Cluster, Grid and Cloud Computing

Jazz Wang
Yao-Tsung Wang
jazz@nchc.org.tw

Powered by **DRBL**

Brief History of Computing

Source: http://pinedakrch.files.wordpress.com/2007/07/

Mainframe Super Computer

Brief History of Computing

Source: http://www.nchc.org.tw

Mainframe *Super Computer*   **PC / Linux** *Cluster Parallel*

Brief History of Computing

Source:
http://www.scei.co.jp/folding/en/dc.html

Mainframe	PC / Linux	Internet
Super Computer	Cluster Parallel	Distributed Computing

Brief History of Computing

Source: http://gridcafe.web.cern.ch/gridcafe/whatisgrid/whatis.html
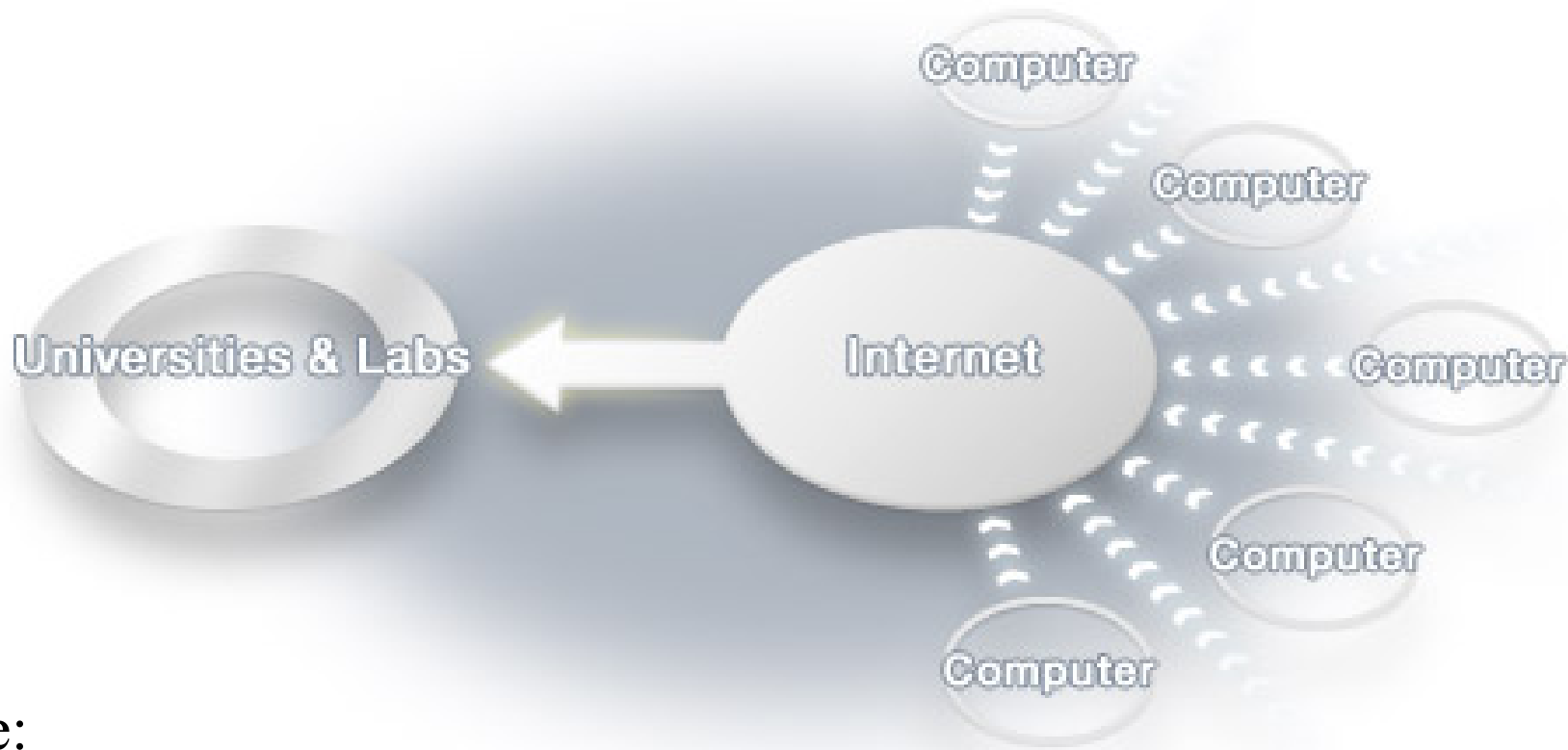
| **Mainframe** Super Computer | **PC / Linux** Cluster Parallel | **Internet** Distributed Computing | **Virtual Org.** Grid Computing |

Brief History of Computing

Grid Infrastructure

Source: http://mmdays.com/2008/02/14/cloud-computing/

nframe / **PC | Linux** / **Internet** / **Virtual Org.** / **Data Explode**

uper / **Cluster** / **Distributed** / **Grid** / **Cloud**

mputer / **Parallel** / **Computing** / **Computing** / **Computing**

Evolution

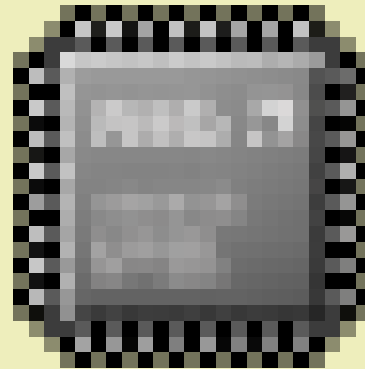# (or is it?)

