1 Overview

Architecture and Configuration

POWER CHALLENGE Architecture

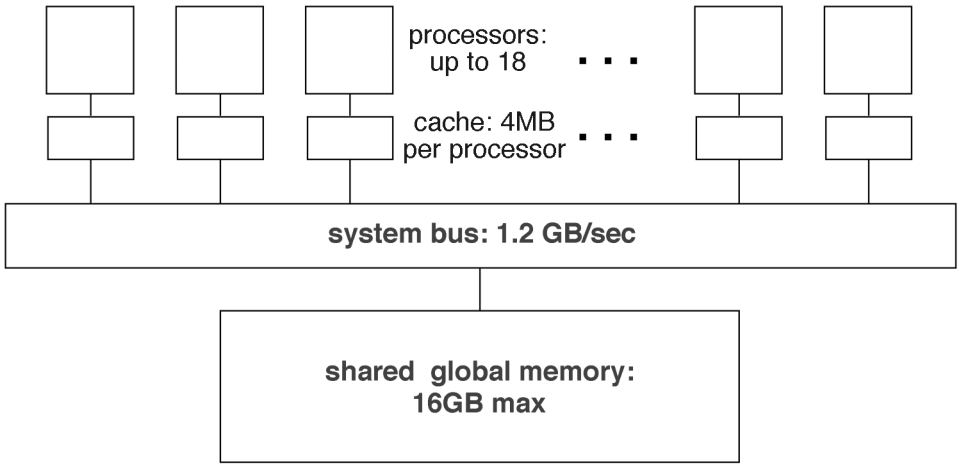
The SGI POWER CHALLENGE is a shared-memory multiprocessor architecture based on the MIPS superscalar RISC R8000 chip. This chip is a 64-bit processor with 64-bit integer and floating-point operations, registers, and virtual addresses. The R8000 uses the MIPS IV instruction set.

The cache system consists of a 16 kilobyte (KB) direct-mapped on-chip instruction cache, a 16 KB direct-mapped on-chip integer data cache, and a 4 megabyte (MB) four-way set associative external cache. The external cache serves as the primary cache for floating-point data and the secondary cache for instructions and integer data. The length of cache lines are:

- on-chip instruction cache: 32 bytes (4 double words)
- on-chip integer data cache: 32 bytes (4 double words)
- external cache: 128 bytes (16 double words)

The processors communicate via a fast shared-bus interconnect. The bus has a bandwidth of 1.2 gigabyte (GB) per second with a 256-bit wide data bus and a separate 40-bit wide address bus that can access up to one terabyte (TB) of physical memory. The bus provides high-bandwidth, low-latency, cache-coherent communication between processors, memory, and I/O.

Figure 1.1 SGI POWER CHALLENGE configuration.



CHALLENGE Architecture

The SGI CHALLENGE is based on the MIPS R4400 chip. The cache system consists of one instruction cache, one primary data cache, and one secondary unified instruction/data cache. The length of the cache lines are:

- primary data cache: 16 bytes (2 double words)
- secondary cache: 128 bytes (16 double words)

	Clock (Mhz)	Peak Performance (Mflops)
CHALLENGE (R4400)	**150**	**75**
POWER CHALLENGE (R8000)	**90**	**360**

NCSA's SGI Configuration

SGI's POWER CHALLENGE array is a distributed parallel processing system consisting of up to 16 POWER CHALLENGE or POWER Onyx systems connected by a High Performance Parallel Interface (HIPPI) interconnect. NCSA's array is comprised of six POWER CHALLENGE machines. The center also has a CHALLENGE machine that serves as a file server. The machines are named for figures from Norse mythology and have the following configurations:

Name	Type	# Processors	Memory	OS
`loki`	**CHALLENGE**	**12**	**1 Gb**	**IRIX 6.2**
`odin`	**POWER CHALLENGE**	**16**	**4 Gb**	**IRIX 6.2**
`thor`	**POWER CHALLENGE**	**8**	**2 Gb**	**IRIX 6.2**
`magna`	**POWER CHALLENGE**	**8**	**2 Gb**	**IRIX 6.2**
`freya`	**POWER CHALLENGE**	**8**	**2 Gb**	**IRIX 6.2**
`sif`	**POWER CHALLENGE**	**8**	**2 Gb**	**IRIX 6.2**
`bragi`	**POWER CHALLENGE**	**16**	**4 Gb**	**IRIX 6.2**

NOTE: Operating system version change effective April 1996.

Software tools are available to effectively use the POWER CHALLENGEarray system as a large parallel machine. The tools include the message passing libraries Parallel Virtual Machine (PVM) and the Message-Passing Interface (MPI) as well as High Performance Fortran (HPF). See "Array Message-passing Standards: PVM and MPI" on page 5–9 and "High Performance Fortran Compiler" on page 5–8 for more information.

