

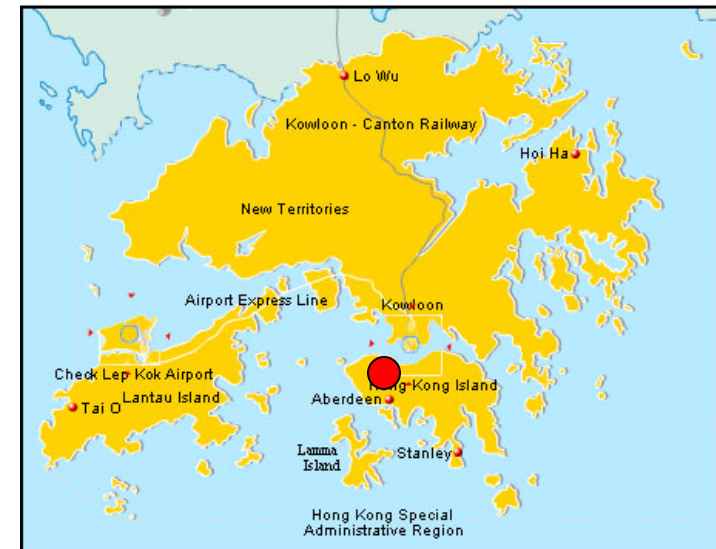
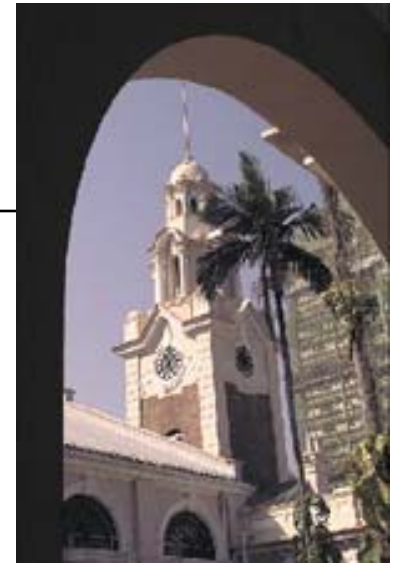
Grid Computing Research in Hong Kong



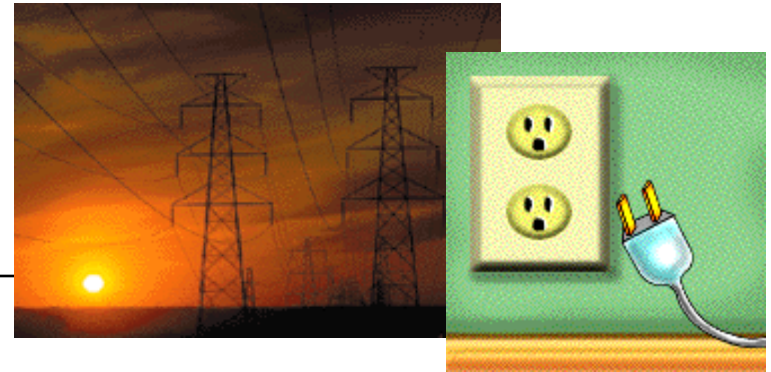
Cho-Li Wang (王卓立)
Systems Research Group (SRG)
Department of Computer Science
The University of Hong Kong
URL: <http://www.cs.hku.hk/~clwang/>

About HKU

- March 16, 1910 : the oldest tertiary education institution in Hong Kong.
- On March 11, 1912, the University launched its official opening with its founding [Faculty of Medicine](#) which had evolved from the Hong Kong College of Medicine, founded in 1887. Of the College's early alumni, the most renowned was Dr Sun Yat-sen, 'the founder of modern China'.
- student population of **19,000**, including **11,700** undergraduate students, **7,300** postgraduate students, of which more than 1,000 international students

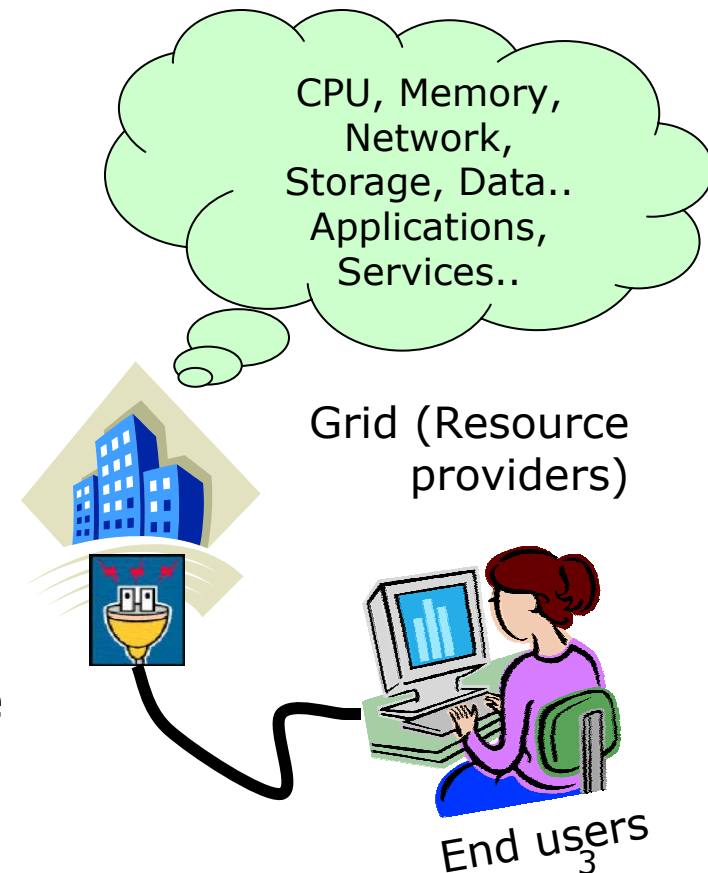


What is Grid ?



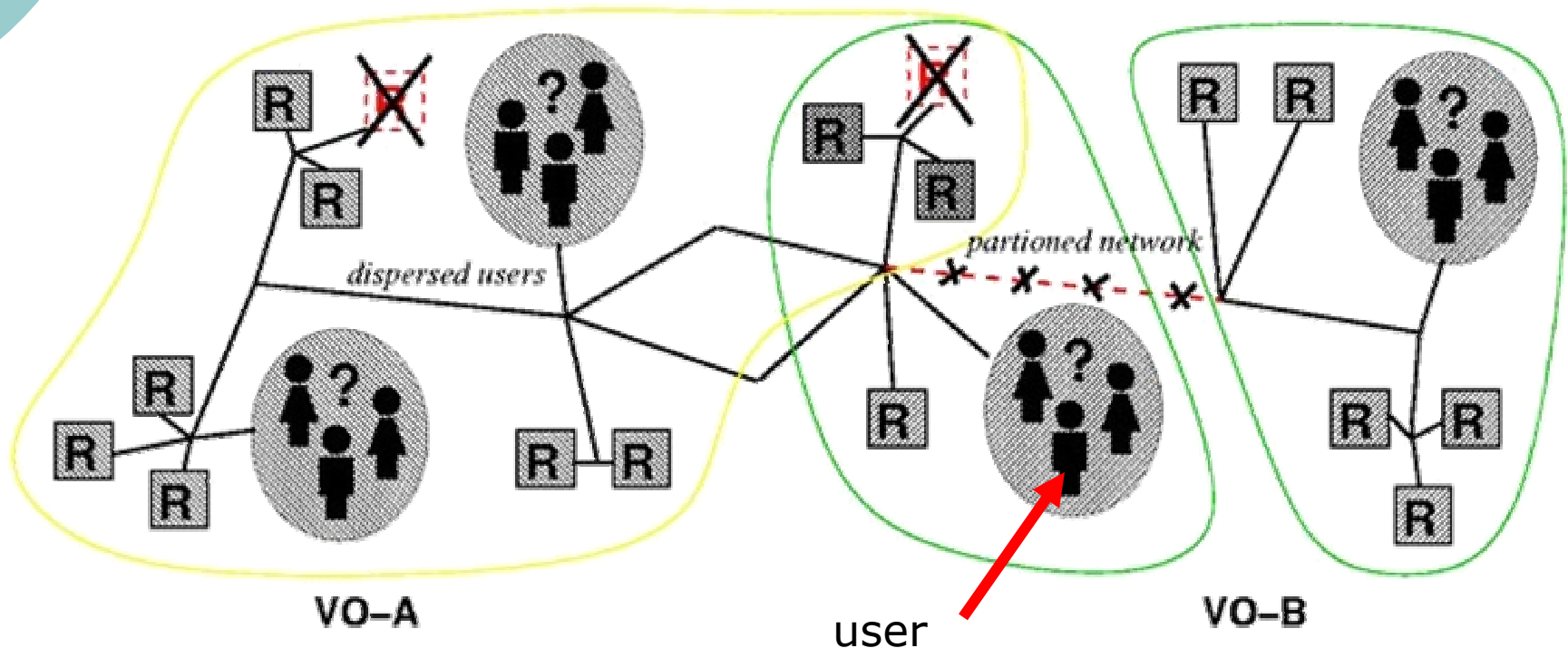
- Grid, or computational grid, is evolving rapidly as a practical extension to distributed computing technology, with the vision of **dynamic and diverse resource sharing across organizations.**
- Helps existing computing resources to be utilized in a more **cost-effective** manner, provides ways to solve **large-scale problems**, and introduces **flexibility in resource planning.**
- Grid resources are usually managed under **different local policies.**
- The types of resources in a grid can be highly diverse and their **availabilities can change dynamically.**

