



SDSC Grids and Clusters

Mason Katz

Group Lead, Cluster Development

San Diego Supercomputer Center

Overview of San Diego Supercomputer Center

- Founded in 1985
 - **Non-military access to supercomputers**
- Over 400 employees
- Mission: Innovate, develop, and deploy technology to advance science
- Recognized as an international leader in:
 - **Data Management**
 - **High Performance Computing**
 - **Grid and Cluster Computing**
 - **Networking**
 - **Visualization**
- Primarily funded by NSF

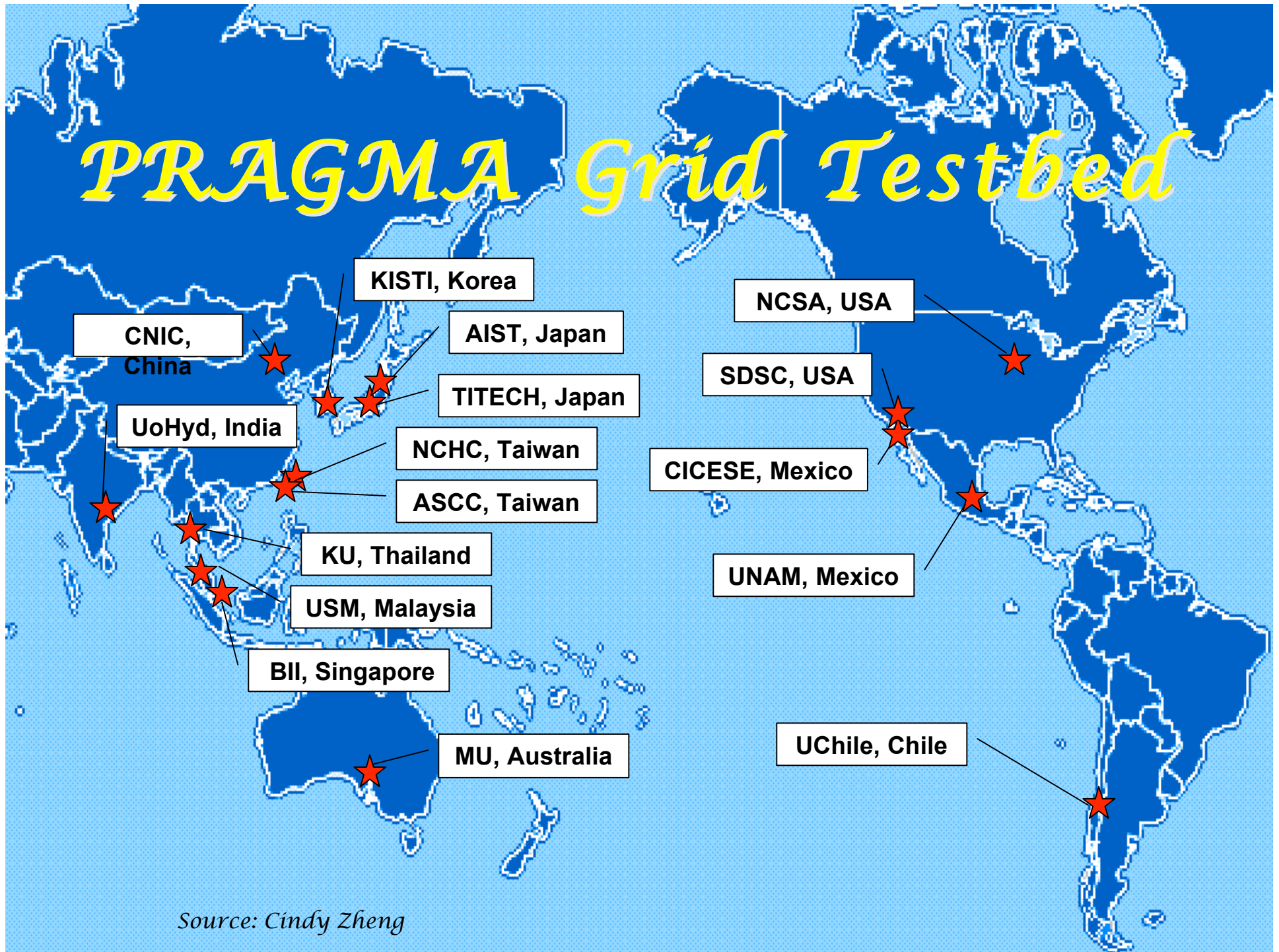


PRAGMA

Pacific Rim Applications and Grid
Middleware Assembly



PRAGMA Grid Testbed



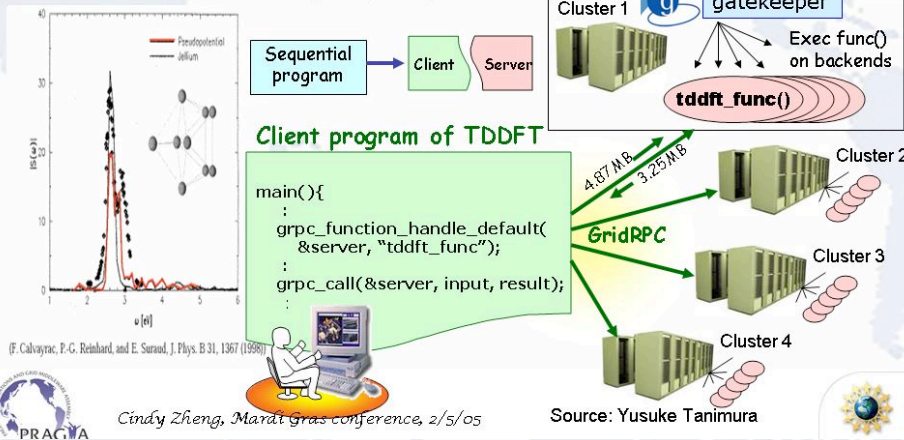
Source: Cindy Zheng

Routine Use Applications

1st application

Time-Dependent Density Functional Theory (TDDFT)

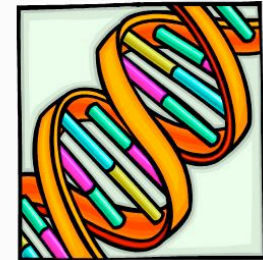
- Computational quantum chemistry application
