Wikiprint Book

Title: File Systems

Subject: DEEP - Public/User_Guide/Filesystems

Version: 36

Date: 05.05.2024 11:29:13

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 2JUST 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 deepv login node exposes the same file systems as the compute nodes, but it lacks a local scratch file system. Since /tmp is very limited in size on deepv please use \$SCRATCH instead (pointing to the project folder) or use e.g. the /pmem/scratch on the dp-dam partition \$LOCALSCRATCH on any other compute node when performing SW installation activities. A quota has been introduced for /tmp on deepv to avoid clogging of this filesystem on the login node which will lead to several issues.

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 YJuDoor page. Alternatively, users can check the projects they are part of with the jutil application:

\$ jutil user projects -o columns

ŀ	laure Paire	==	Channe	Type		NW Yaratan	inipe Paran Innah	=	-	anur .
-	shara	phonejaun	Sin.,	orn.	Comme				AND DIFFE Name dealing and project and project fine	
	-	a project Spring	Server	orns especialists arts	Cintra				ANT DATE AND TO THE	
	-		lagin sala arty (lauge)	SPFE. expensed via MFE.	-				AUST GIPTS Authorized for the party of the p	If you parties from the control of t
	and a	anni ilganja	agicosia ety (marc)	onns especialists Mili	Ginna				ANT CHTS ANTON ANT	if you planed to worker him to be a proper or him to be a proper o
-	den.		MER SE	Base OF S	-	2000 2000			Ξ	
			patricular.	Audit is	-	mandrine 0 2 5			Ξ	
	arant.	_	ann aus	els loud parlition	_					
			lota partition	100 ESS	_				demant for system for sergions; description descriptio	
			==_	m=550 pm1	_				billipiti für speam för semperary den fill für den einer den fill für den fill für Genera SAG Genera SAG Genera SAG Genera SAG Genera SAG Genera SAG Genera SAG General SAG	
			EZ	across:	_			de differ despis de mais de different		A TRUE OF THE PARTY OF THE PART

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

• 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
- 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
```