Electric Power Grid

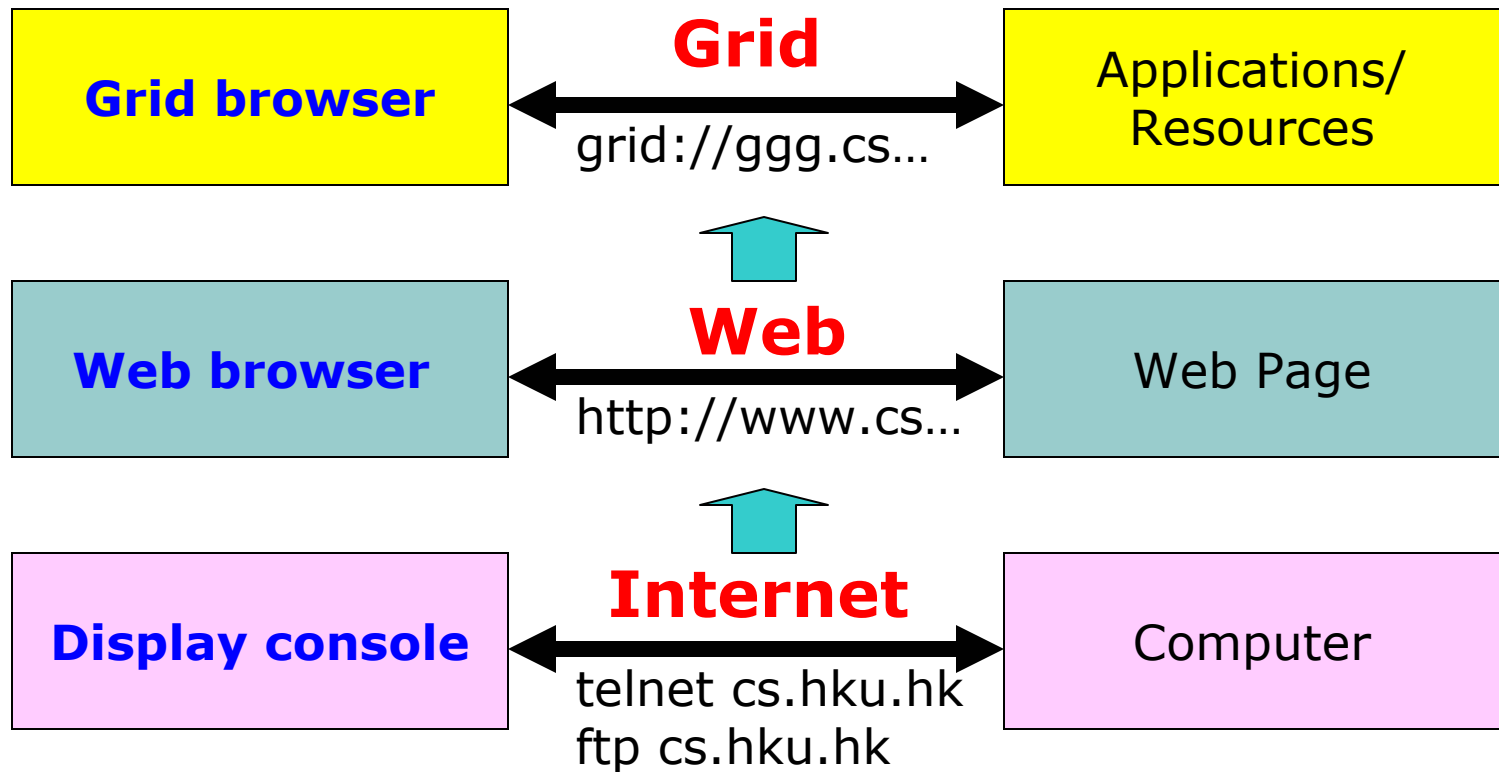


The Grid Problem

Resource sharing & coordinated problem solving in dynamic, multi-institutional virtual organizations



Network Evolution Path



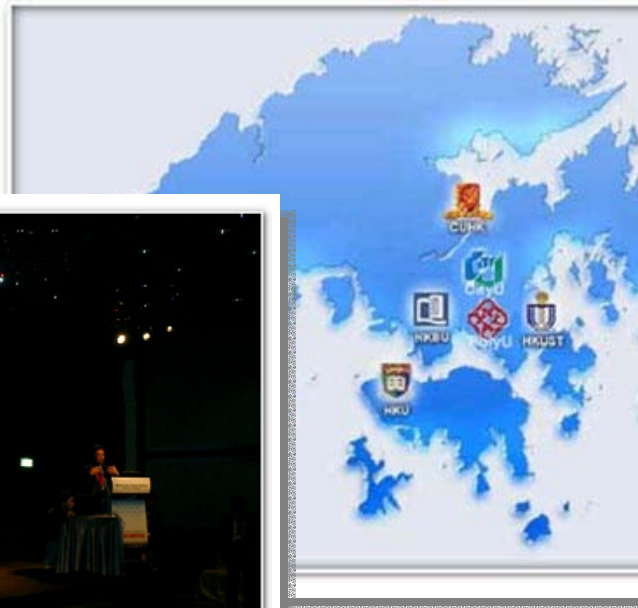


Outline

- Hong Kong Grid Status Report
 - Hong Kong Grid Initiatives
 - HKU CC, HKBU, HKU CS clusters
 - China National Grid Project
 - Asia Pacific Grid Project
- Grid Research Projects in HKU CS
 - SLIM and InstantGrid
 - JESSICA2
 - G-JavaMPI and G-PASS
- Summary and Conclusion

Hong Kong Grid

<http://www.hkgrid.org/>



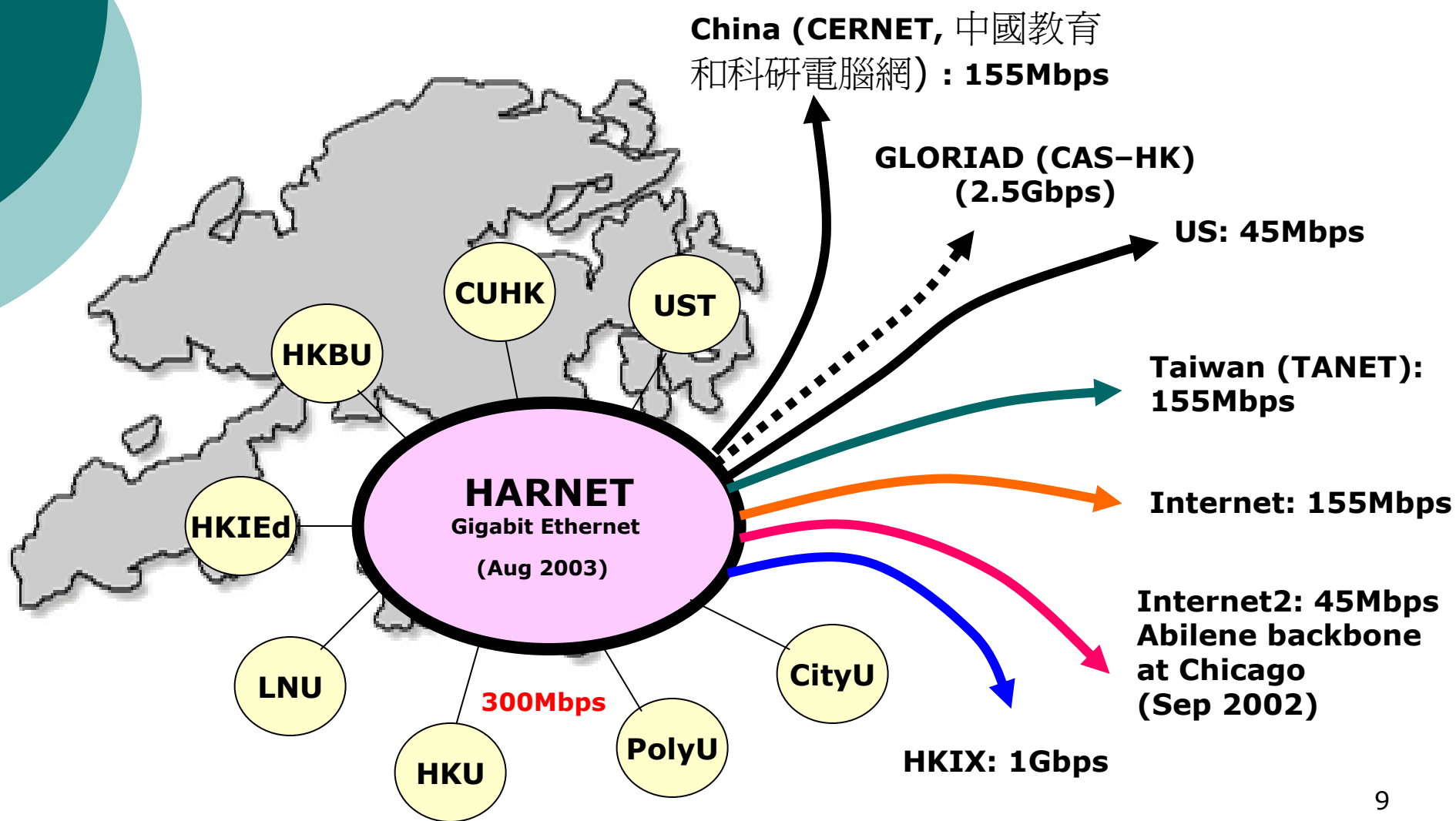
HKGrid Initiatives Launched in Cluster2003 (Dec. 2003)

HKGrid - Current Constituents

Institutions	Computing facilities
香港科技大學 (HKUST)	4-way SMP cluster
香港浸會大學 (HKBU)	2-way Xeon SMP x 64 (#300 in TOP500, 6/2003)
香港城市大學 (CityU)	1 2-way Xeon SMP Service gateway
香港高性能計算所 (HK HPC)	1 2-way Xeon SMP (Service gateway)
香港理工大學 (PolyU)	1 2-way Xeon SMP (Service gateway)
香港大學 (HKU/CC)	2-way Xeon SMP x 128 (#240 in TOP500, 11/2003)
香港大學 (HKU/CS)	Pentium 4 x 300 (#175 in TOP500, 11/2002)

Total computing power (theoretical maximum) = 4 Tflop/s

The Hong Kong Academic & Research Network: HARNET



Grid Research Projects in Hong Kong

- **HKUST**: Incentive scheduling, topology optimization
- **HKBU**: Knowledge grid, autonomous grid service composition
- **CityU**: Agent-based wireless grid computing
- **PolyU**: Peer-to-peer grid, meta-scheduling, fault tolerance
- **HKU**
 - **CC**: Scientific applications running across the ApGrid
 - **CC**: Biosupport project with HKU-Pasteur Research Centre
 - **ETI**: Modeling of Air Quality in Hong Kong (with the Environmental Protection Department, HKSAR)
 - **ETI**: RFID Grid
 - **CS** : China National Grid (CNGrid) project - HKU Grid Point
 -

