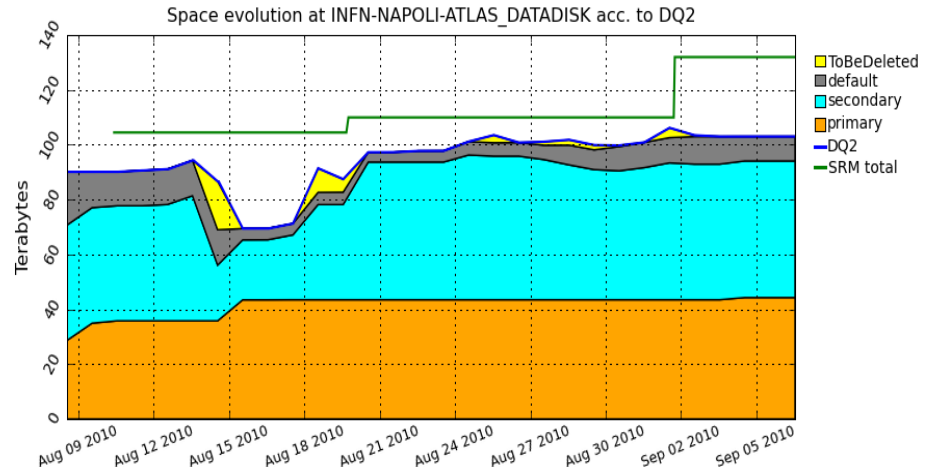# Disk Space publication

Simone Campana

Fernando Barreiro

Wahid Bhimji

- Storage space and usage information collected by different agents and presented in a web frontend
  - Volumes of space, datasets and files registered at sites according to DDM Catalogues
    - possibility to break down volumes by metadata
    - information retrieved once a day
  - Total and Used disk space according to the Storage and retrieved via SRM
    - Information retrieved every hour
  - BDII information
    - Information retrieved twice a day

CERN IT Department

- Total Space Info
 - Used to trigger deletion of redundant replicas in case disk is full

Space evolution at INFN-NAPOLI-ATLAS_DATADISK acc. to DQ2

Legend: ToBeDeleted, default, secondary, primary, DQ2, SRM total

- Free Space Info
  - Used to automatically stop replication of data to full sites

CLOUD	SITE	STATUS							
		f		r		u		w	
None								manual	Reason
CERN	CERN-PROD	manual	Reason	manual	Reason	manual	Reason	manual	Reason
FRANCESITES	GRIF-LAL					DISKSPACE	Reason	DISKSPACE	Reason
	RO-02-NIPNE	manual	Reason					manual	Reason
FZKSITES	HEPHY-UIBK					manual	Reason	manual	Reason
	MAIGRID					DISKSPACE	Reason	DISKSPACE	Reason
ITALYSITES	INFN-FRASCATI					DISKSPACE	Reason	DISKSPACE	Reason
NLSITES	RU-MOSCOW-MEPHI-LCG2	manual	Reason	manual	Reason	manual	Reason	manual	Reason
SPAINSITES	NCG-INGRID-PT					DISKSPACE	Reason	DISKSPACE	Reason
UKSITES	UKI-NORTHGRID-LIV-HEP					DISKSPACE	Reason	DISKSPACE	Reason

There is no issue with all this

Used disk space for INFN-NAPOLI-ATLAS_DATADISK
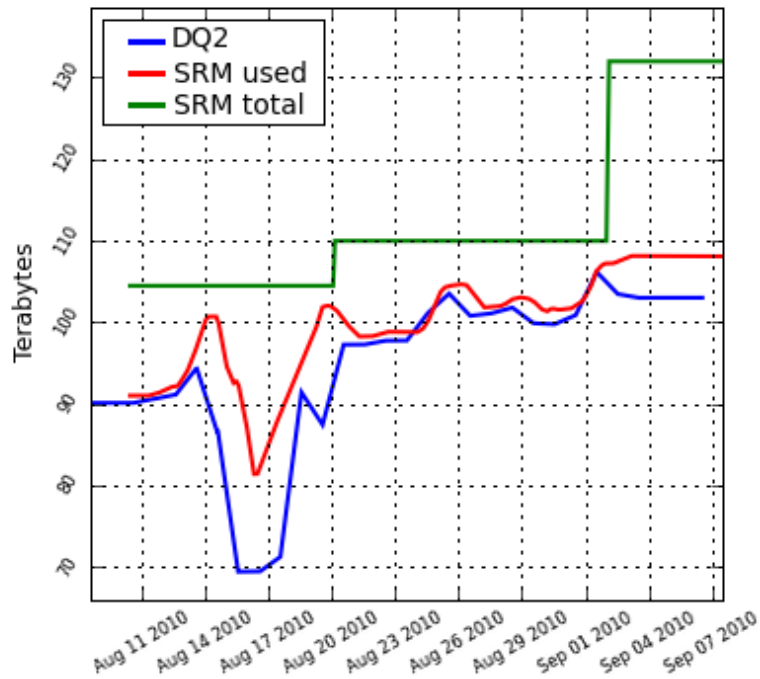
- **Necessary for Consistency Checks**

