Space Organization in DDM groups/user space

Ueda I.

2011.02.02.

Groupdisk

Number of sites for each group space

Group	#	Group	#
DATAPREP	1	PHYS-BEAUTY	6
DET-INDET	3	PHYS-EXOTICS	8
DET-LARG	1	PHYS-GENER	3
DET-MUON	3	PHYS-HI	4
DET-TILE	1	PHYS-HIGGS	10
PERF-EGAMMA	8	PHYS-SM	18
PERF-FLAVTAG	6	PHYS-SUSY	11
PERF-IDTRACKING	5	PHYS-TOP	14
PERF-JETS	10	PROJ-SIT	1
PERF-MUONS	9	SOFT-SIMUL	1
PERF-TAU	5	TRIG-DAQ	7

Groupdisk

The Facts:

- Groups have bits and pieces of spaces (eg. 10 x 50 TB)
- Groups manage each of the many spaces with each quota
- Groups define group tasks specifying the destination(s)
- Problems in accessing data placed at a single site in downtime

Wishes from group space/production managers:

- Less number of spaces to manage
  - 2x 200 TB + 1x 100TB (or 1x 500 TB) rather than 10x 50 TB
- or, Auto-selection of destination
  - rather than checking each free space and changing the destination manually
- Auto-triggering of additional replicas of datasets used often

Some sub-groups would like to have their group spaces at their local sites

What we can/should offer?

Global quota over “Tier-1s”

- each T1 (potentially) hosts every group
- each group has a global quota over the Tier-1s rather than a quota at each Tier-1
- ddm chooses the destination Tier-1 (not the groups)

same for “Tier-2s” ?

- depends on the policies on the usage of Tier-2 disks

Quota at “specific” Tier-2s

- for some (sub-)groups with “local” access to the sites.
- groups choose the destination (no automation in ddm)
- less global quota (over Tier-1s)

same for Tier-1s ?

- local Analysis Facility at some Tier-1(s)...

Option to specify "where" and "how many" replicas

- "where" = {T1s, T2s, specific sites}

What we can/should offer?

PD2P for group data

- when the data is at Tier-1s.
  - not the Tier-2s?
- confined within the cloud?
- global quota per group over “Tier-2s”?
  - or let the popularity decide the share?
- overall quota on sum of group data to protect data/mc spaces?
  - or let the popularity decide the share?

What we would need

DDM endpoints = GROUPDISK

- (N-T1s + N-T2s) x N-Groups = too many
- sub-spaces with quota = new development in ddm

PD2P for group data

- treat AP and GP separately or not?
  - how to treat "group*" data (= without GP flag)
  - pathena should put GP flag for "group*" data ?
-

User spaces

User Space

- ScratchDisk
- LocalGroupDisk

Global user quota?

- For users occupying large space at many sites
  - declined at a S&C workshop
- LocalGroupDisk within a country?
- Less management load than quota at each site

Sub-Space for users?

- Probably useful in managing localgroupdisk

Groupdisk

http://bourricot.cern.ch/dq2/accounting/group_reports2/

Group	Site	Cloud	Used(TB)	Booked(TB)
PERF-EGAMMA	BNL-OSG2_PERF-EGAMMA	USASITES	0.7	27.49
	CERN-PROD_PERF-EGAMMA	CERN	7.03	11.0
	CSCS-LCG2_PERF-EGAMMA	FZKSITES	17.1	54.98
	FZK-LCG2_PERF-EGAMMA	FZKSITES	25.94	27.49
	IN2P3-CC_PERF-EGAMMA	FRANCESITES	34.6	54.98
	INFN-MILANO-ATLASC_PERF-EGAMMA	ITALYSITES	0.68	27.49
	SWT2_CPB_PERF-EGAMMA	USASITES	42.74	54.98
	UKI-NORTHGRID-LIV-HEP_PERF-EGAMMA	UKSITES	0	27.49
	TOTAL	-	128.79	285.9

