

e=MSC<sup>x</sup>

ENGINEERING. EDUCATION. ENTERPRISE.  
**2009 VPD**  
VIRTUAL  
PRODUCT  
DEVELOPMENT  
CONFERENCE

# Sun Grid Engine and MD Nastran

Dale Layfield,  
Staff Engineer, ISV-Engineering  
**Sun Microsystems, Inc.**





# Topics

- Recent series of “How-To” blogs (written in a “Step-by-Step” format) on configuring Sun Grid Engine for MD Nastran
- Background on Sun Grid Engine (SGE)
- More details on the blogs
- Some recent performance data

# <http://blogs.sun.com/4hpcisvs/>



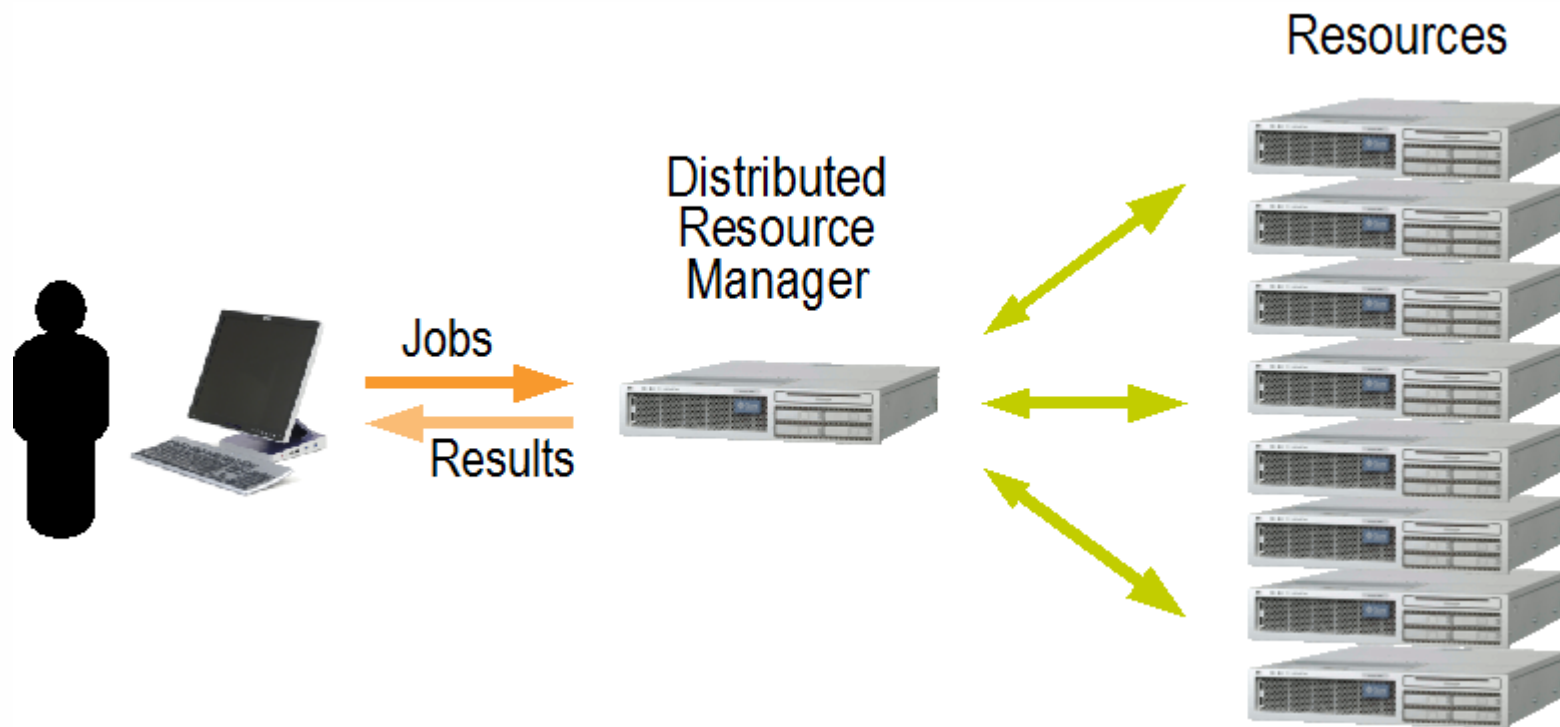
- Recent series of “How-To” blogs (written in a “Step-by-Step” format) on configuring Sun Grid Engine for MD Nastran
  - “Sun Grid Engine and MD Nastran”  
[recommendations: SGE configurations and queues for MD Nastran users...]
  - “Part 1—How to submit “MD Nastran” (serial) jobs with Sun Grid Engine”
  - “Part 2—How to submit “MD Nastran” DMP (Distributed Memory Parallel) jobs with Sun Grid Engine”
  - “Part 3—How to configure consumable resources (Disk, Memory, and License Tokens)”
  - “Part 4—How to Create Queues and Parallel Environments”



# Some Background

# Sun Grid Engine

## A “Distributed Resource Manager” (DRM)



# Sun Grid Engine

## A “Distributed Resource Manager” (DRM)



### Key Feature:

- SGE maximizes resource utilization by matching workload to resources

For example: A job can define Disk, Memory, License, compute time requirements and SGE will find machine(s) in the compute cluster that matches those requirements.