- Simulate how the electronic system evolves in time after excitation
- Grid-enabled by Nobusada (IMS), Yabana (Tsukuba Univ.) and Yusuke Tanimura (AIST) using Ninf-G



2nd Application - mpiBLAST

A DNA and Protein sequence/database alignment tool

- Driver: Hurng-Chun Lee (ASCC, Taiwan)
- Application requirements
 - Globus
 - Mpich-g2
 - NCBI est_human, toolbox library
 - Public ip for all nodes
- Started 9/20/04
- SC04 demo
- Automate installation/setup/testing

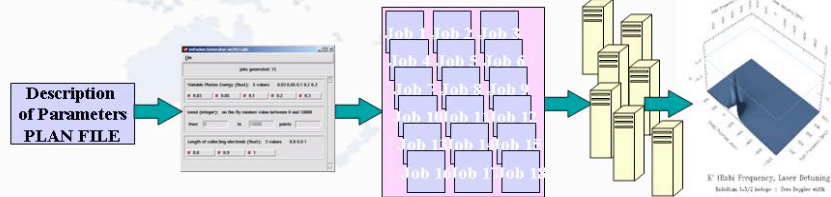
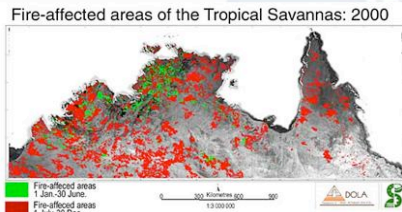


<http://pragma-goc.rocksclusters.org/biogrid/default.html>

3rd Application – Savannah Case Study

Study of Savannah fire impact on northern Australian climate

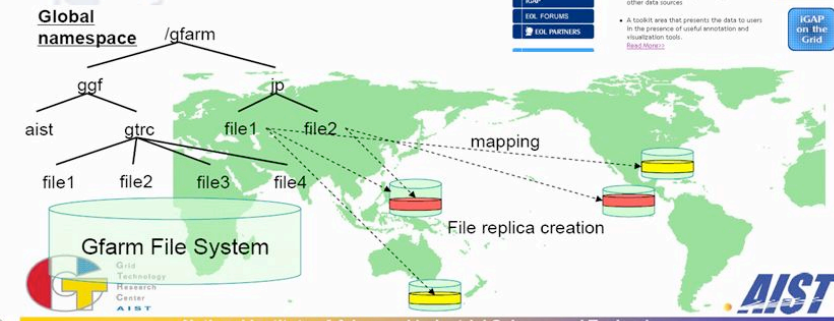
- Climate simulation model
- 1.5 month CPU * 90 experiments
- Started 12/3/04
- Driver: Colin Enticott (Monash University, Australia)
- Requires GT2
- Based on Nimrod/G



