The Grid Engine Open Source Project

What is **Grid Engine** Anyway??

- Job scheduler for HPC clusters
  - *Good for compute farms as well*
- Distributed Resource Manager
- Free and open source replacement for:
  - LSF (Platform Computing)
  - PBSPro (Commercial version of OpenPBS/Torque)
  - LoadLeveler (batch system for IBM SP2, SP3)
- SGE 5.3, released in 2001, was the first open source version. SGE 6.2 is the current stable version

Why Grid Engine??

- Open source – sounds like a must these days!
- Community support
  - Over 300 mailing list messages per month
  - Actively fixing bugs, adding features
- Commercial support
  - *Sun Microsystems, Univa UD, BioTeam*
  - *Rocks cluster distribution*

- It works, it scales, it rules !!!

Supported Platforms

- AIX
- BSD (most architectures)
- HP-UX (PA-RISC & IA64)
- IRIX
- Linux (most architectures)
- Mac OS X (PPC & x64)
- Solaris (SPARC & x64)
- Tru64 (DEC Alpha)
- Windows

- Less common platforms:
  - UNICOS (Cray)
  - SUPER-UX (NEC SX)
- Includes different Linux variants
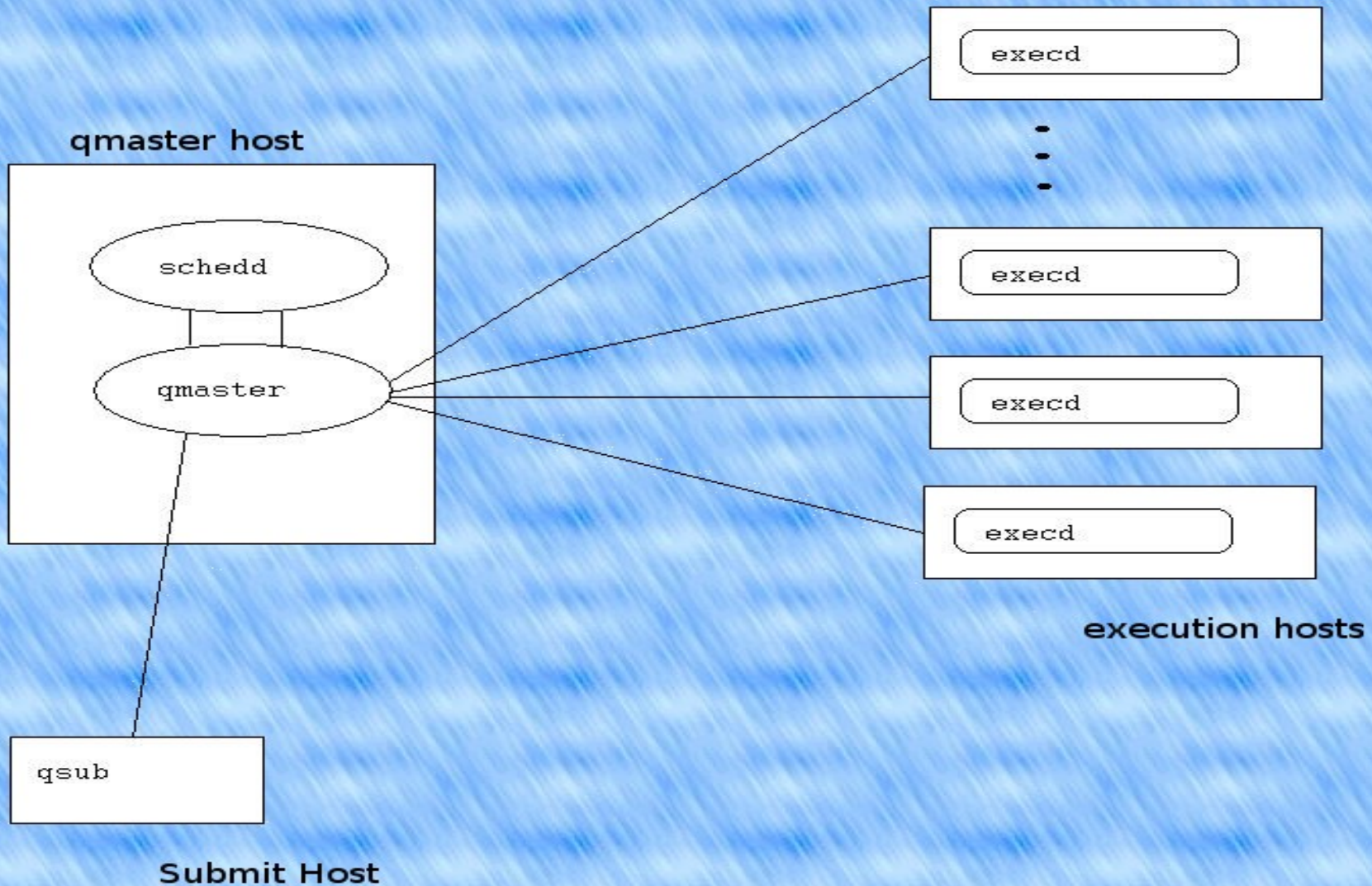 - PS/2 & PS/3
  - zLinux

