PLATFORM USERS CONFERENCE 2000

Building the Best Computer Clusters for Mechanical Design Environments

Daniel Chan and Ed Huott

GE Corporate R&D Center

Key Challenges

- A Highly Diverse Environment
  - Many Different Businesses
  - Engineering, Chemistry, Financial, Services
- Short-Term Technical Support
  - Quick Turnaround
- Costs
  - Downtime, System Admin., Software Licenses, Productivity

Design Requirements

- Web-Centric
  - Simplify Job Submission in a Heterogeneous Environment
  - Access from Anywhere and Anytime
- Minimize Complexity and Cost
- Stable and Highly-Available
  - Isolate Network, NFS and Insufficient Disk Space Problems, Software License Management

Reasons for Having a Heterogeneous Environment

- Performance
- Costs
- Legacy Applications

Let's Do An Experiment

NT 4.0

GE Corporate Standard
PIII 450 MHz
256 MB RAM
$1400

Unix/Solaris 2.6

Sun Ultra 10
UltraSPARC 440 MHz
256 MB RAM
$5000

CFX 5
23,862 Nodes
102 MB Required

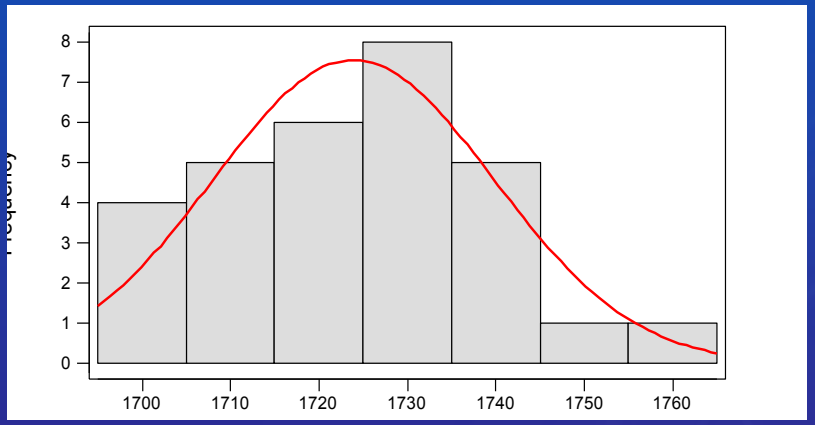
Use Perl script to launch 1, 2 and 3 concurrent jobs, respectively, for 30 times

Extract wall clock times from output files then use Minitab to analyze data

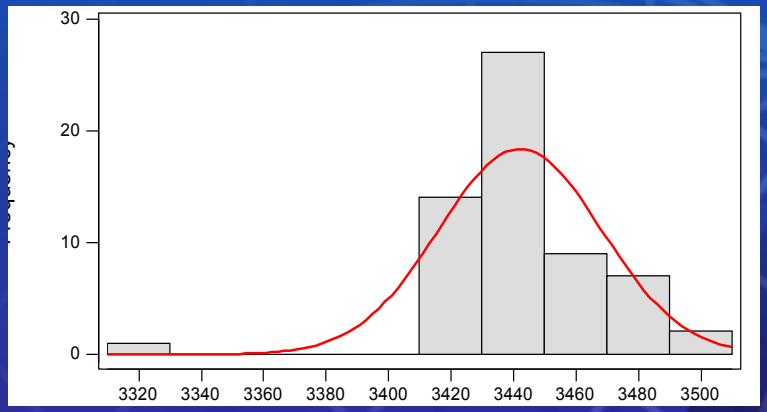
Test stability, multitasking and paging capabilities