HKU Computer Centre



hpcpower: 128 nodes (IBM x335)
of dual Xeon 2.8GHz CPUs GigaEth
connection (CISCO 4506), Linux OS



October 20, 2004 : Inaugural
Ceremony of HPC Cluster on
Windows Platform

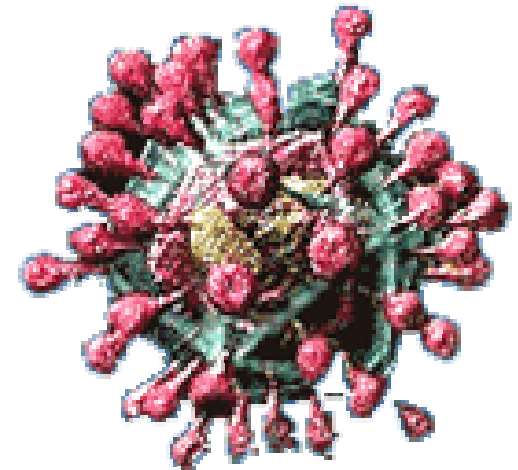
Current Focus:

- Core member of HKGrid
- International collaboration supported by HARNET-Internet2 and HARNET-APAN connections
- More collaborations with Chinese institutions
- Exploring implementation of other forms of GRID computing for various purposes as viewed by different groups and companies.

HKU-Pasteur Research Centre

Biosupport Project

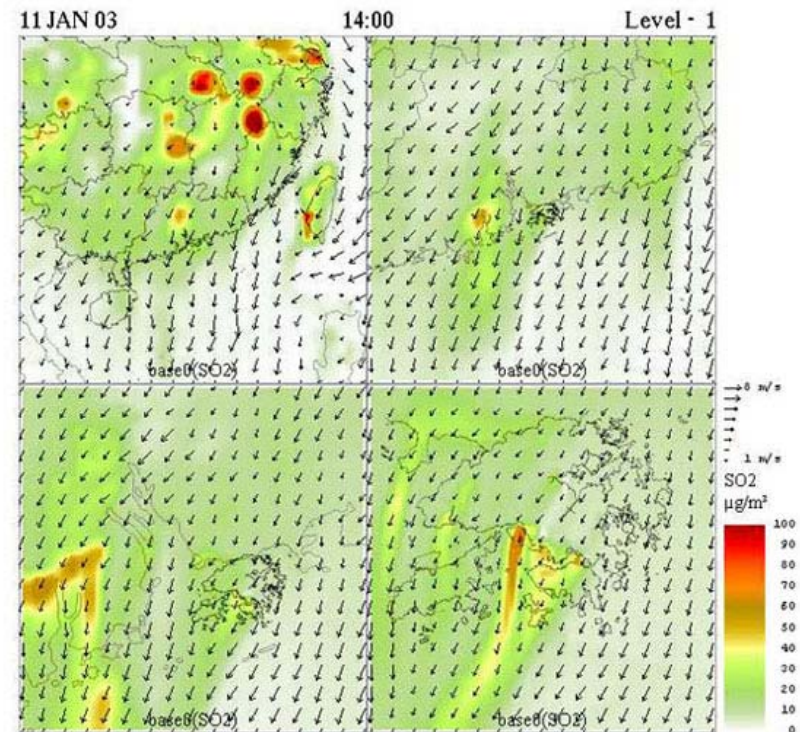
- Collaboration between **HKUCC**, HKU-Pasteur Research Centre and Centre de Ressources INFOBIOGEN (France).
- **Bioinformatics Tools:** The **sequence analysis packages** installed include EMBOSS, NCBI tools, FASTA, STADEN, PHYLIP, READSEQ, ClustalW/ClustalX, DIALIGN2 and the PHRAP/PHRED/CONSED package. Some tools installed also have **on-line web interface**, such as JEMBOSS, EMBOSS-GUI, NCBI-BLAST, FASTA and GenoList



HKU ETI – EPD

Modeling of Air Quality in Hong Kong

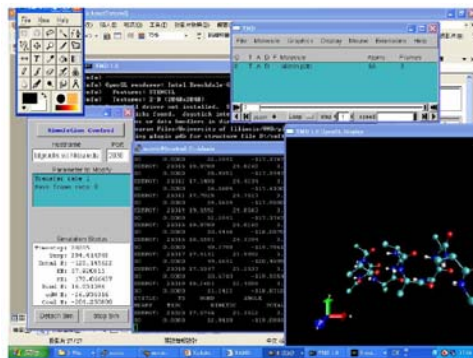
- Collaboration between HKU *E-business Technology Institute* (ETI) and the *Environmental Protection Department* (EPD), HKSAR
- Investigate the inter-connections of the air pollution mosaic through numerical simulation
- Government plans to harness grid technologies to utilize idle PCs during off-hours



Source: <http://www.info.gov.hk/digital21/eng/knowledge/gripapp.html>

Hong Kong Baptist University

High Performance Cluster Computing Centre



Quantum Chemistry

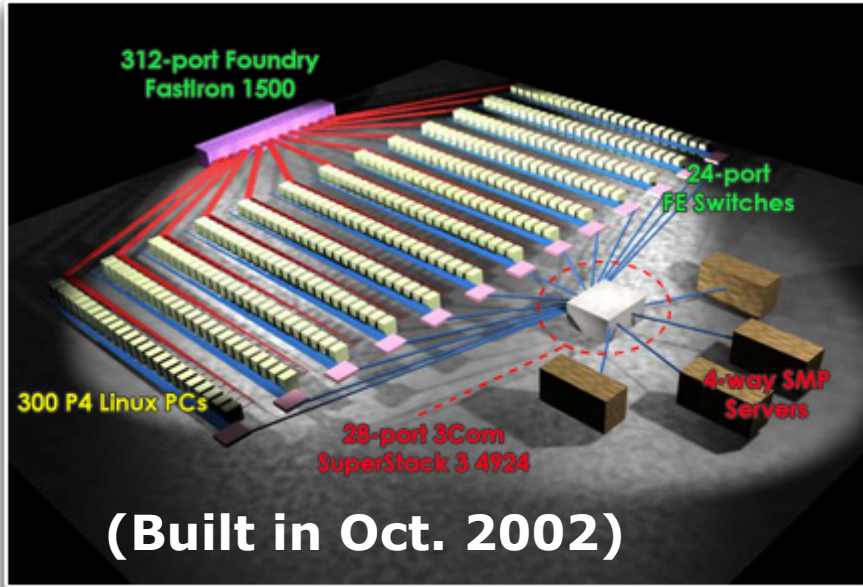


64 nodes (Dual Intel Xeon 2.8GHz with 2GB RAM), Network: 65-port Extreme BlackDiamond 6816 Gigabit Ethernet switch

- **Message Passing Interface**
 - MPICH, LAM/MPI
- **Mathematical:**
 - **fftw** (fast fourier transform)
 - **pblas** (parallel basic linear algebra software)
 - **atlas** (a collections of mathematical library)
 - **sprng** (scalable parallel random number generator)
 - **MPITB** -- MPI toolbox for MATLAB
- **Quantum Chemistry software**
 - gaussian, qchem
 - Molecular Dynamic solver
 - NAMD, gromacs, gamess
- **Weather modelling: MM5**

