

## **Wikiprint Book**

**Title: File Systems**

**Subject: DEEP - Public/User\_Guide/Filesystems**

**Version: 36**

**Date: 06.05.2024 20:10:11**

## Table of Contents

<b>File Systems</b>	<b>3</b>
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 [?JSC GPFS file systems](#), provided via [?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 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 and \$ARCHIVE file systems from GPFS, 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 or software	Cluster	Type	Global / Local	SW Version	Storage Features	Maximum Measured Performance (see benchmarks)	Description	Other
/home	/home/username	SDV, DEEP-EST	GPFS exported via NFS	Global			40MB/s	Home directory, used only for configuration files.	
/opt/jump	/opt/jump/cluster	SDV, DEEP-EST	GPFS exported via NFS	Global			10MB/s	Project directory, GPFS does not store files there, not suitable for performance intensive applications or benchmarking	
/work	/work/cluster	DEEP-EST	BeaGFS	Global	BeaGFS 7.1.2		Work files, system, not backup files. Not meant for permanent data storage	Also available in the SDV but only through 1 Gb network connection	
/noarch	/noarch	DEEP-EST	4fs local partition	Local*				Scratch file system for temporary data. Will be cleaned up after job finished	to use for long for storing temporary files
/home/cluster	/home/cluster	DAM partition	local SSD (NV)	Local*				Scratch file system for temporary data. Will be cleaned up after job finished	"1.5 TB local Optane SSD Data Center (DC) P4800X160GB, FC3x at 2.5T-30T (SFPs)"
/home/cluster	/home/cluster	DAM partition	local SSD (NV)	Local*				Scratch file system for temporary data. Will be cleaned up after job finished	"1.5 TB local Optane SSD Data Center (DC) P4800X160GB, FC3x at 2.5T-30T (SFPs)"
/scratch/cluster	/scratch/cluster	DAM partition	DCPM in replicated mode	Local*		2.2 GB/s stripe size 16-20MB		1 TB in dc-dmep1-10 TB in dc-dmep2-10 TB in dc-dmep3-10 TB in dc-dmep4-10 TB in dc-dmep5-10 TB in dc-dmep6-10 TB in dc-dmep7-10 TB in dc-dmep8-10 TB in dc-dmep9-10 TB in dc-dmep10-10 TB in dc-dmep11-10 TB in dc-dmep12-10 TB in dc-dmep13-10 TB in dc-dmep14-10 TB in dc-dmep15-10 TB in dc-dmep16-10 TB in dc-dmep17-10 TB in dc-dmep18-10 TB in dc-dmep19-10 TB in dc-dmep20-10 TB in dc-dmep21-10 TB in dc-dmep22-10 TB in dc-dmep23-10 TB in dc-dmep24-10 TB in dc-dmep25-10 TB in dc-dmep26-10 TB in dc-dmep27-10 TB in dc-dmep28-10 TB in dc-dmep29-10 TB in dc-dmep30-10 TB in dc-dmep31-10 TB in dc-dmep32-10 TB in dc-dmep33-10 TB in dc-dmep34-10 TB in dc-dmep35-10 TB in dc-dmep36-10 TB in dc-dmep37-10 TB in dc-dmep38-10 TB in dc-dmep39-10 TB in dc-dmep40-10 