Wikiprint Book

Title: File Systems

Subject: DEEP - Public/User_Guide/Filesystems

Version: 36

Date: 19.04.2025 17:17:45

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 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

Baueri Parint	User can	Challer	Туре	COURSE!		Stripe	Maximum Measured Performance (see			
	User can write/read toffrom	Charac	Туро	Clubal / Local	SW Tersion	Stripe Pattern Details	Performance (see (see)	Description	COLAN	
-							fournoses)	AGT GPF3	-	
phone phon	D.Doonalises	Sir.	OPFS	Chana				Huma drectory; used only for configuration files.		
	the sections	SEEP-Lut	eponerva MFS					word only for configuration		
								JUST GPPS Proped deedlay; GPPS main stronge the system; not suitable for		
								CPF3 man		
	p)ropestP*	W.	cerro exported via NPS					storage the system;		
- September				Shaw				nor substate for		
								petomanoe		
								applications		
								performance retevant applications or benchmarking		
								AGT GPFS	f you plan to	
								directory;	anziore,	
								storage	cortact to	
			OPFE equated via MFE					data not	administration	
	000000000	Sign rade any (deeps)						used in a ting time:	(e.g. via the support	
_	au or process	only (deepu)						Data regused to	making but. You can find	
								tipe - not consted to	turber eterminer	
								AUST OFFEE Author dealory; Long-team shouse substantive data not used in a long-team; Data sing-team; Data ting-team; Data tin	If you plan to use the authors, and the authors and to the submitted to th	
								Recovery	using the	
								days.	300	
\neg								Walk Sile		
	hondPROJ	DESP-137	BeeGFS	COURSE	BeeSFS			backup.		
in.	NonSPROJ	CHEP-EST		Chana	7.1.2			means for		
								Work file system, me backup, tense not meant for permanent data storage		
-								Ė	Due to offerent retrieval. someolivity to the APSM stronge the performance regist offer on the offerent modules.	
			BeeGF3					Work file	network.	
		sev.			BeeSFS			System, no backup,	to the AFSM	
No.	Man	SDV. DEEP-EST	BeeGF3	Chana	8mc9F3 7.1.2			Work the system, no backup, tenso not meant for permanent data storage	storage the performance	
								personers data stora	regreative	
									attenus modules	
-	_	<u> </u>				<u> </u>	_	Scient to sydem for temporary data. Will be cleaned up after job Snigherall "Reconstrated to use	-	
			alls tocall partition					system for temporary		
	Australia	DEEP-EST						data William		
								atter job		
one				Louis				Resources		
								Distance of		
								Sing for storing		
								to use repeal of top for soring semporary ties		
	_	\vdash		-		\vdash	-	Scotto Se		
								system for temporary		
								data WE be cleaned up		
								ater you tropped ***		
vmehoons	Dymencon?	DAM	Sect 230 (49)	Local				TR YOU		
		patton			1	1		Data Cwiav		
		patton	100							
		patton	124					POSIN POSIN		
		patton						PCINS NS. 2.67, 3D		
		patton						PCING PARKET PCING MI, 2.87, 3D 3Panel		
		patton						pt/de pt/de PCB3 st, 2.57, 3D 3Pant() Scanh the system for		
		patton	144					pt/site PCIO3 xt, 2.87, 3D 3Punit! Scumh the system for temporary data. Will be		
								(DC) Politics (PCNd) x4, 237, 3D 3Punit) Screen for langurary data, Will be cleaned up after job		
wmetucated:	Divineholation			Louir				(DC) Policies (PATRIE PCIAS AL 237, 3D 3Punit) Screen for sequent for sequency data. Will be cleaned up after job Screen of T. S Till your		
emeticana);	Divine husball	DAM partition	Secritor Secritor	Lear				(DC) Fellows (NYMe PCBA's K, 237, 3D 3Punit) Suran the lenjorary data Will be cleaned up after job Snahes) "1.5 TS Soot Optare SSD Data Center		
**********	Shvite hurand			Lear				(DC) Policion (NYMB PCInC M, 237, 3D 3Pund) South Tib system for lemporary data. Wil be shared up after job finales of "1.8 "It your Cylone SSD Data Center (DC) Policion (NYMB		
m+1,03030	Sharma hairiidad			Louir				DCF PRIOX POME TO POME		
en et caso X	Shvina husabil			Lear				Scottish Ste- system for system for system for system for sweepings data, VMI for shoulded Type data you for sold that Contain (PCT PERSON PY-Mo Security System Security Syst		