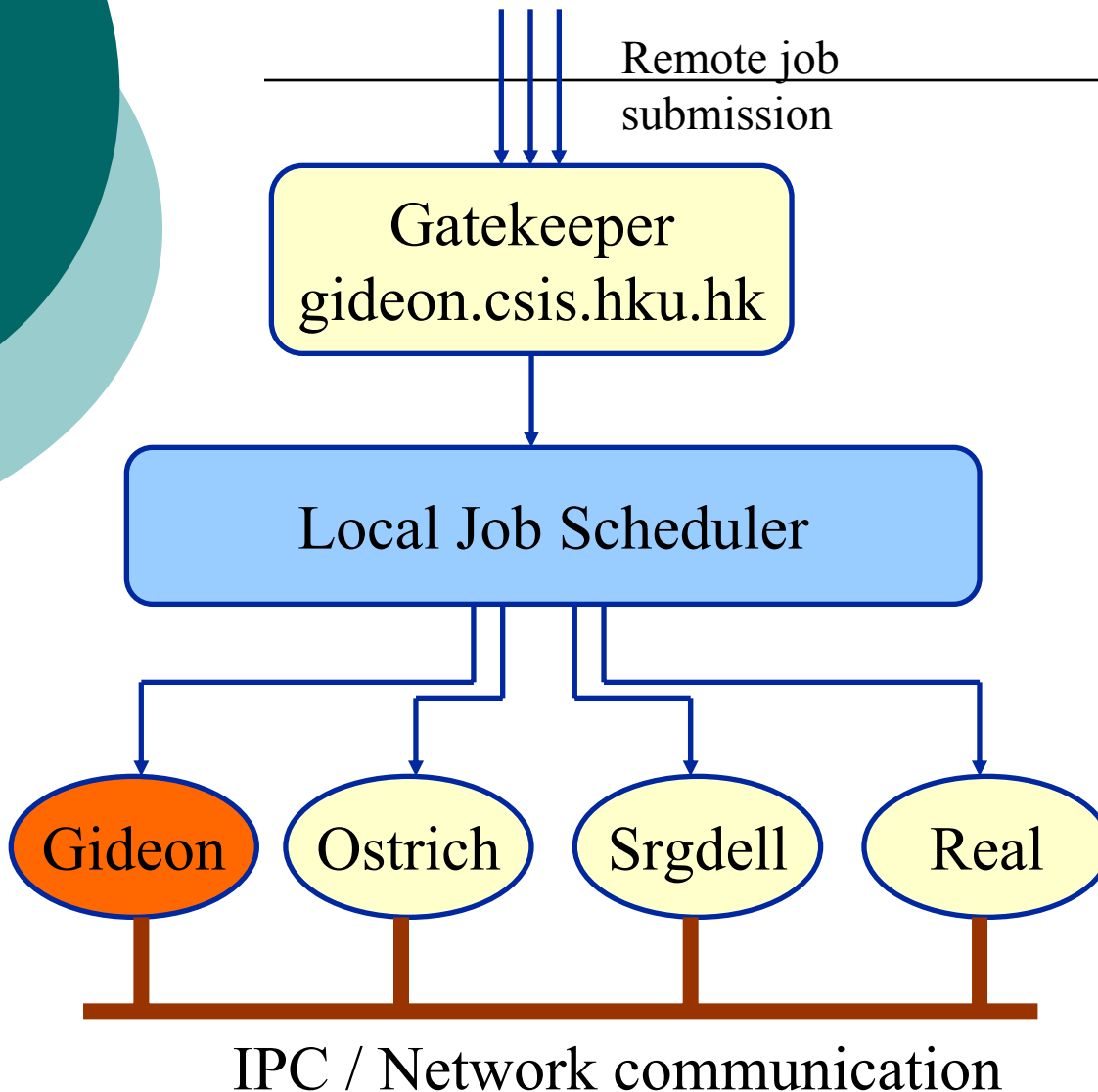
HKU Computer Science

“Self-Made” Gideon 300 Linux cluster



300 Pentium 4 PCs @355 Gflops; Ranked #175 in TOP500 (11/2002)

HKU CS Grid Point: Grid and Cluster Software



Grid middleware

- **Globus Toolkit (GT) 2.0, 2.4, 3.0.1**

Job scheduling

- **OpenPBS 2.3.16**
- **Maui 3.2.5**

Programming

- **HPF, Fortran 90**
- **C, C++, Java with MPI**
- **JESSICA2 (HKU)**
- **WireGL, MatlabMPI**

Communication Lib

- **MPICH-G2**

Performance Monitoring with Ganglia

HKU-CSIS Grid > Gideon cluster > GD269B

GD269B Overview



This node is up and running

Time and String Metrics

Name	Value
boottime	Sat, 30 Aug 2003 01:46:09 +0800
genec	OFF
machine_type	x86
os_name	Linux
os_release	2.4.18-14custom
sys_clock	Sat, 30 Aug 2003 16:28:12 +0800
uptime	4 days, 9:57

Constant Metrics

Name	Value
cpu_wide	97.9%
cpu_num	1
cpu_speed	2000 MHz
mem_total	505664 KB
mbs	1500 B
swap_total	9772552 KB



HKU-CSIS Grid Report for Wed, 3 Sep 2003 11:32:13 +0800

Get Fresh Data

Last Sorted

HKU-CSIS Grid >

Name / Info

HKU-CSIS Grid (4 sources) [\(raw view\)](#)

Hosts up: 83
(83 CPUs Total)

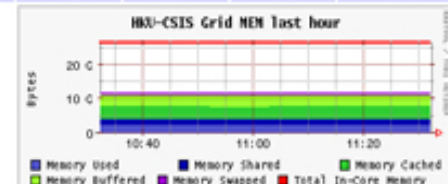
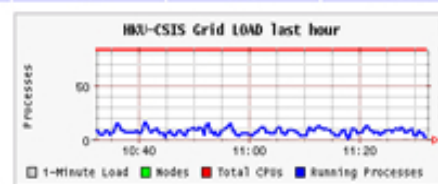
Hosts down: 0

Load Averages

0.54 0.43 0.20

%CPU User, Nice, System, Idle

0.1 0.0 0.1 99.7

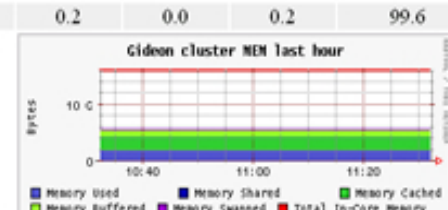
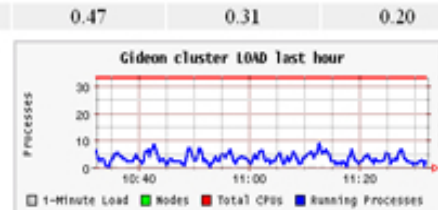


[Gideon cluster \(physical view\)](#)

Cluster Localtime:
September 3, 2003, 11:32 am

Hosts up: 33
(33 CPUs Total)

Hosts down: 0

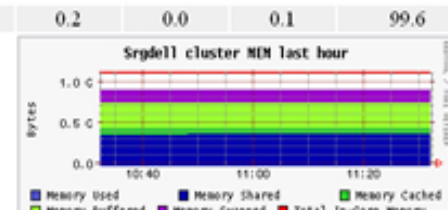
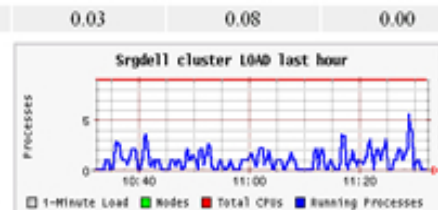


[Srgdell cluster \(physical view\)](#)

Cluster Localtime:
September 3, 2003, 11:32 am

Hosts up: 9
(9 CPUs Total)

Hosts down: 0

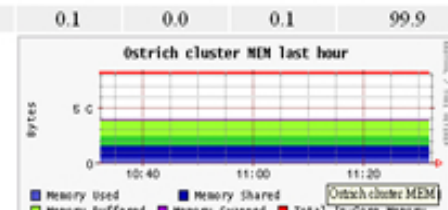
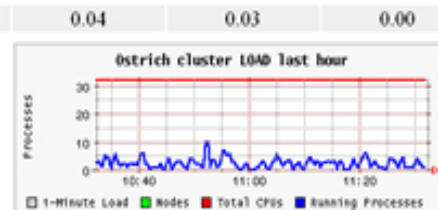


[Ostrich cluster \(physical view\)](#)

Cluster Localtime:
September 3, 2003, 11:31 am

Hosts up: 32
(32 CPUs Total)

Hosts down: 0

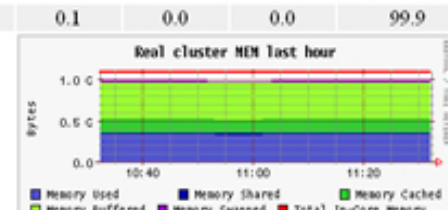
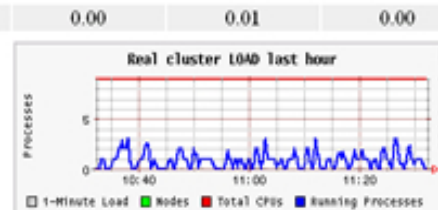


[Real cluster \(physical view\)](#)

Cluster Localtime:
September 3, 2003, 11:32 am

Hosts up: 9
(9 CPUs Total)

Hosts down: 0



URL: <http://gideon.cs.hku.hk/hkgrid/> ^{1/}

China National Grid : HKU Grid Point



上海超级计算中心

中科院计算所

香港大学 (Gideon300)

西安交通大学

中国科技大学

国防科技大学

中科院应用物理所

清华大学

DAWNING TC4000A, SUPER
copyright by Dawning internal
2003

Dawning4000A (2560
Opteron proc, now
17th in TOP500)

DeepComp 6800
(1024 I2 proc, now
38th in TOP500)

First test run
on Dec. 27,
2004



Supporting software:

Vega (织女星) GOS: dynamic service deployment, single-sign-on, data replication, and performance monitoring.
Developed by Institute of Computing Technology, Chinese Academy of Sciences (中科院计算所)

(2004. Nov. 28) : HKU supports G-JavaMPI, JESSICA2, WireGL, MatlabMPI

China National Grid - 欢迎使用中国国家网格 - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://147.8.179.124:8080/index.jsp

Applet Tutorial: Backbuffers

China National Grid - 欢迎使...

 **中国国家网格** China National Grid www.cngri

网格系统

- 退出
- 个人信息
 - 浏览/修改
- 组管理
 - 组用户审批
- 资源管理
 - 浏览/使用资源
 - 添加资源
- 任务管理
 - 查看任务状态
 - 全部列表
- 系统监控
 - 查看系统状态
- 记账信息
 - 统计
 - 明细

可用资源列表 【刷新】 【后退】 【前进】

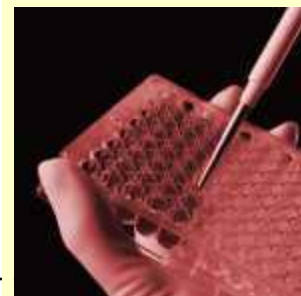
ICT2 组

资源名	地址	删除
WireGL	http://147.8.179.124:8081/ogsa/services/HKUServices/WireGL	删除
Mpp960	http://147.8.179.124:8081/ogsa/services/gos/Mpp960Service	删除
1	http://147.8.179.124:8081/ogsa/services/gos/BatchService	删除
CalPI	http://147.8.179.124:8081/ogsa/services/HKUServices/CalPI	删除
JmpiBLAST	http://147.8.179.124:8081/ogsa/services/HKUServices/JmpiBLAST	删除
01000001	http://147.8.179.124:8081/ogsa/services/mymath/MathService	删除
MatMPI	http://147.8.179.124:8081/ogsa/services/HKUServices/MatMPI	删除

Drug Discovery Grid (DDGrid)

新药研发网格

<http://202.127.19.33/>



- **Shanghai Institute of Materia Medica** (上海药物所)
- **Shanghai Jiao Tong University** (上海交通大学)
- 江南计算技术研究所
- **University of Hong Kong** (香港大学)