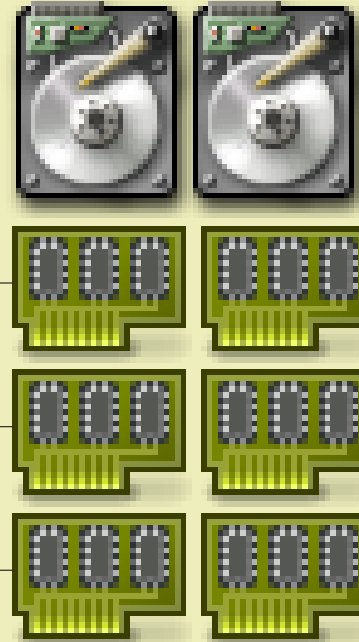# Evolution of Computing Architecture
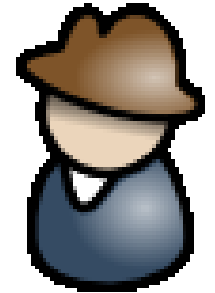

Multiple Users

Single CPU

Shared Memory

One Admin.
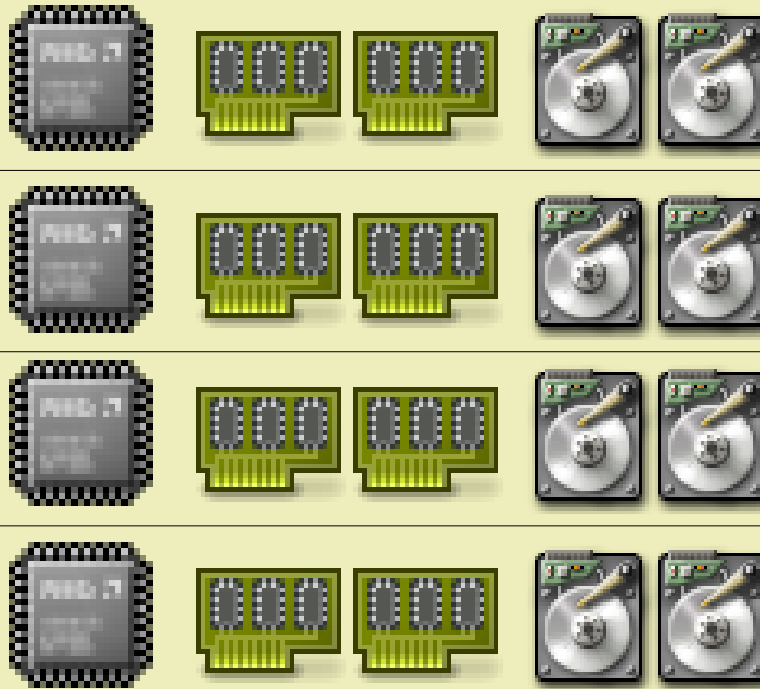
Mainframe Super Computer

Single Super Computer

Evolution of Computing Architecture

Multiple Users

Separate CPU

Separate Memory
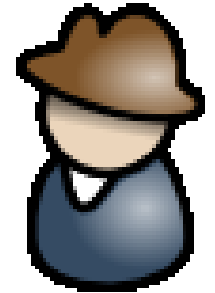
One Admin.

rame
er
uter

PC / Linux Cluster Parallel

Multiple PC in One Location

Evolution of Computing Architecture

Multiple Users

One Admin.

Single CPU **Shared Memory**

Single Powerful Server

Single CPU **Shared Memory**

Single Powerful Server

Multiple Users

One Admin.

Network

Single Broker

Linux ter lel

Internet Distributed Computing
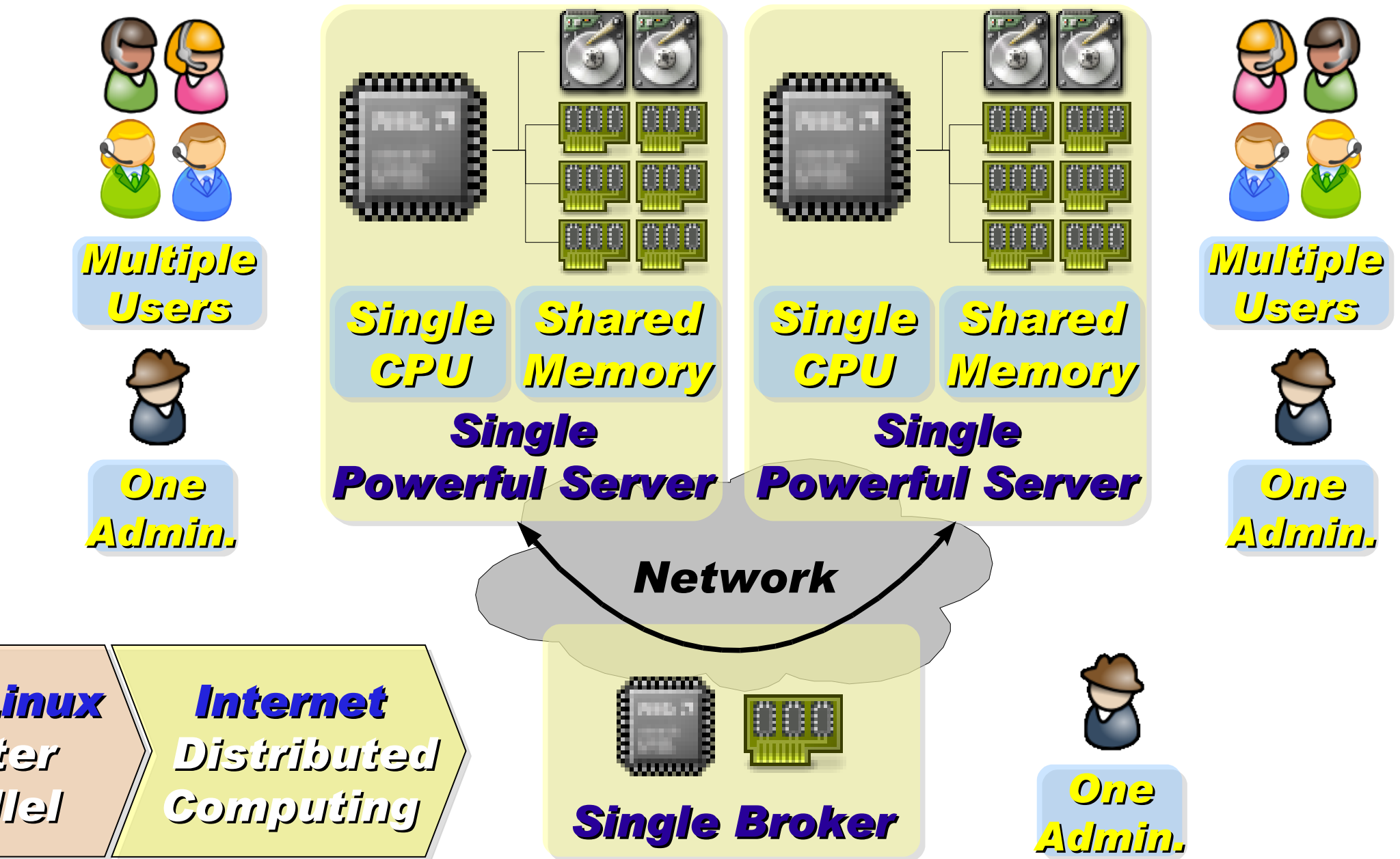
One Admin.