<http://pragma-goc.rocksclusters.org/savannah/default.html>

4th Application – iGAP/Gfarm

- iGAP and EOL (SDSC, USA)
 - Genome annotation pipeline
- Gfarm – Grid file system (AIST, Japan)
- Demo in SC04 (SDSC, AIST, BII)
- Plan to start in testbed February 2005



Cindy Zheng, Mardi Gras conference, 2/5/05

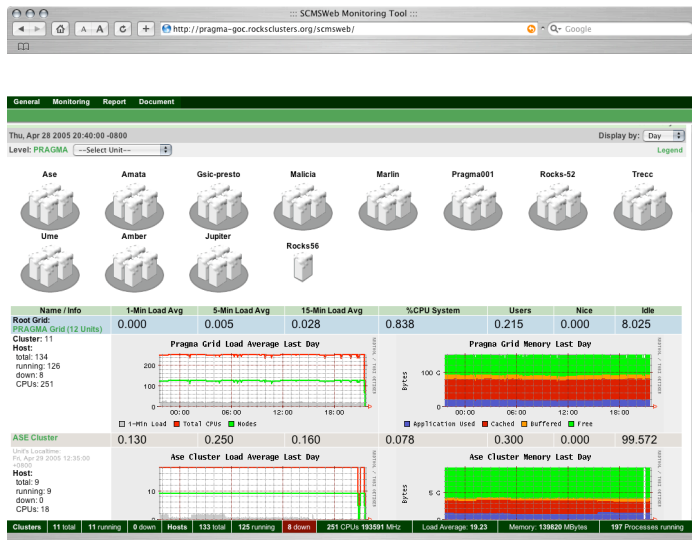
Resources

| Country / Region | Institution Name | System Name | Submit System Info | Provide Users Guide | Setup Monitor Account | SCMSWeb Ready | Joined Testbed Since | Status |
|------------------|------------------|-------------|--------------------|---------------------|-----------------------|---------------|----------------------|-----------|
| Japan | AIST | ume | Done ☺ | Done ☺ | Done ☺ | Done ☺ | 2004/06/01 | active |
| Korea | KISTI | jupiter | Done ☺ | Done ☺ | Done ☺ | Done ☺ | 2004/06/01 | active |
| Thailand | KU | amada1 | Done ☺ | Done ☺ | Done ☺ | Done ☺ | 2004/06/02 | active |
| Malaysia | USM | aurora | Done ☺ | Done ☺ | Done ☺ | Ganglia | 2004/06/02 | active |
| USA | SDSC | rocks-52 | Done ☺ | Done ☺ | Done ☺ | Done ☺ | 2004/06/17 | active |
| Japan | TITECH | gsic-presto | Done ☺ | | Done ☺ | Done ☺ | 2004/06/30 | active |
| Taiwan | NCHC | ase | Done ☺ | Done ☺ | Done ☺ | Done ☺ | 2004/07/05 | active |
| USA | NCSA | ccg-login | Done ☺ | Done ☺ | Done ☺ | Done ☺ | 2004/07/07 | In active |
| Singapore | BII | viper | Done ☺ | Done ☺ | Done ☺ | Ganglia | 2004/08/13 | active |
| Mexico | UNAM | malicia | Done ☺ | | Done ☺ | Done ☺ | 2004/08/25 | active |
| Taiwan | ASCC | Pragma001 | Done ☺ | Done ☺ | Done ☺ | Done ☺ | 2004/10/01 | active |
| India | UoHyd | amber | Done ☺ | Done ☺ | Done ☺ | Done ☺ | 2004/10/12 | active |
| Singapore | BII | marlin | Done ☺ | Done ☺ | Done ☺ | Done ☺ | 2004/10/13 | active |
| Mexico | CICESE | solaris | Done ☺ | Done ☺ | Done ☺ | Ganglia | 2004/11/30 | active |
| Australia | MU | mahar | Done ☺ | | Done ☺ | | 2004/12/17 | active |
| USA | SDSC | rocks-47 | Done ☺ | | Done ☺ | Done ☺ | 2004/12/21 | active |
| USA | NCSA | tgc | Done ☺ | Done ☺ | Done ☺ | Done ☺ | 2005/02/01 | active |
| Malaysia | USM | hawk | Done ☺ | | Done ☺ | | 2005/02/16 | active |
| China | CNIC | Pragma | Done ☺ | Done ☺ | Done ☺ | | 2005/04/19 | active |

