



財團法人國家實驗研究院

國家高速網路與計算中心

NATIONAL CENTER FOR HIGH-PERFORMANCE COMPUTING

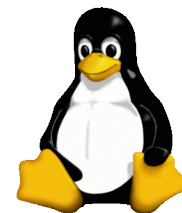
# Diskless Remote Boot in Linux (DRBL) & Rocks

Steven Shiau, Kuo-Lien Huang, Ceasar Sun, Jazz Wang,  
Fang-Pang Lin

<http://drbl.nchc.org.tw>, <http://drbl.sf.net>



Nov, 2005



# Outline

- ✓ **Introduction to Diskless Remote Boot in Linux (DRBL)**
- ✓ **Embedded system - DRBL-based mobile sensors**
- ✓ **Cluster computing - scalable cluster management**
- ✓ **Education - smart classroom**
- ✓ **Q&A**

# Outline

- ✓ **Introduction to Diskless Remote Boot in Linux (DRBL)**
- ✓ **Embedded system - DRBL-based mobile sensors**
- ✓ **Cluster computing - scalable cluster management**
- ✓ **Education - smart classroom**
- ✓ **Q&A**

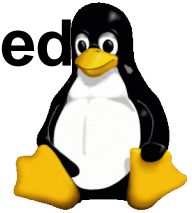
# What is DRBL, Clonezilla ?

- Free software projects developend by NCHC Free Software Labs
- Our goal
  - FSSC (Free Software Smart Classroom)
    - DRBL, Clonezilla
    - Multicast VNC
    - Access Grid, E-learning
    - Cluster computing ready classroom
    - Grid computing ready classroom



# DRBL - Diskless Remote Boot in Linux

- Developed by NCHC Free Software Labs, based on



- **Network boot mechanism**



- PXE (Preboot Execution Environment)
    - Etherboot

- **NFS file system**

- **Other management programs**


- Major Features

- Centralized management
  - Integrate distributed system, either embedded or cluster, into a virtually single system

# Diskfull, diskless or systemless

- **diskfull** - client nodes have dedicated disks
- **diskless** - client nodes have no disks
- **systemless** - client nodes have dedicated disks, but they don't contain a disk bootloader and they are boot from the network, disk is for swap, tmp
- Why diskless ?
  - It's easier to manage one image than many individual installations
  - save budget
  - a disk is a mechanical part that is subject to failure
  - lesser mechanical parts, greater reliability

# Diskfull, diskless or systemless

- Using diskless, systemless or diskfull for clients ?
  - depends on applications
- DRBL provides diskless and systemless mode for clients, we also have a program called "Clonezilla" to deploy a diskfull system 
- Clonezilla has been used in many schools to clone the M\$ Windows or Linux system image

# Development and tuning in DRBL

Package  
Installation  
Live CD



NIS, SSH, NTP...

initrd, busybox  
pcitable

init, booting,  
runlevel

---

kernel space

NBD

KNFS

DEVFS

TMPFS



# Schematic figure for DRBL



server



switch



client nodes



DHCP

———— IP ———▶

pxe/etherboot

192.168.0.1

pxe/etherboot

192.168.0.40

TFTP

———— kernel ———▶

boot

boot

NFS

———— file system ———▶

/, /usr, /home ...

/, /usr, /home ...

NIS

———— account ———▶

user login

user login



# Adopted areas

- **Embedded system**
  - **DRBL-based mobile sensors**
- **Cluster computing**
  - **Scalable cluster management**
- **Education**
  - **Smart classroom**
- **Grid environment**
  - **DRBL-G**

# Hierarchical integration

Fu Shan



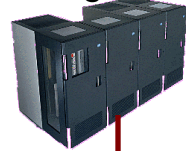
SMPS

三民國小

DRBL-G



Storage/Data



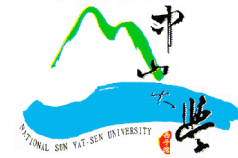
Computers  
- DRBL



NCTU



NSYSU



THU



Data logger  
(CR10X,,campbell)

**Embedded system**

NCHC

Network Backbone

DFES



LTES



Control system

Observation  
Station-DRBL



Tele-robotics -  
DRBL

Domain  
Knowledge  
Center

End Users



**Cluster  
computing**

Computers  
- DRBL



# Outline

- ✓ Introduction to Diskless Remote Boot in Linux (DRBL)
- ✓ **Embedded system - DRBL-based mobile sensors**
- ✓ Cluster computing - scalable cluster management
- ✓ Education - smart classroom
- ✓ Q&A

