



## Finnish Material Sciences Grid (M-grid)

Arto Teräs

[arto.teras@csc.fi](mailto:arto.teras@csc.fi)

Nordic-Sgi Meeting October 28, 2004



## Contents

- M-grid overview
- Partners
- Hardware and software
- System administration
- Grids and NorduGrid
- M-grid vs. Swegrid
- New challenges, open questions



## M-grid Overview

- Joint project between seven Finnish universities, Helsinki Institute of Physics and CSC
- Jointly funded by the Academy of Finland and the participating universities
- First large initiative to put Grid middleware into production use in Finland
- Coordinated by CSC, users and local site admins from the physics and chemistry labs of the partner universities
- Based on Linux clusters, targeted for serial and pleasantly parallel applications





## Partners

- Laboratory of Physics at Helsinki University of Technology
- Department of Physical Sciences and Department of Chemistry at the University of Helsinki
- Helsinki Institute of Physics (HIP)
- Departments of Physics and Chemistry at the University of Jyväskylä
- Department of Electrical Engineering at Lappeenranta University of Technology
- Department of Physics and the Institute of Advanced Computing at Tampere University of Technology
- Department of Physics at the University of Turku
- Departments of Biochemistry, Chemistry and Physical Sciences at the University of Oulu



## Hardware Acquisition

- Starting points: PC cluster, Gigabit Ethernet, Linux OS
- Joint acquisition project
  - Total value approximately 700 000 euros (incl. VAT)
  - Each group needed to come up with 25% of the money, 75% funded by the Academy of Finland
  - RFP April-June 2004
  - Vendor chosen July 2004 (Hewlett Packard)
  - Hardware delivered September 2004
- Oulu decided not to participate in the joint acquisition
  - Preferred the same hardware as in their existing cluster

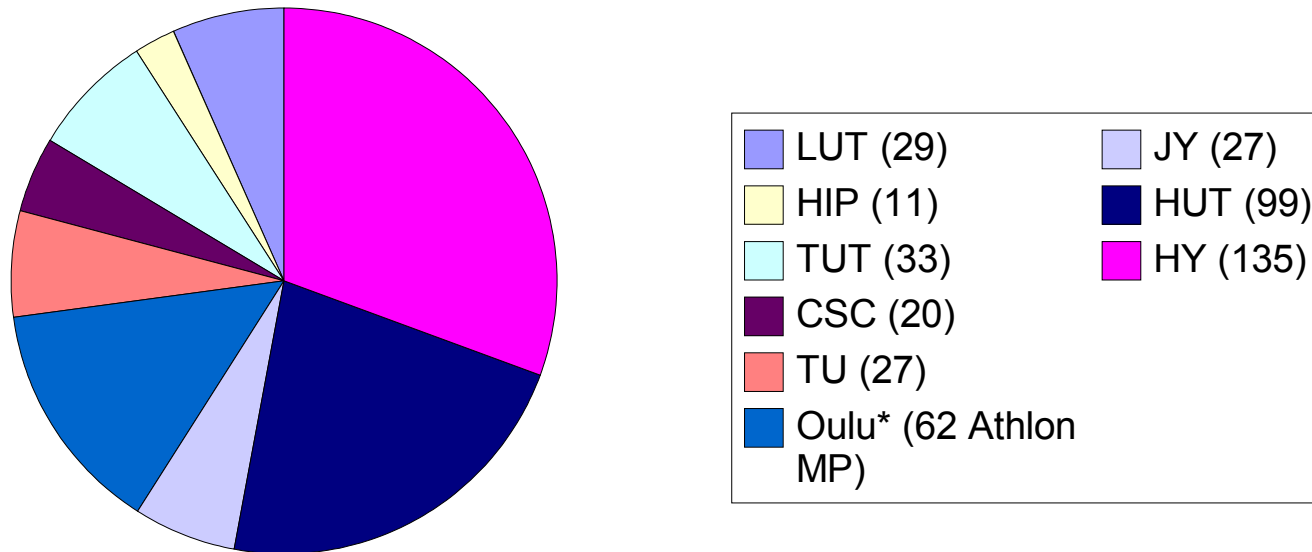


## Hardware

- Hardware of each cluster (7 sites)
  - Frontend: HP ProLiant DL585 (2 x Opteron 1.8-2.2 GHz, 2-8 GB RAM, upgradable up to 4 CPUs and 64 GB RAM)
  - Computing nodes: HP ProLiant DL145 (2 x Opteron 1.8-2.2 GHz, 2-8 GB RAM, 80-320 GB local disk)
  - Admin node: HP ProLiant DL145 (1 x Opteron 1.6 GHz)
  - Storage: SCSI RAID-5, typically 1 - 2 TB per site, cluster frontend used as NFS server for the nodes
  - Separate gbit ethernet networks for communication and NFS
  - Remote administration network and management boards
- University of Oulu (1 site): AMD Athlon MP based solution