matcasis.	Startin harabil			Loor					73 Till IN 49-84490 JUI	
ematucano.sc	Street, his said.			Loor					73.7% in 4p-daught.pg 2.7% in 4p-daugh-105	
metanak	Situation had statute			Louis					"2 TB In dy-danger, co 2 TB In dy-danger to lene Cytime	
mehoanX	Shrene hirakili	DAM partition	Social SSD (excit)	Lesser					"3 TB in dy-danglo, I/2 3 TB in dy-danglo, I/2 3 TB in left Cytains Dr. Personnel To Honorous Committee Cytains and Cytains an	
emericanic N	Seven-haraba	DAM partition	Social SSD (excit)	Louir					"2 TR si di-dango (2 T R) di-dango (3 T R) di-dango (4 T R) di-dango (4 T R) di-dango (4 T R) di-dango (4 T R)	
emarkusandi:	Division bucklished			Local			2.2 GBIN SWIPP AD WATER 49-SANCE		dp-dand01,02 2 TB in dp-dand03 16 blet Optione DC Personers Memory (DCPMM) 20008	
errantussisti.	Division buckshild	DAM partition	Social SSD (excit)	Laur Laur					dp-dand01,02 2 TB in dp-dand03 16 blet Optione DC Personers Memory (DCPMM) 20008	
mekussadi.	Sharina harishidi Jamenukasahid	DAM partition	Social SSD (excit)	Laur Laur					dp-dand01,02 2 TB in dp-dand03 16 blet Optione DC Personers Memory (DCPMM) 20008	
mehlussist.	Theree backed d	DAM partition	Social SSD (excit)	Local*					dy-Bangon, cz 2 TB in dy-Bangon 16 Boel Optaine DC Persosaut Mansay (DCPMB) 20608 DMMN Based on Boel'n 3D 3Phas son-volable mentaly	
ere erkussistä.	Divini Access	DAM partition	Social SSD (excit)	Lease"					dy-Bangon, cz 2 TB in dy-Bangon 16 Boel Optaine DC Persosaut Mansay (DCPMB) 20608 DMMN Based on Boel'n 3D 3Phas son-volable mentaly	
ere ere la constitución de la co	Divine Accided	DAM partition	Social SSD (excit)	Lesse					dy-Bangon, cz 2 TB in dy-Bangon 16 Boel Optaine DC Persosaut Mansay (DCPMB) 20608 DMMN Based on Boel'n 3D 3Phas son-volable mentaly	
HINANGOOD	The weeks could be a second of the second of	DAM partition	Social SSD (excit)	Lastr					dy-Bangon, cz 2 TB in dy-Bangon 16 Boel Optaine DC Persosaut Mansay (DCPMB) 20608 DMMN Based on Boel'n 3D 3Phas son-volable mentaly	
ere er kussisti.	Divine located	DAM partition	Social SSD (excit)	Lesse			2.7 data simple as less in q-dased		dy-Bangon, cz 2 TB in dy-Bangon 16 Boel Optaine DC Persosaut Mansay (DCPMB) 20608 DMMN Based on Boel'n 3D 3Phas son-volable mentaly	
ere erkussistät.	Divini historia	DAM partition	Social SSD (excit)	Lear			2.7 data simple as less in q-dased		dy-Bangon, cz 2 TB in dy-Bangon 16 Boel Optaine DC Persosaut Mansay (DCPMB) 20608 DMMN Based on Boel'n 3D 3Phas son-volable mentaly	
meeticalast	The weeks could be a second of the second of	DAM partition	Social SSD (excit)	Lesser			2.7 data simple as less in q-dased		dy-Bangon, cz 2 TB in dy-Bangon 16 Boel Optaine DC Persosaut Mansay (DCPMB) 20608 DMMN Based on Boel'n 3D 3Phas son-volable seenby	alang was spale samulanan ku syang sa
mension	printerland	DAM partition	Social SSD (excit)	Lear-	Brod/93	Black Size: Of	2.7 data simple as less in q-dased		dy-Bangon, cz 2 TB in dy-Bangon 16 Boel Optaine DC Persosaut Mansay (DCPMB) 20608 DMMN Based on Boel'n 3D 3Phas son-volable seenby	denge se sinde standardina kan specific
senikojosti	general cond	DAM partition	Social SSD (excit)	Later	BardF3 7.12	Elich Size:	2.7 data simple as less in q-dased		49-Take(\$1.50) 49-Take(\$1.50) 50-Cipane 50-Cip	lang or later humanakan specia
mentional Control	финальная	DAM partition	Social SSD (excit)	Leon*	Ber073 7.12	Risk Six of			49-Take(\$1.50) 49-Take(\$1.50) 50-Cipane 50-Cip	lange or inter-leaning lands a
ennikussus	principal de la constante de l	DAM partition	Social SSD (excit)	Later	Bee073	Elick Size	2.7 data simple as less in q-dased		49-Take(\$1.50) 49-Take(\$1.50) 50-Cipane 50-Cip	
em Ahusala.X	рененализата до предостава на	DAM partition	Social SSD (excit)	Lease"	BardFS 7.13	Blok Gar of	2.7 data simple as less in q-dased		dy-Bangon, cz 2 TB in dy-Bangon 16 Boel Optaine DC Persosaut Mansay (DCPMB) 20608 DMMN Based on Boel'n 3D 3Phas son-volable seenby	

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
```