**INSTITUT DU DÉVELOPPEMENT
ET DES RESSOURCES
EN INFORMATIQUE SCIENTIFIQUE**

**CENTRE NATIONAL
DE LA RECHERCHE
SCIENTIFIQUE**

IDRIS Environment

Beginner's Guide

February 29, 2012

This document gives very important recommendations to use efficiently IDRIS environment.
Warning: IDRIS environment changes continuously so consult regularly the updates of this document on our Web server (`http://www.idris.fr` –> English Corner –> User Support –> IDRIS environment).

Contents

1 About IDRIS

1.1 Missions and objectives

Created in 1993, IDRIS (Institute for Development and Resources in Intensive Scientific computing) is CNRS's national supercomputing centre. IDRIS works in close partnership with other national supercomputing centres CINES at Montpellier, funded by the french Ministry for National Education and CCRT at Bruyere le Chatel (supercomputing centre of CEA). Together, they provide high end supercomputing resources and services to academic research laboratories in France.

To cope with demanding supercomputing requirements arising from computational sciences, IDRIS, CINES and CCRT deploy world class supercomputing environment and assist research scientists through their highly efficient User Support teams. n and Engineering Sciences

IDRIS is both a national server for high performance computing and a centre of excellence on information technologies applied to Computational Sciences. IDRIS operates like other organizations operating large scientific instruments, and reports to an Administration Committee and to Information and Engineering Sciences and Technologies CNRS's department (ST2I). The Administration Committee is the authority that supervises the way in which IDRIS fulfils its missions. IDRIS direction also relies on the Scientific Council for all what concerns the scientific management of its HPC resources.

Actually, director of IDRIS is Denis GIROU

2 IDRIS resources

2.1 HPC at IDRIS

Computer	Number of cores	Memory	Peak Performance
IBM BG/P *Babel*	40960	20 Tbytes	139 Tflops
IBM Power6 *Vargas*	3584	17.5 Tbytes	67.3 Tflops

2.1.1 IBM Blue Gene/P (Babel)

IBM Blue Gene/P logins are available only on an IBM Power5+ front-end (`Babel`). There is no access on the compute nodes (IBM PowerPC 450).

- IBM Power5+ front-end:
  - 16 Power5+ cores (1.85 GHz),
  - 64 Gbytes of memory,
  - Operating system Linux Suse.
- 10 racks Blue Gene/P (10240 compute nodes, 40960 PowerPC 450 cores):
  - 2 midplanes per rack,
  - 16 node cards per midplane,
  - 32 compute nodes per node card,
  - Each compute node is a quad-core PowerPC 450 (850 MHz, 32 bits, 2 Gbytes of memory, 13.6 Gflop/s),

- There is one I/O node for each partition of 64 compute nodes (400 Mbytes/s per I/O node maximum for writing and reading),
- Operating system lightweight Linux kernel.
- Disk space:
  - 800 Tbytes of disk shared between Blue Gene/P (front-end and compute nodes) and Power6 nodes for `HOME`, `WORKDIR` and `TMPDIR` (8 Gb/s maximum flow rate).
- 3 main networks:
  - 3D Torus: 6 bidirectional links (total bidirectional bandwidth 5.1 GB/s per node, MPI latency of 3 to 10 microseconds). Optimised for point-to-point communications and multicast. Full torus only if job partition size is a number of midplanes (512 compute nodes),
  - Tree Global Collective (one-to-all): 3 bidirectional links (total bidirectional bandwidth 5.1 GB/s per node, MPI latency of network traversal less than 2 microseconds with an additional 2 microseconds for broadcasting). Optimised for collective communications (reductions, broadcasts,...),
  - Global Interrupt allowing very fast process synchronization with a low latency: used by global MPI barrier (in MPI_Comm_world communicator).

2.1.2   IBM Power6 (Vargas)

- 112 computing nodes with for each 32 IBM Power6 processors (Dual-core p575 IH, 4.7 GHz):
  - 28 nodes with 256 Gbytes of shared memory per node,
  - 84 nodes with 128 Gbytes of shared memory per node,
  - Operating system IBM AIX,
  - InfiniBand x4 DDR network.
- Disk space:
  - 800 Tbytes of disk shared between Blue Gene/P (front-end and compute nodes) and Power6 nodes for `HOME`, `WORKDIR` and `TMPDIR` (8 Gb/s maximum flow rate),
  - 400 Tbytes of disk for all local TMPDIRs (distributed on Power6 nodes, between 500 Mb/s and 1 Gb/s flow rate).

2.2   Centre front-end and pre/post-processing IBM X3950 M2 (Ulam)

- Architecture:
  - 16 quad-core Xeon X7350 (2930 MHz),
  - Peak performance of 187.5 Gflops,
  - 192 Gbytes of memory (PC2-5300 CL5, bus à 1066 MHz),
  - Operating system Linux Suse.
- Disk space: 38 To distributed in (by default):
  - 150 Mo per unix group for `HOME` (daily backups),
  - 150 Mo per unix group for `WORKDIR` (permanent without backup),
  - 2 To for `TMPDIR` (shared by all users).
- `Ulam` is IDRIS centre front-end because it allows:
  - to propagate the same password on all IDRIS machines with IDRIS `Passwd` command,
  - Pre/post-processings of numerical results issued from computing hosts,

– to access to the `HOME` of the file server by means of `HOMEGAYA` environment variable.