# **Embedded system - DRBL-based mobile sensors**

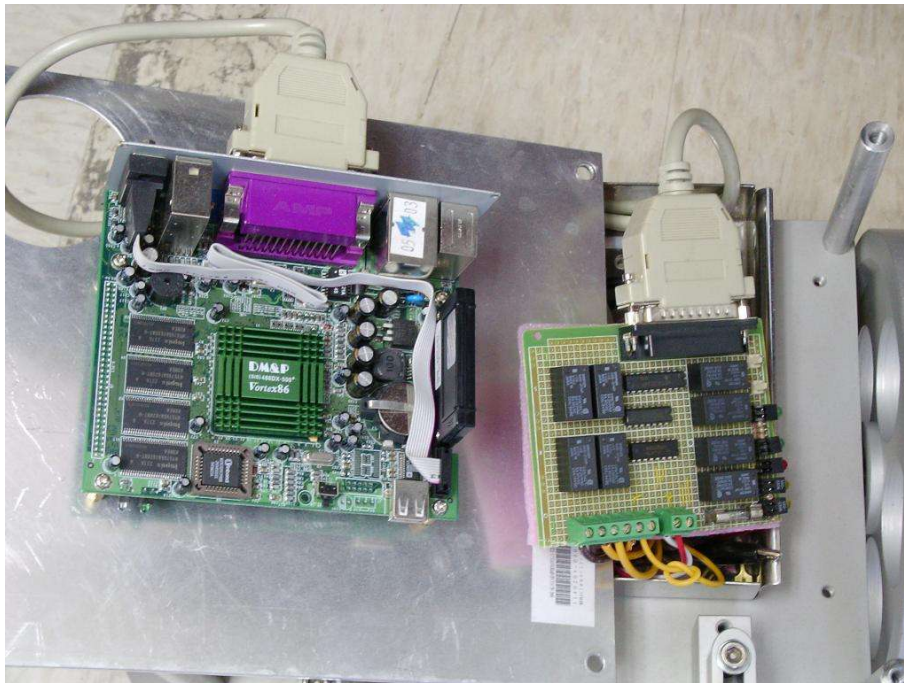
## **From sensors to robot**

**Eco-grid**

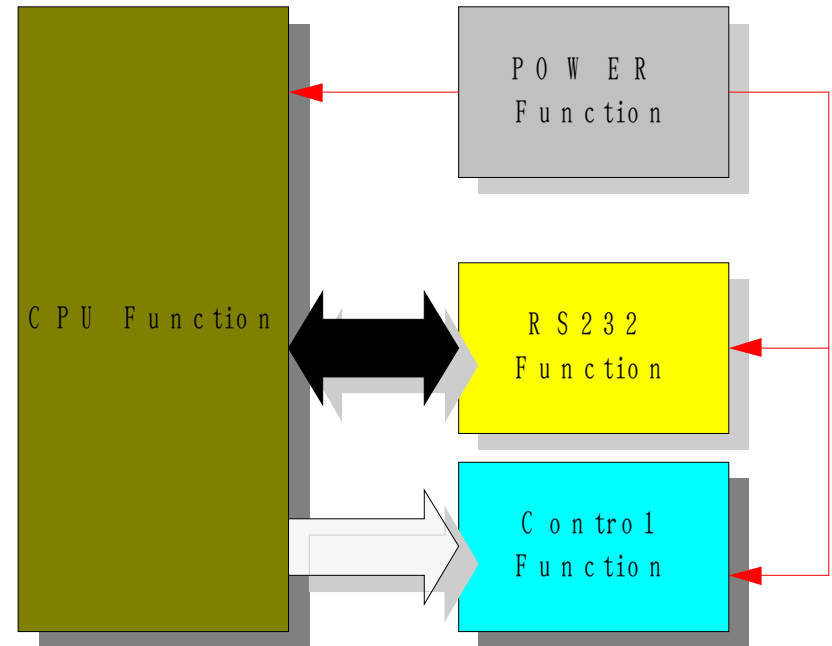
**Agriculture-grid**

# Wireless Robot Net

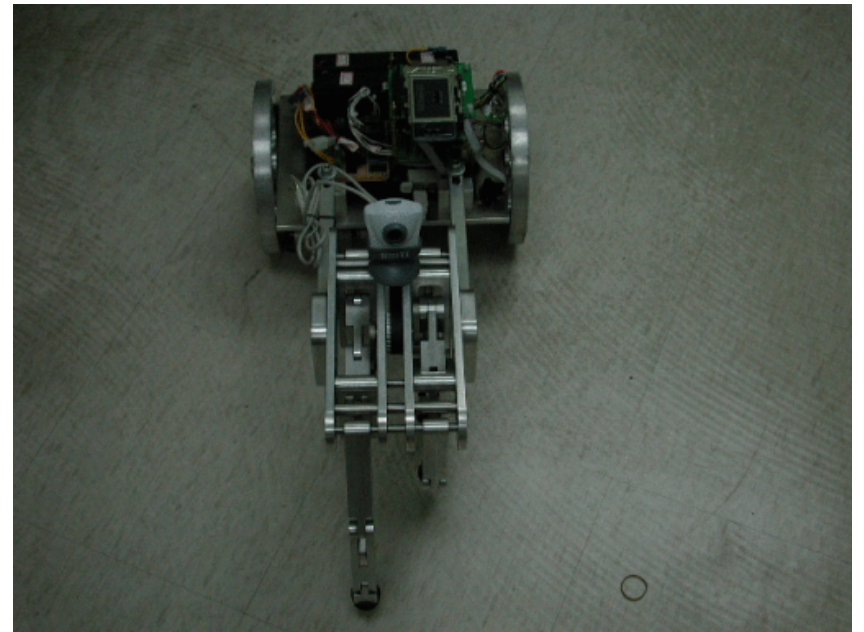
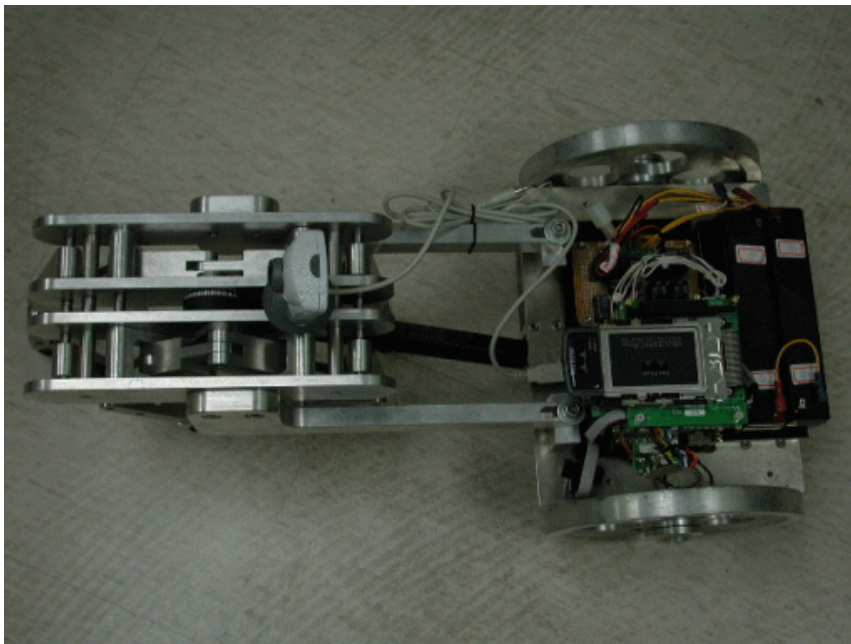
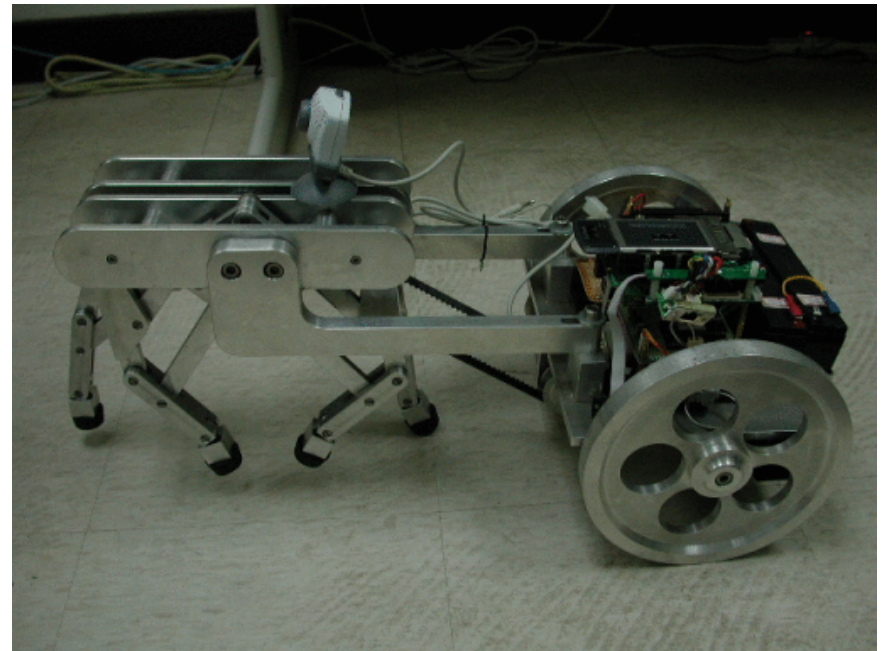
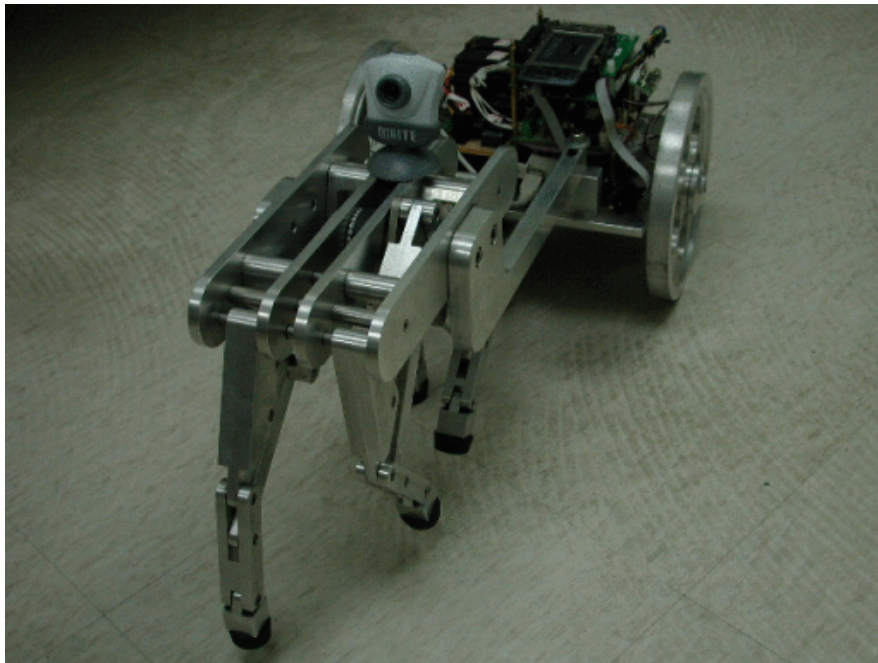
- **Basic idea: intelligent and mobile sensors using Grids.**
- **Ex. IP-driven Horse Robot**
  - Redesign the electronic control panel: PC104 w/ RS232, IDE, Ethernet and USB
  - Use both IDE HD and/or DRBL system setup
  - RS232 & Ethernet for the IP-driven, IPv4 or **IPv6**, wireless robot control
  - USB for webcam



- **Vortex86 200MHz SoC**
- **128MB SDRAM**
- **Setup in DRBL, easy to deploy different systems**









# Outline

- ✓ **Introduction to Diskless Remote Boot in Linux (DRBL)**
- ✓ **Embedded system - DRBL-based mobile sensors**
- ✓ **Cluster computing - scalable cluster management**
- ✓ **Education - smart classroom**
- ✓ **Q&A**

# DRBL in cluster computing

- The advantages come with diskless/systemless environment – **management, budget, reliability**
- Manage one system image only
  - **homogeneous** system
- Client machines can be workstations at daytime, and become cluster computing nodes at night
  - These nodes can be very quickly integrated into a cluster without any alteration of the main OS stored on their disks
- Plug & Play for client nodes

# NPACI Rocks & OSCAR

- NPACI Rocks
- An Open Source High Performance Linux Cluster Solution
- <http://rocks.npaci.edu>
- OSCAR – Open Source Cluster Applications Resources
- <http://oscar.sourceforge.net/>



# DRBL, Rocks & OSCAR

	DRBL	Rocks	OSCAR
diskless/systemless	Y	N	N (*1)
diskfull	Y(*2)	Y	Y
cluster packages	Y(*3)	Y	Y
scale up	good(*4)	excellent	excellent