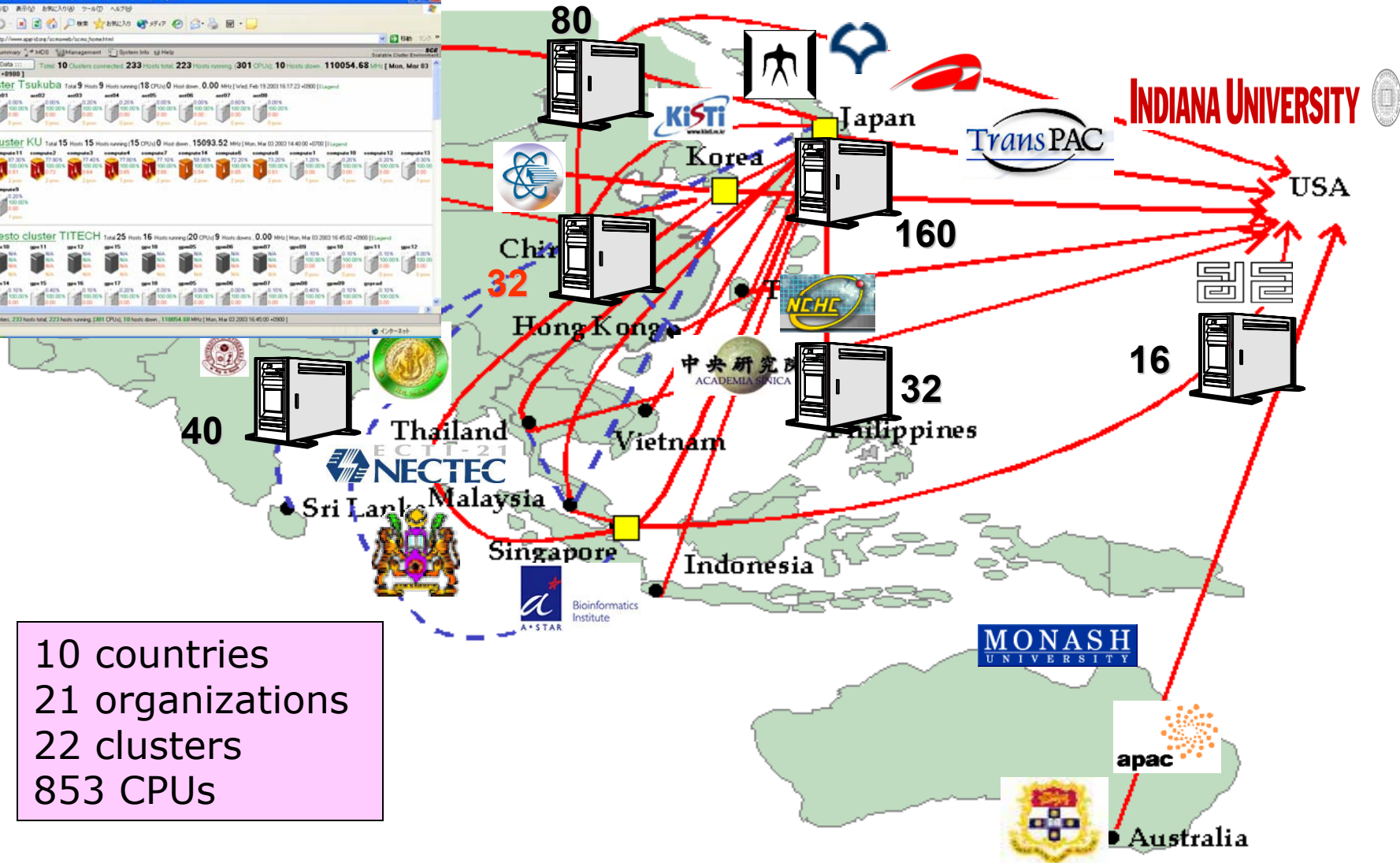
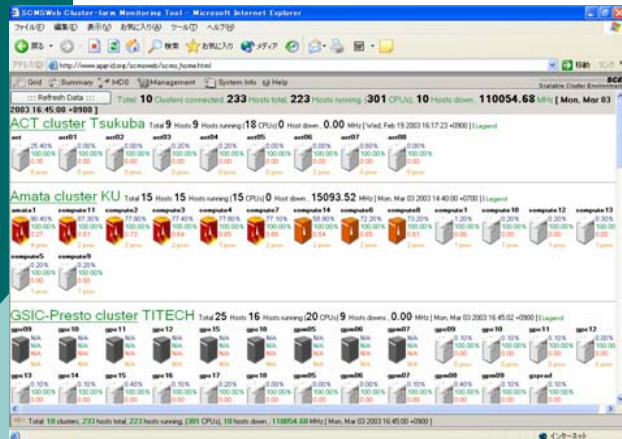
Database: 中国天然产物（中草药）分子数据库、合成化合物分子数据库，化合物毒性数据库、

Computing Resources: 上海药物所神威32A集群、北京军事医学科学院神威256P集群、香港大学Gideon 300集群、上海超级计算中心神威64P集群、曙光4000A、大连理工大学等多个网格结点

The screenshot displays the DDGrid website in a Microsoft Internet Explorer browser window. The address bar shows <http://202.127.19.33/p-grid/>. The website features a blue header with the '新药研发网格' (Drug Discovery Grid) logo and title. Below the header is a navigation bar with links for '用户登录' (User Login), '最新消息' (Latest News), '网络介绍' (Network Introduction), '使用网络' (Use Network), '软件下载' (Software Download), and '联系我们' (Contact Us). The main content area includes a '用户登录' section with fields for '邮箱' (Email) and '密码' (Password), and a '最新消息' section with a list of recent news items. A large 3D molecular model is prominently displayed on the right side of the page.

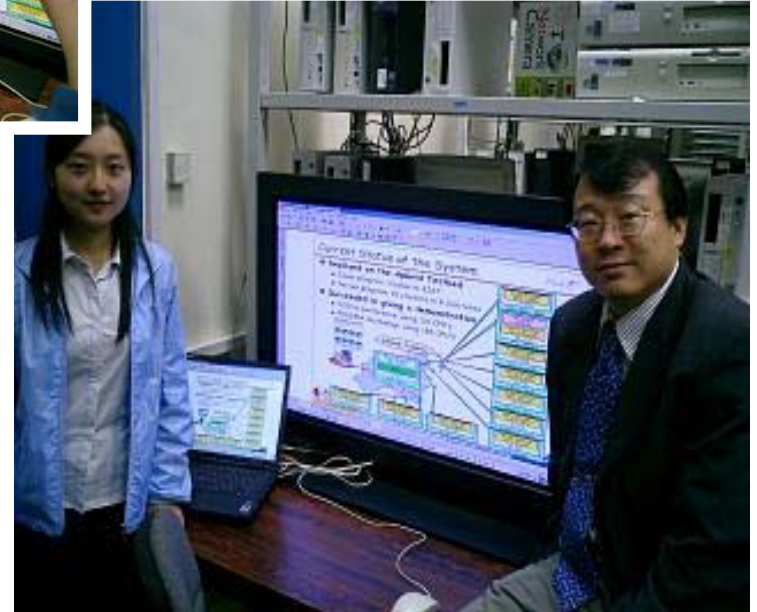
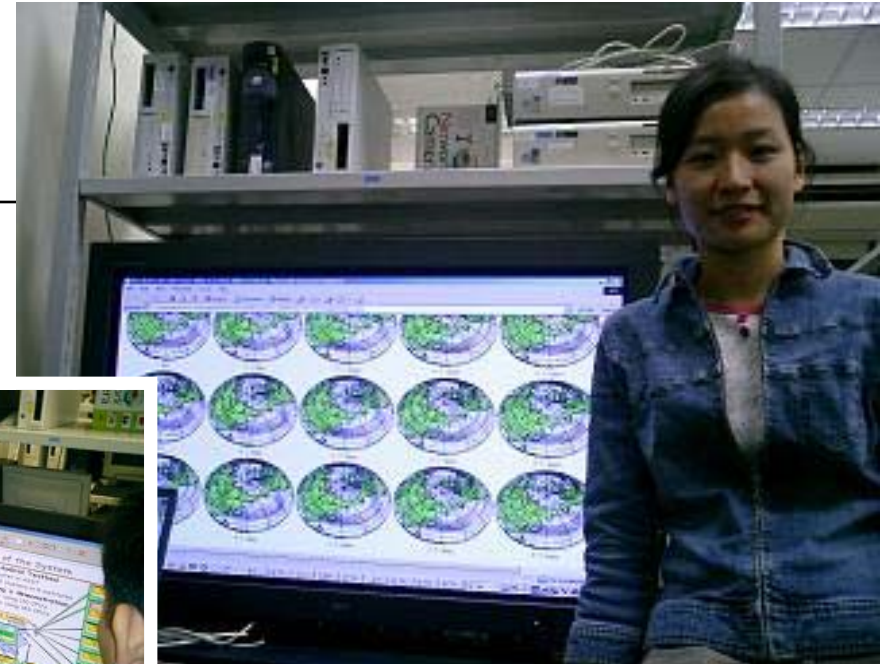
化合物数据库筛选 20

Asia Pacific Grid (APGrid)



10 countries
21 organizations
22 clusters
853 CPUs

Weather Forecast Demonstration on HKU Open Day – (Oct 2003)





Grid Research at HKU SRG

Selected Projects

- SLIM + InstantGrid
- JESSICA2
- G-JavaMPI + G-PASS

Acknowledgement



HKU Systems Research Group (SRG)

Our Goal

To construct an advanced grid computing platform to accommodate **utility-like computing** via **traditional** and **“pervasive” means**

- **Utility computing**: to aggregate and make use of distributed computing resources **transparently**
- **Traditional means**: to utilize the dedicated HPC facilities distributed across institutions
 - **Performance** and **reliability** are key
- **Pervasive means**: any user can be resource provider (e.g., idle PCs, etc.) or consumer, or both
 - **Convenience** and **security** are key

An Advanced Grid Computing Platform

Objectives

(Programming Environment)

(Execution Environment)

User's convenience

system administrator's convenience

Performance and Reliability

Grid point construction

AGP

G-JavaMPI

JESSICA

SLIM

InstantGrid

Research Issues

Load
balancing

Single-
system
image

On-demand Grid point
construction (ODGPC)

SLIM

Single Linux Image Management

文/攝影：Vincent

香港大學發明快速 Linux 部署方案

SLIM 專案即將開放源碼

供全球電腦用戶自由使用



SLIM 專案的兩位發明人，分別是香港大學計算機科學及資訊系統系電腦師孔慶輝(右)及助理電腦師李俊明(左)。

香港大學作為本地歷史最悠久的專上學府，一直以培育世界級的科研、人文人才為使命，在全球開放源碼運動上，他們即將有一項震撼世界的貢獻，Linux Pilot 讀者將可率先了解這項劃時代的開放源碼專案 SLIM，將如何在 Linux 的教育、科研及企業應用上發揮巨大影響力。

