

FARMA GOLIÁŠ

Jiří Chudoba

ATLAS seminář, FZÚ

6.2.2014

Výpočetní středisko

- Provozuje výkonné výpočetní klastry a datové servery
 - Goliáš (4176 jader), Dorje (512 jader), Kalpa (48 jader), Luna2013 (752 jader)
 - DPM (3400 TB RAW), xrootd (1100 TB RAW), pásková knihovna (100 TB)
- Správci ze SAVT, jen JCh ze sekce I
 - na správě se příležitostně podílejí i další pracovníci



Nová výpočetní kapacita - Malva

- 12 serverů v 6 šasi o velikosti 1U
- 2xCPU Intel E5-2650 v2 @2.6 GHz, 64 GB RAM a 2x SATA HDD 500 GB každý
- Zapnutý hyperthreadingu: každý server výkon 352 HSo6
 - 32 HT jader, 11 HSo6 každý
- Celkem 4224 HSo6
 - jako superpočítač Dorje, spotřeba $\frac{1}{4}$ (4.5 kW)
- Dodávka hw 30.11.2013, v provozu od 4.12.2013
- 87% pro ATLAS, 13% pro NOvA



sléz lesní (malva sylvestris).

Nová úložná kapacita – 2 PB

- 3 komplety, každý 1 server a 3 JBOD
- 1 server poskytuje přístup k 171 HDD, každý disk 4 TB
 - $4 * 171 = 684$ TB
- celkem 2052 TB hrubého prostoru, 1308 TiB využitelného (**1437 TB**)
 - $1 \text{ TiB} = 1024^4 \text{ B} = 2^{40} \text{ B} = 1.099 \text{ TB}$
- xrootd₄, dpmpool₁₆, dpmpool₁₇
 - každý server má 15 svazků po 32 TB
- v produkci od 3.2., resp. 4.2.2014

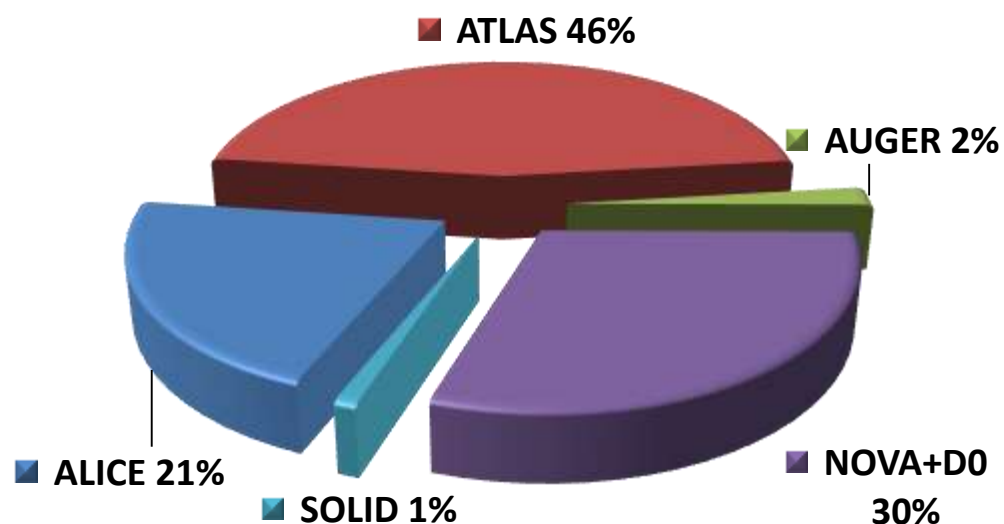
Virtualizační servery, domovské adresáře

- 3 servery pro virtuální stroje
 - virto1, virto2, virto3 (dosavadní ha1, ha2, ha3)
 - KVM (Kernel-based Virtual Machine)
 - postupný přenos **UI**, CREAM, VOBOX, ARGUS, MONBOX, ...
 - ui6 již dostupný – 16 GB RAM, 4 CPU
- Domovské adresáře
 - dosud na poli HP EVA, stáří 6 let, mimo záruku
 - nový server 2.5 TB čistá kapacita (RAID6, 5+2)
 - síťová karta 10 Gbps
 - zátěžové testy OK
 - další testy pomocí uživatelů (svazek /raidr/atlas, /raidr/auger přes NFSv4)