Grid Status Book Keeping

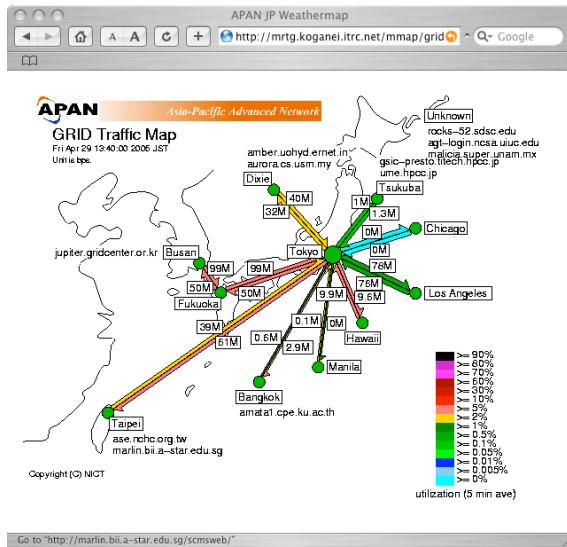
| Institution | Hostname | Compute Nodes | Application requirements | | | SCMSWeb upgrade | Notes |
|-------------|-------------|---------------|--------------------------|--------------------|--------------------|--------------------|---------------------------------------|
| | | | BioGrid | Savannah | QM-MD | | |
| AIST | ume | 32 | Done ☺ | Done ☺ | Done ☺ | Done ☺ | All done ☺ |
| ASCC | pragma001 | 3 | Done ☺ | Done ☺ | Done ☺ | Done ☺ | All done ☺ |
| BII | marlin | 4 | Exempt | Done ☺ | Done ☺ | Done ☺ | All done ☺ |
| CICESE | solaris | 7 | Exempt | Exempt | <i>in progress</i> | Exempt | <i>Working on hardware problem</i> |
| CNIC | pragma | 8 | Exempt | Exempt | Done ☺ | Exempt | All done ☺ |
| KISTI | jupiter | 16 | Done ☺ | Done ☺ | Done ☺ | Done ☺ | All done ☺ |
| KU | amata1 | 14 | Exempt | Done ☺ | Done ☺ | Done ☺ | All done ☺ |
| MU | mahar | 50 | Exempt | Done ☺ | <i>in progress</i> | <i>in progress</i> | |
| NCHC | ase | 8 | Done ☺ | Done ☺ | Done ☺ | Done ☺ | All done ☺ |
| NCSA | tgc | 12 | Done ☺ | Done ☺ | Done ☺ | Done ☺ | All done ☺ |
| SDSC | rocks-52 | 15 | Done ☺ | Done ☺ | Done ☺ | Done ☺ | All done ☺ |
| SDSC | rocks-47 | 3 | Exempt | Exempt | Done ☺ | Exempt | All done ☺ |
| TITECH | gsic-presto | 8 | Done ☺ | Done ☺ | Done ☺ | Done ☺ | All done ☺ |
| UNAM | malicia | 5 | Exempt | Done ☺ | Done ☺ | Done ☺ | All done ☺ |
| UoHyd | amber | 8 | <i>in progress</i> | <i>in progress</i> | <i>in progress</i> | <i>in progress</i> | |
| USM | hawk | 16 | <i>in progress</i> | Done ☺ | <i>in progress</i> | <i>in progress</i> | |
| USM | aurora | 16 | Done ☺ | Done ☺ | Done ☺ | <i>in progress</i> | <i>Pending OS upgrade in May 2005</i> |
| Done | | | 8 | 11 | 13 | 10 | |
| Exempt | | | 7 | 3 | 0 | 3 | |
| In progress | | | 2 | 3 | 4 | 4 | |