Results For Sun Ultra 10

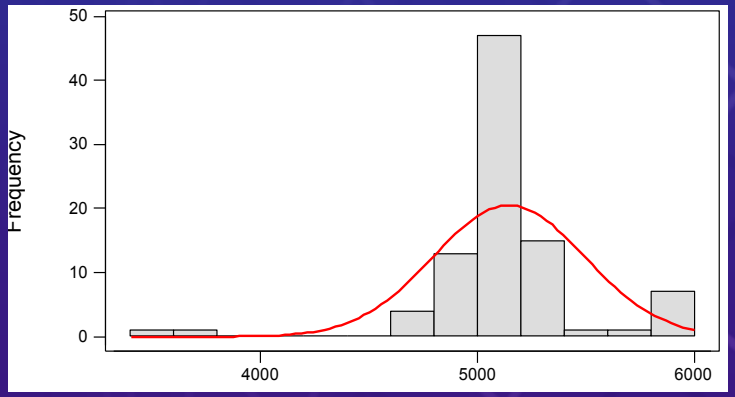
1 job



2 jobs

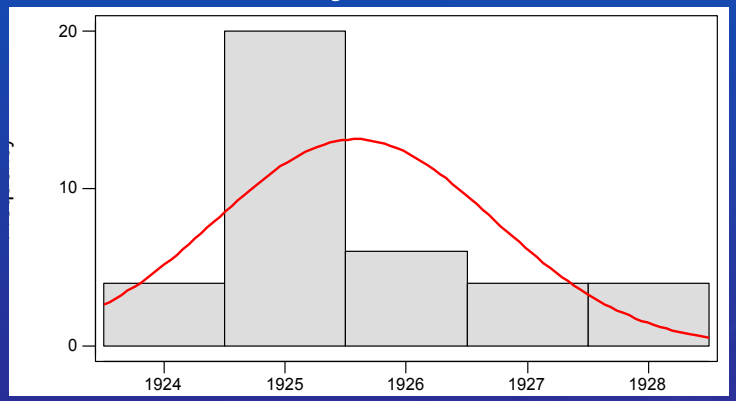


3 jobs

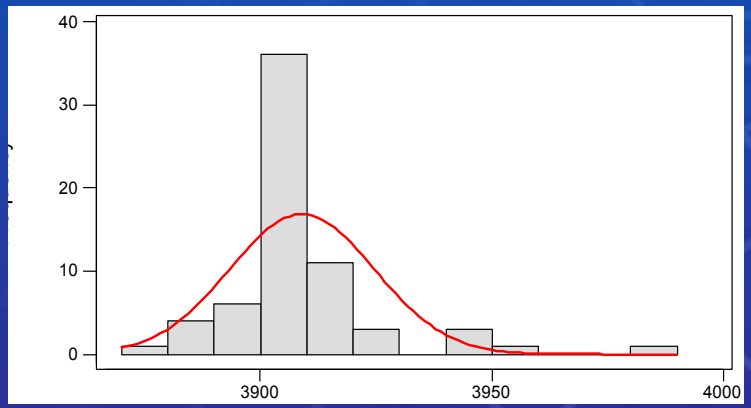


Results for Dell OptiPlex G1x

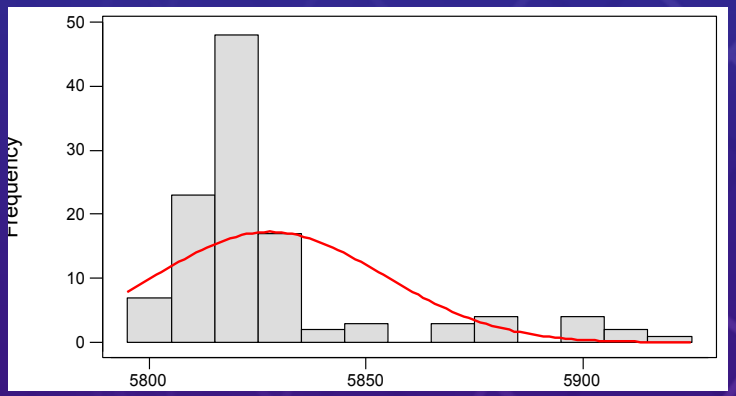
1 job



2 jobs



3 jobs



Scorecard for Sun Ultra 10

$$Z = \left| \frac{USL - \mu}{\sigma} \right| = \frac{0.1\mu}{\sigma}$$

	Mean	Standard Deviation	Z
1 job	**1724**	**15.9**	**10.8**
2 jobs	**3442 (2X)**	**26.1**	**13.2**
3 jobs	**5144 (3X)**	**350.2**	**1.5**