Celková kapacita

- 3200 TiB, z toho **ATLAS 2280 TiB (využitelné)**
- 309 WN (+113 Dorje a Luna2013), 4176 jader (i s HT)
- 35 000 HSo6, z toho **ATLAS 15800 HSo6**

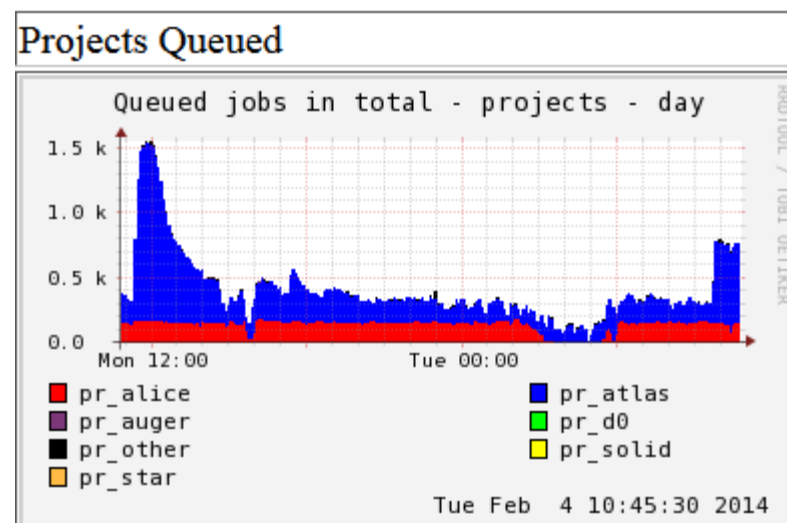
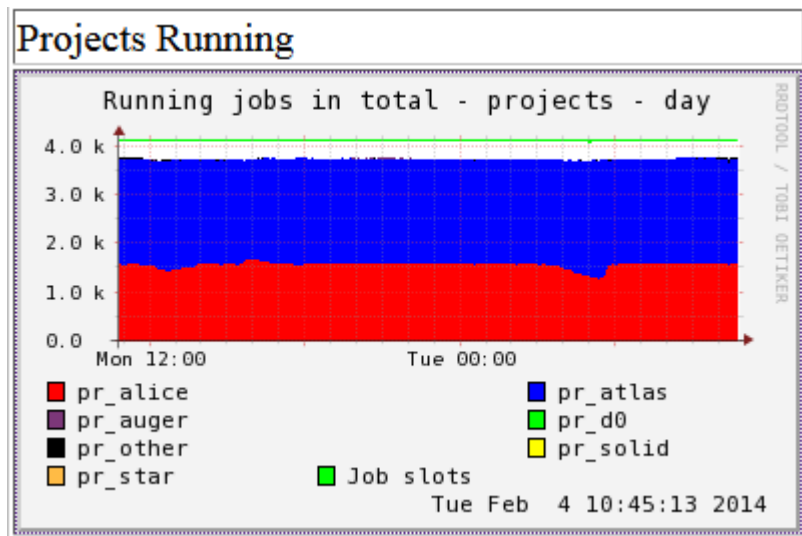
ne vše pod zárukou, možné snížení v průběhu tohoto roku



Rozdělení výpočetních kapacit v roce 2014

Přehledy

- Obecné informace a dokumentace na Twiki
 - <http://www.farm.particle.cz/twiki/bin/view/VS/VsPrehled>
- Monitorování
 - https://monitor.farm.particle.cz/~rrdmonitor/total_overview.html