Evolution of Computing Architecture

Multiple Users

One Admin.

Multiple PC in one location
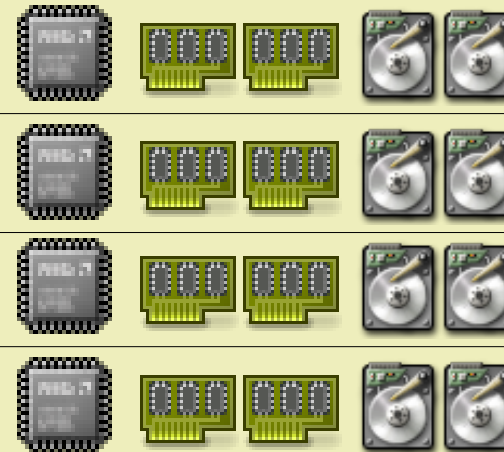
Multiple PC in other location

Grid Middleware

Network

Multiple Users

One Admin.
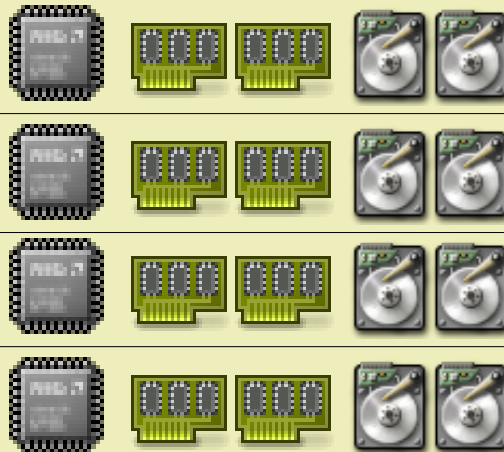
ternet
tributed
mputing

Virtual Org.
Grid Computing

Virtual Organization

Heterogeneous

CyberInfrastructure

Evolution of Computing Architecture

Each User

||

Virtual Admin.

Access any time, any where with mobile device

Multiple PC in different locations

Multiple Admin.

Virtual World

Physical World

tual Org. *Grid* *mputing*

Data Explode **Cloud Computing**
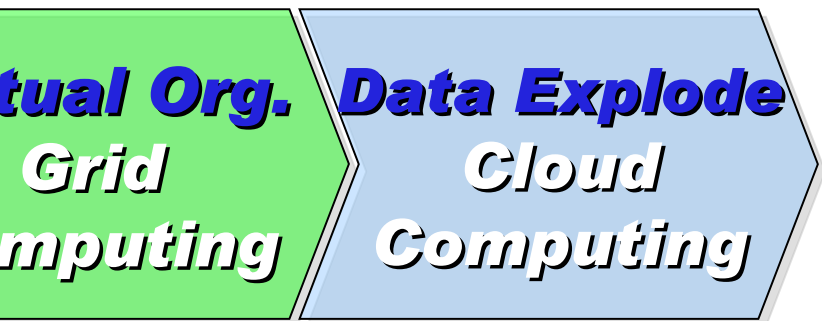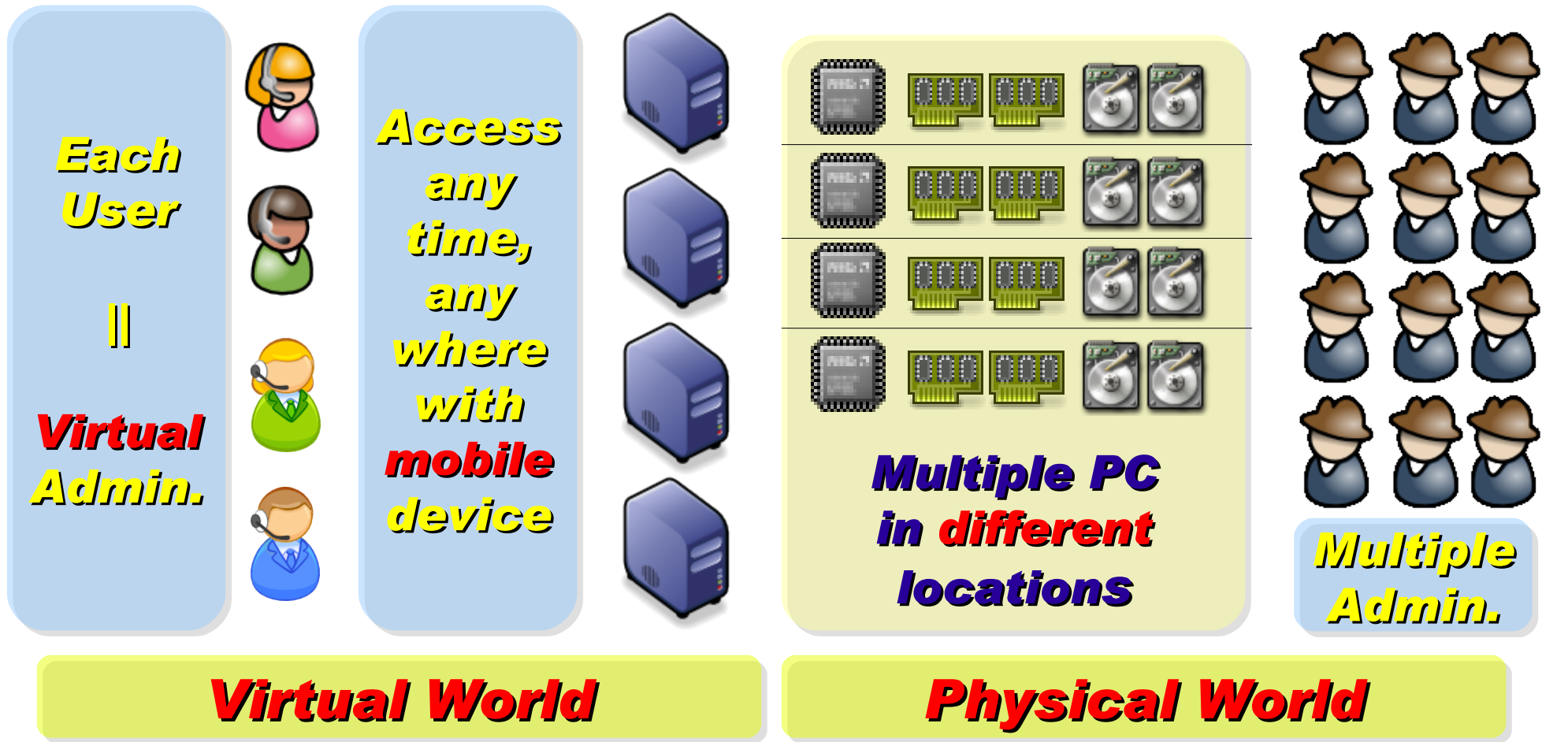
What is NEXT ?!?