Paging

Scorecard For Dell Optiplex GX1

$$Z = \left| \frac{USL - \mu}{\sigma} \right| = \frac{0.1\mu}{\sigma}$$

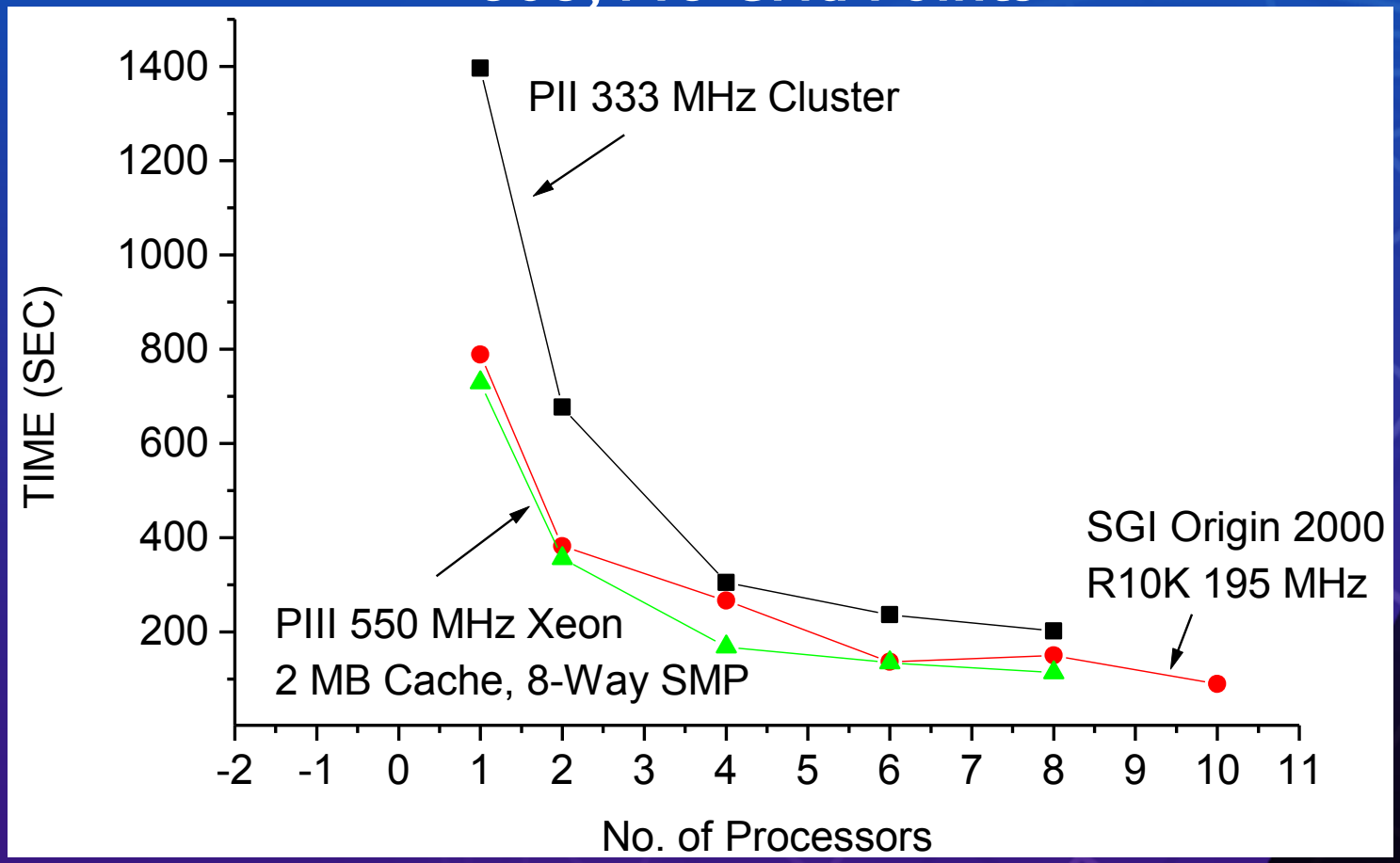
	Mean	Standard Deviation	Z
1 job	**1926**	**1.2**	**161**
2 jobs	**3909 (2.1X)**	**15.6**	**25**
3 jobs	**5828 (3X)**	**26.4**	**22**

Summary of Results

- The SUN workstation is about 10% faster, but nearly 5 times as expensive
- Both systems are just as stable for a period of one week
- NT can multitask
- NT can page and appears to have “better” memory management scheme