Přehled pracovních uzlů

Generated: Tue Feb 4 10:50:05 2014

State	Description
Down	Node is inaccessible by the torque server
Offline	Node is intentionally out of production
Offline with jobs	Offline with some jobs still running
Empty	Node is ok, but with no jobs
Partially full	There are free job slots on the node
Full	Node is ok and running maximum jobs
Unknown	State is unknown - see output of pbsnodes

goliass143	goliass144	goliass145	goliass146	goliass147	goliass148	goliass149	goliass150	goliass151	goliass152
goliass153	goliass154	goliass155	goliass156	goliass157	goliass158	goliass159	goliass160	goliass161	goliass162
goliass163	goliass164	goliass165	goliass166	goliass167	goliass168	goliass169			

ib01	ib02	ib03	ib04	ib05	ib06	ib07	ib08	ib09	ib10
ib11	ib12	ib13	ib14	ib15	ib16	ib17	ib18	ib19	ib20
ib21	ib22	ib23	ib24	ib25	ib26				

iberis01	iberis02	iberis04	iberis05	iberis06	iberis07	iberis08	iberis09	iberis10	iberis11
iberis12	iberis13	iberis14	iberis15	iberis16	iberis17	iberis18	iberis19	iberis20	iberis21
iberis22	iberis23	iberis24	iberis25	iberis26	iberis27	iberis28	iberis29	iberis30	iberis31
iberis32	iberis33	iberis34	iberis35	iberis36	iberis37	iberis38	iberis39	iberis40	iberis41
iberis42	iberis43	iberis44	iberis45	iberis46	iberis47	iberis48	iberis49	iberis50	iberis51
iberis52	iberis53	iberis54	iberis55	iberis56	iberis57	iberis58	iberis59	iberis60	iberis61
iberis62	iberis63	iberis64	iberis65	iberis66	iberis67	iberis68	iberis69	iberis70	iberis71
iberis72	iberis73	iberis74	iberis75	iberis76					

ibis01	ibis02	ibis03	ibis04	ibis05	ibis06	ibis07	ibis08	ibis09	ibis10
ibis11	ibis12	ibis13	ibis14	ibis15	ibis16	ibis17	ibis18	ibis19	ibis20
ibis21	ibis22	ibis23	ibis24	ibis25	ibis26	ibis27	ibis28	ibis29	ibis30
ibis31	ibis32	ibis33	ibis34	ibis35	ibis36	ibis37	ibis38	ibis39	ibis40
ibis41	ibis42	ibis43	ibis44	ibis45	ibis46	ibis47	ibis48	ibis49	ibis50
ibis51	ibis52	ibis53	ibis54	ibis55	ibis56	ibis57	ibis58	ibis59	ibis60
ibis61	ibis62	ibis63	ibis64	ibis65					

malva01 malva02 malva03 malva04 malva05 malva06 malva07 malva08 malva09 malva10
malva11 malva12

rubuk01 rubuk02

rubul01 rubul02

rubus01 rubus02 rubus03 rubus04 rubus05 rubus06 rubus07 rubus08 rubus09 rubus10
rubus11 rubus12 rubus13 rubus14 rubus15 rubus16 rubus17 rubus18 rubus19 rubus20
rubus21 rubus22 rubus23

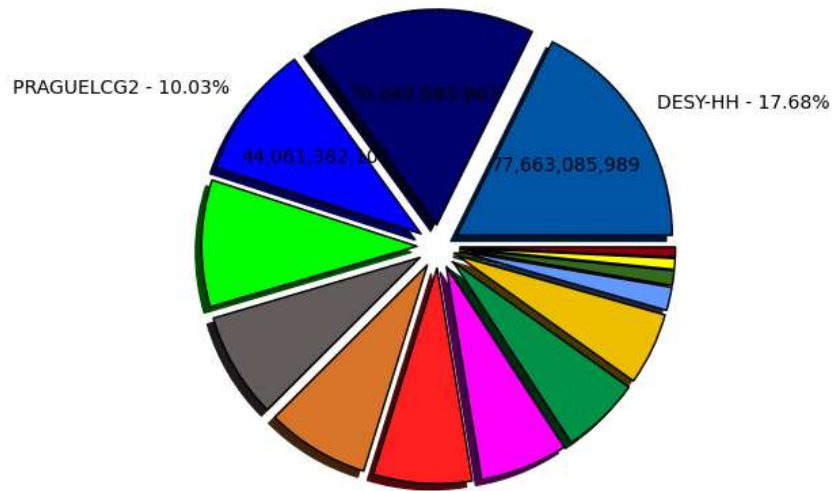
salix01 salix02 salix03 salix05 salix06 salix07 salix08 salix09 salix10 salix11
salix12 salix13 salix14 salix15 salix16 salix17 salix18 salix19 salix20 salix21
salix22 salix24 salix25 salix26 salix27 salix28 salix29 salix30 salix31 salix32
salix33 salix34 salix35 salix36 salix37 salix38 salix39 salix40

