Wikiprint Book

Title: File Systems

Subject: DEEP - Public/User_Guide/Filesystems

Version: 36

Date: 19.04.2025 17:18:00

Table of Contents

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

\$ jutil user projects -o columns

Mauri Palei	User can write/read safrom	Cluster	Туре	Chokat / Local	SW Version	Stripe Patients Details	Maximum Measured Performance (see footnates)	Description	Other	
phone	Shomebuser	SET/ SESP-EST	GPFS expoted via NFS	Chilled				JUST GPFS Home directory; used only for configuration Sec.		
(pigasajes)	ipopospos di BPP	Wir ear	GPFS exponed via NFS	dated				JUST GPPS Project directory: GPPS main disciple file system; not substite for performance introduced applications or projections or projections		
lash	миляном	jągie node Siły (desp.)	GFFS expoted via NFS	Children				JUST CIPPS Author deadiny) Long-term disable solution for disable solution soluti	If you plan to use the actions of the actions of the actions of the system of the syst	
/work	/worksprency	OBSP-63T	BeeGF3	Chillest	BeeGFS 7.1.2			Work tile system, se beckup, hence not meant for permanent data storage. Deprecated		
iatun	isten	SDV, DESP-63T	BeeGFS	Challed	BeeGF3 7.1.2			Fast work the system, ear tackup, hence not meant for permanent data storage		
/borasin	no seno	DESP-EST	ath total partition	Locar				Scools Se system for imporary size. Still be cleaned up after pilo frontes. Size offices on the modules! "Recommend to size included! "Recommend to size moting lemporary the	4	
ihumehaa akih	c (hwherlacolad)	DAM partition	No.00 238D (s44)	Locar				Soudh Se system for lempurary data. SSE be deared up after pilo Southers (T.S. TE base Optare SSD Data Center (DC) PAROX (NAME PCHS AS, 237, 30 XPung)		
Shumahaa salah	STATE OF STREET	DAM partition	Notes SSED (MED)	Locar				SPunit) Boston the system for temporary one. Titl be desired up after pio trusted; "1.8 trust (DER CHIEF (DE) PARCOX (DAME PARCOX (DAME PARCOX (DAME)) THE STATE OF THE STATE		
ijene mini ostal	рменьши	DAM partition	DCPMM is applied and	Locar			2.2 GBVs simple ad less in sip-danich		"A TIS IN dp-dan(01,02 2 TE IN dp-dan(01,02 2 TE IN dp-dan(02,02 INC Options DC DAMA DAMA DAMA DAMA DAMA DAMA DAMA D	
State	Sweeting	SDV	NVIMe device	Local	BardF3 7.1.2	Misch size: of	T1-05 Million write, 3708 Million shed 1281-08 opain centale, 1281-7 opain sentious*	1 NVAN- device available at each SEV compute code	Test media and parameters wand dated in ARE. ware interp. 5 oil year/lanel, ware interp. 5 year/lanel, ware interp. 5 years	dang-we jude-kanadanaka legushirak dang-we jude-kanadanaka legushirak

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
```