1. A project called "Thin-OSCAR" provides the diskless/systemless OSCAR cluster
2. Using Clonezilla, DRBL can deploy diskfull nodes
3. Packages are same with those in Rocks and OSCAR, but some are not included
4. From good to excellent
  - ✓ Good: NFSRoot, however many-to-1 problem in large scale
  - ✓ Excellent: (1) ramdisk and multicast model or (2) hierarchical network topology , the scale can be excellent

# PC clusters using DRBL in NCHC

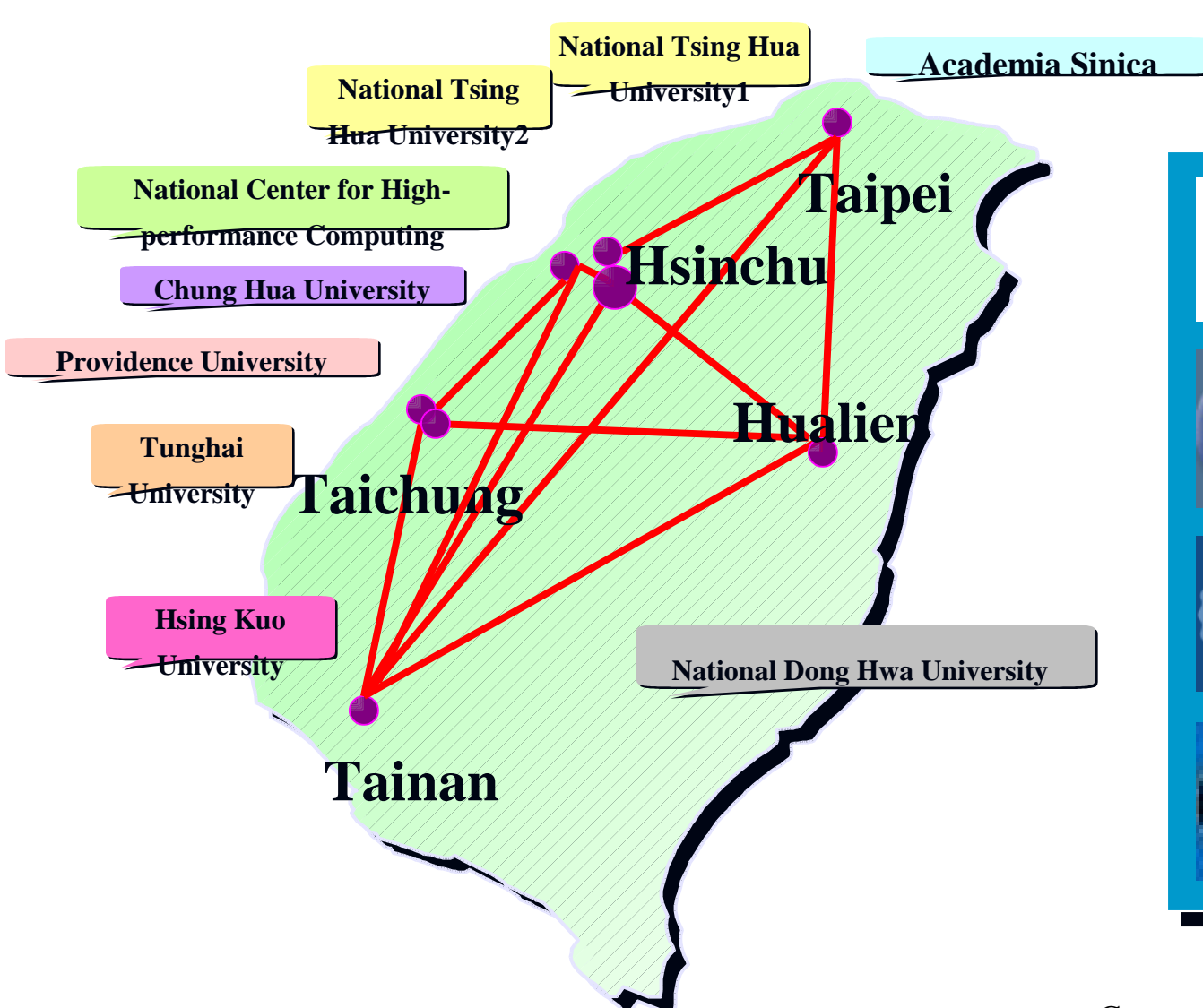
- **Two PC clusters are deployed with DRBL**
  - **ASE cluster : 8 + 1 nodes**
  - **GT3 cluster : 4 + 1 nodes**
- **Deploying DRBL to large scale cluster computing**
  - **Formosa I diskfull approach by clonezilla**
  - **DRBL extention to Formosa I and Taiwan Unigrid**