**On-demand construction of
customized execution
environments**

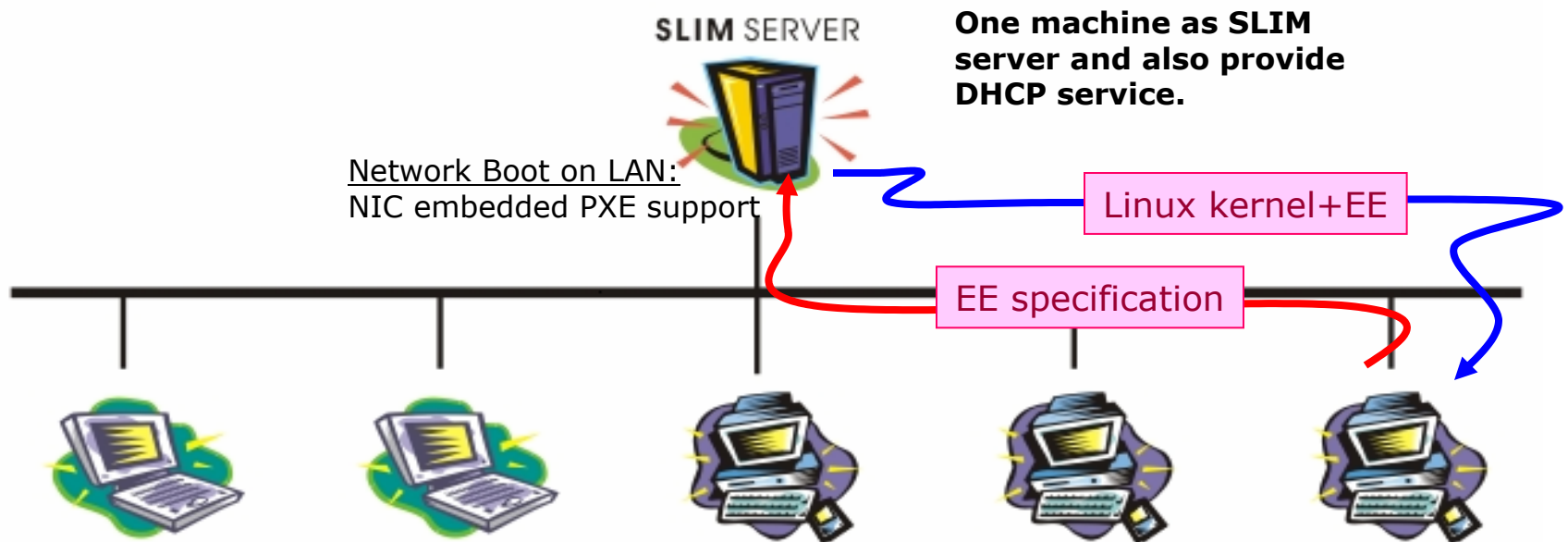
(LinuxPilot 2004/04)

SLIM

- Utility computing **decouples** computing platforms (resources) and computing logic (applications)
- I.e., a single platform can run completely different applications
- **Problem**: different applications demand different execution environments (OS, shared libraries, supporting apps, etc.)
- SLIM is a network service for **managing** and **constructing** EE's, and **disseminating** them to remote computing platforms

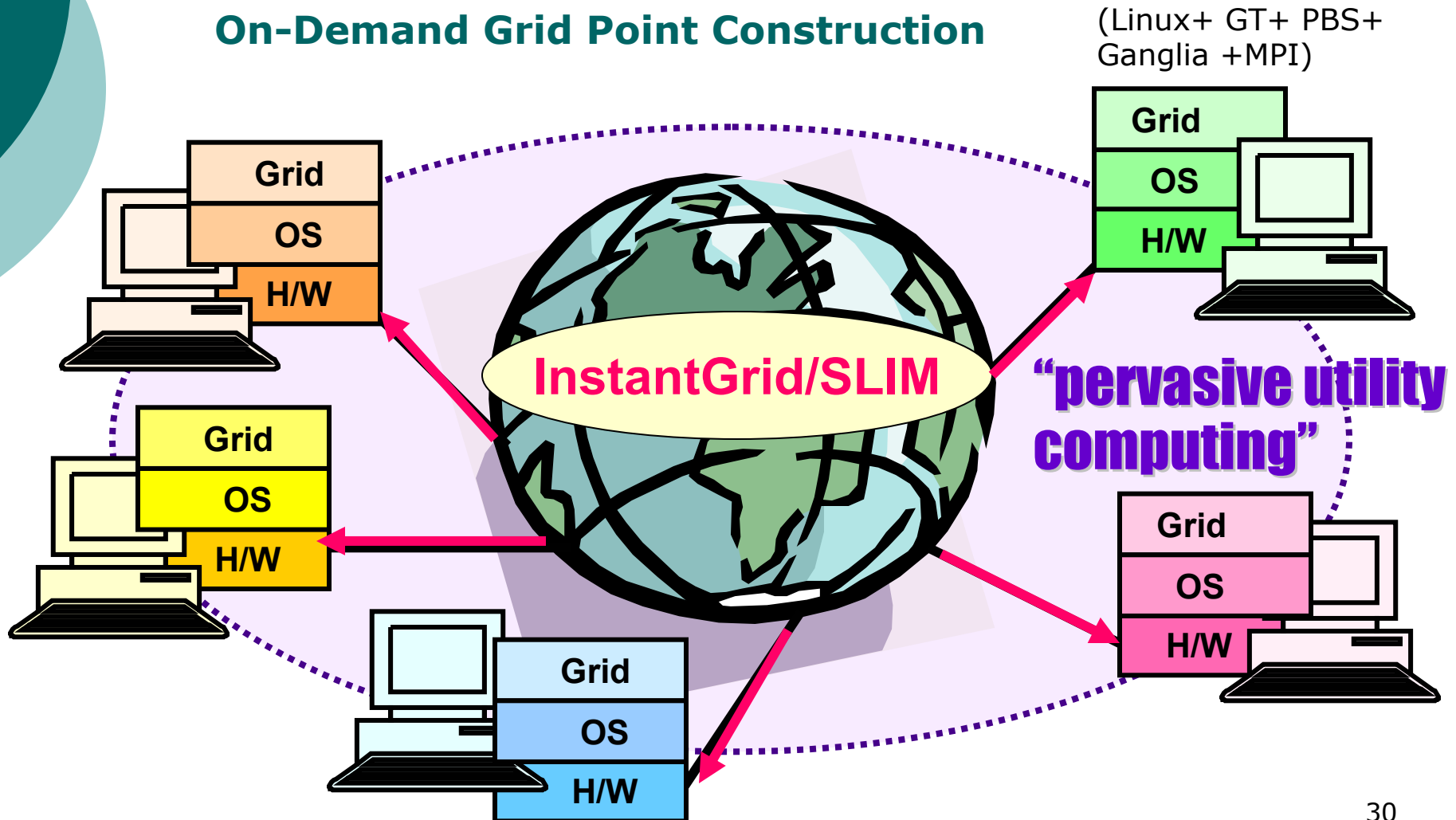
SLIM

- **On-demand construction of customized execution environments (EE)**
 - Single Linux system image shared across the network.
 - Centralized system administration and software management. *Do once, and use by all.*