saltix01 saltix02 saltix03 saltix04 saltix05 saltix06 saltix07 saltix08 saltix09 saltix10
saltix11 saltix12 saltix13 saltix14 saltix15 saltix16 saltix17 saltix18

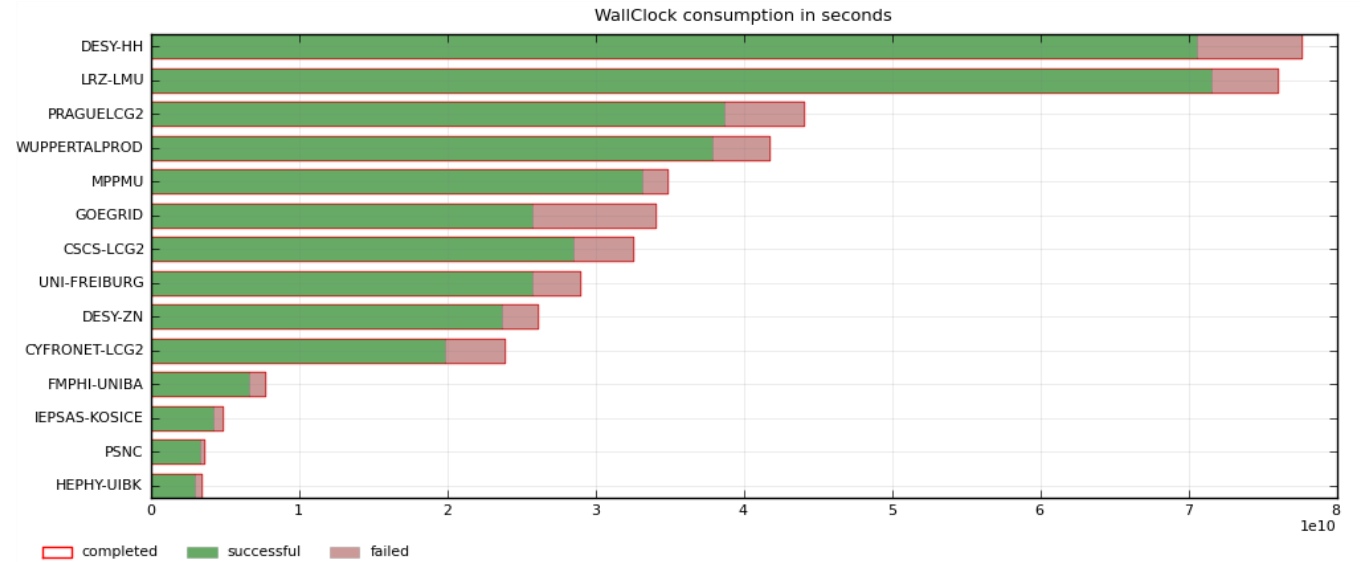
ATLAS - Propočítaný čas za rok 2013



Wall Clock consumption All Jobs in seconds (Sum: 439,173,897,909)
LRZ-LMU - 17.32%