## CPU Distribution



- Size of sites varies greatly
- Number of CPUs: 410 (computing nodes only, 443 total)
- Total theoretical computing power: 1.5 Tflops (CSC IBM SC 2.2 Tflops)



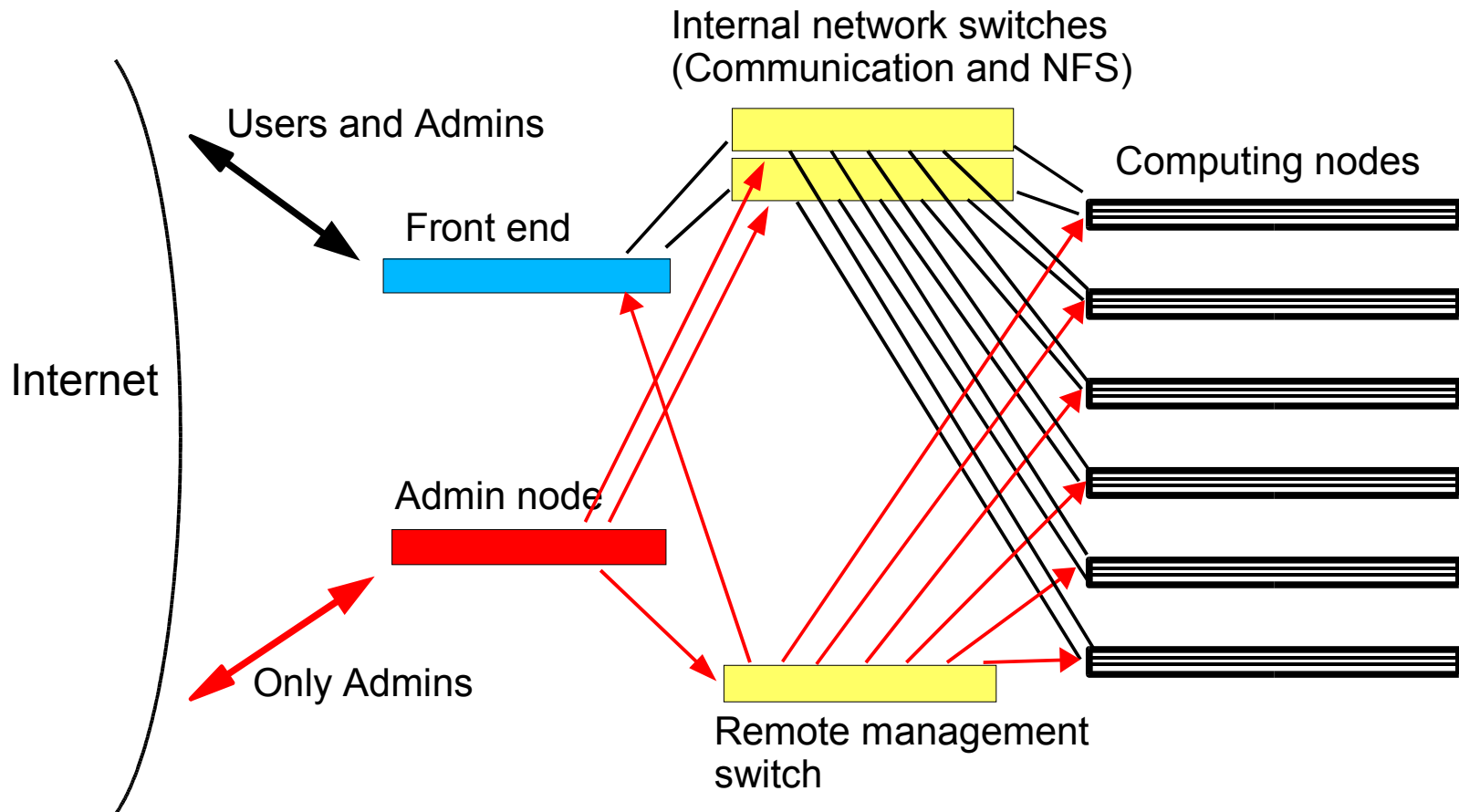
## Software

- Operating system: Rocks Cluster Distribution (<http://www.rocksclusters.org>)
  - Linux distribution designed for clusters, based on Red Hat Enterprise Linux 3.0 (but not an official Red Hat product)
  - Main developer San Diego Supercomputing Center (U.S.A)
  - Includes a set of open source cluster management tools
- Local batch queue system: Sun Grid Engine (<http://gridengine.sunsource.net/>)
- Grid middleware: NorduGrid ARC (<http://www.nordugrid.org>)
  - One of the few middleware packages which has already been used in production with real users and applications