Grid Engine Architecture

Main Features

- Multiplatform Support

- Scheduler Fault Tolerance

- Advance Reservation

- Resource Reservation

- Resource Quotas

- Fairshare Scheduling

- Multi-Clustering

- Interactive Job Support

- Calendar Aware

- Job Arrays
  - Array interdependency

- Job API: **DRMAA**
  - C/C++
  - Java/Javascript
  - Perl, Python, Ruby
  - New language? SWIG

Basic Grid Engine Commands

User Commands

- **qsub** : submit a job
- **qstat** : show status of all the jobs
- **qdel** : delete a job
- **qhold/qrls** : stop and release a job
- **qhost** : show status of all the hosts in the cluster
- q**alter** : modify an existing job

Admin Commands

- **qconf** : configure parameters of a cluster
- **qmod** : modify a queue

Creating a Job

<u>Write a job script</u>

```
% cat first_job.sh
#!/bin/sh
echo "My First Job.
echo "Time now is : `date`
echo "Executed on: `hostname`"
```
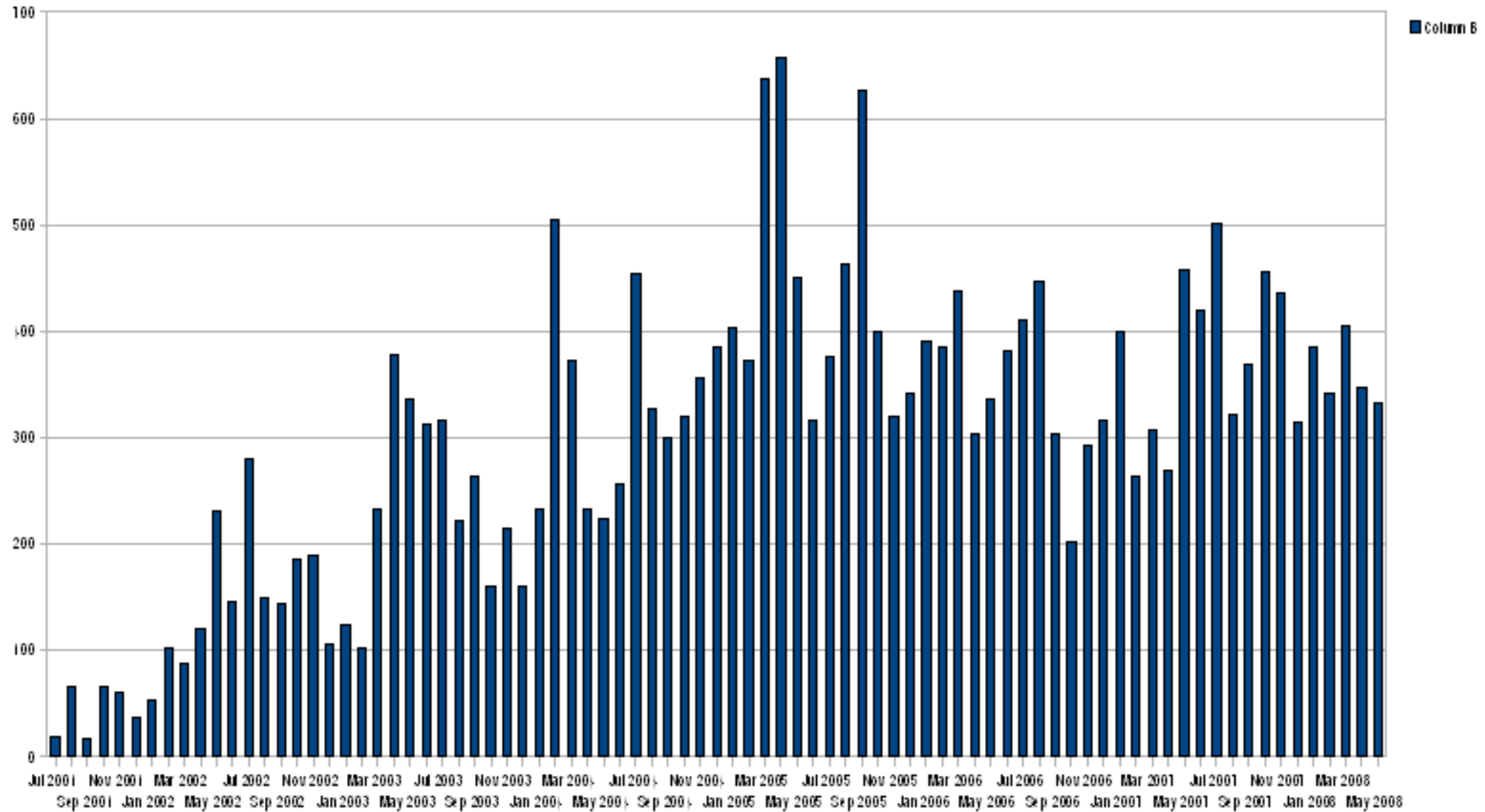
<u>Submit to Grid Engine</u>

```
% qsub first_job.sh
Your job 1 ("first_job.sh") has been submitted
%
```