Challenges of Cluster Computing

- **Hardware**
  - *Ethernet Speed | PC Density*
  - *Power | Cooling | Heat*
  - *Network and Storage Architecture*

- **Software**
  - *Job Scheduler (Cluster level)*
  - *Account Management*
  - *File Sharing | Package Management*

- **Limitation**
  - *Shared Memory*
  - *Global Memory Management*

Common Method to deploy Cluster

1. Setup one Template machine

2. Cloning to multiple machine

3. Configure Settings

↓

4. Install Job Scheduler

↓

5. Running Benchmark

Challenges of Common Method

Add New User Account ?

Upgrade Software ?

How to share user data ?

Configuration Syncronization

Advanced Methods to deploy Cluster

- **SSI (Single System Image)**
  - *Multiple PCs as Single Computing Resources*
  - **Image-based**
    - *homogeneous*
    - *ex. SystemImager, OSCAR, Kadeploy*
  - **Package-based**
    - *heterogeneous*
    - *easy update and modify packages*
    - *ex. FAI, DRBL*
- **Other Deploy tools**
  - *Rocks : RPM only*
  - *cfengine : configuration engine*

Introduction to Diskless Remote Boot Linux

Jazz Wang
Yao-Tsung Wang
jazz@nchc.org.tw

Powered by **DRBL**

At First, We have " 4 + 1 " PC Cluster

It'd better be 2^n

Manage Scheduler

Then, We connect 5 PCs with *Gigabit Ethernet* Switch

GiE Switch

10/100/1000 MBps

WAN

Add 1 NIC for WAN
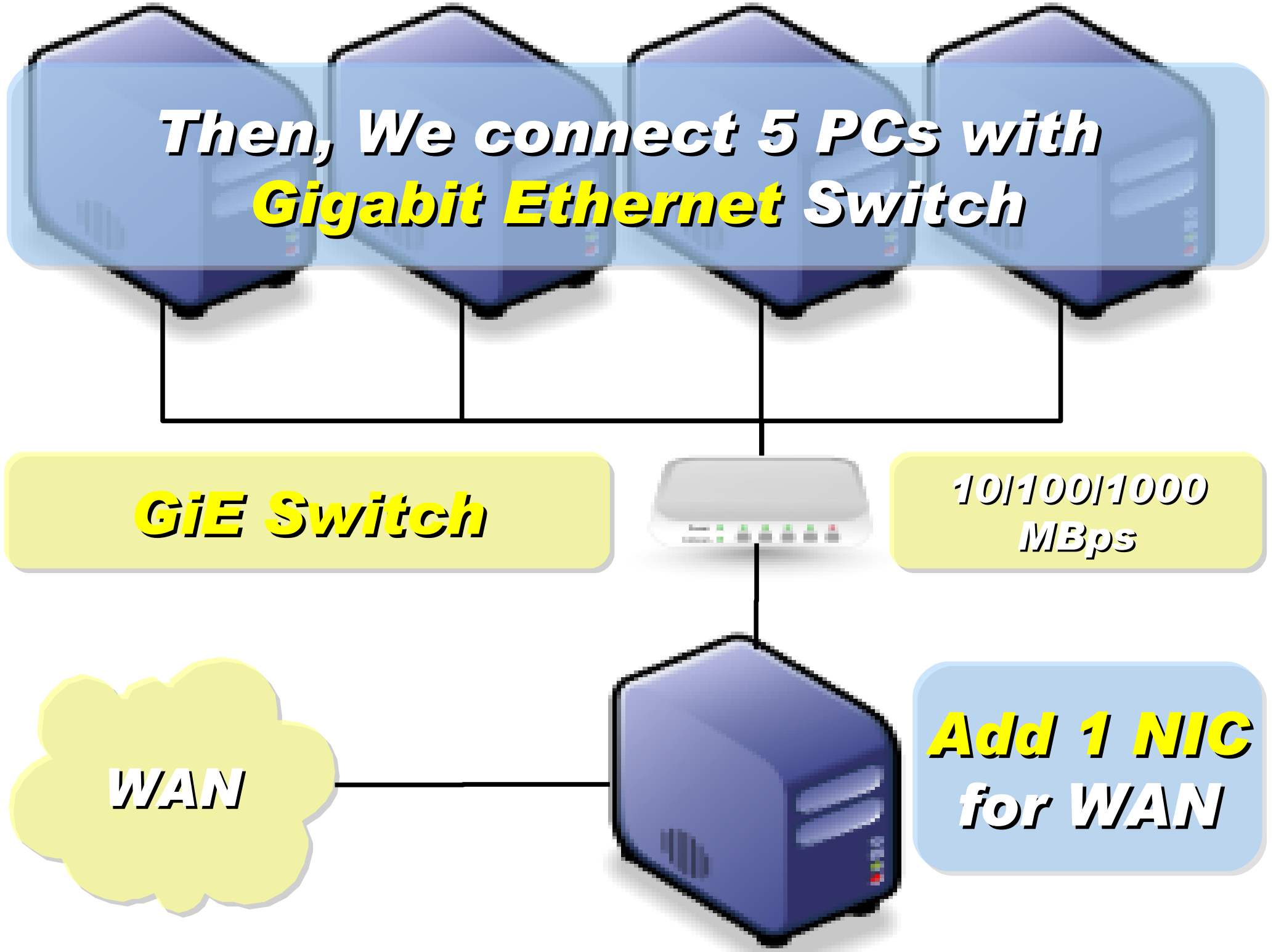
Compute Nodes

4 *Compute Nodes* will communicate via *LAN Switch*. Only *Manage Node* have *Internet Access* for Security!