# NCHC PC Cluster - Formosa I

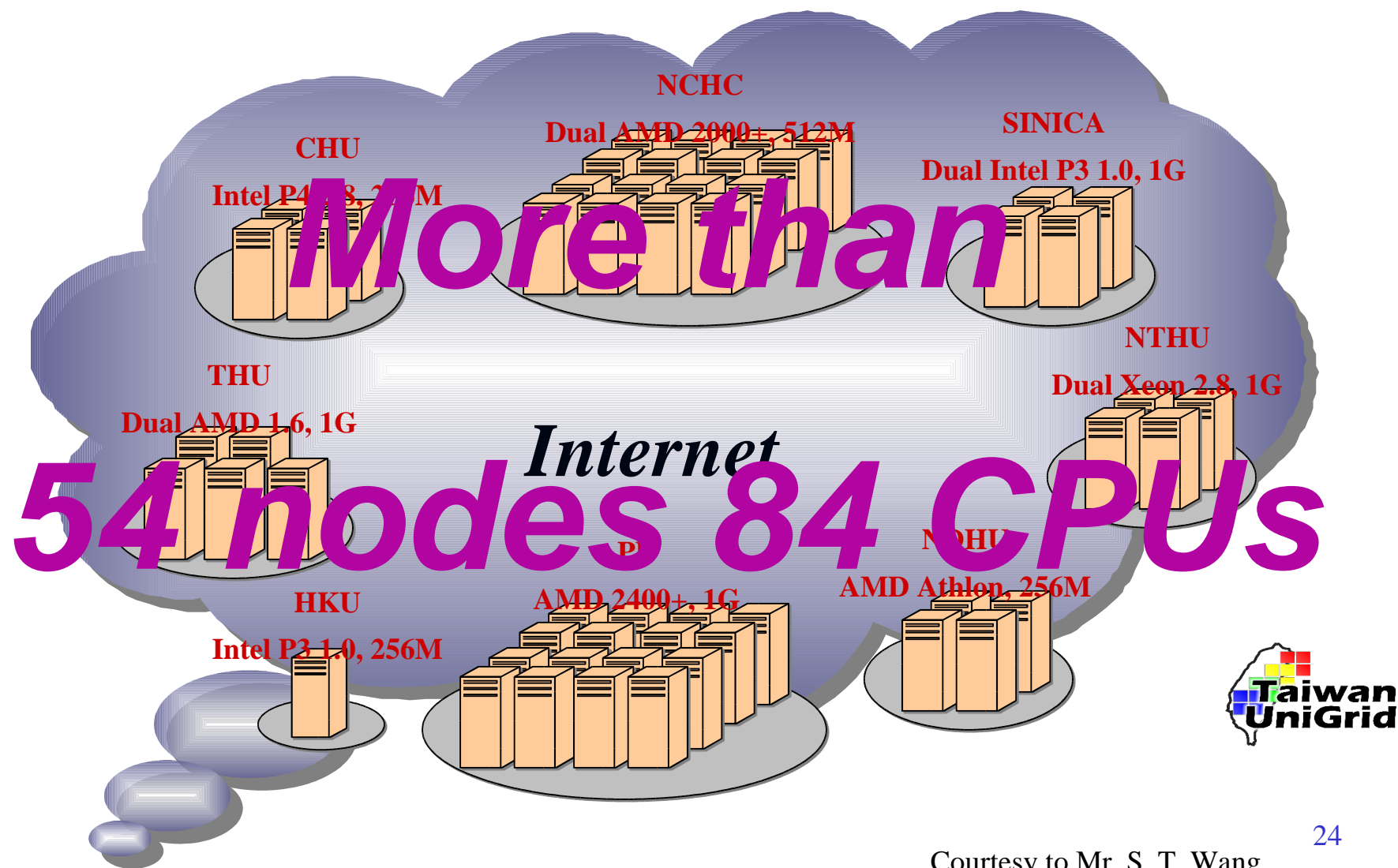


- Built in Oct 2003
- Rank 242 in Top500 on Jun/2004
- CPUs : 300 (2 CPUs X 150 nodes)
- Memory : 384GB (4 GB X 32 + 2 GB X 128)
- Gigabits switch & NICs
- $R_{peak}$  : 1.68 TFLOPS
- $R_{max}$  : 1.002 TFLOPS