## One M-grid cluster



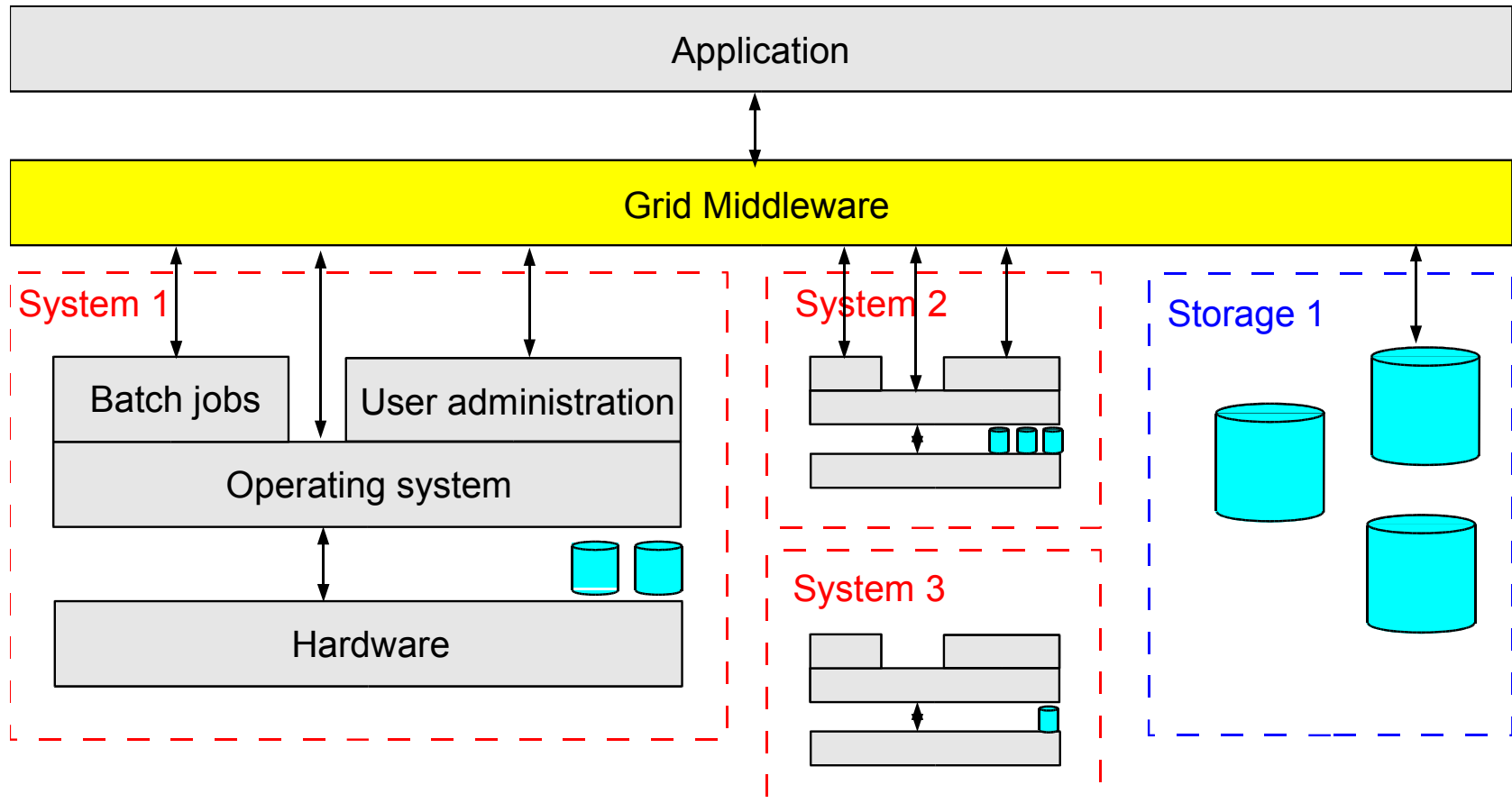


## System Administration

- Tasks divided between CSC and site administrators
- CSC
  - Maintains the OS, LRMS, Grid middleware, certain libraries centrally for all sites except Oulu
  - Provides tools for system monitoring, integrity checking, etc.
  - Separate mini cluster for testing new software releases
- Site administrators
  - Local applications and libraries, system monitoring, user support
  - Admins are typically working for the department or lab, not I.T.
- Regular meetings of administrators, support network



## Role of Grid Middleware





## Connecting Through Grids - What Changes?

- Users need personal certificates in addition to their local user account
- New interface(s) to learn, but the same can be used for heterogeneous platforms
  - library dependency problems are not magically solved by Grid  
=> one reason to maintain certain uniformity within M-grid
- Both the set of users and connected resources vary dynamically
- Grids go across **multiple administrative domains!**
- Collaboration and social networking, not only technology



## NorduGrid Collaboration



- Started in 2001, originally for connecting resources in Scandinavia and Finland
- Later joined by groups in Estonia, Germany, Slovakia, Slovenia and Australia
  - Currently about 5000 CPUs total (mostly Linux clusters)
- NorduGrid ARC middleware
- Open for anyone to participate
- <http://www.nordugrid.org>



## Facts on NorduGrid

- Production Grid available 24/7 since July 2002
  - Serves researchers and consists of academic resources
  - Built from "bottom to top", connecting already existing resources to the Grid
  - Real users and applications
- Resources do **NOT** need to be dedicated to the Grid
  - Also the case in M-grid: users have both a local user account for their own cluster and a Grid identity for the whole M-grid (and beyond...)
- Possibility to take just the middleware and run your own private Grid



## M-grid Compared to Swegrid

- Common characteristics
  - Approximately same scale (computing power, number of sites)
  - NorduGrid ARC middleware
- Differences
  - System administration