Group	Site	Cloud	Used(TB)	Booked(TB)
PERF-JETS	BNL-OSG2_PERF-JETS	USASITES	13.18	27.49
	CERN-PROD_PERF-JETS	CERN	13.8	21.99
	MPPMU_PERF-JETS	FZKSITES	0	27.49
	MWT2_UC_PERF-JETS	USASITES	0	27.49
	NCG-INGRID-PT_PERF-JETS	SPAINSITES	22.19	27.49
	RU-PROTVINO-IHEP_PERF-JETS	NLSITES	24.05	27.49
	SLACXRD_PERF-JETS	USASITES	14.3	27.49
	TOKYO-LCG2_PERF-JETS	FRANCESITES	5.98	54.98
	TRIUMF-LCG2_PERF-JETS	CANADASITES	42.25	54.98
	UKI-LT2-QMUL_PERF-JETS	UKSITES	26.14	27.49
	TOTAL	-	161.89	324.38

Groupdisk

Group	Site	Cloud	Used(TB)	Booked(TB)
PHYS-TOP	CA-SCINET-T2_PHYS-TOP	CANADASITES	41.47	54.98
	CERN-PROD_PHYS-TOP	CERN	7.72	11.0
	DESY-ZN_PHYS-TOP	FZKSITES	42.12	54.98
	GRIF-IRFU_PHYS-TOP	FRANCESITES	37.1	54.98
	IFAE_PHYS-TOP	SPAINSITES	28.48	54.98
	IN2P3-LPC_PHYS-TOP	FRANCESITES	28.51	49.48
	INFN-ROMA1_PHYS-TOP	ITALYSITES	23.55	27.49
	NDGF-T1_PHYS-TOP	NDGF	24.49	54.98
	NET2_PHYS-TOP	USASITES	35.35	54.98
	NIKHEF-ELPROD_PHYS-TOP	NLSITES	43.66	54.98
	PRAGUELCG2_PHYS-TOP	FZKSITES	23.65	54.98
	SWT2_CPB_PHYS-TOP	USASITES	33.41	54.98
	UKI-LT2-QMUL_PHYS-TOP	UKSITES	26.26	54.98
	WEIZMANN-LCG2_PHYS-TOP	NLSITES	26.47	54.98
	TOTAL	-	422.24	692.75

Groupdisk

Group	Site	Cloud	Used(TB)	Booked(TB)
PHYS-SM	AGLT2_PHYS-SM	USASITES	51.57	54.98
	AUSTRALIA-ATLAS_PHYS-SM	CANADASITES	0	27.49
	BNL-OSG2_PHYS-SM	USASITES	21.7	27.49
	CERN-PROD_PHYS-SM	CERN	23.2	21.99
	DESY-HH_PHYS-SM	FZKSITES	42.45	54.98
	GRIF-LPNHE_PHYS-SM	FRANCESITES	18.66	27.49
	IL-TAU-HEP_PHYS-SM	NLSITES	0	27.49
	IN2P3-LAPP_PHYS-SM	FRANCESITES	13.87	27.49
	INFN-MILANO-ATLASC_PHYS-SM	ITALYSITES	22.88	27.49
	JINR-LCG2_PHYS-SM	NLSITES	11.6	21.99
	LRZ-LMU_PHYS-SM	FZKSITES	11.63	27.49
	NDGF-T1_PHYS-SM	NDGF	3.71	16.49
	NIKHEF-ELPROD_PHYS-SM	NLSITES	33.22	54.98
	PIC_PHYS-SM	SPAINSITES	12.79	27.49
	SFU-LCG2_PHYS-SM	CANADASITES	18.39	27.49
	SLACXRD_PHYS-SM	USASITES	15.65	27.49
	TW-FTT_PHYS-SM	TAIWANSITES	0	5.5
	UKI-SCOTGRID-GLASGOW_PHYS-SM	UKSITES	25.44	54.98
	TOTAL	-	326.76	560.79