TB in dc-dmep41-10 TB in dc-dmep42-10 TB in dc-dmep43-10 TB in dc-dmep44-10 TB in dc-dmep45-10 TB in dc-dmep46-10 TB in dc-dmep47-10 TB in dc-dmep48-10 TB in dc-dmep49-10 TB in dc-dmep50-10 TB in dc-dmep51-10 TB in dc-dmep52-10 TB in dc-dmep53-10 TB in dc-dmep54-10 TB in dc-dmep55-10 TB in dc-dmep56-10 TB in dc-dmep57-10 TB in dc-dmep58-10 TB in dc-dmep59-10 TB in dc-dmep60-10 TB in dc-dmep61-10 TB in dc-dmep62-10 TB in dc-dmep63-10 TB in dc-dmep64-10 TB in dc-dmep65-10 TB in dc-dmep66-10 TB in dc-dmep67-10 TB in dc-dmep68-10 TB in dc-dmep69-10 TB in dc-dmep70-10 TB in dc-dmep71-10 TB in dc-dmep72-10 TB in dc-dmep73-10 TB in dc-dmep74-10 TB in dc-dmep75-10 TB in dc-dmep76-10 TB in dc-dmep77-10 TB in dc-dmep78-10 TB in dc-dmep79-10 TB in dc-dmep80-10 TB in dc-dmep81-10 TB in dc-dmep82-10 TB in dc-dmep83-10 TB in dc-dmep84-10 TB in dc-dmep85-10 TB in dc-dmep86-10 TB in dc-dmep87-10 TB in dc-dmep88-10 TB in dc-dmep89-10 TB in dc-dmep90-10 TB in dc-dmep91-10 TB in dc-dmep92-10 TB in dc-dmep93-10 TB in dc-dmep94-10 TB in dc-dmep95-10 TB in dc-dmep96-10 TB in dc-dmep97-10 TB in dc-dmep98-10 TB in dc-dmep99-10 TB in dc-dmep100-10 TB in dc-dmep101-10 TB in dc-dmep102-10 TB in dc-dmep103-10 TB in dc-dmep104-10 TB in dc-dmep105-10 TB in dc-dmep106-10 TB in dc-dmep107-10 TB in dc-dmep108-10 TB in dc-dmep109-10 TB in dc-dmep110-10 TB in dc-dmep111-10 TB in dc-dmep112-10 TB in dc-dmep113-10 TB in dc-dmep114-10 TB in dc-dmep115-10 TB in dc-dmep116-10 TB in dc-dmep117-10 TB in dc-dmep118-10 TB in dc-dmep119-10 TB in dc-dmep120-10 TB in dc-dmep121-10 TB in dc-dmep122-10 TB in dc-dmep123-10 TB in dc-dmep124-10 TB in dc-dmep125-10 TB in dc-dmep126-10 TB in dc-dmep127-10 TB in dc-dmep128-10 TB in dc-dmep129-10 TB in dc-dmep130-10 TB in dc-dmep131-10 TB in dc-dmep132-10 TB in dc-dmep133-10 TB in dc-dmep134-10 TB in dc-dmep135-10 TB in dc-dmep136-10 TB in dc-dmep137-10 TB in dc-dmep138-10 TB in dc-dmep139-10 TB in dc-dmep140-10 TB in dc-dmep141-10 TB in dc-dmep142-10 TB in dc-dmep143-10 TB in dc-dmep144-10 TB in dc-dmep145-10 TB in dc-dmep146-10 TB in dc-dmep147-10 TB in dc-dmep148-10 TB in dc-dmep149-10 TB in dc-dmep150-10 TB in dc-dmep151-10 TB in dc-dmep152-10 TB in dc-dmep153-10 TB in dc-dmep154-10 TB in dc-dmep155-10 TB in dc-dmep156-10 TB in dc-dmep157-10 TB in dc-dmep158-10 TB in dc-dmep159-10 TB in dc-dmep160-10 TB in dc-dmep161-10 TB in dc-dmep162-10 TB in dc-dmep163-10 TB in dc-dmep164-10 TB in dc-dmep165-10 TB in dc-dmep166-10 TB in dc-dmep167-10 TB in dc-dmep168-10 TB in dc-dmep169-10 TB in dc-dmep170-10 TB in dc-dmep171-10 TB in dc-dmep172-10 TB in dc-dmep173-10 TB in dc-dmep174-10 TB in dc-dmep175-10 