Grid/Cluster Monitoring



- SCMSWeb
 - Kasetsart University, Thailand
 - SCE Group
- Ganglia
 - UC Berkeley
 - Millennium Group
- Working to support Ganglia data inside SCMSWeb

Network and Grid Monitoring



PRAGMA GRID STATUS
 http://rocks-52.sdsc.edu/pragma-grid-status.html

PRAGMA GRID STATUS (Fri Apr 29 1:05:00 GMT 2005)

| Site | Authenticate in Gatekeeper | GRIS | GSFTP Server | Simple Test of jobmanager | GSFTP Job Test of jobmanager | Simple Test of jobmanager: fork | GSFTP Job Test of jobmanager: fork | Simple Test of jobmanager: pbs | GSFTP Job Test of jobmanager: pbs | Simple Test of jobmanager: sge | GSFTP Job Test of jobmanager: sge | Simple Test of jobmanager: sums | GSFTP Job Test of jobmanager: sums |
|-----------------------------|----------------------------------|-----------|-----------------|---------------------------------|------------------------------------|--|---|---|--|---|--|--|---|
| amata1.cpe.ku.ac.th | Pass | Pass | Pass | | | Pass | Pass | Pass | Fail | | | Pass | Pass |
| ase.nchc.org.tw | Pass | Pass | Pass | | | Pass | Pass | Pass | | | | | |
| aurora.cs.usm.my | Timed Out | Timed Out | Timed Out | | | | | | | | | | |
| globus.csse.monash.edu.au | Pass | No Server | Fail | Pass | | | | | | | | | |
| gsic-presto.titech.hpcc.jp | Pass | Pass | Pass | Timed Out | | | | | | | | | |
| jupiter.gridcenter.or.kr | Pass | Timed Out | Pass | Pass | Fail | | | | | | | | |
| malicia.super.anam.mx | Pass | Pass | Pass | Pass | Pass | | | | | | | | |
| marlin.bii.a-star.edu.sg | Pass | Pass | Fail | Pass | | | | | | | | | |
| pragma.sdg.ac.cn | Pass | No Server | Pass | Fail | | | | | | | | | |
| pragma01.grid.sinica.edu.tw | Pass | No Server | Pass | Pass | Pass | | | | | | | | |
| rocks-47.sdsc.edu | Pass | Pass | Pass | Pass | Pass | | | | | | | | |
| rocks-52.sdsc.edu | Pass | Pass | Pass | Pass | Pass | | | | | | | | |
| solaris-1.cicese.mx | Fail | Pass | Fail | | | | | | | | | | |
| tgc.trecs.org | No Server | No Server | Pass | Pass | Pass | | | | | | | | |
| ume.hpcc.jp | Pass | Pass | Pass | | | Pass | Pass | | | Pass | Pass | | |
| 202.41.85.81 | Fail | No Server | Fail | | | | | | | | | | |

- Real-Time Network Traffic
 - **APAN**
 - **Custom map created for PRAGMA**
- Testing Grid Services
 - **NCSA (Originally used on Tera-Grid)**
 - **Actively probes all Globus services**



