Wikiprint Book

Title: File Systems

Subject: DEEP - Public/User_Guide/Filesystems

Version: 36

Date: 29.04.2024 13:40:41

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 (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. 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

Name Pains	ilar san artistraal artism	Course	Type	-	SW tenier	Stripe Pattern Strain		Generation	Citiza	
			-	_		Details		L.		
phone	phonoloon	Eller.	GPFS expended da NFS	Grand				riana dantery materially for		
		_		_		_		day.	_	
								Project descript		
			GMS					strate in		
e)-reject	priprojent Spro	Siraer Siraer	experted de N/S	Grand				or substa		
								rate and		
								-		
								ANT GATE Antice	if you plan to one the	
								Long-term strange	man prin	
								date or	na spilan administration	
-	and temper	ingin nada miy (Magra)	GATS equator de	Grand				ing inc	matery late.	
								mpaces conductor	unter common	
								Gen.	and same lank on	
								ranchalas days		
								State Size	rides eradelità in	
and .	and broken	DEEP-SET	Bearing S	Orași	5m075 1142			Tanana Tanana	mig through 1 Gig	
								nament nament	union mountain	
								Companie (in system for		
								data Will be described up		
arash.		0007-007	eta bend perilina	-				oter jan marken) Terretori	L	

_										
	-	\vdash	-	-	-	-	-	Gen Screent Se	-	
								separate by series will be		
				1				alter job		
	nematurant	DAN partition	terminal SEED (MA)	-				Skinst Open 650		
								Gene Gener (GC) Pelason (GC) Pelason		
								POMENA JULY NO		
_	_	\vdash		-				APatro) Smant la	-	-
								opposed for semposely side. Will have shared top other job. Colonia (** 1 d.) Colonia (** 1 d.) Coloni		
								charactup after job		
	nematurant	netten	(mark)	-				Skinst Open 650		
								Gene Gener (GC) Pelason (GC) Pelason		
								PCM0 s4, 240,10		
	-	\vdash		-	-	-	-	nhan)	194	
									arangona o Tibin arangona	
									Distriction of the Control of the Co	
	manhorati	CAN	DOMESTA applicant made				22 GBm ample 66 seat in disclared?		Manay (CCPMA)	
		·	- and				dydamir		CMA.	
									intelle III Iillian	
									-anay mining	
								State Sia spatient, na		ĺ
				1					L	
				1				income.		
				1				fee SCOAL	- ARE	
		SOV Shoper sets		1		Type RADIA	1821.BS	n markets n markets migrations		l
		ICO phospar sale contact on EXTELL, antique on Galf enty, DEEP dEET or Galf enty,	Seed/S	Orași.	5m075 110	Type RADII, Oranbaine 0-24, Number of straigs targets desired if	1501.65 Billio sette, 1308.60 Billio seal 1500 spalo mass, 5711 spalo	marker (- Jakes	mag. ar salti handmada siyedar id
		Galf and C			Γ~	arapa arapa	man, D11	manusian, manusian	in.	
		r Gall anigs					-	Carriero C		
				1				out ten	=	magican radio handinasha rajeshan ist
								Amendal professional profession		
_	-	\vdash		\vdash	\vdash	-	\vdash	partitions).	Tara sasah	
				1						
inere.				1					- 186	
				1			THE SERVICE	11000	-	Assept our radio describerarios repressor in
	nemarkey	sov.	V Alle desire	-	8m075 114	Stant size:	r tali Million unite, 2013 Million seasi 1 30 million specia mession 62500° appart sementar	r some	1 Janes	
				1			6260 mais	=-		
				1						turn an indi denderada is
				1					T Table	
		\vdash		_	_	_	_	_	nanin nantaurtu	
									and common	
_									- AME	
			1	1			+OLMBA			
								1 See Child	ner inni	Angust sale describering appring in
			Basil/E On Damand	L	See CFS	First size	Million wast	-		
nethaund		acc	Base GPE-On Demonst Serving on the NORth	-	8m075 3142	Black size 0 24	COST man Costs Costs Costs Costs Costs	manager man Wilde derive	t jakes manin	
mahaana			Basil/S On Damand naming on the Nichba	-	8m675 312	Ries son 0.24	r CSC MARRIS write, 2640° Million want r 2011 agrain create, 1 Million agrain writerant*	an Ville	and the part of th	
nthaoni		es.	Basel/15-On Garmani natring on the Nickle	-	Amadria Trial	0.24 0.24	Millio mad 12011 spain made, 18031 spain	namen n namen n namen namen namen namen namen namen namen namen namen na	T Salar Control of Control o	Bang-en radio Janoinsantia regimberi kri

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
```