### Terminology:

Queues:	Defines where and how jobs are executed
Master host:	Contains the distributed resource manager
Execution host:	Machine (associated with a Queue) where jobs will run
SGE's CLI:	The command line interface for administering SGE
QMON:	The gui interface for administering SGE



# Terminology, continued:

- **SGE resource (complex):**

Defines a property of an execution host (disk, memory, licenses, etc.)

Types:

- **Static** (example: OS level on machine)
- **Dynamic** (example: load sensor script that monitors disk space on /export),
- **Consumable** (example: license tokens that are “consumed” by jobs)

# Customers



<b>Health Sciences</b>	<b>One of the largest pharmaceutical companies in USA</b>
<b>Financial Services</b>	<b>Largest Wall Street financial institution, Deutsche Bank</b>
<b>Manufacturing</b>	<b>Ford, Transmeta, Mentor Graphics, Monsanto</b>
<b>Entertainment/Media</b>	<b>Rising Sun Pictures, GlobeXplorer, Inc.</b>
<b>Energy</b>	<b>Landmark Graphics, Amerada Hess</b>
<b>Government/Edu</b>	<b>DoE INEEL, Tokyo Institute of Technology, TACC</b>

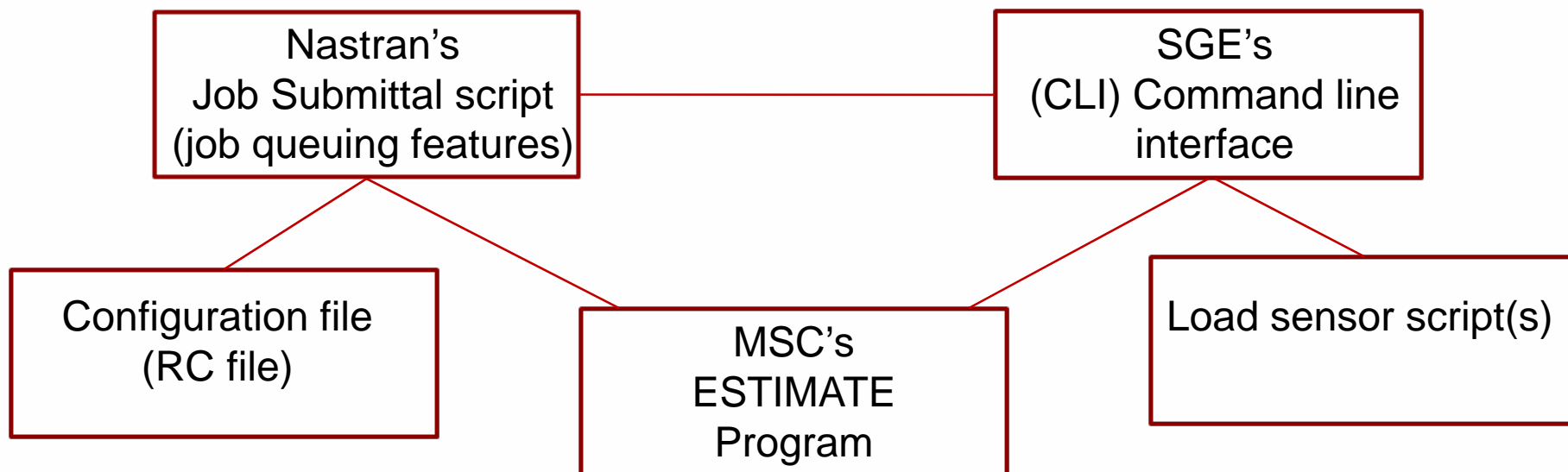


More on the blogs...

<http://blogs.sun.com/4hpcisvs/>



# The moving parts...



# Questions being answered by the blogs...



- **Time limits:** How can you configure SGE to limit the time jobs are allowed to run on certain machines?
- **Disk:** How can you configure SGE to know which machines have enough disk space for the database files (scratch and scr300)?
- **Memory:** What machines have enough physical memory to “avoid spill”?
- **License Tokens:** How can you avoid flex license “time-out” for a job due to insufficient tokens?