Rocks

Linux Clustering Solution



Rocks

Open Source Cluster Distribution

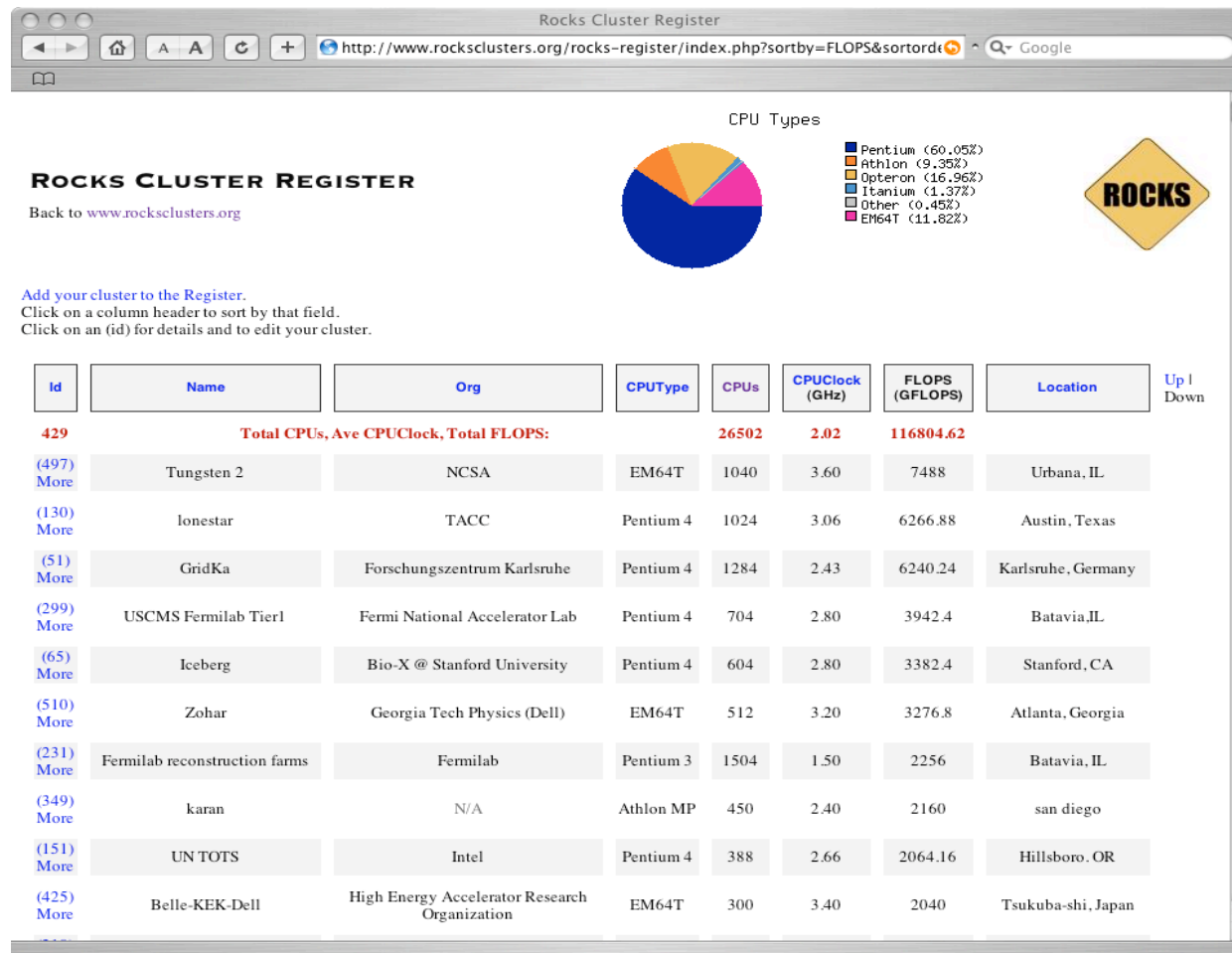
<http://www.rocksclusters.org>



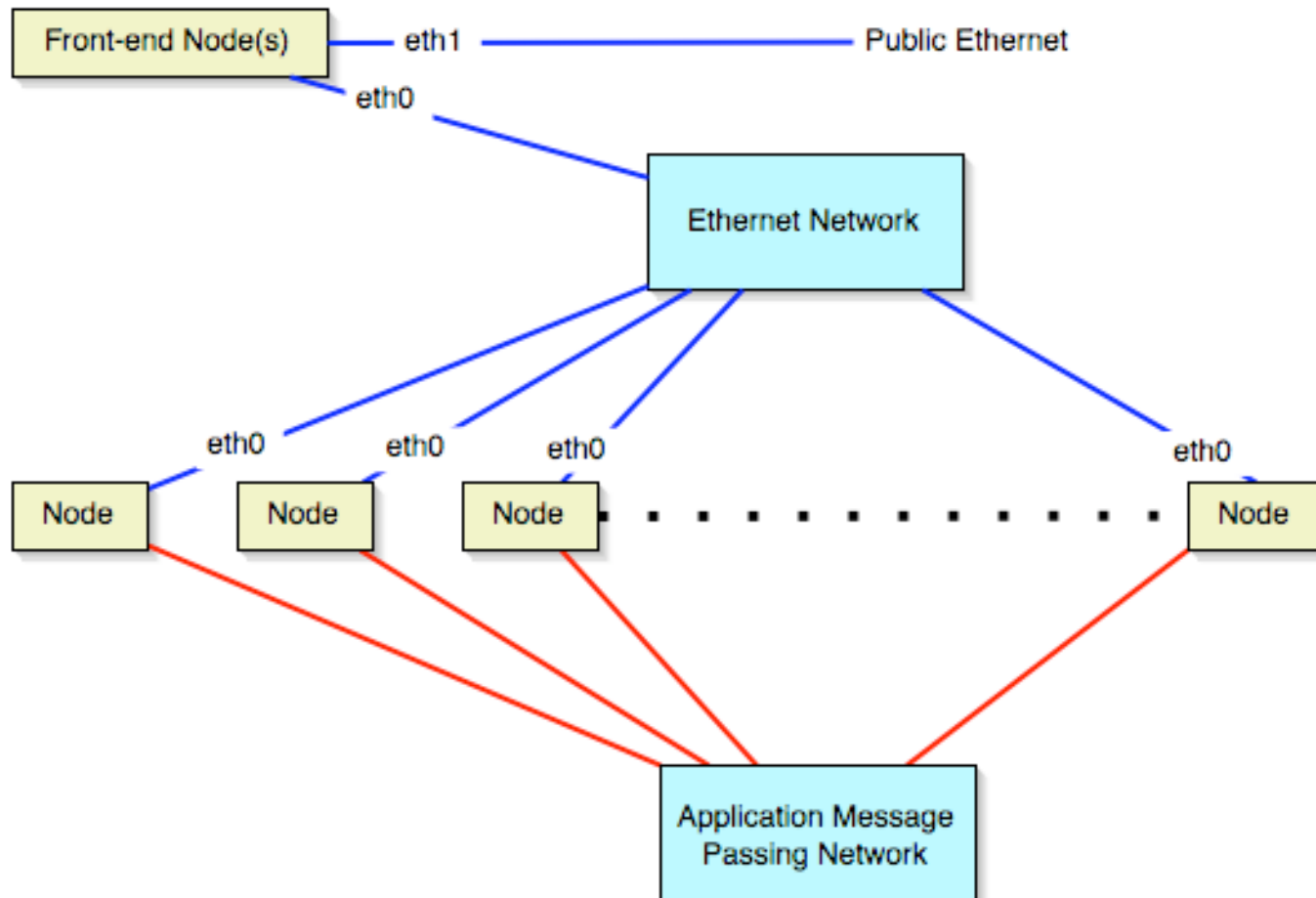
- Make clusters easy. Scientists can do it.
- A cluster on a CD
 - **Red Hat Linux, Clustering software (PBS, SGE, Ganglia, NMI)**
 - **Highly programmatic software configuration management**
 - **x86, x86_64 (Opteron, Nacona), Itanium**
- Korea localized version: KROCKS (KISTI)
<http://krocks.cluster.or.kr/Rocks/>
- Optional/integrated software rolls
 - **Scalable Computing Environment (SCE) Roll (Kasetsart University, Thailand)**
 - **Ninf-G (AIST, Japan)**
 - **Gfarm (AIST, Japan)**
 - **BIRN, CTBP, EOL, GEON, NBCR, OptIPuter**
- Production Quality
 - **First release in 2000, current 3.3.0**
 - **1000+ Worldwide installations**
 - **4 installations in PRAGMA testbed**
 - **Used is several large UCSD projects**
- HPCWire Awards (2004)
 - **Most Important Software Innovation - Editors Choice**
 - **Most Important Software Innovation - Readers Choice**
 - **Most Innovative Software - Readers Choice**



Some of our large users



Rocks Hardware Architecture



Goal of Rocks

- “Make Clusters Easy”
 - **Easy to Build**
 - **Easy to Manage**
 - **Easy to Use**
 - **Easy to Program?**
- Target Users
 - **Application Scientists**
 - **Busy Administrators**
 - **Anyone who wants their own supercomputer**



How Easy?