LAN Switch

WAN

Manage Node

DRBL Client PCs

WAN Switch

For *Effective Management*,
We'll install **DRBL Server** on
Manage Node, and *Compute Nodes*
become **DRBL Client**.

WAN

DRBL Server

1st, We install Base System of *GNU*|*Linux* on *Management Node*.
You can choose:
Redhat, Fedora, CentOS, Mandriva, Ubuntu, Debian, ...

GNU Libc

Kernel Module

Linux Kernel

Boot Loader

2nd, We install _DRBL package_ and configure it as _DRBL Server_. There are lots of service needed: SSHD, DHCPD, TFTPD, NFS Server, NIS Server, YP Server ...

Network Booting

Account Mgnt.

| NFS | TFTPD | DHCPD | SSHD | NIS | YP |

| _Perl_ | _Bash_ | **GNU Libc** |

DRBL Server based on existing Open Source and keep Hacking!

Kernel Module

Linux Kernel

Boot Loader

After running *"**drblsrv -i**"* & *"**drblpush -i**"*, there will be *pxelinux, vmlinux-pex, initrd-pxe* in TFTPROOT, and different *configuration files* for each Compute Node in NFSROOT

NFS	TFTPD	DHCPD	SSHD	NIS	YP

Config. Files
Ex. hostname

GNU Libc

Kernel Module

initrd-pxe

vmlinuz-pxe

Linux Kernel

pxelinux

Boot Loader

3nd, We enable *PXE* function in *BIOS* configuration.

BIOS PXE　　**BIOS PXE**　　**BIOS PXE**　　**BIOS PXE**

| NFS | TFTPD | DHCPD | SSHD | NIS | YP |

Config. Files Ex. hostname

GNU Libc

initrd-pxe

Kernel Module
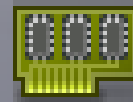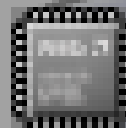
vmlinuz-pxe

Linux Kernel

pxelinux

Boot Loader

While Booting, *PXE* will query IP address from *DHCPD.*

BIOS PXE   **BIOS PXE**   **BIOS PXE**   **BIOS PXE**

NFS   **TFTPD**   **DHCPD**   **SSHD**   **NIS**   **YP**

Config. Files Ex. hostname

GNU Libc

initrd-pxe

Kernel Module
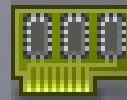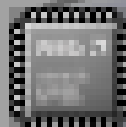
vmlinuz-pxe

Linux Kernel

pxelinux

Boot Loader

While Booting, *PXE* will query IP address from *DHCPD*.

IP 1　　**IP 2**　　**IP 3**　　**IP 4**

NFS	TFTPD	DHCPD	SSHD	NIS	YP

Config. Files Ex. hostname

GNU Libc

initrd-pxe

Kernel Module

vmlinuz-pxe

Linux Kernel

pxelinux

Boot Loader

After PXE get its IP address, it will download booting files from *TFTPD*.