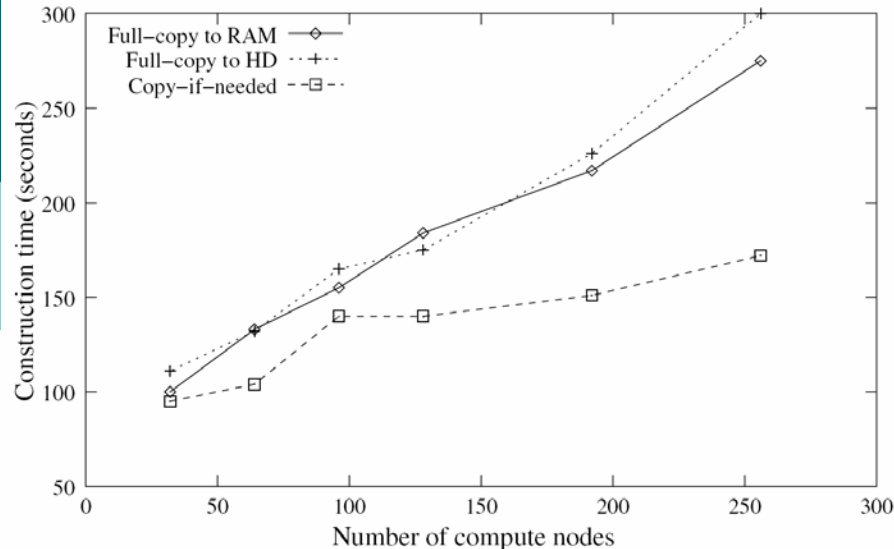


InstantGrid on SLIM

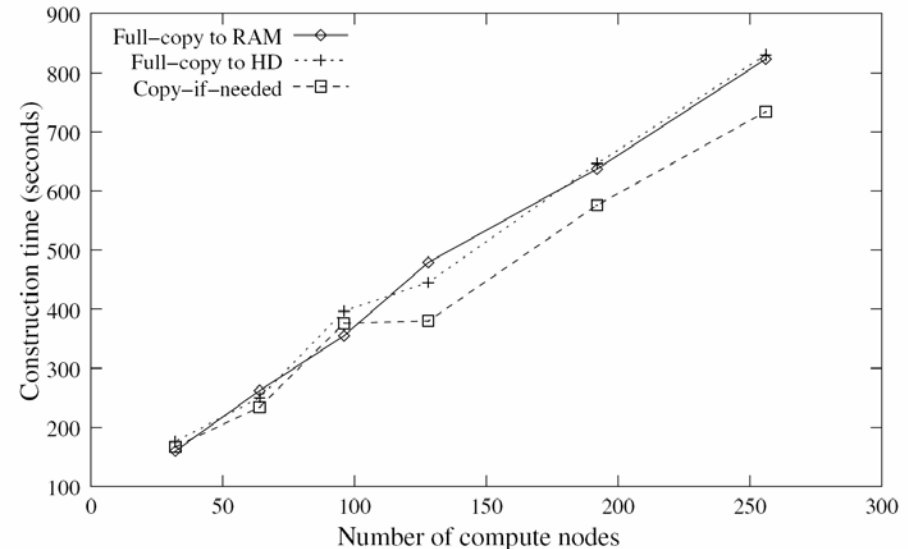
On-Demand Grid Point Construction



InstantGrid Performance



(a) A cluster-based grid point



(b) Standalone grid points

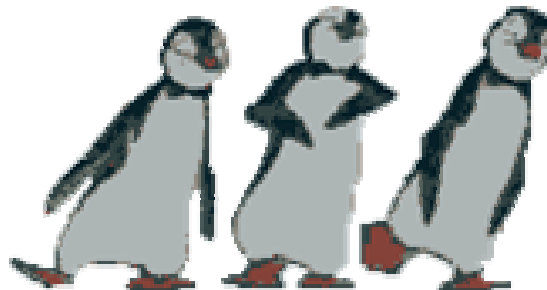
- Construct a 256-node grid point **from scratch** (PXE enabled) through Fast Ethernet in **three** (copy-if-needed) to **five** (full-copy to hard disk) **minutes** using one SLIM server
- Construct 256 standalone grid points take longer time to construct. The overhead mainly lies on the process to generate host certificates

SLIM – Ongoing and future work

- SLIM has been managing:
 - the HKU CS grid point (350 nodes) for various grid research projects
 - an addition 300+ lab machines for teaching purpose (different courses have different requirements)
- Future work
 - To overcome the challenges in deploying SLIM over broadband links for realizing the “pervasive utility computing”

SLIM – Key references

- Download:
<http://slim.cs.hku.hk/>
- SLIM released since April 2004. >150 downloads; from Mainland China, Hong Kong, Macau, Taiwan, USA, and Singapore



如何在中學架設Linux Clusters

香港大學與Linux Pilot合辦 Linux Clusters工作坊

計算機集群 (Clusters) 技術可將多部電腦的資源整合，從而建立一個廉價的超級計算 (Supercomputing) 平台。此技術已被廣泛地應用在不同的行業，例如科學研究、金融運算、生物科技，和電影制作等。除此之外，電腦集群亦是一種理想的學習工具。透過架設和使用電腦集群，使用者能學習到高性能運算 (High Performance Computing) 平台的原理和特性，以及許多既有趣又實用的編程 (Programming) 技巧和應用程式。有見及此，香港大學計算機科學系聯同 Linux Pilot 雜誌舉辦全港首個專為中學而設的 Linux Clusters 工作坊，旨在推廣高性能運算與及有關的開源軟件 (Open Source Software) 在中學的學習和應用。

工作坊內容如下：

- 電腦集群的基本原理和結構
- 如何運用免費的 Linux 操作系統和一些舊電腦架設 Linux Clusters
- 介紹由本系研發的 SLIM 集群管理系統 (<http://slim.cs.hku.hk/>)，及如何運用此系統簡化集群的安裝和管理工作
- 集群在中學的應用範例
- 學員更可親自利用提供的電腦嘗試架設集群，並即場運用它們來作簡單的圖像處理

本工作坊的對象為中學老師和中四至中七的學生，我們並鼓勵師生一同參與。由於名額有限，一間中學只可派出兩名學員參加（一位老師和一位同學，或兩位同學均可）。工作坊將於香港大學開放日期間（2004年10月16和17日）舉辦四次，有興趣者可選擇其中一節參加。

日期和時間：

- 第一節：2004 年 10 月 16 日上午十一時至下午一時
 - 第二節：2004 年 10 月 16 日下午三時至下午五時
 - 第三節：2004 年 10 月 17 日上午十一時至下午一時
 - 第四節：2004 年 10 月 17 日下午三時至下午五時
- （每節可供三十名學員參加）

地點：香港大學綜合大樓（徐朗星文娛中心）地庫 104 室
(Room LG104, Composite Building, HKU)

費用：全免。

報名辦法：

登入以下網址：<http://www.cs.hku.hk/linux-workshop/>
如老師參與，請老師替同學報名，否則同學可自行報名。

查詢：香港大學計算機科學系嚴小姐
電話：2241 5757
電郵：support@cs.hku.hk

名額有限，成功報名的學員將於 10 月 13 日前收到電郵確認。
完成工作坊後學員將獲頒發證書和紀念品乙份。



Linux Cluster Workshop (2004/Oct.)



JESSICA2

Java
Enabled
Single
System
Image
Computing
Architecture



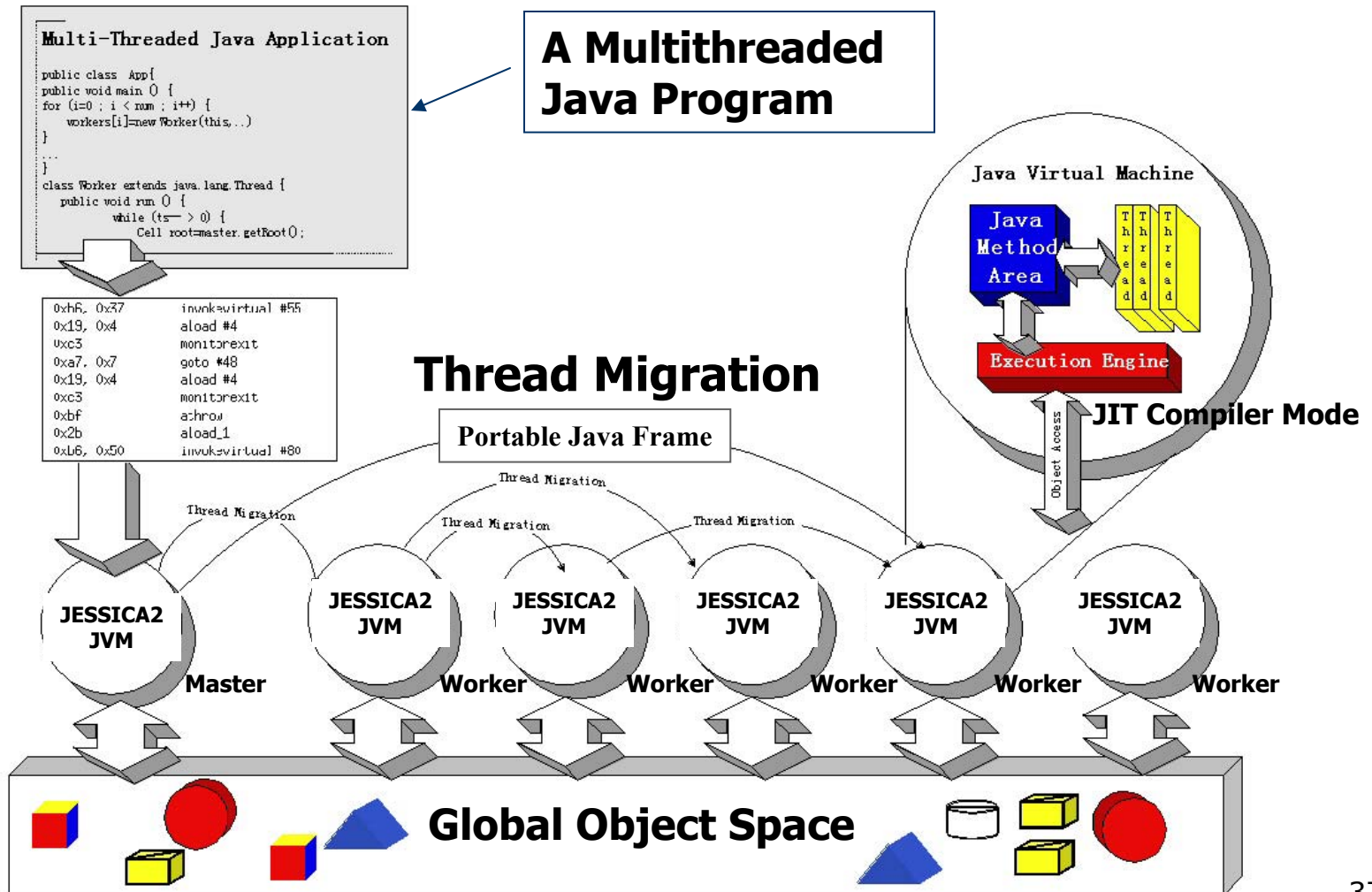
JESSICA Team



JESSICA2

- JESSICA2 is a ***distributed Java Virtual Machine (DJVM)*** which consists of a group of extended JVMs running on a distributed environment to support true parallel execution of a multithreaded Java application.
- Java threads can freely move across node boundaries and execute in parallel to achieve more scalable high-performance computing.
- The JESSICA2 DJVM provides standard JVM services, that are compliant with the Java language specification, as if running on a single machine – **Single System Image (SSI)**.