# MD Nastran's Job Submittal Script



[http://www.mscsoftware.com/support/prod\\_support/mdnastran/cog.pdf](http://www.mscsoftware.com/support/prod_support/mdnastran/cog.pdf)

- Existing MD Nastran job queuing interfaces allow for efficient and convenient integration of MD Nastran and SGE
- You can use the MD Nastran configuration file (RC file) to define the SGE configuration:

```
#cat sgc_rc
```

```
scr=yes
```

```
bat=no
```

```
app=no
```

```
submit=small.q,medium.q,large.q=qsub -q %queue% -l nastran_tokens=%qoption%  
-l export_size=10G -S /bin/ksh %job%
```

- To submit a (serial) MD Nastran job:

```
$MDNAST2008 v10101.dat queue=small.q rcf=sgc_rc qoption=`token_estimate.sh`
```

# MD Nastran's ESTIMATE Program

[http://www.mscsoftware.com/support/prod\\_support/mdnastran/cog.pdf](http://www.mscsoftware.com/support/prod_support/mdnastran/cog.pdf)



- The ESTIMATE program does a quick scan and analysis of the Nastran input file to determine its resource requirements (disk, memory, and license tokens).
- This information can then be used along with a SGE "consumable resource" to ensure that sufficient license tokens are available "before" starting the job-- so that jobs don't start running, later request more licenses, and then "time-out"/stop due to insufficient licenses:

This is useful for the following types of runs:

1. **Custom DMAP**
2. **Superelements**
3. **...other Solution Sequences**

- To submit a (serial) MD Nastran job that uses "ESTIMATE" to ensure enough license tokens are available:

```
$MDNAST2008 v10101.dat queue=small.q rcf=sgc_rc qoption=token_estimate.sh
```



A few SGE configuration examples described in the blogs...



## Time limits:

# How to configure SGE to limit the time jobs are allowed to run on certain machines.

- Only 2 SGE commands are required to create a time-limiting queue
- For example:
  - To create a SGE job queue (called “small.q”) that will limit jobs to **5 minutes** (elapsed time)
  - #qconf -aq small.q**
  - # qconf -mattr queue complex\_values h\_rt=00:05:00 small.q**



# Disk:

## How to configure SGE to know which machines have enough disk space for the database files (scratch and scr300).

- Create a “load sensor” script to monitor the available free disk space on the filesystem of interest

**disk\_space.sh**

- Create a “consumable” resource for this filesystem you want to monitor

**#qconf -mc**

- Tell SGE which host machine to monitor (and point to your “load sensor” script)

**#qconf -mconf global**



# Some Recent Performance Benchmarks on Sun Hardware

## MD Nastran performance (Solaris 10 x86) on new Sun Fire X4270 (using the Intel 5500 Series processor (aka Nehalem))



[more details on <http://blogs.sun.com>]

### ...highlights from the above blog:

- The table below shows the "% reduction in elapsed time" for the MD Nastran MDR3 benchmarks on the Sun Fire X4270 compared to the Sun Fire X4150:

getrag	gm20a_1_1	md0mdf1_1	xl1fn40	vl0sst1	xl0imf1	xl0tdf1_1	xx0cmd2_1
34%	27%	32%	32%	16%	55%	17%	35%

A description of these MD Nastran MDR3 benchmarks can be found on MSC.Software's [performance web site](#).

- Solid State Drive (SSD) performance vs SAS disk:**  
Up to **56% reduction** in elapsed times with SSD's compared to internal SAS
- The ZFS filesystem** in Solaris 10 works well with MD Nastran---"very" easy to configure filesystems and has excellent I/O performance: For example: One simple command creates/stripes/mounts a Raid0 filesystem of 7 disks:

```
zpool create ZFS1 c0t1d0 c0t2d0 c0t3d0 c0t4d0 c0t5d0 c0t6d0 c0t7d0
```



# Summary

- **Sun Grid Engine is being used *today in the world's largest compute clusters***
  - And the smallest...
  - And in the mid-sized...
- **The Grid Engine community.**
  - Grid Engine has a very strong community
  - Join the mailing lists! <http://gridengine.sunsource.net/>
  - <http://gridengine.info>
- **Blogs on “Sun Grid Engine and MD Nastran”**
  - <http://blogs.sun.com/4hpcisvs>



# Contact Details :

- For further information please contact

Dale Layfield  
Sun Microsystems, Inc.

10 Universal City Plaza  
Universal City, CA 91608  
USA

(818) 259-7694  
dale.layfield@sun.com