IP 1 IP 2 IP 3 IP 4

| NFS | TFTPD | DHCPD | SSHD | NIS | YP |

Config. Files
Ex. hostname

GNU Libc

initrd-pxe

Kernel Module

vmlinuz-pxe

Linux Kernel

pxelinux

Boot Loader

initrd **initrd** **initrd** **initrd**

vmlinuz **vmlinuz** **vmlinuz** **vmlinuz**

pxelinux **pxelinux** **pxelinux** **pxelinux**

IP 1 **IP 2** **IP 3** **IP 4**
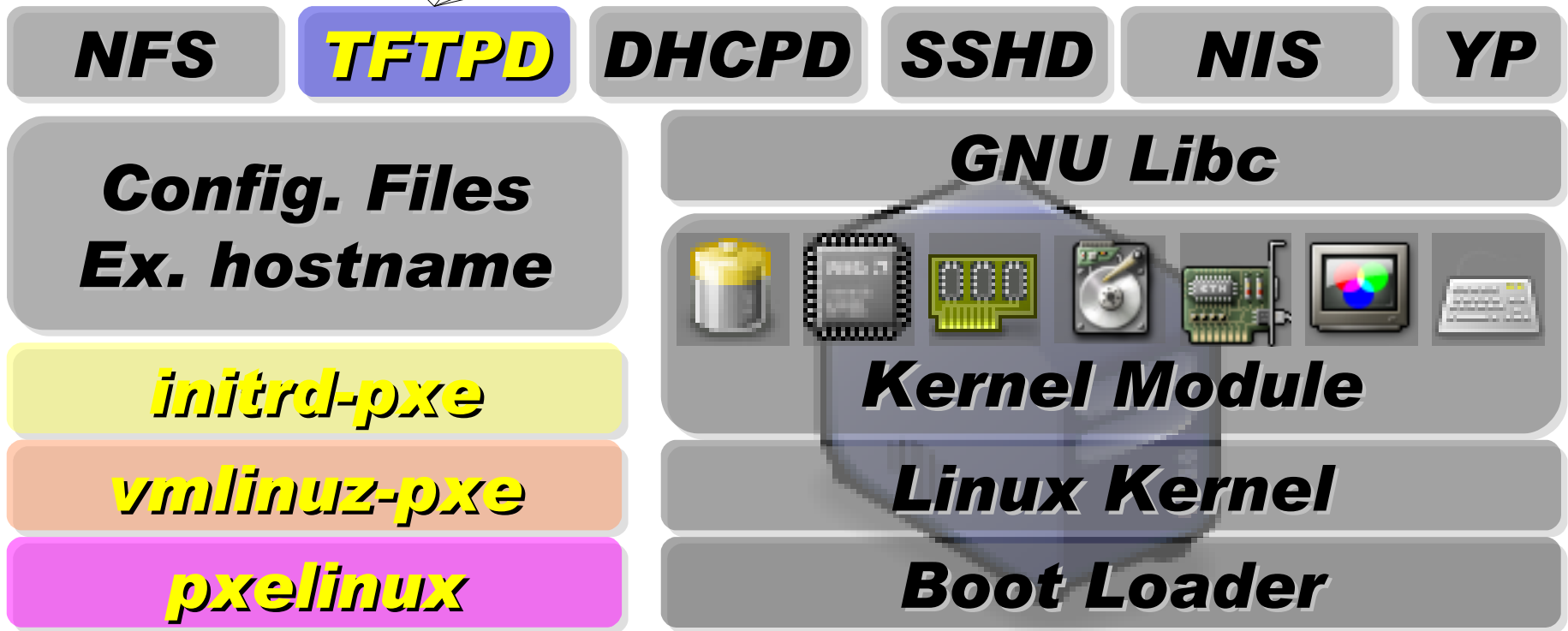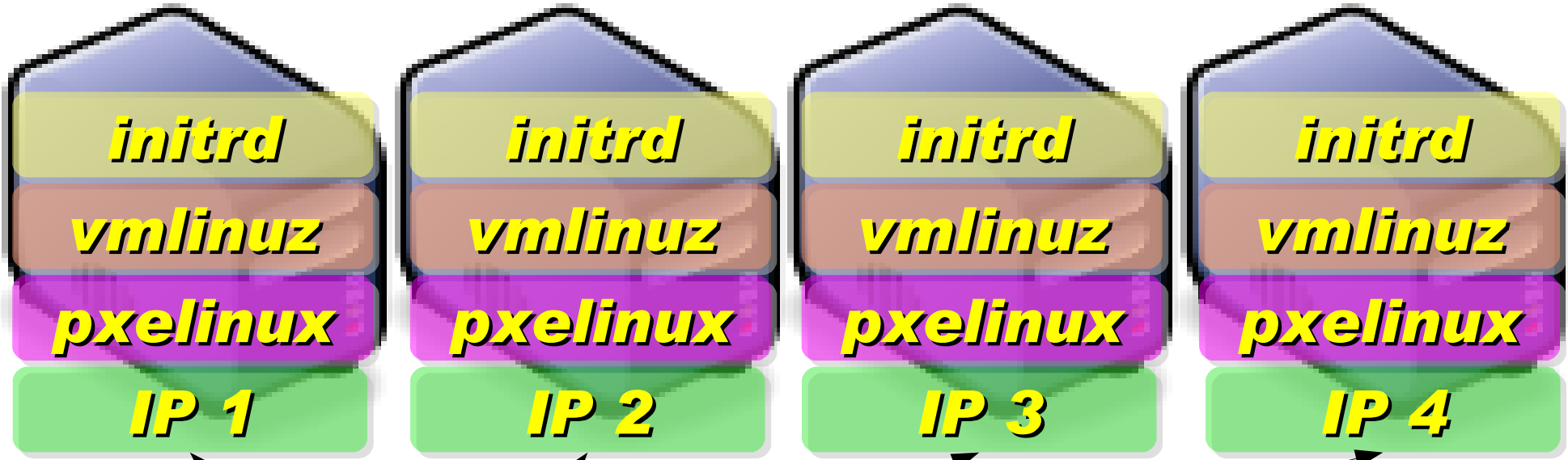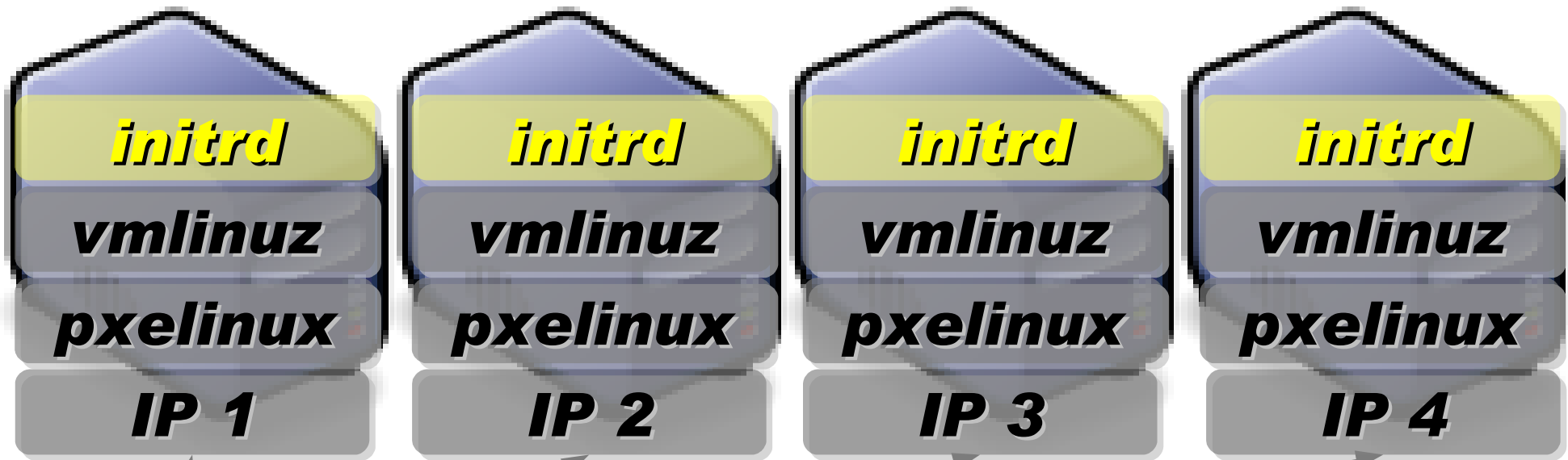
NFS **TFTPD** **DHCPD** **SSHD** **NIS** **YP**

Config. Files *GNU Libc*

*After downloading booting files, scripts in **initrd-pxe** will config **NFSROOT** for each Compute Node.*

