Wikiprint Book

Title: File Systems

Subject: DEEP - Public/User_Guide/Filesystems

Version: 36

Date: 13.05.2025 03:16:41

Table of Contents

File Systems	3
Available file systems	3
Stripe Pattern Details	4
Additional infos	4
Notes	4

File Systems

Available file systems

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

- the <a>?JSC GPFS file systems, provided via <a>?JUST and mounted on all JSC systems;
- the DEEP-EST (and SDV) 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 has a folder within each of the projects he/she is part of. For the DEEP project, such 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 and \$ARCHIVE file systems, as it is expected to provide similar functionalities with its own parallel file systems.

The following table summarizes the characteristics of the file systems available in the DEEP-EST and DEEP-ER (SDV) systems:

Mount Point	User can write/read to/from	Cluster	Туре	Global / Local	SW Version	Stripe Pattern Details	Maximum Measured Performance (see footnotes)	Description	Other	
/p/home	/p/home/jusen	SEXER _{EST}	GPFS exported via NFS	Global				Home directory; used only for configuration files.		
/p/project	/p/project/cdee	PENES,	GPFS exponed via NFS	Global				Project directory; QPFS main stonage file system; not suitable for performance relevant applications or benchmarking		
Jwafk	/work/odeep	DEEP-EST*	BeeGFS	Global	BeeGFS 7.1.2			Work file system	"Also available in the SDV but only through 1 Gig network connection "Recommende	
/scratch	/acratch	DEEP-EST	xfs local partition	Local*				Scratch file system	"Recommenda to use instead of /tmp for storing temporary files	d
./nvme/scratch	/nvme/scratch	DAM partition	xfs local SSD partition	Local*				Scratch file system	Optane SSD Data Center (DC) P4800X (NVMe PCle3 x4, 2.57, 3D XPoint))	
/nvme/scratch	2/nvme/scratch	DAM partition	ext4 local SSD partition	Local*				Scratch file system	"1.5 TB Intel Optane SSD Data Center (DC) P4800X (NVMe PCle3 x4, 2.57, 3D XPoint))	
Jady-work	/adv-workloda	SDV (deeper-adv nodes via EXTOLL, populse mi-gar via GbE only), DEEP-EST (1 GbE only)	BeeGFS	Global	BeeGFS 7.1.2	Type: RAIDO, Churksize: 512K, Number of storage tangent: distired: 4	1831.85 MB/s write, 1308.62 MB/s read 15202 ops/s creates, 5111 ops/s remove*	Work file system	"Test results and parameters used stored in JUBE: usersideep of /war/local, usersideep \$\(\) jube 2 result benchmarks of /war/local, usersideep \$\(\) jube 2 result benchmarks of /war/local, usersideep \$\(\) jube 2 results have "Test results "Test results"	desg-ev/abr-bendinavks/systhetic/tor desg-ev/abr-bendinavks/systhetic/abted
Jeografia	/mvms.tmp	SDV	NVMe device	Local	BeeGFS 7.1.2	Block size: 4K	1145 MMB/s write, 3108 MMB/s read 133148 ops/s create, 62587 ops/s remove*	1 NVMe device available at each SDV compute node	and parameters used stored in JUBE: user@deep \$ cd /var/local, user@deep \$ jube2 result benchmarks user@deep \$ cd /var/local, user@deep \$ cd /var/local, user@deep \$ jube2 result benchmarks	deeg-er/adv-benchmanks/symthetic/ior deeg-er/adv-benchmanks/symthetic/mftant
./mrt/beecnd	/mrabeleond	SDV	BeeGFS On Demand nutning on the NVMe	Local	BeeGFS 7.1.2	Block size: 512K	1130 MiB/s write, 2447 MiB/s read 12511 opsis create, 18424 opsis remove*	1 BesOND instance running on each NVMe device	"Test risults and parameters used stored in JUBE: useradeep \$ od \$ jubez result benchmarks useradeep \$ od \$ jubez result benchmarks useradeep \$ jubez result benchmarks to the pubez result benchmarks to the pubez result benchmarks	desg-er/adv-bendinants/symthetic/ior desg-er/adv-bendinants/symthetic/adtent

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"
         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

- 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 Gig. 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

• 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
```