Archival Project & Gratia

Marco Mambelli – marcom@fnal.gov
Dmitry Litvintsev, Tanya Levshina
Stakeholder meeting 6/25/2014
Fermilab

Desired outcome

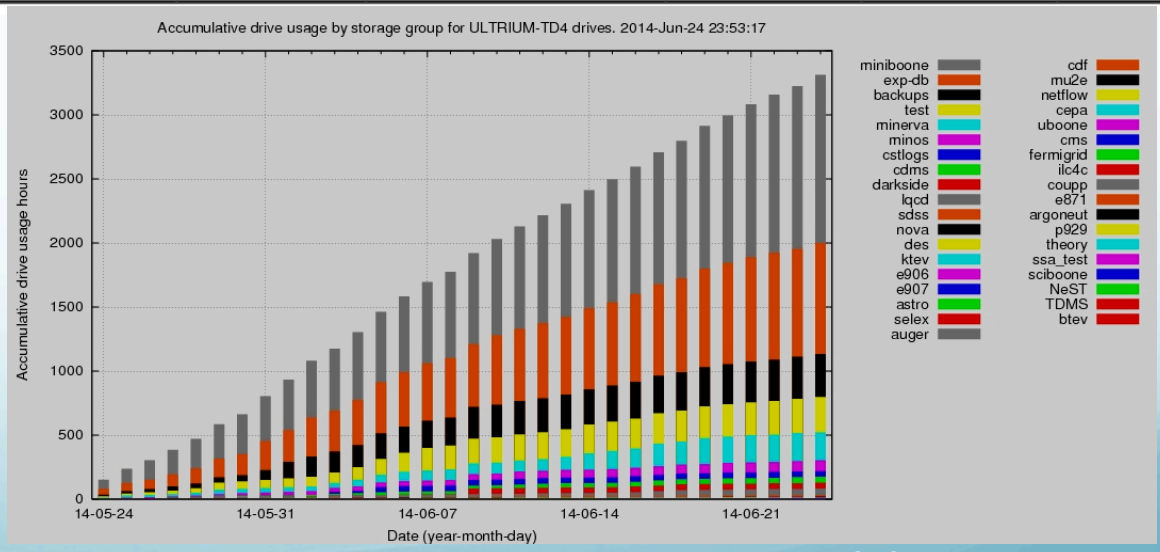
- Record of used and available space in the different storage Facility at Fermilab, partitioned by VO (storage group)
- Separate records of transfers in and out of all Fermilab's storage systems, grouped by VO
- Records of tape drive usage per VO
- Possibility to query the above data and display tables or plots (e.g. via GratiaWeb)
- Monthly email report with storage information for the storage group and the stakeholders (experiments): transfers in and out, space used and available in dCache and Enstore, all grouped by VO

Examples

Archival Report for May, 2014

VO Name	dCache Transfers		Enstore Transfers		dCache Pool Usage	Enstore storage usage		
	Read(GB)	Write (GB)	Read(GB)	Write (GB)	Total (GB)	Active Storage(GB)	All Storage (GB)	Drive Usage (hours)
auger	1,089.50	0.00	0.00	0.00	0.00	7,554.00	7,560.00	x
cdms	0.00	0.00	2.00	2,706.00	15,096.00	184,635.00	186,335.00	x
coupp	0.00	0.00	0.00	0.00	336.00	6,875.00	6,891.00	x
darkside	149,340.00	56,802.70	6.00	57,346.00	310,168.00	303,834.00	304,188.00	x
des		3,495.60	0.00	3,620.00	6,058.00	382,803.00	409,919.00	x
fermilab	0.20	65,232.90	0.00	0.00	8,761.00	367.00	370.00	x
lbne	5,843.40	1,462.80	0.00	0.00	581.00	628.00	629.00	x
lqcd	114.10	267.20	4,504.00	36,035.00	388,253.00	2,163,919.00	2,320,831.00	x
minerva	302,738.70	0.50	0.00	19,736.00	336,621.00	547,681.00	560,966.00	x
minos	4,017.20	6,126.80	549.00	5,889.00	86,232.00	319,028.00	741,101.00	x
mu2e	26.10	8,662.30	0.00	0.00	8,659.00	3,590.00	3,592.00	x
nova	500,031.90	244,724.00	908.00	137,412.00	538,279.00	698,266.00	701,256.00	x
osg	2,048.90	13.30	0.00	0.00	0.00	0.00	0.00	x
uboone	0.00	0.00	0.00	0.00	981.00	3,396.00	3,481.00	x
Total	965,250.00	386,788.10	5,969.00	262,744.00	1,700,025.00	10,950.00	5,247,119.00	x

Archival Report



Tape drive usage plot

Current information

- The information is currently available on some Web pages or information systems and not archived:
  - dCache info page: http://fndca3a.fnal.gov:2288/info
  - dCache disk usage summaries: http://fndca3a.fnal.gov/cgi-bin/sg_usage_cgi.py , http://fndca3a.fnal.gov/cgi-bin/space_usage.py
  - Enstore transfer summary: http://fndca3a.fnal.gov/cgi-bin/sg_transfers_cgi.py
  - Enstore space usage summary: http://fndca3a.fnal.gov/cgi-bin/sg_summary_cgi.py
  - Enstore tape drive usage: http://www-stken.fnal.gov/enstore/drive-hours/plot_enstore_system.html

Next steps

- Collect dCache usage information by integrating Gratia with the script generating the summary Gratia (Marco and Dmitry)
- Query the Enstore DB to collect information about transfers (every 15min) and storage (monthly) (Marco, Dmitry provides DB access and support)
- Collect the tape drive spinning data (Dmitry will make the data available in a structured fashion and Marco will write the probe to load it in Gratia)
- Change the Gratia core to accommodate the new data: database schema, store procedures
- Add GratiaWeb plots to display the above data
- Write email report with an archival report

Some technical details (1)

- Gratia database:
  - Storage Elements and Storage Element Record will store the information about used and available space
  - Tape drive usage information will be added in the JobUsageRecord with the a new ResourceType
  - New summary tables or views need to be added: with used and available space and with tape drive usage
  - Store procedures need to be added to generate the summaries
  - Current store procedures need to be verified, specially against the added ResourceType

Some technical details (2)

- GratiaWeb:
  - Add a new field to data transfer queries to allow to filter transfers (in/out/all)
  - Add new queries/plots for available and used space
  - Add new queries/plots for tape drive usage
  - Update summaries with some of the new plots
- A test Gratia collector will test the database changes and collect information from the new probes