pxelinux **Boot Loader**
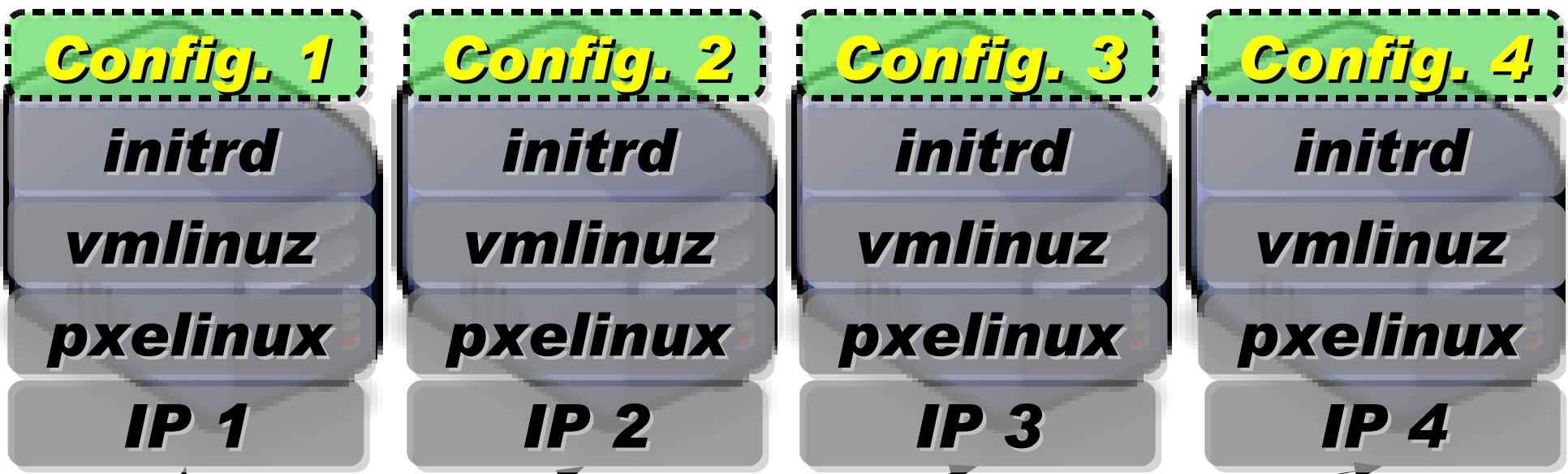
Config. 1	Config. 2	Config. 3	Config. 4
initrd	initrd	initrd	initrd
vmlinuz	vmlinuz	vmlinuz	vmlinuz
pxelinux	pxelinux	pxelinux	pxelinux
IP 1	IP 2	IP 3	IP 4

NFS TFTPD DHCPD SSHD NIS YP

Config. Files
Ex. hostname

GNU Libc

Kernel Module

initrd-pxe

vmlinuz-pxe

Linux Kernel

pxelinux

Boot Loader

Perl **Bash** **SSHD** (×4 compute nodes)
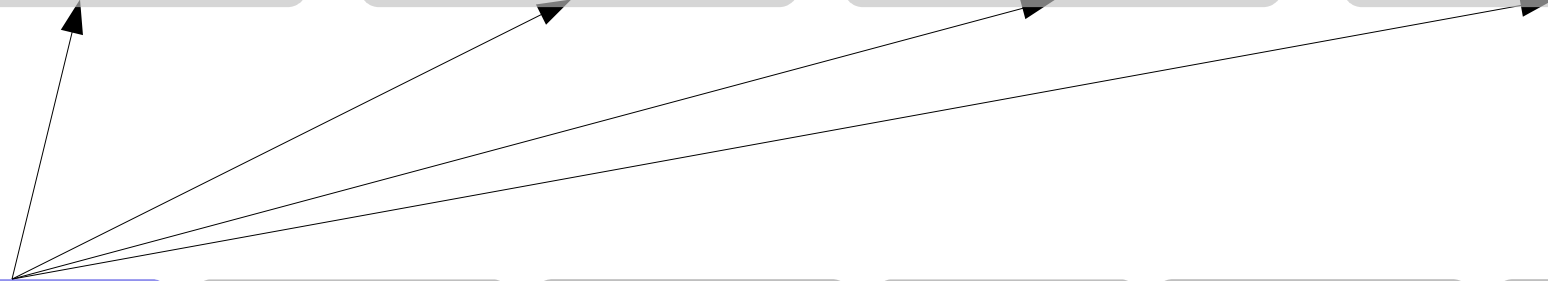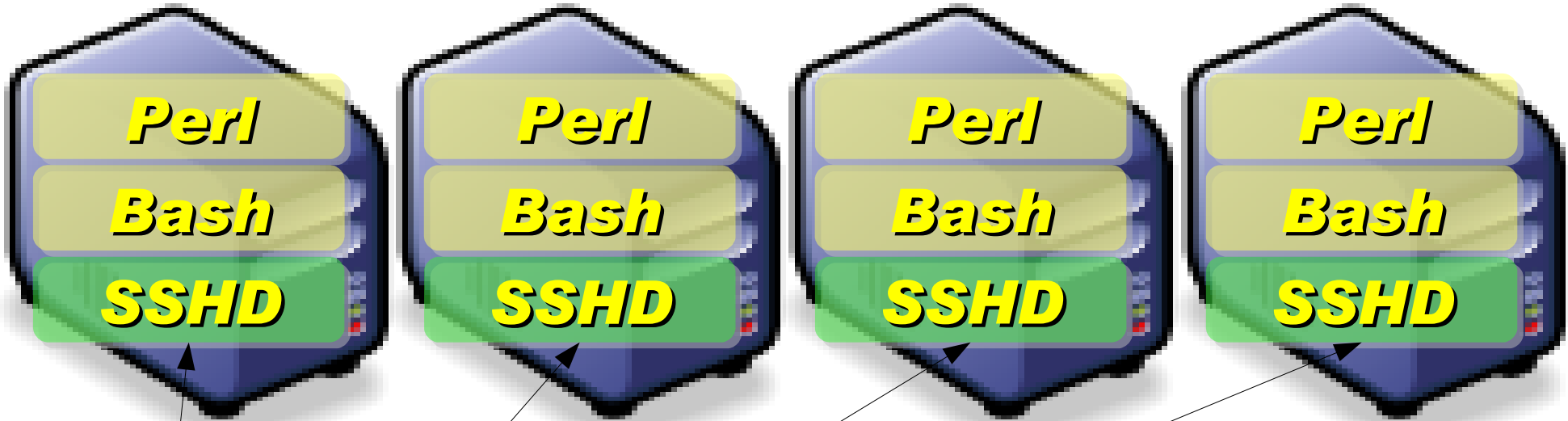
Applications and **Services** will also deployed to each Compute Node via **NFS** ....