2.3 File server SGI ALtix 4700 (Gaya)

- Architecture:
  - 32 dual-core Itanium 2 (1,67 GHz),
  - Peak performance (cores*frequency*4flops): 427,52 Gflops,
  - 128 GBytes of memory,
  - Operating system Linux Suse 10 SP2.
- Devices:
  - 768 TBytes of RAID Disks on line (InfiniteStorage 4600),
  - One tape storage system StorageTek SL8500: since the end of 2010, it manages 3300 tapes offering a total capacity of 6.6 PBytes. With such tapes (about 2 TBytes per tape) maximum capacity is about 10 PBytes (maximum of 5000 tapes),
  - **DMF** (*Data Migration Facility*) to manage file migration between disks and tape storage system,
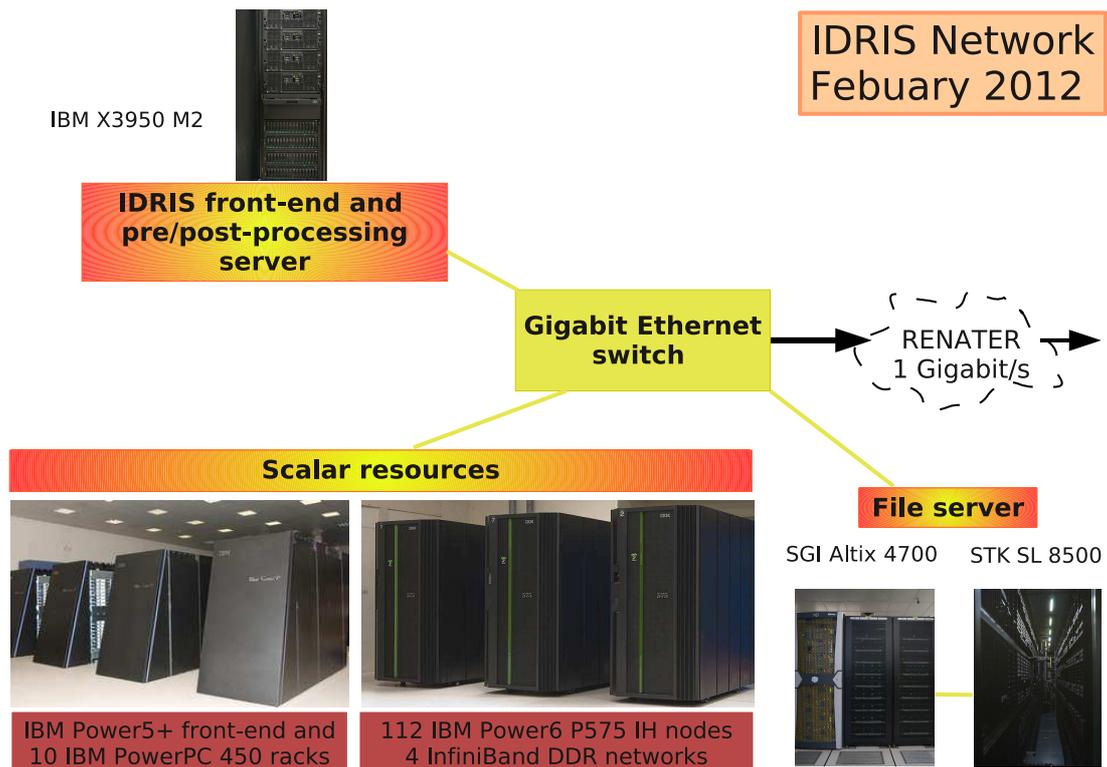
2.4 IDRIS network

Figure 1: IDRIS network

3 Connecting to IDRIS systems

3.1 The scientific project

To have an account on our computing hosts, a scientific project should be presented. For that purpose, the person in charge for this project must fill an annual request for attribution of data-processing resources (DARI form) from our web site (https://www.edari.fr/). Scientific evaluations (of all projects) are carried out by thematic committees (in june and december). They make recommendations which allow the Scientific Council to attribute hours on the computing hosts referred by the project.

Apart from these sessions, requests can be submitted to the IDRIS director (see « au fil de l'eau » request on https://www.edari.fr/), who is authorized to carry out provisional attributions to start or to finish projects.

When the project has received CPU hours, the project leader must ask to open account(s) on the computing machine(s) by sending a FOGEC form to us. Then, each participant to the project will receive a letter specifying his login and the associated initial password (the same on all hosts) to log in our machines (for the very first time).

Note: for each account on an IDRIS computing host, you have also access to our IDRIS centre front-end Ulam and to our file server Gaya.

You will also have to return us administrative forms (see on our Web site: `www.idris.fr -->` `English version --> Forms`:
- **FEPC**: Login Protection Commitment Form,
- **FAIP**: TCP/IP Connection with IDRIS Administration Form,
- **FTIP**: TCP/IP Connection with IDRIS Technical Form.