- DESY-HH - 17.68% (77,663,085,989)
- PRAGUELCG2 - 10.03% (44,061,382,102)
- MPPMU - 7.93% (34,834,472,882)
- CSCS-LCG2 - 7.40% (32,498,710,339)
- DESY-ZN - 5.94% (26,087,882,997)
- FMPHI-UNIBA - 1.74% (7,640,694,805)
- PSNC - 0.81% (3,559,698,192)
- LRZ-LMU - 17.32% (76,043,583,867)
- WUPPERTALPROD - 9.51% (41,754,985,091)
- GOEGRID - 7.75% (34,050,329,337)
- UNI-FREIBURG - 6.59% (28,941,078,154)
- CYFRONET-LCG2 - 5.43% (23,852,915,053)
- IEPSAS-KOSICE - 1.10% (4,829,894,786)
- HEPHY-UIBK - 0.76% (3,355,184,315)



$44 * 10^8 \text{ s} = 1397 \text{ let}$

DE site availability during 2013

Historic view for "panda analy status NEW"
from 00:00 01.01.2013 to 00:00 31.12.2013

Show 100 entries

SITE Name	TIER	CLOUD	History plot time bin - 728 hours	offline		brokeroff		online	
				%	count	%	count	%	count
CSCS-LCG2	T2D	DE		0.5	3	0.14	3	87.76	276
CYFRONET-LCG2	T2	DE		12.11	37	0.05	1	75.5	173
DESY-HH	T2D	DE		0.34	6	0.1	3	96.75	151
DESY-ZN	T2D	DE		0.9	4	0.64	23	96.71	131
FMPH-UNIBA	T2	DE		0.17	3	0.05	1	26.44	33
FZK-LCG2	T1	DE		0.04	1	0.1	3	97.5	138
GoeGrid	T2D	DE		1.91	11	0.14	6	77.62	239
HEPHY-UIBK	T2	DE		1.74	10	3.83	5	77.65	136
IEPSAS-Kosice	T2	DE		7.46	4	0.64	4	19.26	16
LRZ-LMU	T2D	DE		1.48	5	1.99	47	94.22	170
MPPMU	T2D	DE		3.65	6	0.28	6	88.31	105
prague1cg2	T2D	DE		0	0	1.3	24	95.92	141
PSNC	T2	DE		56.48	27	0	0	10.23	61
UNI-FREIBURG	T2D	DE		1.15	8	0.05	1	91.81	131
wuppertalprod	T2D	DE		3.29	6	0.14	3	93.61	140

95.92

- 4 sites >95%: FZK, Desy-HH, Desy-ZN, Prague
- 4 sites 90-95%: LRZ, Wuppertal, Freiburg
- 5 sites 70-90%: MPPMU, CSCS, Innsbruck, Goettingen
- 3 sites only partially included: PSNC, UNIBA, Kosice

DE site availability Jan 2014

Historic view for "panda analy status NEW"
from 00:00 01.01.2014 to 00:00 31.01.2014

Show entries

SITE Name	TIER	CLOUD	History plot time bin - 60 hours	offline		broker off		online	
				%	count	%	count	%	count
CSCS-LCG2	T2D	DE		0	0	0	0	96.05	13
CYFRONET-LCG2	T2	DE		0	0	0	0	97.63	10
DESY-HH	T2D	DE		0	0	0	0	96.96	7
DESY-ZN	T2D	DE		0	0	0	0	99.46	4
FMPH-UNIBA	T2	DE		12.67	1	0	0	85.86	8
FZK-LCG2	T1	DE		0	0	0	0	98	8
GoeGrid	T2D	DE		0	0	0	0	76.12	26
HEPHY-UIBK	T2	DE		0	0	0	0	96.35	13
IEPSAS-Kosice	T2	DE		13.16	3	0.46	1	84.56	6
LRZ-LMU	T2D	DE		0	0	0	0	99.46	4
MPPMU	T2D	DE		0	0	0	0	98.9	5
prague _{lcg2}	T2D	DE		0	0	0	0	99.46	4
PSNC	T2	DE		99.46	4	0	0	0	0
UNI-FREIBURG	T2D	DE		0	0	0	0	99.11	6
wuppertalprod	T2D	DE		0	0	0	0	99.46	4

