others the supervisor field is not required. Please notice that a reason in the form of a short description of the computational problem is required. Furthermore it should be noted that the requested username cannot be guaranteed.

Any employee with a permanent position at AAU can request an account on Fyrkat. People outside AAU should select the type "other" when applying for an account.

Time restrictions: Ph.D. student accounts are active in 3 year or can be prolonged until they get their degree, master students can get a full year and regular students a semester. Employees will need to confirm their need for their account every five year.

Name:

Username:

Email:

Mobile:

Nationality:  E.g. dk

Organization:  E.g. aaau

Computational problem description:

Account type:

Supervisor email:  Needed for all forms of student account types.

# PART 3

# Fyrkat

- `ssh fyrkat.grid.aau.dk`
  - Pretty easy to get access to - Talk to your supervisor and use:
  - <https://fyrkat.grid.aau.dk/useraccount>



# Scout

- `ssh scout.es.aau.dk`
  - Hard to get access to - Talk to Torben Larsen or Josva Kleist



# Demonstration