Wikiprint Book

Title: File Systems

Subject: DEEP - Public/User_Guide/Filesystems

Version: 36

Date: 01.05.2024 15:52:39

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 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:

	Day on						Maximum			1
Mount Point	Over can selfatead tallium	Cluster	Турн	Chahal / Local	EW Sersion	Stripe Fatters Details	Measured Performance (see footnoise)	Description	Other	
phone	phonepase	Sir	CPPS exported via NPS	Chang				JUST GPPS Home directory, used only for configuration Size.		
ponjesi	a project das	anv. Dear-ear	GPFS exposted via NPS	Chang				JUST GPPE Propest directory; GPPE main: shouge the system; exit substite performance extensive applications.		
lank	antilities	login radio orly (despx)	GPFS exposited via	Chiled				or heredenication of AUST GPP2 Autorities of	If you plan to use the author, please get in content to the system administrators (e.g. via the support mading but). You'der information and find further information and the plan with the plan with the plan with the plan with the plan with the plan with the plan with the plan with the plan with the plan with the plan with the plan with the plan with the plan with the plan with	
heart	hearth/calmeg	DEEP-RET	Bed73	China	BedFS 712			Shark the system, no backup, hence not meant for permanent	"Xino available in the 32% but only Brough 1 Gig network connection	
Acceptab	access.	DREP-EST	alle boosel partition					data storage Suratio file spatem for temporary alterpole foruses ("Recommend to use instead of ding for temporary these temporary these		
invine/social-in		DAM partition	lacel SSD (als)	Louir				Straigh file spales for temporary date. Will be cleaned up after pile front-self "Lis Contex City (CC) PERSON (CV) FERSON (CV)		
investigation.		DAM partition	land SSD (mill)	Local				Sunside Size of the spales for temporary date. Will be channed up after you flowing (*1.5 TB load) (*2.5 TB loa		
-		DAM patition	DCPMM in applicat made	Local			2.2 CBNs simple did lesi in sip-danici		"3 TE in alp-dane(01, 52 and only dane(01, 52 and o	
Mirwish	Andrews Andrews	SDV (Stephen sår ESTCL), SSCERR ESTCL) CERPEST (1 CAR only).	Bedří	Chaine	Bee078 712	Type KAIDO Charladae 812K, Number of shringe larges desired d	1831.85 Mills wife, 1338.43 Mills read 1885 read 1886, 111 spak semour*	Work the spities, as backer, hence and need for permanent data stronge	Trediments and parameters used through the parameters used through the parameters of	deng-ner/selv-kennikasaka/synthetid/iden deng-ner/selv-kennikasaka/synthetid/idenda/id
livin	discheding:	sov	NO file device	Local	ResOFE 7.1.2	Black size of	T145 MBN wile, 2358 MBN med T39165 cpch special country of the table country of the table country of the table country of tab	1 NVMIn device available at each 30°C congula code	"Bedineading and parameters, used shired in JUBB: some bilance 6 nd /war /1 nm 1 same bilance 6 yelled 5 yelled 6 yelled	deng-se (sdir-kendinanka/aysi kiri.k/) deng-se (sdir-kendinanka/aysi kiri.k/)
invibrend	milleend	30V	BeedPS On Deniald surrang on the NOMe	Local	ResOFE 7.1.2	Block size: 812K	T130 MMNs with 2657 MMs mad 12311 opin to case on the case of the	1 BeeCHD Indicate surroug on each NUMe device	tear many (f. mi /war / Inna I war dinay f. yaha? results lanchmarks and patanaha and patanaha in Jilliti war dinay f. mi yaar / Inna I war dinay f. mi yaar / Inna I war dinay f. mi yaar / Inna I war dinay f. yaba2 erault hanobmarka erault hanobmarka erault hanobmarka f. yaba2 erault hanobmarka	dweg an (after kansdinarika) igyal kirik ke (dweg an (after kansdinarika) igyal kirik ke (

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
```