TB in dc-dmep176-10 TB in dc-dmep177-10 TB in dc-dmep178-10 TB in dc-dmep179-10 TB in dc-dmep180-10 TB in dc-dmep181-10 TB in dc-dmep182-10 TB in dc-dmep183-10 TB in dc-dmep184-10 TB in dc-dmep185-10 TB in dc-dmep186-10 TB in dc-dmep187-10 TB in dc-dmep188-10 TB in dc-dmep189-10 TB in dc-dmep190-10 TB in dc-dmep191-10 TB in dc-dmep192-10 TB in dc-dmep193-10 TB in dc-dmep194-10 TB in dc-dmep195-10 TB in dc-dmep196-10 TB in dc-dmep197-10 TB in dc-dmep198-10 TB in dc-dmep199-10 TB in dc-dmep200-10 TB in dc-dmep201-10 TB in dc-dmep202-10 TB in dc-dmep203-10 TB in dc-dmep204-10 TB in dc-dmep205-10 TB in dc-dmep206-10 TB in dc-dmep207-10 TB in dc-dmep208-10 TB in dc-dmep209-10 TB in dc-dmep210-10 TB in dc-dmep211-10 TB in dc-dmep212-10 TB in dc-dmep213-10 TB in dc-dmep214-10 TB in dc-dmep215-10 TB in dc-dmep216-10 TB in dc-dmep217-10 TB in dc-dmep218-10 TB in dc-dmep219-10 TB in dc-dmep220-10 TB in dc-dmep221-10 TB in dc-dmep222-10 TB in dc-dmep223-10 TB in dc-dmep224-10 TB in dc-dmep225-10 TB in dc-dmep226-10 TB in dc-dmep227-10 TB in dc-dmep228-10 TB in dc-dmep229-10 TB in dc-dmep230-10 TB in dc-dmep231-10 TB in dc-dmep232-10 TB in dc-dmep233-10 TB in dc-dmep234-10 TB in dc-dmep235-10 TB in dc-dmep236-10 TB in dc-dmep237-10 TB in dc-dmep238-10 TB in dc-dmep239-10 TB in dc-dmep240-10 TB in dc-dmep241-10 TB in dc-dmep242-10 TB in dc-dmep243-10 TB in dc-dmep244-10 TB in dc-dmep245-10 TB in dc-dmep246-10 TB in dc-dmep247-10 TB in dc-dmep248-10 TB in dc-dmep249-10 TB in dc-dmep250-10 TB in dc-dmep251-10 TB in dc-dmep252-10 TB in dc-dmep253-10 TB in dc-dmep254-10 TB in dc-dmep255-10 TB in dc-dmep256-10 TB in dc-dmep257-10 TB in dc-dmep258-10 TB in dc-dmep259-10 TB in dc-dmep260-10 TB in dc-dmep261-10 TB in dc-dmep262-10 TB in dc-dmep263-10 TB in dc-dmep264-10 TB in dc-dmep265-10 TB in dc-dmep266-10 TB in dc-dmep267-10 TB in dc-dmep268-10 TB in dc-dmep269-10 TB in dc-dmep270-10 TB in dc-dmep271-10 TB in dc-dmep272-10 TB in dc-dmep273-10 TB in dc-dmep274-10 TB in dc-dmep275-10 TB in dc-dmep276-10 TB in dc-dmep277-10 TB in dc-dmep278-10 TB in dc-dmep279-10 TB in dc-dmep280-10 TB in dc-dmep281-10 TB in dc-dmep282-10 TB in dc-dmep283-10 TB in dc-dmep284-10 TB in dc-dmep285-	
/data/work	/data/work/cluster	SDV, deeper via EXT4L, 4fs, NVMe via GbE (NV), DEEP-EST (7 GbE only)	BeaGFS	Global	BeaGFS 7.1.2	Type RAID, Character, SIOK, Number of storage targets, Number of storage devices	180T GB, 1600 write, 1500 read, 1500 write, 1511 read, 1500 read		

## 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 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](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
```