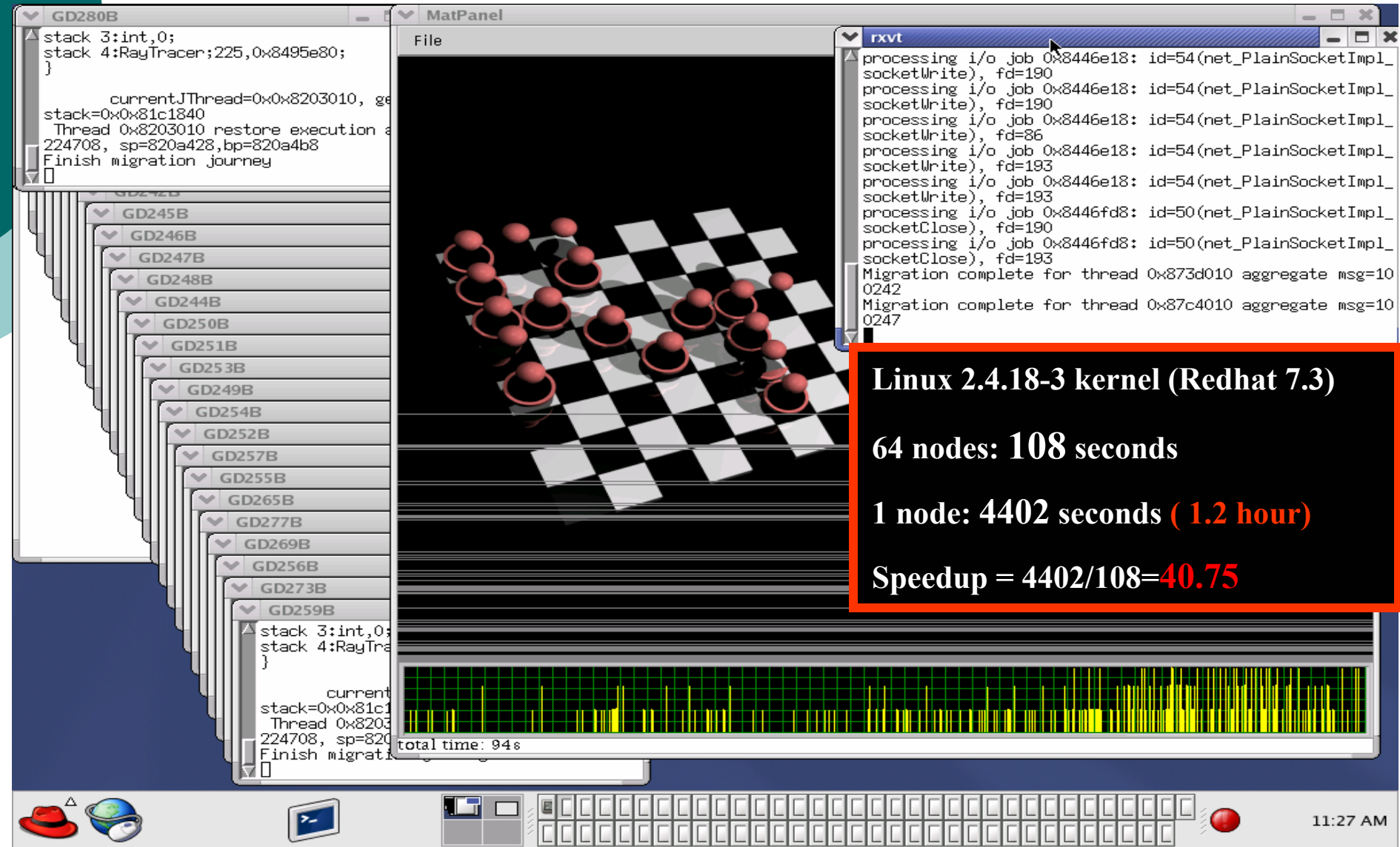
JESSICA2 Architecture



JESSICA2 Main Features

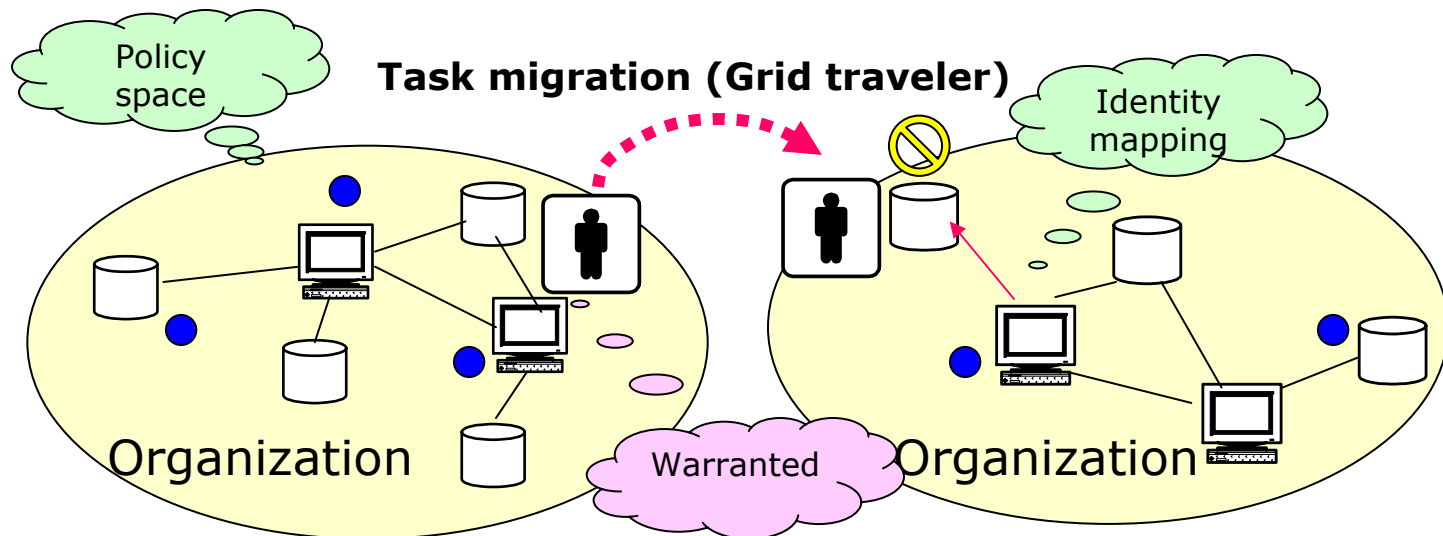
- **Transparent Java thread migration**
 - Runtime capturing and restoring of thread execution context.
 - No source code modification; no bytecode instrumentation (preprocessing); no new API introduced
 - Enable dynamic load balancing
- **Full Speed Computation**
 - JITEE: cluster-aware bytecode execution engine
 - Operated in Just-In-Time (JIT) compilation mode
 - Zero cost if no migration
- **Transparent Remote Object Access**
 - Global Object Space : A shared global heap spanning all running nodes
 - Adaptive migrating home protocol for memory consistency + various optimizing schemes.
 - I/O redirection

Ray Tracing on JESSICA2 (64 PCs)



G-JavaMPI

A grid-enabled Java-MPI system with dynamic load-balancing via process migration



G-JavaMPI

- **A grid middleware that supports portable messaging-passing programming for achieving dynamic **load-balancing** and **non-stop** parallel computing in grid.**
- **Special feature: Transparent Java process migration**
 - State capturing and restoration through JVM Debugger Interface (JVMDI). No modification of JVM
 - Facilitates more flexible task scheduling and more effective resource sharing. Avoid running hotspots.
- **G-PASS: security enhancement for G-JavaMPI**
 - Perform identity mapping and access control while Java processes move across multiple grid points that are under different control policies. Avoid chain-delegation.
- **Migration policies :**
 - Grid point CPU and network workload
 - Application's communication pattern
 - Scheduled down time
 - Data location

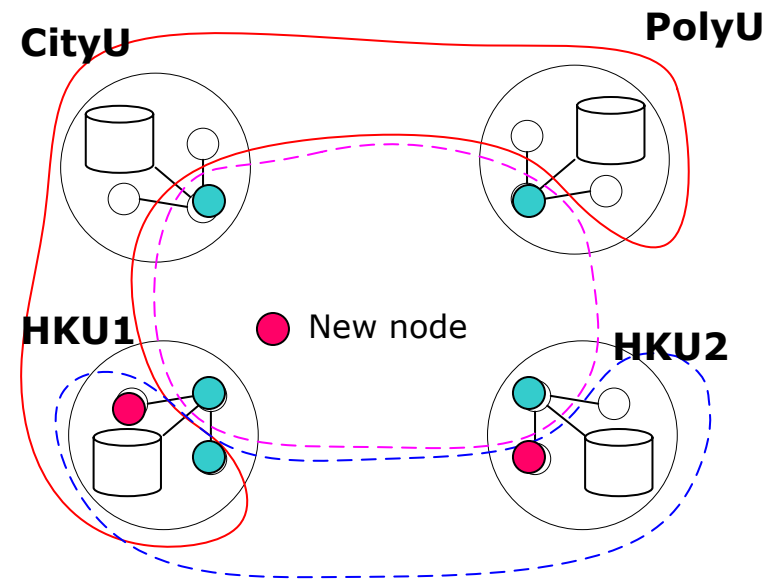
Preliminary Results at HKGrid

- **Parallel BLAST program implemented by G-JavaMPI**
 - Three universities sharing CPU cycles and local bio-databases
 - Executing 3 Blastp programs concurrently, total 18 processes
 - Original no. of nodes: 5; 2 nodes join then 2 nodes quit
- The size of the migrated execution context is about **2.1 Kbytes**.
- Total execution time : **566~911 seconds** under different scheduling policies.

Migration Overhead Analysis

	HKU-PolyU	PolyU-CityU	HKU-CityU
G-PASS	1.21s	0.51s	0.43s
Migration	1.90s	1.67s	0.46s
Total	3.112	2.18s	0.89s

Single process migration is **less than 0.5%** of the total execution time under different CPU load.





Summary

- **Performance**

- SLIM and InstantGrid: for high-speed construction of Grid computing environment, establish extensible grid platforms
- G-JavaMPI and JESSICA : Process/thread migration enables performance optimization and load balancing

- **Reliability**

- Java checkpointing (G-JavaMPI and JESSICA)
- SLIM helps construct platforms for failover

- **Convenience**

- G-JavaMPI and JESSICA enable users to utilize HPC facilities distributed across institutions via traditional means (e.g., message passing, Java)
- SLIM and InstantGrid fulfill on-demand Grid point construction, and simplify Grid point management.



Conclusion

- Grid/utility computing are relatively new paradigms that deserve further investigation
- We address the **performance, reliability,** and **user convenience** issues in grid/utility computing
- Our advanced grid computing platform (consisting of G-JavaMPI/G-PASS, JESSICA2, and SLIM/InstantGrid) is geared to deploy in the HKGrid for easy adoption of Grid technologies.



Thanks!

For more information:

The HKU Systems Research Group

<http://www.srg.csis.hku.hk/>

Hong Kong Grid

<http://www.hkgrid.org/>

Grid Computing Research Portal

<http://grid.csis.hku.hk/>