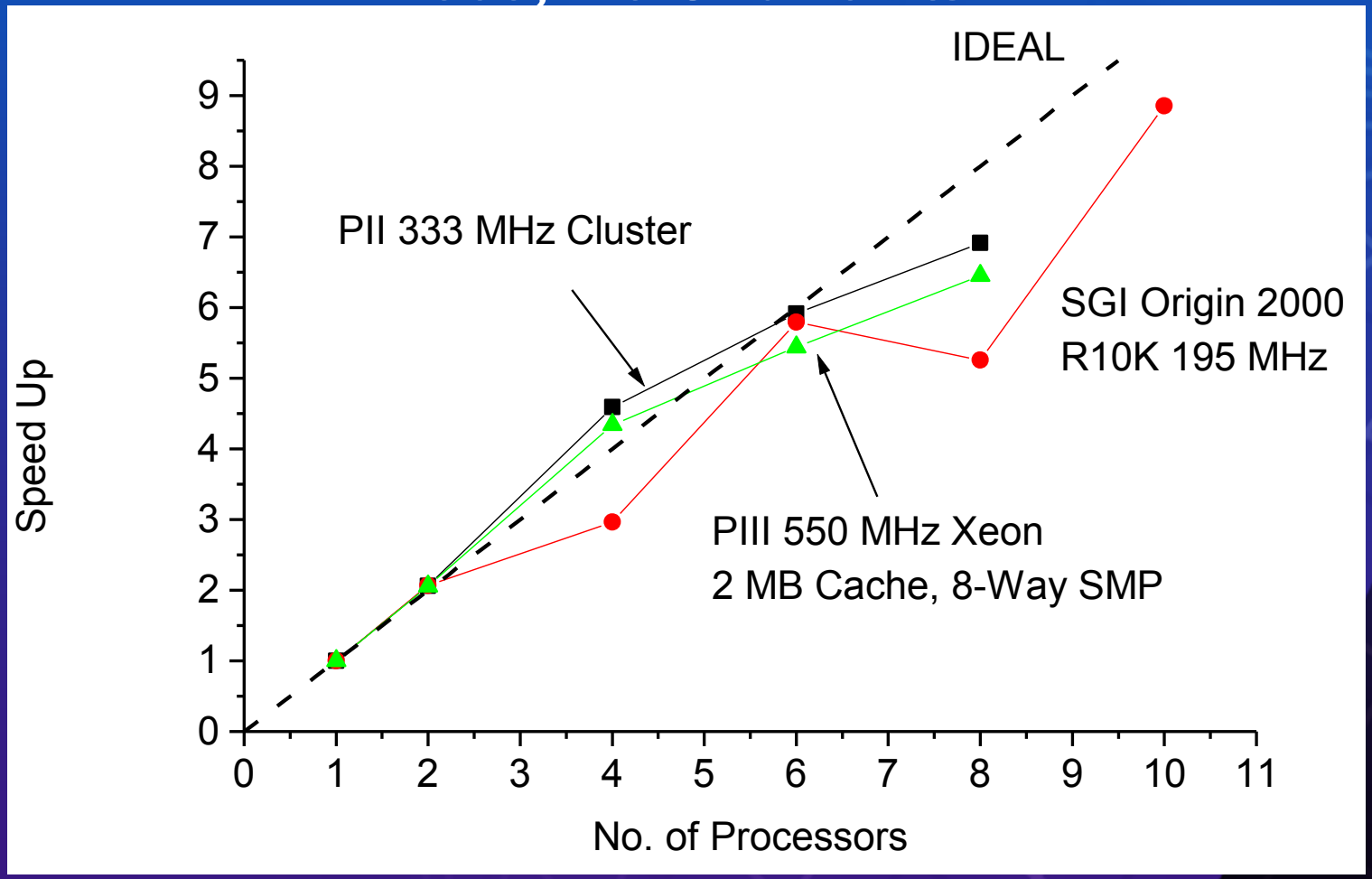
APNASA Parallel Performance

308,115 Grid Points

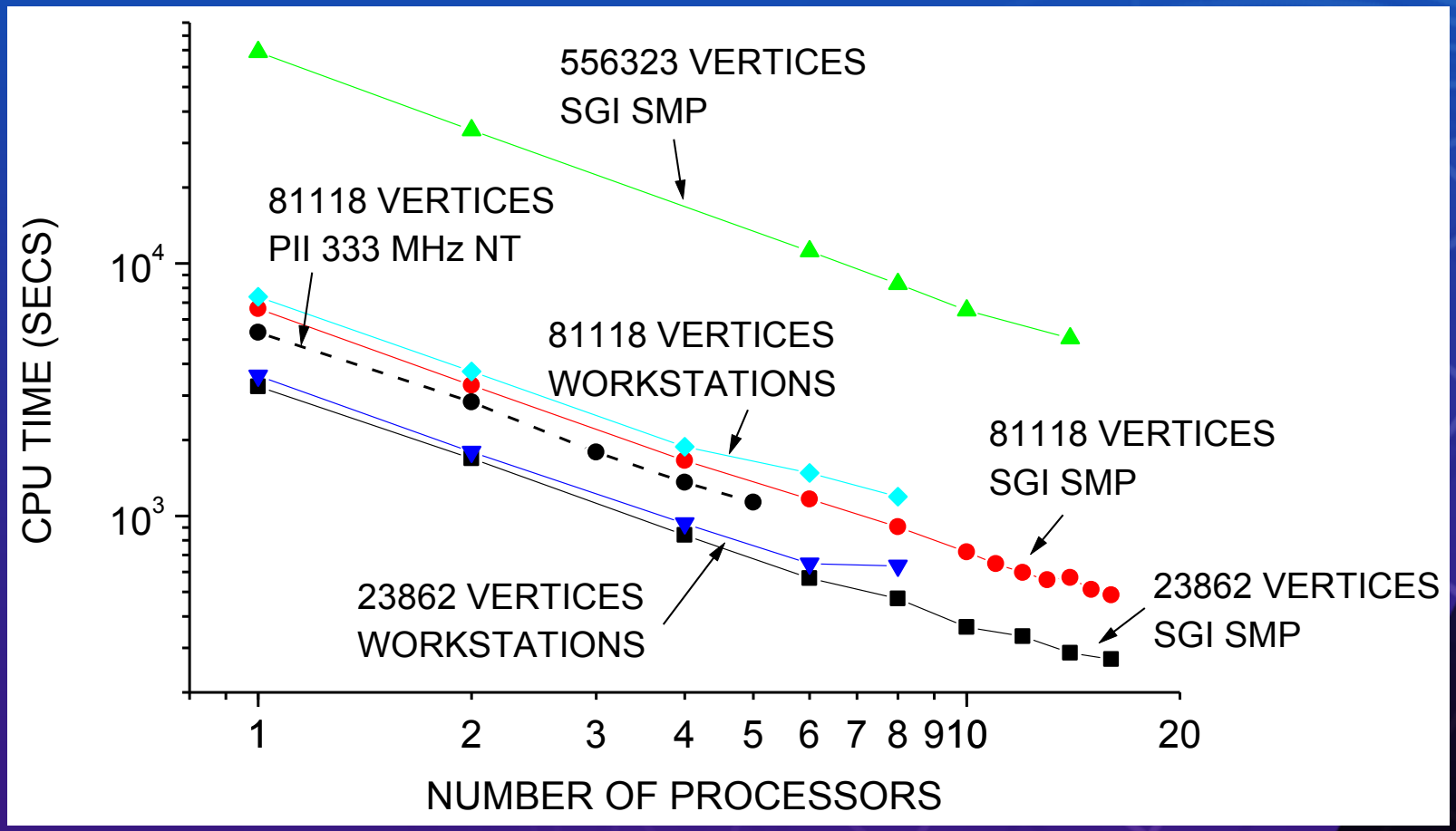


APNASA Parallel Performance

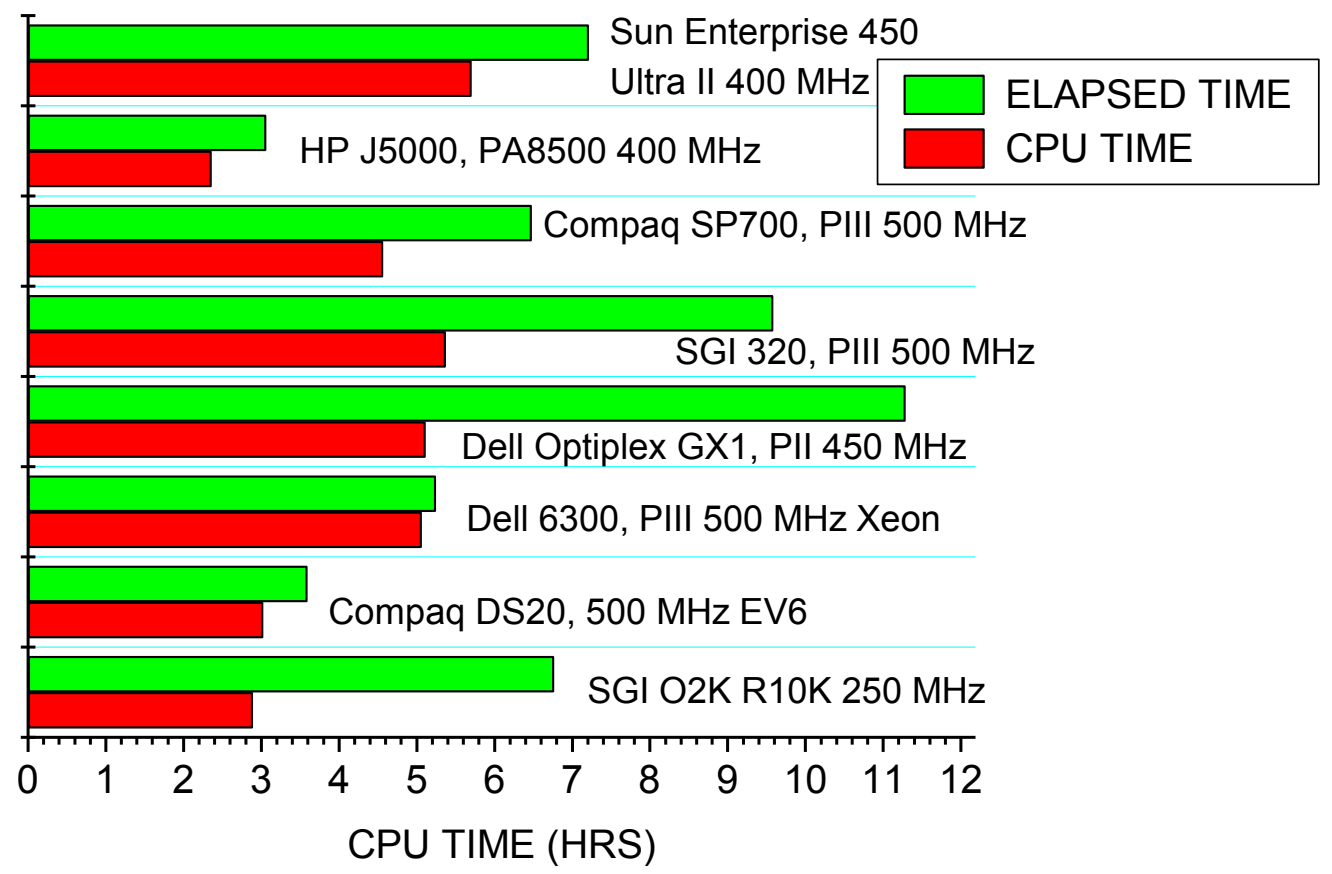