Configuration Information Online

The most current information about the center's configuration is available online by entering **`lshosts`** at the system prompt or reading the `/usr/news/configuration` file.

Access

The CHALLENGE/POWER CHALLENGEarray machines are accessed in interactive, batch, or dedicated mode. See Chapter 6, "Running Jobs," for more information.

Interactive Mode

Only `loki` and `sif` are available for interactive use. The `/usr/news/interactive-limits` file contains information on CPU time and memory limits when running interactive jobs. If your job needs more time or memory than is available in interactive mode, use the batch system.

Batch Mode

The CHALLENGE and the POWER CHALLENGEarray run the load sharing batch utility lsbatch. lsbatch distributes (load balances) jobs between the array machines based on the job requirements and the load on the machines. Jobs can be submitted to the batch system from the interactive machines `loki` and `sif`.

Dedicated Mode

Jobs in the dedicated queues are run one at a time. Dedicated mode is used primarily for benchmarking code. The dedicated queues run on the weekends.

See Chapter 6, "Running Jobs," and the `/usr/news/dedicated-requests` file for more information.

File Systems

Directories

Home Directories

A *home directory* is the default directory you are placed in when you log on. NCSA maintains a disk quota of 25 megabytes (MB) for your home directory. Home directories are backed up daily.

All user home directories are contained in the `/u` file system. Home directories are NFS mounted from `loki`. Enter **`quota -v`** to see your disk usage and limits.

Scratch Directories

Scratch file systems are intended for short term use and should be considered volatile. Each scratch file system contains approximately 30 gigabytes (GB) and is named `/scratch-`*machinename* (e.g., `/scratch-odin`). Scratch directories are cross-mounted on all machines. Cross-mounted directories are not guaranteed to be available at all times. Additionally you cannot access files larger than 2 GB on cross-mounted directories.

To change to a scratch directory enter:

`%` **`cd /scratch-`***machinename* **`/`***loginname*

where *machinename* is the appropriate machine name (e.g., `loki`, `sif`) and *loginname* is your login name.

You can also use the command:

`% cd /scratch/`*loginname*

to change to the machine-local scratch directory.

No scratch disk quotas are currently in place. No guarantees are made as to how much scratch space will be available at any time. Files may be removed at any time should disk space become critically low.

Files in the scratch file systems are removed on the basis of size and time since the last access:

File Size	Removed after
> 1 megabyte	3 days
< 1 megabyte	14 days

Please do not attempt to circumvent this removal scheme (e.g., with `touch`). Such attempts may result in the loss of access to the scratch file systems.

Please note that backups are *not* performed on the scratch directories. In the event of a disk crash, files on the scratch directories are unlikely to be recoverable. In the event of a file purge, files are *not* recoverable. Therefore, you should make sure to back up your files to permanent storage as often as significant changes are made (at least daily). Refer to Chapter 4, "Mass Storage," for more detailed information.

Reserved Scratch Directories

The reserved scratch directories are intended for users who require larger amounts of scratch space than exist in the normal scratch file systems. NCSA has several reserved scratch directories available. These directories are intended for short term use and are *not* backed up. For information on requesting this reserved disk space, read the `/usr/news/reserved_scratch` file.

More Information

For the most recent information about the file systems, read the `/usr/news/Filesystems` file.

Permanent File Storage

Permanent storage of your files is available using the UniTree archival storage system. There are no limits on the number or size of files that may be saved to Unitree. Unitree files are read accessible for one year after your account has been deleted. Refer to Chapter 4, "Mass Storage," for more detailed information.

Application Software

See "Application Software" on page 5–8 for information on available software on the NCSA SGI systems.

Your Account

Applying for an Account

Projects requiring a large-memory, symmetric multiprocessor architecture are allocated on the POWER CHALLENGEarray. A limited number of education accounts are accommodated.

Application forms are available online on the NCSA anonymous FTP server (see "Anonymous FTP" on page 2–3) in the `/ncsapubs/ApplicationInfo` directory and at the following URL:

`http://www.ncsa.uiuc.edu/General/Allocations/ApplyTop.html`

If you are not able to view and print PostScript files, you can request copies of the forms from NCSA Allocations. Send your electronic mail request for hard-copies to allocations@ncsa.uiuc.edu.

When you are given an account, you are allocated a specific number of service units (SU) on the CHALLENGE/POWER CHALLENGEarray.

Passwords

When you receive a form with your account information in the mail, a default password is included. Save this form, and keep it in a safe place. You have the same password on all of the machines in the CHALLENGE/POWER CHALLENGEarray. If you forget your password, contact the Consulting Services (see "Consulting Services" on page 2–4). For security purposes, the consultant will reset your password to your default password.

NOTE: Your password can only be changed on `loki`.

Never share your password with another user or make your password known to anyone else.