3.2 Connexion to IDRIS Systems

To log in an IDRIS machine, your local machine must be registered within the IDRIS filters (send FTIP and/or FAIP forms).

To access IDRIS production systems over the Internet, you can use the classical `rlogin` and `ssh` commands. But if the domain name of your local host is not suffixed by `.fr` then you can access only to `Ulam` and `Gaya` via `ssh` protocol:
- **rlogin**:
    `local_machine: rlogin -l logname_idris machine.idris.fr`
- **ssh**:
    `local_machine: ssh -X machine.idris.fr -l logname_idris`

Note: the `-X` flag of `ssh` is not valid on `Gaya` (file server).

3.2.1 Using ssh through a RSA/User Authentication

The RSA/User Authentication method is recommended if you need to bypass password checking. Ronald Rivest, Adi Shamir, and Leonard Adleman developed the RSA system in 1977; RSA stands for the first letter in each of its inventors' names. As long as `ssh` is made available on your local machine, you can simply run `ssh` and bypass password checking by using the following steps:

1. **Setting up ˜/.ssh/known_hosts file**:
   This step authenticates a client machine. Each server host has a `˜/.ssh/known_hosts` file that identifies all possible client hosts. To set up this file, log in an IDRIS machine (with any method) and run ssh to the client host (your local machine):

   ```
   machine.idris.fr: ssh login@local_machine.domaine.fr pwd
   The authenticity of host 'local_machine.domaine.fr (xxx.xx.xxx.xx)'
   can't be established.
   RSA key fingerprint is 52:9c:53:09:f4:b4:f6:f1:bb:ca:5d:11:03:03:f2:94.
   Are you sure you want to continue connecting (yes/no)? yes
   Warning: Permanently added 'local_machine.domaine.fr,xxx.xx.xxx.xx' (RSA)
   to the list of known hosts.
   login@local_machine.domaine.fr's password:
   /home/local_machine/login
   ```

2. **Creating public keys with ssh-keygen**:
   Log in the client machine (your local host) and run ssh-keygen with a choosen authentication mode (RSA or DSA) to create pair of keys (one public and one private). When prompted for a passphrase, you can simply enter an empty passphrase. Normally each user willing to use `ssh` with RSA or DSA authentication mode goes through this procedure only once to create the authentication key in `$HOME/.ssh/id_rsa` (or `$HOME/.ssh/id_dsa`).
   Important: if `ssh` is initialized through an `RSA/User Authentication`, for file transfers (`sftp`, `bbftp`) in batch jobs you must initialize `ssh` with an empty passphrase or run `ssh-agent` to avoid blocking the job with prompt for the passphrase.

   ```
   client_machine: ssh-keygen -t rsa
   Generating public/private rsa key pair.
   Enter file in which to save the key (/home/client_machine/login/.ssh/id_rsa):
   Enter passphrase (empty for no passphrase): ************
   Enter same passphrase again: ************
   Your identification has been saved in /home/client_machine/login/.ssh/id_rsa.
   Your public key has been saved in /home/client_machine/login/.ssh/id_rsa.pub.
   The key fingerprint is:
   9e:84:04:ed:cb:d8:01:04:21:d4:f4:4d:a5:46:74:1b login@client_machine.domain.fr
   ```

3. **Copy public keys to server**:
   Log in the client machine (your local host) and add the public keys (`˜/.ssh/id_rsa.pub` file) from the client machine to the server machine (`˜/.ssh/authorized_keys` file). The list of public keys stored in this file will enable the server to grant a connection request from the client.

   ```
   client_machine:
      cat ˜/.ssh/id_rsa.pub | ssh logname_idris@machine.idris.fr \
        "cat - >> ˜/.ssh/authorized_keys"
   ```

4. **Run ssh-agent to control private keys**:
   From the client machine (where you run `ssh`), type:

   ```
   client_machine: eval 'ssh-agent'
   ```

7

```
Agent pid xxxx
```

The purpose of running `ssh-agent` is to avoid typing the passphrase interactively. `Ssh-agent` provides the passphrase automatically.

Remarks:
- `ssh-agent` is not running as a separate process: user must run `ssh-agent` and `ssh` from the same window/process.
- To avoid running `ssh-agent`, you have just to provide an empty passphrase to `ssh-keygen` (but it is less secure).

5. **Add keys to agent with ssh-add**:
   From the process under `ssh-agent` on the client machine (where you run `ssh`), type:

```
client_machine: ssh-add ~/.ssh/id_rsa
Enter passphrase for /home/client_machine/login/.ssh/id_rsa:
Identity added: /home/client_machine/login/.ssh/id_rsa
(/home/client_machine/login/.ssh/id_rsa)
```

You can also type `ssh-add -l` to display the private keys in the agent or `ssh-add -d <file>` to delete keys from the agent.
Once the private key of a client machine is held by the agent, the interactive passphrase dialog will be bypassed.

3.3 Managing your account

3.3.1 Passwords

A default initial password is sent to you as a part of the new user package that you receive via postal mail after your account is created. It is important that you keep this letter in a secure place for your further reference.