308,115 Grid Points



CFX PERFORMANCE ON A VARIETY OF COMPUTER PLATFORMS

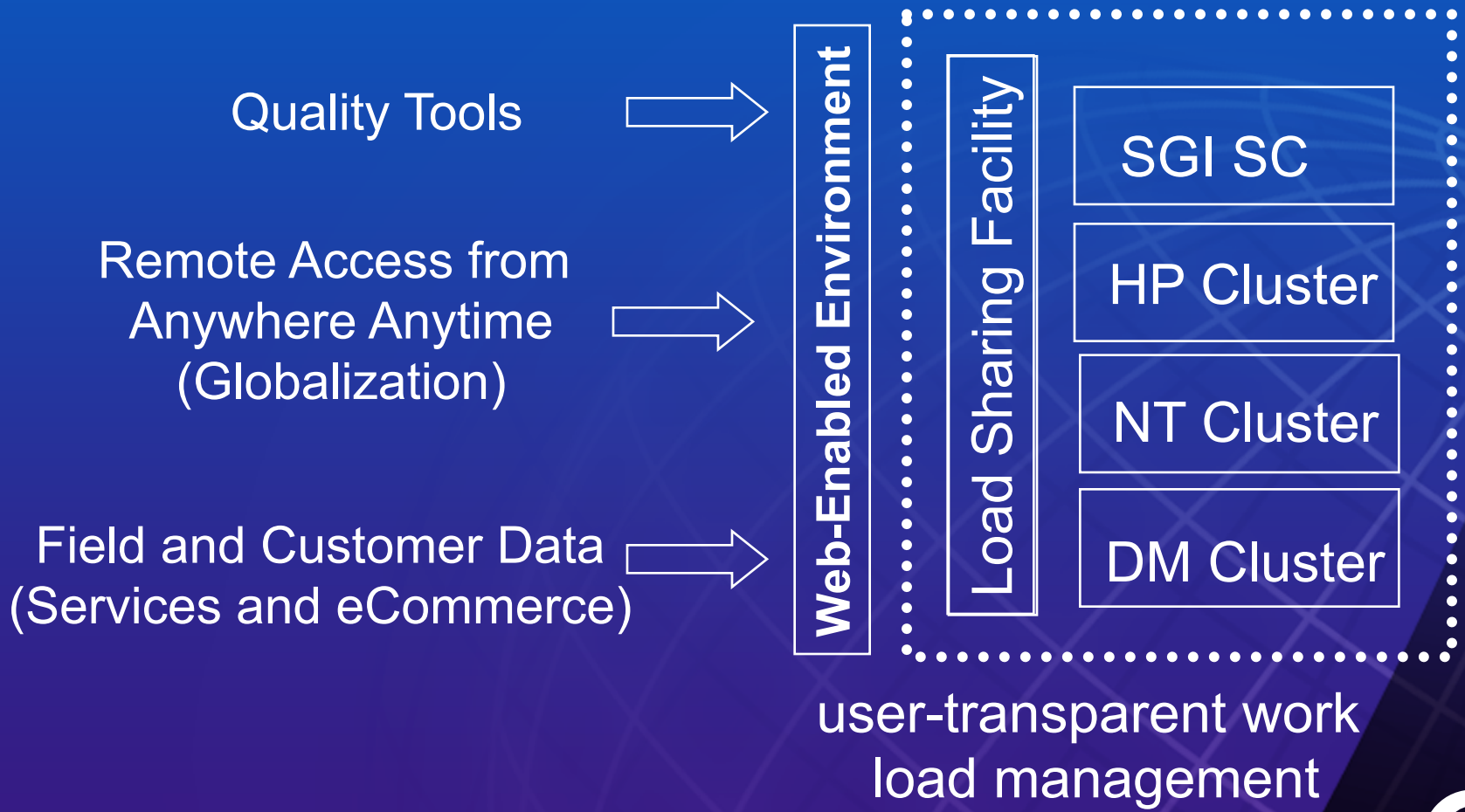


ANSYS Performance on a Variety of Computers



160,000 ELEMENTS WITH 2D CONTACT