- **Compare**
  - Used Space as known by DDM (blue)
  - Used Space as advertized by the site (red)

Here comes the issue

Example for CERN-PROD_DATADISK

From SRM

```
[atlddm16] /tmp > lcg-stmd -b -e httpg://srm-
atlas.cern.ch:8443/srm/managerv2 -s ATLASDATADISK
+ Space Token: atlas:ATLASDATADISK
    - Owner: (null)
    - Total Size: 956661451997184
    - Guaranteed Size: 956661451997184
    - Unused Size: 351091464821760
    - Assigned Lifetime: 0
    - Left Lifetime: -1
    - Retention Policy: OUTPUT
    - Access Latency: UNKNOWN
```

Total space considers **both online and offline** disks

Unused (FREE) space considers **only online** disks

Calculating Used=Total-Unused is considering the offline disks as used

Used disk space for CERN-PROD_DATADISK
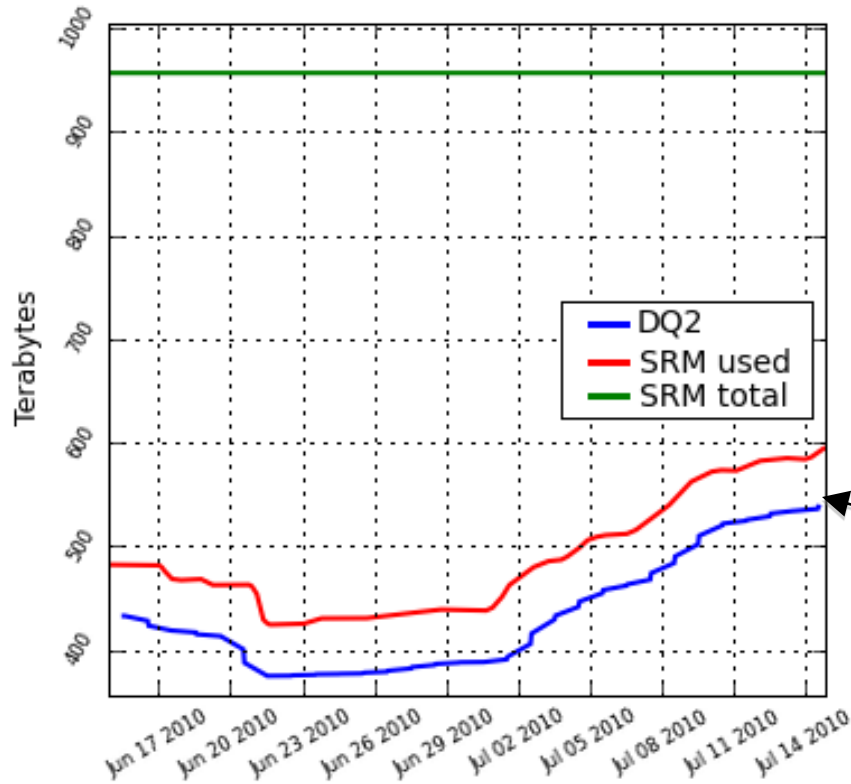
DQ2 confirms the value obtained from the pool dump

Calculating the used size using a CASTOR pooldump: Used=553TB, this means ~50TB difference to the SRM value

- ATLAS would like to know how much data is stored in TxD1 space tokens
  - Regardless the fact that they are accessible or not
- If this is not possible, at least it would be good to know the size of the offline disks
  - So that one can compare if the DQ2vsStorage discrepancy is compatible with it
- Info can be "quasi static"
  - Once per day would do

- According to the specifications, the information provided by Information System is:

  - **Total: Total space available at a given moment.** *"It SHOULD not include broken disk servers, draining servers, etc."*