Notes:
- machines require you to change the initial password the first time you log in,
- IDRIS requires you to change your password every ten weeks. You will be notified by email of your password expiration (30, 14 and 1 days before it expires),
- expiration means only that you will be prompted to change your password when you log in after the deadline (your account isn't closed and your current password remains valid),
- If you change your password more frequently, you should never be notified.

Warning: Never share your password with another user nor make your password known by anyone else: you are responsible for your account use against the law.

3.3.2 Changing your password

We advise you to change your password from the centre front-end with the IDRIS command named `Passwd` (note the `P` uppercase letter). This IDRIS command will change password on centre front-end and then automatically propagate the new one to all IDRIS computers you are allowed to connect.

Alternatively, you can use the unix command `passwd` on each machine.

Password recommendations: never use
- login name in any form,
- family member's first name,
- other information easily obtained about you, such as license plate numbers, phone numbers,
- only digits or only letters,
- word of English or foreign language dictionaries, spelling lists, or other lists of words,
- password shorter than six characters,
- alphabetic characters replaced with non-alphabetic characters that are look-a-likes: '0' for 'o' or '1' for 'l',
- `#` or `@` character.

3.3.3 If you have forgotten your password

1. If this problem does not apply to the centre front-end `Ulam`, you can use IDRIS command `Passwd -f` (only on `Ulam`) to transmit your current `Ulam` password to all other IDRIS machines. If your account was locked (after a great number of unsuccessfull login test) on computing host, it will be automatically unlocked. Notes that password propagation can take few minutes to be done. So wait a little time before trying again to connect you on computing host.

2. if this problem applies to `Ulam`, you have to contact the IDRIS Helpdesk (`assist@idris.fr` or (33) (0)1 69 35 85 55) which can reset your password to the default (see default initial password letter). If you did not keep this letter handy, you will need to send a signed fax ((33) (0)1 69 85 37 75) to require a new one (specify your login, hosts, project number, name and your phone number).

3.3.4 Customize your environment

When your account is first activated your default shell is `Bash`. Enter "man bash" at the prompt for details about `Bash`. To change your login shell, use the unix command `chsh`.

You can customize your sessions by modifying your default "dot" files (`$HOME/.bash_profile`, `$HOME/.bash_login`, `$HOME/.profile` and/or `$HOME/.bashrc`).

For example, you may set the environment variable `TERM`:

```
export TERM=xterm        # Ksh/Bash syntax
setenv TERM xterm        # Csh/Tcsh syntax
```

3.3.5 Set display to your local workstation

Some tools (Totalview, XEmacs, etc..) need to display a window on your local host. Thus, you may obtain the following error messages:

```
Error: Can't open display
```

or

```
Xlib:  connection to "machine:0.0" refused by server
Xlib:  Client is not authorized to connect to Server
```

1. In the first case, you must set the environment variable `DISPLAY`:

   ```
   idris_machine: export DISPLAY=machine.locale.fr:0    # Ksh/Bash syntax
   idris_machine: setenv DISPLAY machine.locale.fr:0    # Csh/Tcsh syntax
   ```

2. In the second case, you must allow your local machine `local.machine.fr` to accept openning of a window comming from IDRIS machine. On your local host, type the command:

   ```
   local_machine: xauth list $DISPLAY
   ```

   this command returns a line like this:

   ```
   local.machine.fr:0  MIT-MAGIC-COOKIE-1  529f63e074742901d058539c0149691a
   ```

   and then type this command (on IDRIS machine):

   ```
   idris_machine: xauth add machine.locale.fr:0  MIT-MAGIC-COOKIE-1  \
                      529f63e074742901d058539c0149691a
   ```

   Note: Remember that your local machine must be registered within the IDRIS filters.

3.3.6   `.rhosts` file

The use of remote commands (*rlogin,rsh,rcp, etc.*) requires a valid `$HOME/.rhosts` file: this file defines which remote hosts (computers on a network) can invoke certain commands on the local host without supplying a password.
Here is an example of a `.rhosts` file on `Ulam`:

```
Ulam-rlabxxx: cat .rhosts
  vargas.idris.fr rlabyyy            (1)
  local_machine.domain.fr logname    (2)
```

(1) allows `rlabyyy` to execute remote commands on `Ulam` under `rlabxxx` from `Vargas`.
(2) allows `logname` to execute remote commands on `Ulam` from `local_machine.domain.fr`.
Notes:
  • access permissions of the `.rhosts` file must be set to 600: read and write for the owner only (`chmod 600 .rhosts`).
  • `rcp` and `rsh` fail if your user password has expired.

3.4   Managing your time allocation

You can monitor your elapse time usage with the `cpt` command on each IDRIS HPC machines. On `Ulam`, you can use the `Cpt` IDRIS command to display cumulative information about your elapse time usage on all applicable IDRIS HPC machines at the same time.
These two commands give:
  • your project number with the list of logins of your group,
  • the number and elapse time usage of jobs for each login,

- the total number and elapse time usage of jobs for your group,
- elapse time allocation and the consumed percentage for your group.

On a computing machine, when project reachs:
- 90% of its time attribution, a postal mail is sent to the project leader,
- 100% of its time attribution, all accounts of the project are blocked on this computing host:
  - you cannot submit batch job any more,
  - but you can again log in this host during 3 weeks to recover your files,
  - project accounts will be closed on this host after these 3 weeks.
  - project accounts stay open on all other computers until the end of the year.