NFS **TFTPD** **DHCPD** **SSHD** **NIS** **YP**

Perl **Bash**

DRBL Server

SSHD   **SSHD**   **SSHD**   **SSHD**
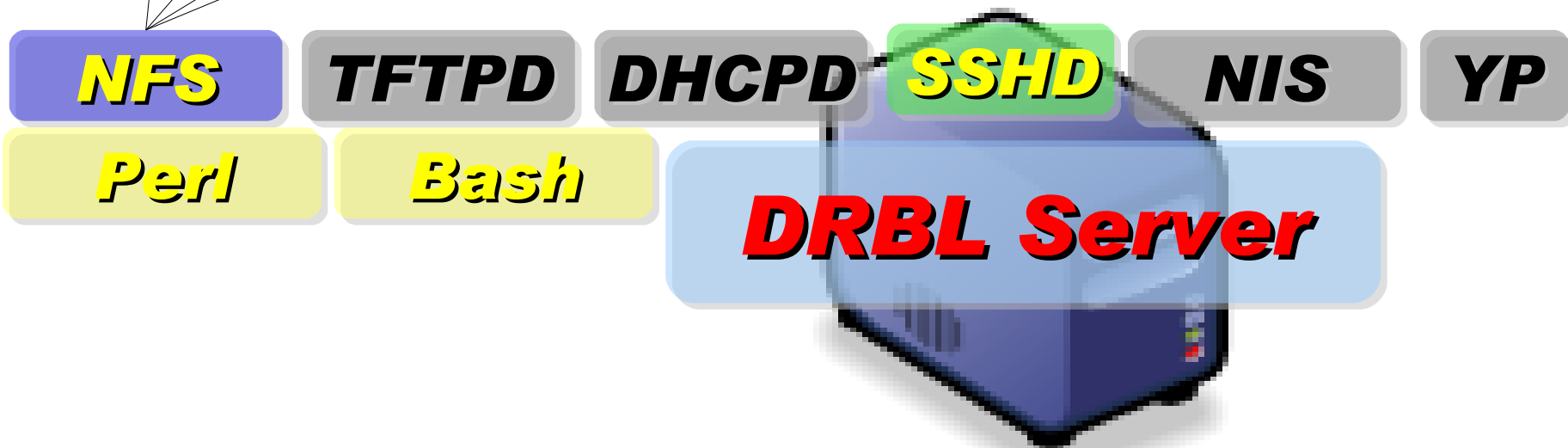
With the help of *NIS* and *YP*,
You can login each Compute Node
with the *Same ID / PASSWORD*
stored in DRBL Server!

SSH Client

NFS   **TFTPD**   **DHCPD**   **SSHD**   **NIS**   **YP**

DRBL Server

Questions?

Jazz Wang
Yao-Tsung Wang
jazz@nchc.org.tw

Powered by **DRBL**

Appendix A

Jazz Wang
Yao-Tsung Wang
jazz@nchc.org.tw

Powered by **DRBL**

自由軟體 (*Free Software*)

站在巨人的肩膀上，是自由軟體發展的理念。其靈活、可自由複製、分享的價值，將有效解決資訊教育的管理成本及商業軟體高成本負擔的問題。

降低資訊教育管理成本

需要「化繁為簡」的解決方案！

一般國內小學的電腦教室

☑ **人力、時間成本高**

　教師 1 人維護管理多組設備

　教學同時分派或收集作業

☑ **設備維護成本高**

　需分別處理設定 (每班約 40 台)

　如：電腦中毒、環境設定

　系統操作問題、開關機、

　備份還原等

平衡商業軟體與知識教育

知識和軟體都需要讓孩子「帶著走」！

老師，可以拷貝Word回家練習嗎？

☑ **商業軟體授權高成本**
在校學習，也需回家複習
學校每台（平均）2萬
學生家用（平均）4萬

☑ **知識與法治的學習**
教育知識，也需教育尊重
尊重智財權觀念

國網中心自由軟體開發

多元化資訊教學的新選擇！

以個人叢集電腦 (PC Cluster) 經驗發展 DRBL&Clonezilla

企鵝龍 DRBL
(Diskless Remote Boot in Linux)

適合將整個電腦教室轉換
成純自由軟體環境

再生龍 Clonezilla

Clonezilla

適用完整系統備份、裸機
還原或災難復原

是自由！不是免費…

分送、修改、存取、使用軟體的自由。免費是附加價值。

企鵝龍 DRBL &再生龍 Clonezilla

電腦教室管理的新利器！

■ 以每班 40 台電腦為估算單位

DRBL&Clonezilla	未 使 用	使 用
管理簡化	分別管理40台	管理 1台 伺服器
硬體設備成本	每台都需配備周邊硬體	伺服器控制，節約每台學生機之周邊硬體
軟體授權成本	40台:3000*40=120,000 (MS Windows授權1台電腦之授權費NT$3,000)	軟體授權 NT$0
合法複製、分享	需負擔授權費	複製合法 NT$0
多元化電腦教學	不同系統無法並存	Linux 與M$ Windows可並存

降低成本，提升形象

教育單位採用 DRBL

擴至全國各單位

高速計算研究
資料儲存備援

降低管理維護成本
帶動自由軟體使用
節樽軟體授權成本（估計）
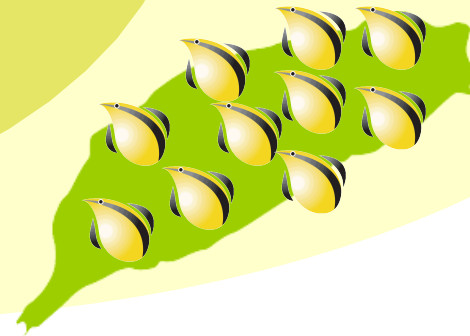
節省龐大軟體授權費

降低台灣盜版率

提升台灣形象

NT. 98,595,000 元
以某商業獨家軟體每機 3000 元授權費計
，每班 35 台電腦 (3000*35*939)

從台灣到國際

DRBL 與 Clonezilla 無遠弗屆

☑ 2007 年 1 月 1 日 -5 月

下載人次: 27,346	DRBL	Clonezilla
台　灣	939人次	783人次
國　外	3,220人次	22,404人次

☑ 2003 年 -2007 年

下載人次：超過 40,000	使 用 單 位
台　灣	超過400個單位
國　外	超過20個國家

由於自由軟體可自由複製與散播，所以仍有未納入記錄之眾多國際與國內隱性的使用者。

Appendix B

Jazz Wang
Yao-Tsung Wang
jazz@nchc.org.tw

Powered by **DRBL**

Research topics about PC Cluster

Cluster Computing

- **System Architecture**
  - Process Architecture
  - Network Architecture
  - Storage Architecture
  - System-level Middleware

- **Parallel Computing**
- **Parallel Algorithms And Applications**
  - Share Memory Programming
  - Distributed Memory Programming
  - Application-level Middleware Programming

Ref: Cluster Computing in the Classroom: Topics, Guidelines, and Experiences
http://www.gridbus.org/papers/CC-Edu.pdf