# Taiwan UniGrid Topology



# Hardware Infrastructure





# Unigrid Project Webpage

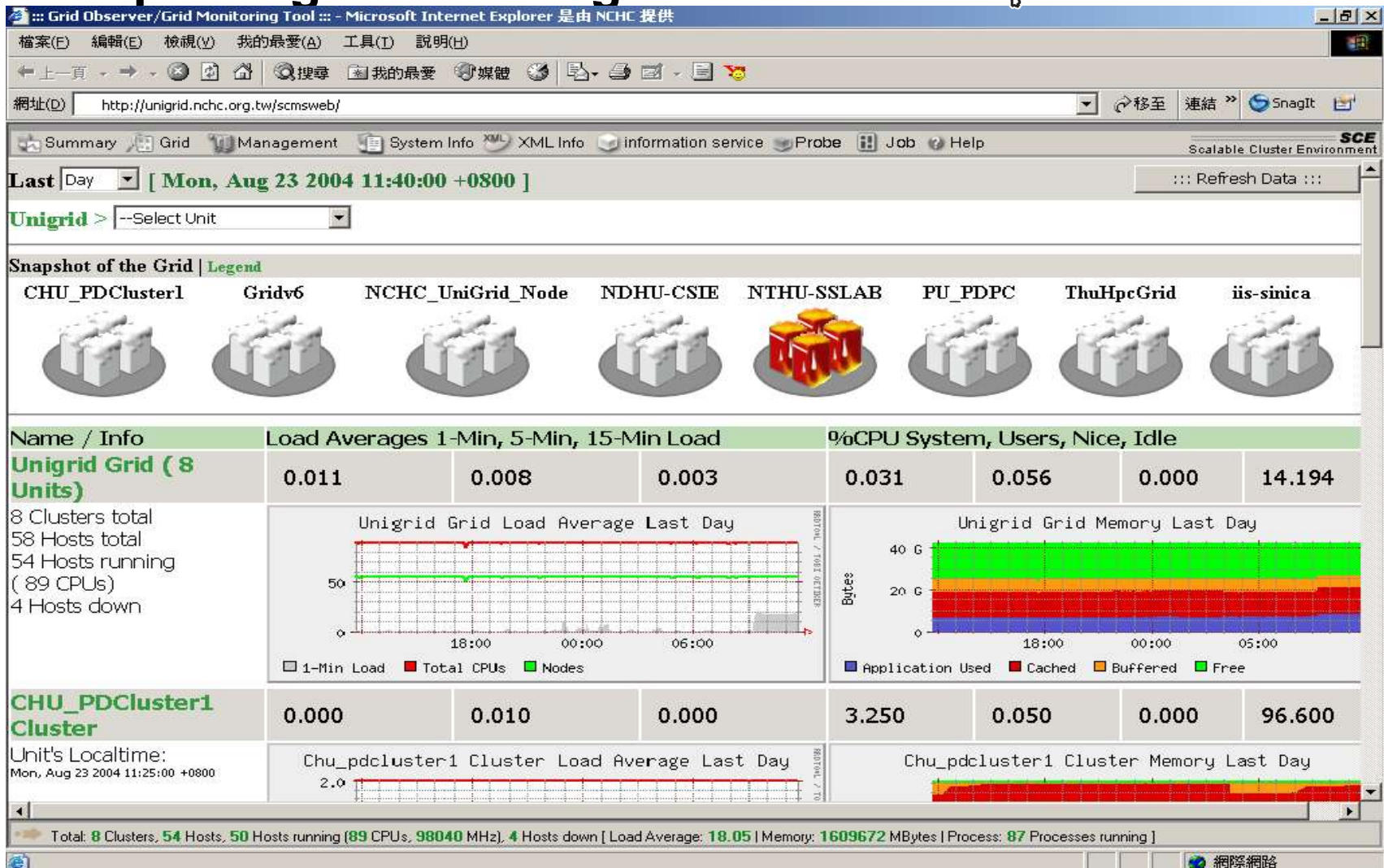


- <http://unigrid.org.nchc.tw/>



# System Monitoring Webpage

• <http://unigrid.nchc.org.tw/scmswe>



Courtesy to Mr. S. T. Wang

# Outline

- ✓ **Introduction to Diskless Remote Boot in Linux (DRBL)**
- ✓ **Embedded system - DRBL-based mobile sensors**
- ✓ **Cluster computing - scalable cluster management**
- ✓ **Education - smart classroom**
- ✓ **Q&A**

# A different kind of light-weight grid computing – computer classroom and office

- DRBL users

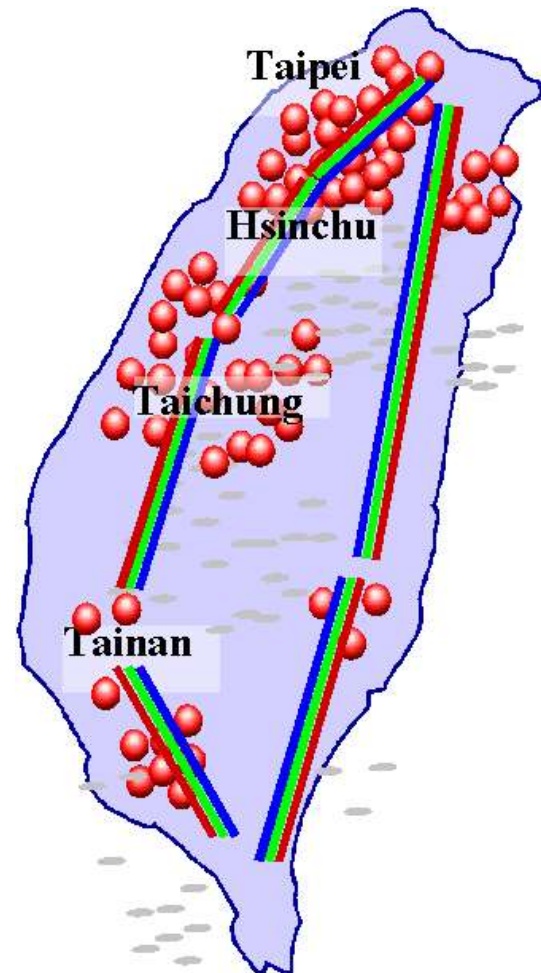
- > 200 sites, > 6000 PCs

- Public sector

- Primary/high school 79
    - University/college 43
    - Hospital 2
    - TV station 1
    - Government 7
    - NPO 14

- Small & Medium Business 44

- Others 38



• Data on 2005/4/13

# Outline

- ✓ **Introduction to Diskless Remote Boot in Linux (DRBL)**
- ✓ **Embedded system - DRBL-based mobile sensors**
- ✓ **Cluster computing - scalable cluster management**
- ✓ **Education - smart classroom**
- ✓ **Q&A**