We also created an extranet web service allowing you to know detailled time usage of your project host by host, login by login and for one login, job by job: (`www.idris.fr --> Espace utilisateurs --> Extranet`).

3.5 Managing your disk usage

It is important to supervise your disk usage because some disk spaces are subjected to quotas. If you reach your quota, you will not be able to create any new file in these spaces. We invite you to consult the section 4.6 which details the principle of the quotas.

To check your quota for an IDRIS machine, use `quota_u [-w]` IDRIS command. The `-w` flag gives the usage and quota on the `WORKDIR` filesystems. By default (without flag), you have information about `HOME` quota:

```
idris_machine: quota_u
    *************************************************************************
            Quotas du group: lab dans le file systeme /home
    *************************************************************************
                                         Mo   Atteint à      Nb.fics
    -----------------------------------------------------------------

    Quota hard  :                       6160     87.10%            0
    Quota soft  :                       5600     95.90%            0
    Utilisation :                       5367                    6815


    Login                   Espace occupé en Mo         %      Nb.fics
    -----------------------------------------------------------------

     rlab011 nom1                       1067     19.88%         1581
     rlab005 nom2                        932     17.36%         1248
     rlab043 nom3                        501      9.33%          130
     rlab054 nom4                       2867     53.43%         3856
```

This is what it means:
- `file systeme` is the file system on which the quota is given (here `HOME` space of lab group),
- `Quota hard` line is the "hard limit" for group. If your group reachs this limit, each member (or batch job) will be unable to create any new file,
- `Quota soft` line is the "soft limit" for group (approximately 90% of hard quota). Beyond this limit, each member is warned by email to clean up his account in the next 14 days,

- there can be a limit on the number of files you are authorized to create (Nb fics),
- `Utilisation` line is the total current group's usage,
- And the list of usage for each login (with corresponding ratio of total usage).

If you are blocked, you can:
- clean your directory by deleting some files,
- move some files on another filesystem or on the file server,
- If you need extra disc space to undertake your work, send a fully justified request to the IDRIS director to ask for special arrangements. For this, we created an extranet web service allowing your project leader to request more inodes or disk space: (`www.idris.fr --> Espace utilisateurs --> Extranet`).

4 All about file systems

The following filesystems are available on all IDRIS machines through environment variables:
- `HOME`,
- `WORKDIR` (not on file server),
- `TMPDIR` (not for interactive sessions on IBM BlueGene/P).

4.1 `HOME` Directory

This is the default directory where you are in when you log in and it is referenced by `$HOME` environment variable . You should use this space to store files that you want to keep for a long term, such as source code, scripts, etc. By default, IDRIS allocates a small disk quota for your HOME directory because it is backed up daily. But project manager can ask more `HOME` quota via our web site: (`www.idris.fr --> Espace utilisateurs --> Extranet`).
On HPC computers, this file system is mounted on all compute nodes.
On file server `Gaya`, `HOME` is a set of disks:
- this disk (RAID) is a buffer between tape storage system and HPC machines. Initially, files moved to the `HOME` of `Gaya` stay on disk where they can be reached quickly.
- files on disk will be automatically selected by **DMF** software (*Data Migration Facility*) for archiving to tapes (copy based on size and time of last access to file). However, copy of migrated files stay also on disk until disk space isn't full.
- access to an offline file (file on tape) is longer than access to an on-line file (file on disk): file must be automatically first copied from tape to disk (except if there is still an on-line copy of the migrated file).

4.2 `WORKDIR` Directory

The `WORKDIR` file system is intended for short term use and should be considered as volatile. The size of these file systems depends on host. Project manager can ask more `WORKDIR` quota via our web site: (`www.idris.fr --> Espace utilisateurs --> Extranet`).
Please note that **backups are not performed on** `WORKDIR` directories. In the event of a disk crash or file purge, files on these directories cannot be recovered. Therefore, you should make sure to back up your files to permanent storage (for example on our file server Gaya) as often as significant changes

are made.

On all hosts, WORKDIR is referenced by $WORKDIR environment variable.

On HPC computers, this file system is mounted on all compute nodes.

4.3 TMPDIR Directory

TMPDIR file system is intended for short term use (in most cases, the duration of a batch job) and are volatile. The size of TMPDIR file systems depends on host but is always wider than HOME and WORKDIR.

- a TMPDIR directory is available (and different) for each interactive work (except on IBM BG/P) and each batch job (referenced by $TMPDIR environment variable),
- files in a TMPDIR directory associated with a batch job are purged as soon as the batch job is completed. Thus, you should take care to save all output files of your run to the mass storage server before the end of the batch job,
- When it's possible, TMPDIR is defined as a local disk space on the compute node. Thus, I/O operations are faster in the TMPDIR than in the HOME or WORKDIR. If it's not possible (runs using a great number of processors distributed on many compute nodes), TMPDIR is defined on global file system shared by all compute nodes.