  - Used: Space occupied by available and accessible files

  - Guaranteed

  - Free: Total-Used

  - **Installed: Size of the physical space of a storage area**

CERN**IT** Department

CATEGORIES

SITE	SPACETOKEN	SRM free	BDII free	SRM used	BDII used	SRM total	BDII total	BDII guaranteed	BDII installed
AUSTRALIA-ATLAS	DATADISK	57.3	57.3	25.2	25.2	82.5	82.5	82.5	82.5
	GROUPDISK	66.0	66.0	0.0	0.0	66.0	66.0	66.0	66.0
	HOTDISK	1.7	1.7	0.5	0.5	2.2	2.2	2.2	2.2
	LOCALGROUPDISK	14.5	14.5	7.5	7.5	22.0	22.0	22.0	22.0
	MCDISK	21.1	21.1	58.1	58.1	79.2	79.2	79.2	79.2
	PRODDISK	8.2	8.2	0.6	0.6	8.8	8.8	8.8	8.8
	SCRATCHDISK	17.5	17.5	4.5	4.5	22.0	22.0	22.0	22.0
CA-ALBERTA-WESTGRID-T2	DATADISK	13.8	None	16.2	None	30.0	None	None	None
	GROUPDISK	24.4	None	0.6	None	25.0	None	None	None
	HOTDISK	0.7	None	0.3	None	1.0	None	None	None
	LOCALGROUPDISK	0.5	None	0.0	None	0.5	None	None	None
	MCDISK	0.5	None	0.0	None	0.5	None	None	None
	PRODDISK	0.7	None	0.3	None	1.0	None	None	None
	SCRATCHDISK	0.4	None	1.5	None	1.9	None	None	None

Example of site publishing correct results: SRM and BDII values are consistent

Example of site not publishing BDII values at all

…

NLSIT

SITE	SPACETOKEN	SRM free	BDII free	SRM used	BDII used	SRM total	BDII total	BDII guaranteed	BDII installed
NIKHEF-ELPROD	DATADISK	216.1	172.1	379.8	379.8	595.9	552.0	595.9	595.9
	GROUPDISK	115.2	71.3	33.9	33.8	149.1	105.1	149.1	149.1
	HOTDISK	10.3	-33.7	0.7	0.7	11.0	0.7	11.0	11.0
	LOCALGROUPDISK	0.5	0.5	9.4	9.4	9.9	9.9	9.9	9.9
	MCDISK	0.0	-44.0	0.0	0.0	0.0	0.0	0.0	0.0
	PRODDISK	20.2	-23.8	1.8	1.8	22.0	1.8	22.0	22.0
	SCRATCHDISK	9.0	-35.1	13.0	13.1	22.0	13.1	22.0	22.0

Sites publishing wrong or negative values for part of their spacetokens…

Disk Space in the IS

CERN

Unfortunately, BDII publishes the same value for total and installed

SITE	SPACETOKEN	SRM free	BDII free	SRM used	BDII used	SRM total	BDII total	BDII guaranteed	BDII installed
CERN-PROD	DATADISK	351.2	357.4	605.5	595.1	956.7	952.6	952.6	952.6
	DATATAPE	65.8	65.8	583.9	581.9	649.7	647.7	647.7	647.7
	GROUPDISK	85.7	85.7	135.4	135.4	221.1	221.1	221.1	221.1
	LOCALGROUPDISK	118.9	119.5	89.1	88.5	208.0	208.0	208.0	208.0
	MCDISK	497.0	497.9	437.4	426.6	934.4	924.5	924.5	924.5
	MCTAPE	10.0	10.0	98.0	88.0	108.0	98.0	0.0	98.0
	SCRATCHDISK	63.9	67.4	49.1	45.5	113.0	113.0	113.0	113.0
	SPECIALDISK	19.4	19.4	1.6	1.6	21.0	21.0	21.0	21.0

- http://www2.ph.ed.ac.uk/~wbhimji/SRMMonitoringAtlas/

  (click on "Space usage table")

- **Big** advantages of this view:

  - This table includes information about storage type and release

  - The columns are sortable – This really helps to track down issues

- Back to requirements:
  - ATLAS would like to know how much data is stored in TxD1 space tokens
  - At least it would be good to know the size of the offline disks
- At the moment this seems not possible via SRM or BDII
- The BDII seems to offer more flexibility
  - But it needs to be agreed which info should be published and work on the information providers
- What is the right forum?