# Grid Computing

## **DRBL-G**

**Implemented by**

**Chien-Lin Huang, Gary Wu, Julian Yu-  
Chung Chen, Weicheng Huang**

**NCHC**



# DRBL-G

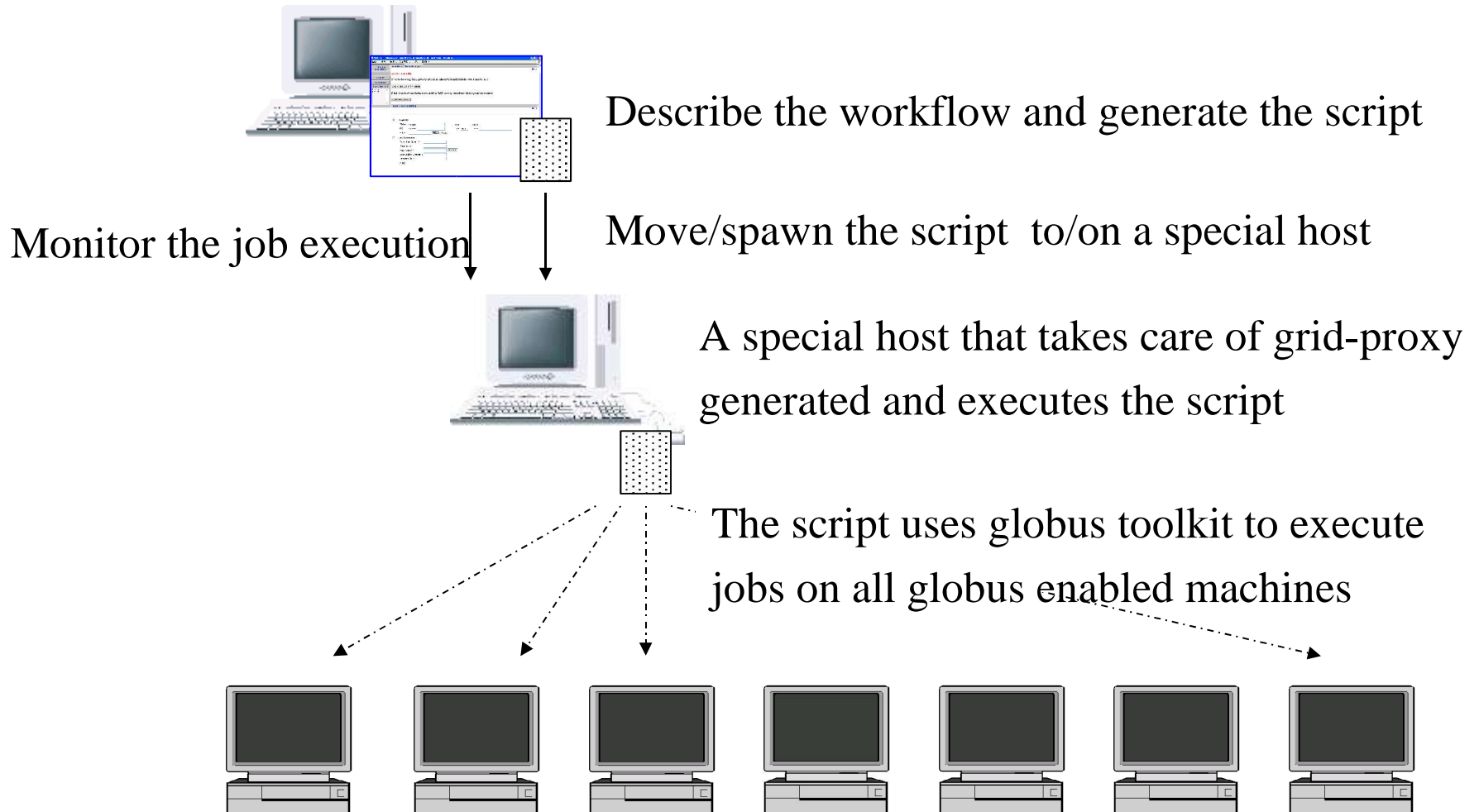
- **Use DRBL to setup grid-enabled clusters**
  - **Two testing PC clusters are grid-enabled with DRBL**
    - **ASE cluster : 8 + 1 nodes**
    - **GT3 cluster : 4 + 1 nodes**
- **DRBL-G environment**
  - **Use globus + shell script + DRBL to integrate the infrastructure**
  - **Web portal**

# DRBL-G

- A web portal is designed to
  - Describe the workflow
  - Monitor the job execution
  - Generate the script based on the workflow description
- Use DRBL to provide a pool of grid-enabled machines
  - DRBL Server
    - Load sharing between DRBL clients
    - Provide a 'sandbox' environment for DRBL clients
      - no sensitive data
      - only /tmp have read-write permission
  - DRBL Clients
    - A pool of grid-enabled machines
    - Take care of grid proxy generation and the script execution



# DRBL-G



# DRBL-G

- Use the Web Portal to describe the work flow

The screenshot displays the DRBL-G web portal interface, which is used for configuring and executing jobs. The interface is divided into several sections:

- Left Sidebar:** Contains navigation links such as "NCHC Eco-Grid", "NCHC LDAP Browser", "Logout", "Customize", and "Users Present".
- Top Section:** Features a "Get New Proxy" button and a "NCHC : Job execution" header with a "Help" link.
- Main Content Area (Left):** Displays the job execution pipeline configuration. It includes a table with columns for "Job" and "Description". The "Job" column contains "jobexecution" and "RSL: &(executable=/bin/date)(directory=/home/globus/)(stdout=datetest.stdout)(stderr=datetest.stderr)". Below the table, there are fields for "Select Destination Host" (palm141.nchc.org.tw) and "Output File Name" (datetest.stdout). There are also buttons for "Run", "Clear All", and "Save as Draft".
- Main Content Area (Right):** Shows the "Grid FTP" configuration section. It includes fields for "FROM \* Host" (palm141.nchc.org.tw), "TO \* Host" (palm141.nchc.org.tw), and "FILE" (datetest.stdout). Below this, there is a "Job Execution" section with fields for "Execution Name \*", "Parameters", "Host Name \*", "Executable Directory", and "Output File \*". There are also buttons for "Submit Option (1): Add after the last Job" and "Submit Option (2): Job # [ ] Insert".
- Right Sidebar:** Contains a "Resources" section with links to "NCHC Job Execution", "NCHC JobStatusReport", "CFD Application", "NCHC Web Link", "NCHC Eco-Grid", "NCHC LDAP Browser", and "Logout". It also has a "Customize" button and a "Users Present" section.
- Bottom Section:** Displays the "Proxy Manager" section, which shows the loaded GSI proxy credentials. It includes a "Get New Proxy" button and a "NCHC : Job execution" header with a "Help" link. Below this, there is a "Status" section showing the current job status.

The interface is designed to be user-friendly, with clear labels and intuitive navigation. It allows users to configure job execution parameters, manage proxy credentials, and view job status in real-time.