4.4 /tmp, /usr/tmp and /var/tmp Directories

The /tmp, /usr/tmp and /var/tmp directories are intended for temporary files that are used during the execution of a process or job. Please **do not use these directories since there are very limited spaces**. Do not attempt to use these directories for permanent storage. Files placed in /tmp, /usr/tmp or /var/tmp may be purged at any time.

4.5 NFS mounts on IDRIS center front-end

HOME directory of file serveur (disk on line) is NFS (*Network File System*) mounted on Ulam and reachable with an environment variable HOMEGAYA.

4.6 Disk quotas

Because of the finite amount of disk space, we have had to create a policy for creating and upgrading volumes. Each user on IDRIS machine has a disk quota and a file quota on the HOME and WORKDIR filesystems.

Quotas are defined per unix group with the following rules:

- a "hard limit" sets the maximum disk space reserved for the group. Cumulative disk usage of all members of the group can not exceed this limit. Once this limit is reached, each member of the group will be locked: nobody will be able to create new files any more,
- a "soft limit" sets at about 90% of hard limit. Beyond this limit, each member of the group is warned by email to reduce his disk usage in the next 14 days. Once this deadline has expired, if you have done nothing then you will be locked.

On each IDRIS machine, you can know your disk usage (and limit) with IDRIS command `quota_u` (see `man quota_u` or the subsection 3.5 above for more information).

If you are locked or if you have received a warning message, you can:
- clean your directory by deleting some files,
- move some files on another filesystem or on the mass storage system `Gaya`,
- If you need extra disk space to undertake your work, send a fully justified request to the IDRIS director to ask for special arrangements. For this, we created an extranet web service allowing your project leader to request more inodes or disk space: (`www.idris.fr --> Espace utilisateurs --> Extranet`).

5   File transfers

You can use several methods to move data between IDRIS and other hosts via the network:

	From IDRIS		To IDRIS
	Ulam Vargas/Babel	Gaya	Ulam/Gaya/ Vargas/Babel
Command	Interactive/batch	Interactive	Interactive mode
mfput/mfget	YES	NO	NO
bbftp	YES	YES	YES
ftp	YES	YES	YES
sftp	YES	YES	YES
scp	YES	YES	YES

Table 1: Supported file transfer methods

5.1   mfput/mfget

- `mfget` IDRIS command is a **secure file transfer command** which allows to retrieve files from the mass storage system `Gaya`:

    ```
    idris_machine-rlab000: mfget gaya_path/gaya_file local_path/local_file
    ```

    where `gaya_file` is a filename stored on `Gaya` in `gaya_path` directory (full path or relative path from `Gaya` HOME). `local_path` can be full path or relative path from the current directory on `idris_machine`.

    Warning: local or remote paths must exist if specified.
- `mfput` IDRIS command is a **secure file transfer command** which allows to store a local file on the mass storage system `Gaya`:

    ```
    idris_machine-rlab000: mfput file gaya_path/.
    ```

where `file` is a file on `idris_machine` to be saved on `Gaya` in `gaya_path` directory (relative path from `HOME`).

Warning:
  - if the remote path does not exist on `Gaya` it is created,
  - if this command fails, contact IDRIS helpdesk to retrieve your files.

5.2 bbftp

`bbftp` is one of several file transfer methods that is supported by IDRIS. `bbftp` only encrypts usernames and passwords, it does NOT encrypt the data being transferred. `bbftp` is a non-interactive FTP-like system that uses parallel tcp streams for data transfers, allowing to achieve bandwidths greater than simple FTP. Because of these characteristics, it is the best way to transfer large data files to and from IDRIS.

In order to use `bbftp`, you will need to obtain a `bbftp` client and server. The BBFTP source code can be obtained from the home `bbftp` web site (`http://doc.in2p3.fr/bbftp/`). A `bbftp` server is running on each IDRIS machine.

Use this for bulk data transfer (i.e. large files) as it is very efficient, and passwords are encrypted. For a complete description of `bbftp`, enter `man bbftp` on an IDRIS machine.

5.2.1 Using bbftp in interactive mode

The standard `bbftp` command takes an input command file containing a list of instructions (`cd`, `put`, `get` and so on). Options are available to enable compression, to specify the number of parallel transfer streams and so on. This command is also suitable for command line use: the user name and password are supplied when the command is invoked.

  - the following command means that `bbftp` is going to connect to remote host `machine.idris.fr` using `rlabxxx` username. If the connection is successful then the command `put file1 file2` will be executed (transfer the local file `file1` to the remote host with the name `file2`).

    ```
    local_machine: bbftp -e 'put file1 file2' -u rlabxxx machine.idris.fr
    Password:
    put file1 file2 OK
    ```

  - this other example means that `bbftp` is going to connect to remote host `machine.idris.fr` using `rlabxxx` username. If the connection is successful then the command `mput file1 file2` will be executed (transfer the local files `file1` and `file2` to the remote host with the same names).