2.1 GB OF OUTPUT AND 615 MB PAGE FILE

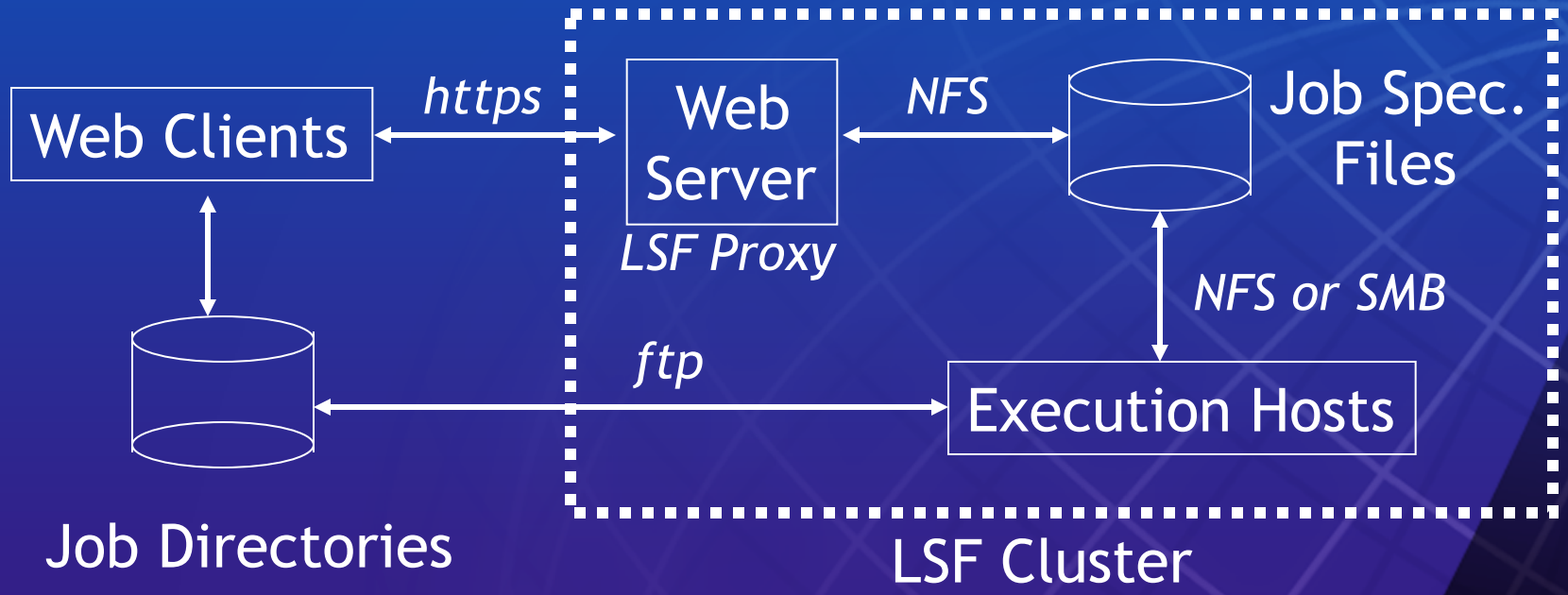
eComputing Architecture



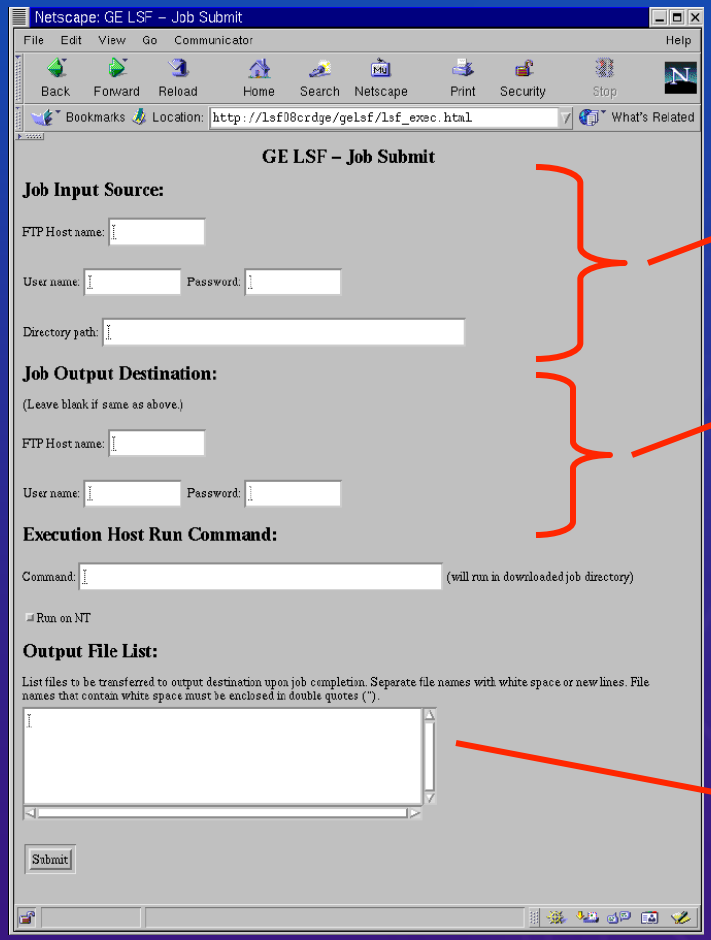
Key LSF Challenges

- Implicit Assumptions of A Homogeneous Environment
  - Uniform View of NFS
  - Uniform Mounting of User Home Directory
  - Unix Bias
- Wholesale Migration of User Environment Variables from Submit to Execution Hosts