- Install Frontend Machine
 1. **Insert Rocks Base CD**
 2. **Insert Optional Roll CDs**
 3. **Answer 7 screens of configuration data**
 4. **Drink coffee for 30 minutes**
- Install Compute Nodes
 1. **Login to Frontend**
 2. **Execute insert-ethers to find compute nodes**
 3. **Boot Compute nodes (PXE)**
 4. **Goto step 3**
- Add user accounts
- Start computing



Rolls are Optional CDROMs

- Ordering a New Car
 - Sports package**
 - Heated seats**
 - Automatic Transmission**
 - Cool Color**
- Choose Cluster Component
 - Schedulers**
 - PBS
 - SGE
 - Monitoring System**
 - Ganglia
 - SCMSWeb (SCE)
 - Interconnect**
 - Myrinet
 - Infiniband
- Choose from well known combinations
 - Compute**
 - Grid**
 - Visualization**
- Create your own combination
 - Expert mode**



PICK PACKAGES

- > COMBO #1: PREMIUM
- > COMBO #2: SPORT
- > COMBO #3: COLD WEATHER
- > NEXT STEP

MINI COOPER S

CLICK IMAGE TO ADD THE SPORT PACKAGE TO YOUR LIST.

THE SPORT PACKAGE WILL ADD:
Dynamic stability control (DSC), bonnet stripes, xenon headlamps with powerwashers, front fog lamps, 17-inch alloy S-lite wheels with 205/45 R17 performance or all-season run-flat tires.

Sport Package (\$1350)

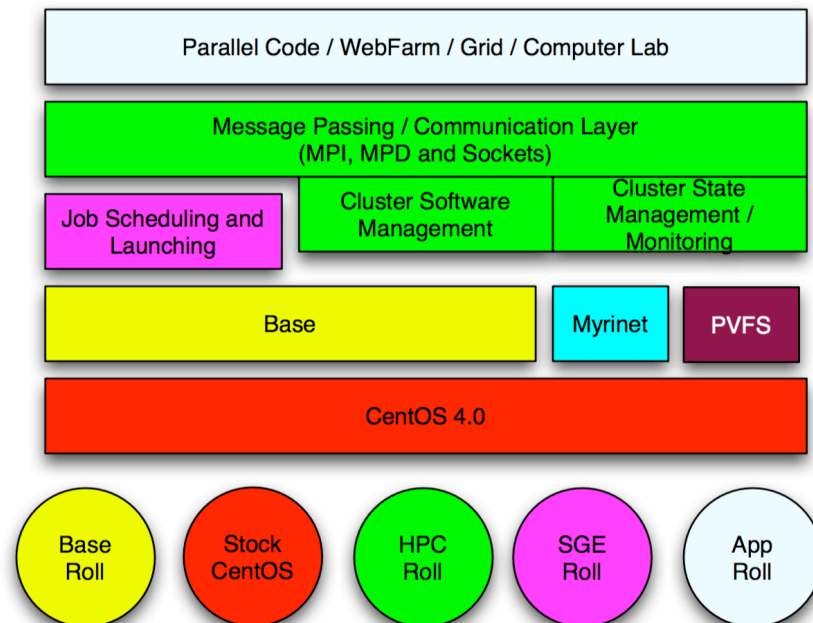


Sample of Available Rolls

- Area51
 - **Tripwire related tools**
- Condor
 - **High-throughput computing grids**
- Java
 - **JRE from Sun**
- Infiniband
 - **IB drivers & MPI (Infinicon)**
 - **IB drivers & MPI (Voltaire)**
- Intel
 - **Compilers and libraries (Scalable Systems)**
- Lava
 - **Scheduler (Platform Computing)**
- Gfarm
 - **Grid Filesystem (AIST)**
- Grid
 - **Globus, etc (NMI R5)**
- PBS
 - **Portable Batch System (University of Tromso)**
- RxC
 - **Cluster Management GUI (Scalable Systems)**
- SGE
 - **Sun Gridengine**
- SCE
 - **Scalable Clusters Environment (Kasetsart University)**
- Viz
 - **Create Visualization Clusters**

Current Status

- Rocks 4.0.0
 - **In Beta**
 - **Release in May**
- Supports
 - **x86, x86_64, and Itanium**
 - **Myrinet, IB, GigE**
- Base OS
 - **Ships with CentOS**
 - **Or user provided RHEL (commercial)**
- Sample Roll Stack on Right





Thank you

www.sdsc.edu

www.pragma-grid.net

www.rocksclusters.org