    ```
    local_machine: bbftp -e 'mput file1 file2' -u rlabxxx machine.idris.fr
    Password:
    mput file1 file2 OK
    ```

  - the following command transfers the remote file `file1` to the local host with the name `file2`. If the local file already exists it is overwritten (only in case of successful transfer and if the

setoption tmpfile has been set). Under some circumstances (No space left on device, Access denied, File is a directory ...), no retry is done and the next command is processed.

```
local_machine: bbftp -e 'get file1 file2' -u rlabxxx machine.idris.fr
Password:
get file1 file2 OK
```

If the remote file is given in relative mode (not beginning by a /), it is supposed to be relative to the current directory on the remote host (which is set to the home directory of the RemoteUsername at the beginning). If the local file is given in relative mode (not beginning by a /) the file is created relatively to the directory where the `bbftp` command is running (which may have been changed during the session with the `lcd` command).

5.2.2   Using `bbftp` in batch mode

When used in a script, `bbftp` would normally ask you for your password. This is obviously locking for a batch job. To avoid this problem, use `ssh` authentication and run bbftp as usual with the `-s` flag:

```
machine.idris.fr: bbftp -e 'put file1' -u login -s local_machine.domaine.fr
put file1 OK
```

This will force the `ssh` program to connect to the remote machine and start the bbftp daemon. The username/password prompt will be skipped if ssh has already taken care of that: section 3.2.1 explains how to use `ssh` and bypass password checking through a RSA/User Authentication. To use `bbftp` in batch mode from an IDRIS machine to your local machine, you must generate public keys with `ssh-keygen` and enter an empty passphrase (to avoid passphrase checking), then copy it to your local machine.
Control messages will be transmitted via the encrypted ssh channel, but data will be sent using the normal bbftp stream(s).

5.3   ftp

`ftp` is one of several file transfer methods that is supported by IDRIS. In order to use `ftp`, you will need a `ftp` server running on your system. `ftp` requires (but does not encrypts) usernames and passwords; as a consequence, IDRIS does not recommend this method of data transfer. In more, `ftp` can not be used in your batch jobs.

5.4   sftp

`sftp` is one of several file transfer methods that is supported by IDRIS. Since it encrypts the transferred data, it should only be used to transfer small files (1-10KB) containing sensitive datas. For large files that contain no sensitive information it's better to use methods that **don't** encrypt datas (like `bbftp`).
Since `sftp` is an Ssh 2 command, you need one of the `ssh 2` clients that provides the `sftp` client program.

`sftp` is a functional clone of `ftp` so if you know how to use `ftp` you know the rudiments of `sftp`. The Unix version of `sftp` is a command line program, the commercial Windows version of `sftp` provides a GUI front end to the program.

Again, since `sftp` encrypts all datas, the load on both the source and destination systems will be high. Therefore it is only available in interactive mode on HPC IDRIS machines (see section tab 5).

As for `bbftp` through `ssh`, the username/password protocol will be skipped if ssh has already taken care of that (see . section 3.2.1 for more details).

5.5 scp

`scp` is one of several file transfer methods that is supported by IDRIS. Since it encrypts the transferred datas, it should only be used to transfer small files (1-10KB) containing sensitive datas. Large files that do not contain sensitive information should be transferred via methods that **don't** encrypt datas (like `bbftp`).

6 Work environment

6.1 Recalls and recommandations

- To log in an IDRIS machine, your local host must be registered within the IDRIS security filters (send FTIP and/or FAIP forms);
- The project leader must ask to open account(s) on computing machine(s) (by sending FOGEC form) for each participant in the project;
- From an IDRIS machine, news can be read through the unix command `news`. For example, the command `news planning` will give you planned stops (cf. section 8);
- The default shell for IDRIS users is `Bash`. To change to another shell, type `chsh` at the command prompt (see `man chsh`).
- To work on BG/P host, you must log in front-end (Babel) for:
  - editing files, moving files, packing tar archives, gzipping files, etc
  - cross-compilation; by using:
    * `mpixlf2003_r, mpixlf95_r, mpixlf90_r, mpixlf77_r, mpixlc_r, mpixlcxx_r`,
    * `bgld, bgranlib, bgar, make, man, size`.
  - batch job management (using LoadLeveler).

    Remarks:
    * all new projects are opened with a development mode status (32,768 computing hours and a maximum of 4096 computing cores) for porting applications and doing scalability tests. After these tests, project leader must ask to go to production mode status via our web site (https://ebabel.idris.fr/) to be able to use all project hours attribution on all host configuration.
    * all informations can be founded on our web site `http://www.idris.fr`. Just follow the links: English Corner –> IBM Blue Gene/P

If you have any problem then contact IDRIS Helpdesk (`assist@idris.fr` or (33) (0)1 69 35 85 55).

6.2 Interactive and batch mode

You have two possible working methods: interactive sessions and batch jobs. In both cases, you will be constrained to respect the maximum limits, in CPU/Elapse times, number of processors and memory, designed by IDRIS for a best managing of data-processing resources. You will find further information concerning these limits by typing the command `news class` on machine which you are interested in.

6.2.1 Interactive access

User limits (elapse time, number of cpus, available memory, ...) for all active login sessions depend on the IDRIS machine. Jobs exceeding limit or IDRIS policy are terminated or rejected. Violation messages are sent to the user's active terminal windows along with the list of processes killed. In general, interactive work should be used to compile or cross-compile code, for all development tasks (such as debugging or file editing) and batch jobs management. To submit and control yours batch requests, a batch processing system is available on each computing cluster - see . section 6.2.2 for information.