    - M-grid: Partly centralized, tasks shared between CSC and local sites
    - Swegrid: Handled by the individual sites, different Linux distributions being used
  - Resource dedication
    - M-grid: Users can submit jobs both locally and through Grid interface
    - Swegrid: Local job submission not allowed or deprecated



## New Challenges for CSC

- Geographically distributed system
  - Systems can be fully remotely administered
- Collaborative system administration
  - Needs careful planning, best practises are being formed
  - Security and trust issues need special attention
- Linux as a high performance computing environment new to CSC
  - The small cluster which CSC acquired as part of the project will be extended later and opened to all CSC customers
- Grid technology





## Current Status of M-grid

- Hardware delivered, operating systems installed, acceptance tests in progress
- User environment (compilers, MPI libraries) and system monitoring usable but needs work
- NorduGrid middleware not yet installed
  - A few problems related to 64 bit environment and Sun Grid Engine as local batch queue system, mostly solved
  - To be deployed Dec 2004 - Jan 2005 (estimate)



## Future

- Development of the Grid environment continues
  - Still large tasks ahead in security, user administration, accounting, user training and other areas
- Possibly opening our customized Rocks package to others (without central administration by CSC), contributing mathematical library packages to mainline Rocks
  - All software freely distributable with the exception of commercial Fortran compiler (Portland Group)
- Open questions:
  - Will the scientists start really using the Grid middleware?
  - Should HPC resources in Finland be distributed or centralized?



## M-grid Contact People

- Acquisition project: Ville Savolainen <[ville.savolainen@csc.fi](mailto:ville.savolainen@csc.fi)>
- System administration project: Arto Teräs <[arto.teras@csc.fi](mailto:arto.teras@csc.fi)>
- Coordinator Kai Nordlund <[kai.nordlund@helsinki.fi](mailto:kai.nordlund@helsinki.fi)>  
(University of Helsinki)
- Web pages: <http://www.csc.fi/proj/mgrid/>

Thank you! Questions?