A Web-Centric Job Management System



Screen Shot of Web Interface



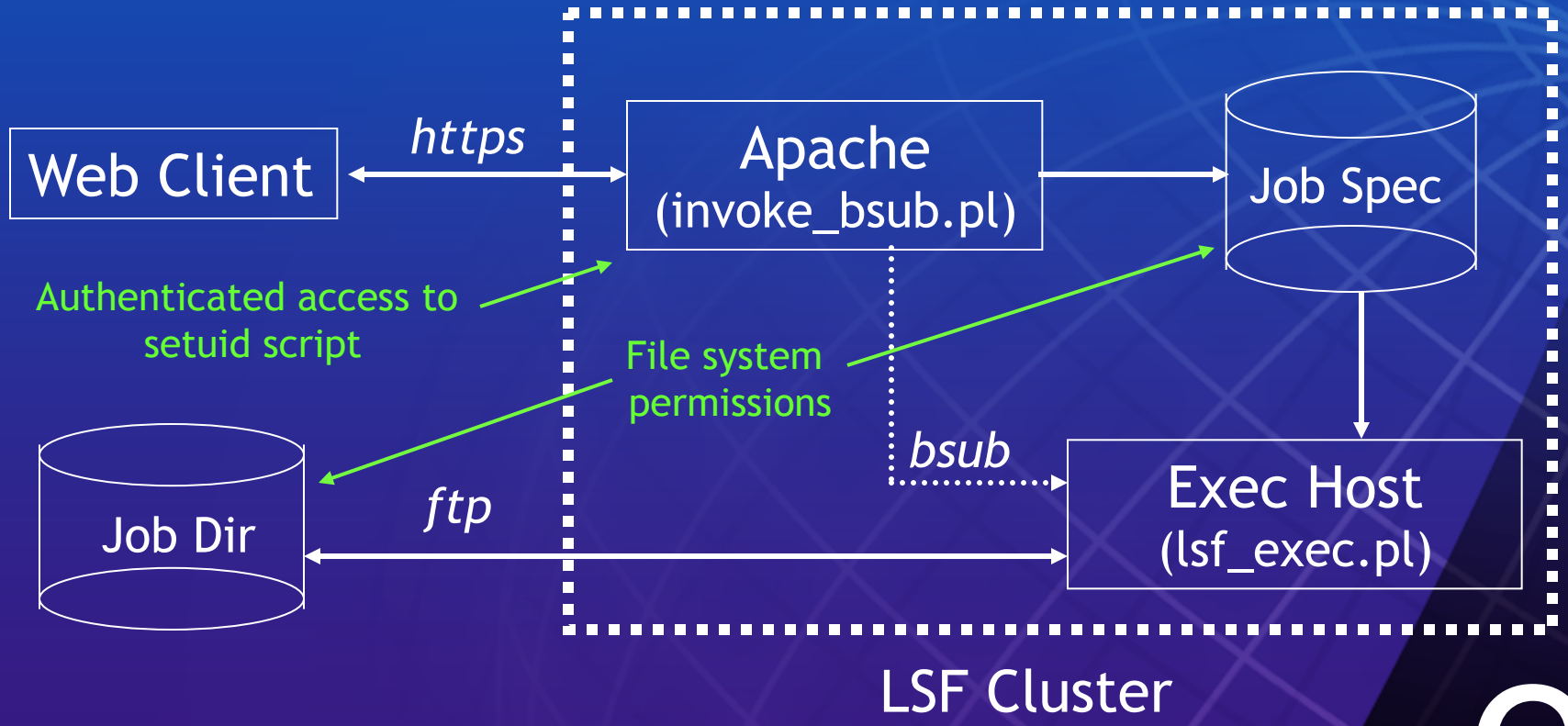
Job Source Files

Job Output Destination

Command to Run on Execution Host

Files Sent to Output Destination

Process, Data Flow and Security for Web-Based Jobs



Normalized Job Execution Wrapper/ Environment

- Standard job execution wrapper that runs on all execution hosts.
- Key "ingredients": Perl (with Net::FTP package), wget, subset of "standard" Unix utilities (e.g. /bin/sh, cp, mv, etc.)
  - (Note: Makes use of Cygwin tools on Windows NT. See http:// sources.redhat.com/cygwin/)
- Reads Job Specification File to determine:
  - Source of remote job directory
  - Command to run o Destination for job output file(s)
- Pre-determined top level (local) directory for jobs on each execution host.
- Remote jobs are migrated by FTP (wget) to unique sub-directories under top level.
- Specified command is run in migrated job directory.
- Output files are transferred by FTP to specified destination.

Summary

- Leveraging Low-Cost NT Solution
- Shield Users from Complexity
  - productivity gain
  - better centralized resource management
- Enterprise Resource Planning
  - compute cycles
  - software licenses