Charging Algorithm

The charging algorithm for the POWER CHALLENGEarray is:

$$SUs = \left[1 - 0.28\left[\frac{(concurrency - 1)}{(concurrency + 1)}\right]\right] \bullet CPUnodehours$$

where *SUs* is the number of service units charged for the job, *CPUnodehours* is the sum of CPU time in hours across all the processors used, and *concurrency* is the number of processors used for the job. (At the time this guide was prepared, concurrency information was not available and was set to *3*.)

The most current information about the charging algorithm is available online in the file named `/usr/news/Charging_Algorithm`.

The charging algorithm for the CHALLENGE machine is 1/4th the rate for the POWER CHALLENGEarray.

Job Statistics

You should investigate any dramatic increase in your account charges. To investigate, try to establish the job information associated with the increased charges. Once a short window of time is established, contact Consulting Services by electronic mail (consult@ncsa.uiuc.edu) if the charge is still unexplained.

Account information is entered at midnight into the accounting database. Therefore, numbers for the current day's activity are not available until the following morning.

Verifying Your Account Balance

You can monitor usage of your allocation using the `usage` utility. Entering **`usage`** with no options displays cumulative information about the usage of all your allocations on all applicable NCSA machines. For example:

```
% usage
Proj  Mach    Login   Usage   Status  P_alloc  P_usage  P_expire         Name
 aad    cm5     lahti     0.00    Active         50       0.00    09/30/96    Lahti, Tracy
 non   loki     lahti     0.02    Active         50       4.93    12/31/96    Lahti, Tracy
```

If the account or project name is specified using the `-p` option, information is returned for all users of that account or project on all applicable machines. The data displayed by `usage` is normally updated once each day.

For a more complete description of the `usage` options, enter **`man usage`**.

Changing Accounts for Charging

If you have more than one account, you can charge to various accounts within a logon session. To begin charging to a different account, use the `newacct` command. The syntax is:

`% newacct` *account*

where *account* is another valid account for your userid. *account* is then charged until you log off or reissue the `newacct` command with a different account number. To report your currently active account, enter:

`% newacct -l`

NOTE: In the line above, enter a lower-case L, not the number 1.

To find your valid account names, enter:

`% newacct -a`

The `defacct` command lets you define a default account to be charged during the logon process. Enter `defacct` at the prompt to set your default.

```
% defacct
You currently do not have a default account.
Enter new default account (or none): your_default_account
Default account set to your_default_account .
```

Where *your_default_account* is the account or project name.

To reenable the prompt for account at login, enter **none** in response to the `defacct` prompt. To leave your default account unchanged, press `RETURN` in response to the `defacct` prompt. `defacct` *is local to the machine on which it is executed.*

NOTE: Most users have only one account. You do *not* need to define a default account unless you have multiple accounts.

Refunds

NCSA considers a request for a refund for service units when one of the following is true:

- output was lost because of a system crash or hardware failure
- a bug in system software caused incorrect output
- NCSA staff gave misleading advice, causing incorrect output

Refund requests are denied when one of the following is true:

- fewer than 0.5 SUs were consumed (academic users)
- the request was made more than two weeks after the charge
- no charges were recorded for the job in question (e.g., the system crashed before a job finished so the CPU usage was not recorded in the accounting file)
- output was lost because the job did not use the machine-local scratch file system

Submit the following information to consult@ncsa.uiuc.edu to request a refund:

- user ID
- reason for the request
- machine name
- lsbatch request ID (see "lsbatch Commands" on page 6–3)
- date the job finished

Industrial users contact your assigned consultant.

The list above shows the minimum information required. Additional information that may be requested includes error messages or a copy of your batch script. In such cases, you will be contacted by a member of the consulting staff who will let you know what additional information is needed to review your request.

Remember that job charge information is *not* entered into the accounting records until a batch job finishes. Therefore, you are not charged for the time used by your job if it aborts due to a system crash. (lsbatch does not restart jobs after a system crash.)

See the `/usr/news/Refund_Policy` file for the most current information on this topic.

Extensions

If your project is approaching its expiration date and you want to request an extension of time, send electronic mail at least two weeks prior to deactivation to the NCSA Allocations Group (allocations@ncsa.uiuc.edu). Include a brief report of progress to date, justification for the extension, and a plan for completing the project. Your request will be reviewed prior to deactivation, and you will be contacted via electronic mail with the extension decision.

The extension policy is the same for all NCSA systems. No extensions are granted without a written request received in advance of the expiration. At least 75 percent of your account must have been used by the termination date. A maximum of one three-month extension period can be requested.

Training

NCSA presents bimonthly training sessions on the POWER CHALLENGE . The training is held at NCSA facilities on the campus of the University of Illinois at Urbana-Champaign. The current training schedule is available on NCSA's World Wide Web server (see "World Wide Web" on page 2–3) at the following URL:

```
http://www.ncsa.uiuc.edu/General/Training/training_homepage.html
```

Send electronic mail to ncsa-train@ncsa.uiuc.edu if you have questions about training or if you would like to register for one of our training events.