Mailing List Traffic

Large Scale Users

Supercomputers

- Ranger at TACC
- TSUBAME (fastest Top500 computer in Asia)

Grid Computing

- Sun Grid

Compute Farms

- EDA
  - AMD, Mentor Graphics
- Movie Rendering
  - The Ant Bully
- Drug Discovery
- Manufacturing
- Weather Simulation
  - NOAA, Environment Canada

Ranger at TACC

Peak Performace = 504 TFLOPS

• Largest at the Texas Advanced Computing Center
• Largest among the TeraGrid systems, and currently the largest Grid Engine cluster!!

Hardware

➔ ~ 4000 SunBlade nodes (62,976 cores), InfiniBand
➔ Sun Fire x4500 Data Servers **~1.7PB of storage**

Software

• Rocks cluster distribution, Lustre CFS, Open MPI, Sun/Intel/PGI Compilers, AMD ACML & GotoBLAS, Grid Engine

qstat XML

File   Edit   View   History   Bookmarks   Tools   Help

http://xml-qstat.org/xmlqstat-demo/qstat.html   Google

xml qstat

Cluster Queue Status 0%

	Type	Slot Usage	Load Avg.	Load Ratio	System Type	State
alarm.q@test.gridengine.info	BIP	0%	0.11000	11000%	lx24-amd64	⚠ a
all.q@test.gridengine.info	BIP	0%	0.11000	6.3%	lx24-amd64	✓
disabled.q@test.gridengine.info	BIP	0%	0.11000	6.3%	lx24-amd64	✕ d

- There are no active jobs

Pending Jobs: 2

Priority	Job ID	Job Owner	Job Name	Slots Requested	Array Tasks	Submission Time	State
0.56000	1	dag	impossibleJob.sh	1		05:20:45 PM, May 04	qw
			Job 1 Hard Request: arch=solaris64				
0.55500	2	dag	hostname	1		08:15:15 PM, Jun 29	Eqw

Rendered: Thu, 27 Dec 2007 01:31:59

XHTML Sony PSP   RSS Available   XML VALIDATE

Done

Load Sensors & License Integration

- Monitors the amount of available resources
  - Example: software license

- Starts job only when the required license is available

- Very useful in EDA environments, where licenses are usually expensive!!

- Grid Engine 6 FlexLM License Integration
  - *Also know as the Olesen-FLEXlm-Integration*

Power Saving & Green DataCentre

- Power and cooling cost

  $$$

- Power aware scheduling
  - Processor Temperature
  - IPMI
- Greenhouse heating
  - University of Notre Dame
    - ➔ Heating a greenhouse using computers
    - ➔ Ref: *Grid Heating: Managing Thermal Loads with Grid Engine*

Documentation &
Further Information

- http://gridengine.sunsource.net
- http://gridengine.info
- http://en.wikipedia.org/wiki/Sun_Grid_Engine

- Documentation
  - Available in hardcopy, PDF, or wiki
  - HOWTOs on project website
  - Sun BluePrints