

Table of Contents

File Systems	2
Available file systems	2
Stripe Pattern Details	3
Additional infos	3
Notes	3

File Systems

Available file systems

On the DEEP-EST system, three different groups of file systems are available:

- the [?JSC GPFS file systems](#), provided via [?JUST](#) and mounted on all JSC systems;
- the DEEP-EST parallel BeeGFS file systems, available on all the nodes of the DEEP-EST system;
- the file systems local to each node.

The users home folders are placed on the shared GPFS file systems. With the advent of the new user model at JSC ([?JUMO](#)), the shared file systems are structured as follows:

- \$HOME**: each JSC user has a folder under `/p/home/jusers/`, in which different home folders are available, one per system he/she has access to. These home folders have a low space quota and are reserved for configuration files, ssh keys, etc.
- \$PROJECT**: In JUMO, data and computational resources are assigned to projects: users can request access to a project and use the resources associated to it. As a consequence, each user can create folders within each of the projects he/she is part of (with either personal or permissions to share with other project members). For the DEEP project, the project folder is located under `/p/project/cdeep/`. Here is where the user should place data, and where the old files generated in the home folder before the JUMO transition can be found.

The DEEP-EST system doesn't mount the \$SCRATCH file systems from GPFS, as it is expected to provide similar functionalities with its own parallel and local file systems.

The following table summarizes the characteristics of the file systems available in the DEEP-EST and DEEP-ER (SDV) systems. **Please beware that the `$project` (all lowercase) variable used in the table only represents any JuDoor project the user might have access to, and that it is not really exported on the system environment.** For a list of all projects a user belongs to, please refer to the user's [?JuDoor page](#). Alternatively, users can check the projects they are part of with the `jutil` application:

```
$ jutil user projects -o columns
```

Mount Point	User can write/read without	Cluster	Type	Global / Local	SW Version	Native Pattern Details	Maximum Measured Performance (per node)	Description	Other
<code>/p/home</code>	<code>/p/home/juser</code>	DEEP-EST	GPFS	Global				JUST GPFS mount, used only for configuration files	
<code>/p/project</code>	<code>/p/project/cdeep</code>	DEEP-EST	GPFS	Global				JUST GPFS Project directory, GPFS mount, used for project files, for performance reasons applications or benchmarking	
<code>/scr</code>	<code>/scr/\$project</code>	High node only (DEEP)	GPFS	Global				JUST GPFS Project directory, GPFS mount, used for project files, for performance reasons applications or benchmarking	If you plan to use the partition, please join in the system administration (e.g. via the request mailing list). You can find more info about the partition and some instructions using the JUDO tool.
<code>/data</code>	<code>/data</code>	DEEP-EST	BeeGFS	Global	BeeGFS 2.1.5			Fast write file system, no backup, hence not persistent data storage	
<code>/work_001</code>	<code>/work_001</code>	DEEP-EST	BeeGFS	Global	BeeGFS 2.1.5			Work file system, no backup, hence not persistent data storage (deprecated)	
<code>/scratch</code>	<code>/scratch</code>	DEEP-EST	File local partition	Local				Work file system for temporary data, will be deleted after job. System. Some effect on the node(s). Recommended to use instead of any other writing temporary files.	
<code>/home/scratch</code>	<code>/home/scratch</code>	DATA partition	Local SSD (PCI)	Local				Scratch file system for temporary data, will be deleted after job. Partition 1-5 TB local Optimus SSD Data Center (DC) PABOX (PCIe) as 2.5, 3D (XFS)	
<code>/home/scratch</code>	<code>/home/scratch</code>	DATA partition	Local SSD (PCI)	Local				Scratch file system for temporary data, will be deleted after job. Partition 1-5 TB local Optimus SSD Data Center (DC) PABOX (PCIe) as 2.5, 3D (XFS)	
<code>/scratch/scratch</code>	<code>/scratch/scratch</code>	DATA partition	DCPABA as scratch node	Local			2.5 GB/s latency 40-60 ms	Scratch file system for temporary data, will be deleted after job. Partition 1-5 TB local Optimus SSD Data Center (DC) PABOX (PCIe) as 2.5, 3D (XFS)	5 TB in 10-20 ms 2 TB in 10-20 ms 1 TB in 10-20 ms

Stripe Pattern Details

It is possible to query this information from the deep login node, for instance:

```
manzano@deep $ fhgfs-ctl --getentryinfo /work/manzano
Path: /manzano
Mount: /work
EntryID: 1D-53BA4FF8-3BD3
Metadata node: deep-fs02 [ID: 15315]
Stripe pattern details:
+ Type: RAID0
+ Chunksize: 512K
+ Number of storage targets: desired: 4

manzano@deep $ beegfs-ctl --getentryinfo /sdv-work/manzano
Path: /manzano
Mount: /sdv-work
EntryID: 0-565C499C-1
Metadata node: deeper-fs01 [ID: 1]
Stripe pattern details:
+ Type: RAID0
+ Chunksize: 512K
+ Number of storage targets: desired: 4
```

Or like this:

```
manzano@deep $ stat -f /work/manzano
File: "/work/manzano"
ID: 0      Namelen: 255      Type: fhgfs
Block size: 524288      Fundamental block size: 524288
Blocks: Total: 120178676 Free: 65045470 Available: 65045470
Inodes: Total: 0      Free: 0

manzano@deep $ stat -f /sdv-work/manzano
File: "/sdv-work/manzano"
ID: 0      Namelen: 255      Type: fhgfs
Block size: 524288      Fundamental block size: 524288
Blocks: Total: 120154793 Free: 110378947 Available: 110378947
Inodes: Total: 0      Free: 0
```

See <http://www.beegfs.com/wiki/Striping> for more information.

Additional infos

Detailed information on the **BeeGFS Configuration** can be found [?here](#).

Detailed information on the **BeeOND Configuration** can be found [?here](#).

Detailed information on the **Storage Configuration** can be found [?here](#).

Detailed information on the **Storage Performance** can be found [?here](#).

Notes

- dd test @dp-dam01 of the DCPMM in appdirect mode:

```
[root@dp-dam01 scratch]# dd if=/dev/zero of=./delme bs=4M count=1024 conv=sync
1024+0 records in
1024+0 records out
4294967296 bytes (4.3 GB) copied, 1.94668 s, 2.2 GB/s
```

- The /work file system which is available in the DEEP-EST prototype, is as well reachable from the nodes in the SDV (including KNLs and ml-gpu nodes) but through a slower connection of 1 Gb/s. The file system is therefore not suitable for benchmarking or I/O task intensive jobs from those nodes
- Performance tests (IOR and mdtest) reports are available in the BSCW under DEEP-ER → Work Packages (WPs) → WP4 → T4.5 - Performance measurement and evaluation of I/O software → Jülich DEEP Cluster → Benchmarking reports:
[?https://bscw.zam.kfa-juelich.de/bscw/bscw.cgi/1382059](https://bscw.zam.kfa-juelich.de/bscw/bscw.cgi/1382059)
- Test results and parameters used are stored in JUBE:

```
user@deep $ cd /usr/local/deep-er/sdv-benchmarks/synthetic/ior
user@deep $ jube2 result benchmarks

user@deep $ cd /usr/local/deep-er/sdv-benchmarks/synthetic/mdtest
user@deep $ jube2 result benchmarks
```