# DRBL-G

- Use the Web Portal to monitor the job execution

 財團法人國家實驗研究院  
國家高速網路與計算中心  
NATIONAL CENTER FOR HIGH-PERFORMANCE COMPUTING  
Copyright © 2000-2001 National Center for High-performance Computing

Aug 13, 2004 04:41 pm

My Workspace

Resources

NCHC Job Execution

NCHC JobStatusReport

CFD Application

NCHC Web Link

NCHC Eco-Grid

NCHC LDAP Browser

Logout

Customize

Users Present

Chien-Lin Huan

Proxy Manager

Help

The following GSI proxy credentials are loaded into your account:

(default proxy) /C=tw/O=nchc/OU=Grid/OU=nchc.org.tw/CN=eric View Remove

Click the button below to add another proxy:

Get New Proxy

NCHC : Job execution

The following is the result of the Job execution:

Launched the gram job successfully  
Launched the gram job successfully

Status \*\* Current Job Status = DONE

☐ Grid FTP

FROM \* Host:

TO \* Host:

FILE

 財團法人國家實驗研究院  
國家高速網路與計算中心  
NATIONAL CENTER FOR HIGH-PERFORMANCE COMPUTING  
Copyright © 2000-2001 National Center for High-performance Computing

Aug 13, 2004 04:35 pm

My Workspace

Resources

NCHC Job Execution

NCHC JobStatusReport

CFD Application

NCHC Web Link

NCHC Eco-Grid

NCHC LDAP Browser

Logout

Customize

Users Present

Chien-Lin Huan

Proxy Manager

Help

no proxy available!

The following GSI proxy credentials are loaded into your account:

No proxies currently loaded

Click the button below to add a GSI proxy credential to your account:

Get New Proxy

NCHC JobStatusReport

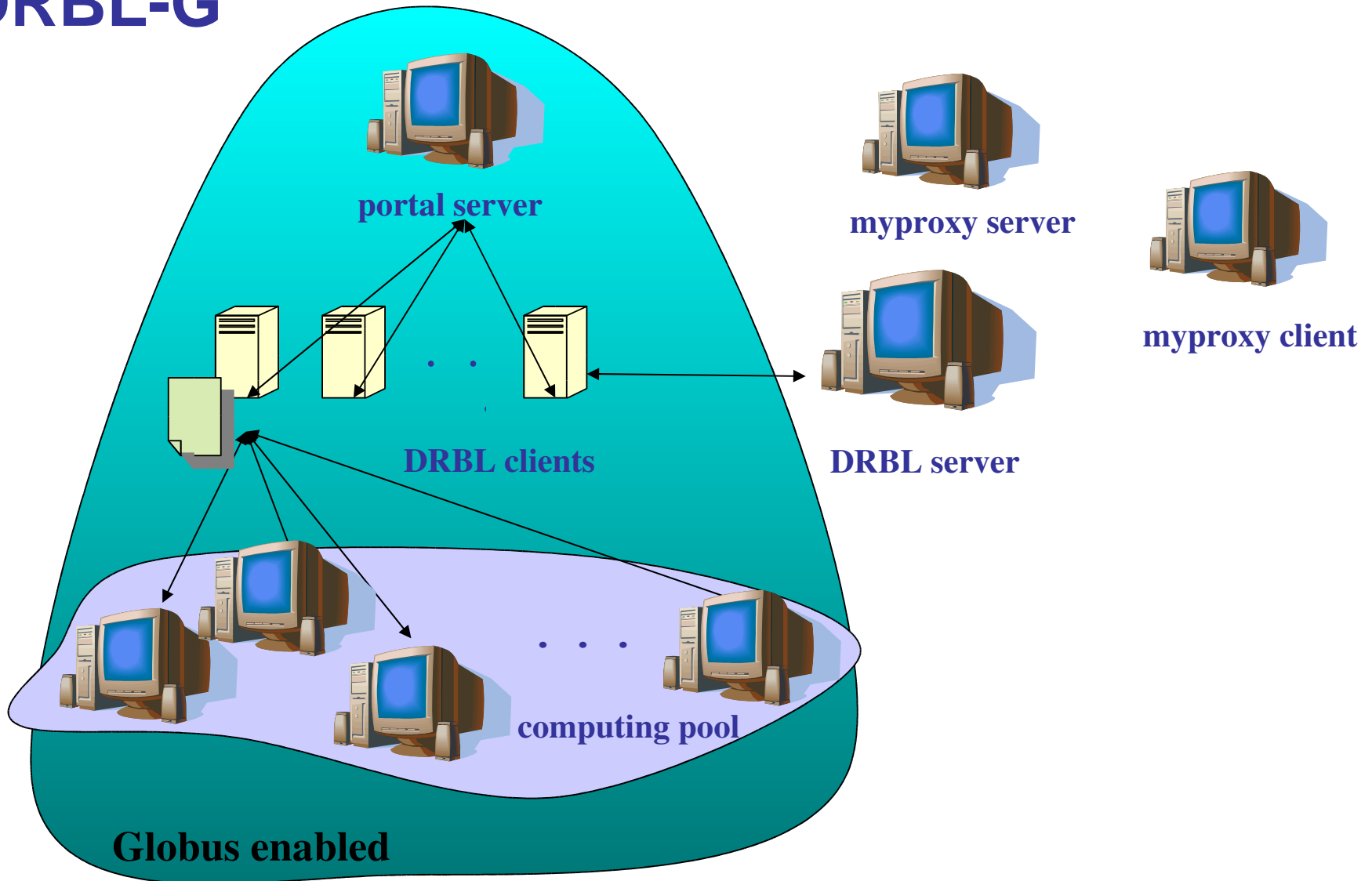
Help

NCHC JobStatus Report Portlet

List out the Job\_ID in the GCE:

no.	Job ID	Job cancellation	Status Report
1	eric_08112004_151854.sh	Submit	CHECK

# DRBL-G



# Reference

- **DRBL Project,**
  - <http://drbl.sf.net>; <http://drbl.nchc.org.tw>
- **EtherBoot Project,**
  - <http://www.etherboot.org>
- **Preboot Execution Environment,**
  - <ftp://download.intel.com/labs/manage/wfm/download/pxespec.pdf>
- **ECOGRID, Grid For Long Term Ecological Research**
  - <http://ecogrid.nchc.org.tw>
- **NPACI Rocks Cluster Distribution,**
  - <http://rocks.npaci.edu>
- **OSCAR : Open Source Cluster Application Resources,**
  - <http://oscar.openclustergroup.org>
- **thin-OSCAR : systemless clients support for OSCAR**
  - <http://thin-oscar.ccs.usherbrooke.ca/>