To run a binary interactively, enter the name of executable and all necessary arguments at the system prompt:

```
idris_machine: ./a.out < input-file > output-file
```

Be aware that this command runs binary on front-end Babel and not on compute nodes. Thus your binary can complain if it is cross-compiled. In this case, you can test your binary by using the command below:

```
babel: bgrun -np 256 -mode VN -exe ./a.out
```

This command runs binary with 256 MPI tasks distributed on compute nodes of our BG/P. Please note that your binary will run only when requested resources will be available.

To bypass limits of interactive sessions, you must use batch system available on the host as explain below.

6.2.2 Running jobs under batch system

The main reasons to use batch mode are:
- time limits and memory limits are higher than for interactive sessions,
- it is possible to close interactive session after submission,
- it allows a better distribution of data-processing resources between the users according to the required resources (for example, a very consuming work will be carried out during the night or the weekend).

At IDRIS, job manager is **LoadLeveler** on `Ulam`, `Vargas` and `Babel`. It controls run of jobs according to the requested resources (memory, Elapse time, number of processors, space disk) and to the number of active jobs at a given moment (total number and number per user).

You have to create a "batch job" (any script that can be run in an unattended way under Unix).

First part of script will define the environment within your job will run, specifying which amounts of memory, Elapse time and number of processors the job requires. If you don't set the job limits, then you will get some low default limits. These environment is set by means of directives which can specify:

- shell,
- request name,
- output filename,
- Elapse time limit for the job,
- memory limit per processus of the job,
- number of process to use (for MPI and/or OpenMP).

Second part of script is a set of unix commands or programs that make up the job.
Thus, general structure of a batch script should be:

- set of directives (to fixe resources);
- go to `TMPDIR` space disk (`cd $TMPDIR`);
- copy all necessary files with:
  - either `cp` command from your `HOME/WORKDIR` filesystem,
  - or `bbftp, scp` (cf. section 5),
  - or `mfget` if your files are stored on `Gaya`;
- run your binary;
- copy output files with:
  - either `cp` command to your `HOME/WORKDIR` filesystem,
  - or `bbftp, scp`,
  - or `mfput` if your files must be saved on `Gaya`.

To submit a batch job (file called `my_job`) , you just have to use the command below:

- LoadLeveler (IBM): `llsubmit my_job`,

Jobs are run in batch queues when the requested resources are available (see `news class` on each host): the higher the limits are, the lower the job priority is.

Important:

- The username/password protocol must be skipped for file transfers (cf. section 5);
- Sample job scripts are available for each job manager system on our web site (`http://www.idris.fr`). Just follow the links: Support technique –> IBM Power6 (or IBM Blue Gene/P or Frontale ulam) –> Exécution/contrôle d'un code...

You can also manage your batch jobs:

1. Getting batch job status:
   - LoadLeveler (Ulam, Babel, Vargas): `llq`,

2. displaying input or ouputs of a running batch job:
   - LoadLeveler (Vargas): `qcat -i job_id` or `qcat -eo job_id`
   - LoadLeveler creates outputs of jobs in the directory where you submit the job from.

3. Terminating (or deleting) batch jobs:
   - LoadLeveler (IBM): `llcancel job_id`

7 Getting help from IDRIS consultants

Members of the IDRIS Support Team are responsible for providing support to all users of IDRIS products and facilities.

They assists users to fully take advantage of IDRIS environment (help to optimize and to parallelize code). They also assume "hot line" services: they answer your questions, sometimes with the help of other IDRIS technical personnels and/or on-site engineers from hardware vendors.

The staff provide assistance by:
- Electronic mail to `assist@idris.fr` (preferred method),
- Phone at `(33) (0)1 69 35 85 55`.

IDRIS assistance team can be reached during the hours of 9:00 a.m. through 6 p.m. (CT), Monday through thursday and Friday from 9:00 a.m through 05:30 p.m.

Information about IDRIS's current high performance computing environment is available on web site: http://www.idris.fr (English version menu).

User Support team also takes part in training activities about languages (Fortran, C) and parallel libraries (MPI, OpenMP).

8 Documentation

- **News** can be read through the unix command `news`:
  - `news -a` gives **all** available news,
  - `news` gives only unread news (take into account the `.news_time` file),
  - `news -n` gives titles of unread news
  - `news subject`, where `subject` is the subject of an unread news, allows you to read the selected news.

  Important :
  - Each IDRIS machine provides its own news,
  - **be careful**, news can be out-of-date!
- **Vendor documentation**: some vendor documents are available through the IDRIS website (http://www.idris.fr).
- **Manuals**: all unix manuals are available through the unix command `man`.

A   Lexicon

Mot de passe	*Password*
Changement du mot de passe	*Change of password*
Transmission du mot de passe	*Password transmission*
Nouveau mot de passe	*New password*
Pages protegees web IDRIS	*Password protected web pages*
Espace occupé	*Disk usage*
Frontale	*Front-end*
Stockage	*File server*
Machines actuellement indisponibles	*List of unavailable machine*
Aucune	*None*
Documentations constructeurs	*Vendor documentation*