99.46

Storage Overview

Site	Space Token (Total / Used in TB)			
	DATA	PROD	LOCALGRP	SCRATCH
CSCS-LCG2	703 / 544	10 / 1	10 / 8	38 / 19
CYFRONET	300 / 273	11 / 1	11 / 5	28 / 14
DESY-HH	1189/1137	28 / 14	1637 / 1248	70 / 43
DESY-ZN	707 / 659	25 / 7	570 / 522	50 / 40
FMPHI-UNIBA	174 / 151	8 / 0	16 / 3	10 / 5
GOEGRID	900 / 756	30 / 13	80 / 74	40 / 23
HEPHY-UIBK	111 / 83	5 / 0	1.1 / 0.8	3 / 2
IEPSAS-KOSICE	75 / 66	19 / 0	13 / 11	27 / 9
LRZ-LMU	725 / 683	19 / 3	270 / 246	49 / 43
MPPMU	625 / 531	20 / 8	175 / 149	40 / 20
PRAGUELCG2	1540 / 953	41 / 4	220 / 178	41 / 29
PSNC	55 / 13	5 / 1	-	22 / 8
DRESDEN	7 / 1	4 / 0	10 / 5	9 / 4
UNI-FREIBURG	800 / 752	35 / 5	270 / 231	65 / 35
WUPPERTAL	716 / 670	70 / 5	135 / 111	45 / 24

Přenosy na LOCALGROUPDISK - DaTRI

- Data Transfer Request Interface
 - http://panda.cern.ch/server/pandamon/query?mode=ddm_req

Information about Data Transfer Request pages [here](#)

- **Data pattern:** dataset name, container name [`<containerName/>`] or some dataset/container pattern [`<dataPattern>[/]`]. container can be used if one pattern do not match all desired datasets (see [TWiki](#))
- **Data type:** None is used only by user datasets and evgen input datasets
- **Destination:** [TiersOfAtlas](#) name for destination endpoint (ATLAS policy states this must be a *DISK* area)

IMPORTANT

- You must be registered to use this interface (for registration go [here](#)). You may check your registration status [here](#)
- If you have any problems with your request (approval or data replication), please place [DDM Ops Savannah ticket](#)
- Additional contact information (after the ticket is placed): [Mikhail.Titov\[at\]cern.ch](mailto:Mikhail.Titov[at]cern.ch)

Request Parameters	Data pattern*	isa_Egamma.merge.NTUP_COMMON.r4065_p1278_p1562_id01316235_00
	Data type*	check type NTUP *
	Destination*	DE * group * PRAGUELOG2_LOCALGROUPDIS *
Control Parameters	Justification*	Replication to local resources *
	Extra emails	
	Comments	
Volume Parameters (to request a part of data)	<input type="checkbox"/> Volume	100 % (of original data)
	<input type="checkbox"/> Files list	
Estimated parameters	<input type="checkbox"/> Sum size limit	10000 GiB
	<input type="checkbox"/> Num files limit	10000
Addon Parameters	Request type	user *
	Is Periodic	False *
	Priority	High *

For request submitting certificate is needed. Please follow this [link](#) to register certificate.
* - required fields

Configuration Production Clouds Incidents DDM PandaMover AutoPilot Sites Releases Analysis Stats Users Physics data ProdTask DDMDash SSB

Panda monitor Times are in UTC
Data Transfer Request Interface (DaTRI) - User Requests
Panda info and help

Data Transfer Request #91727 - Data monitoring: [Dashboard](#) | [DSInfo](#)

Request owner	Michal Svatos (michal.svatos@cern.ch)
Creation date	2014-01-28 12:52:30
Data pattern	data12_8TeV.00207113.physics_Egamma.merge.NTUP_COMMON.r4065_p
Data type	NTUP
Destination	PRAGUELOG2_LOCALGROUPDIS (cloud: DE)
Volume	100 % of original data
Files list	-
Num subscribed datasets	1
Num datasets	1
- Num files	2
- Sum size	0.002953 GiB
Estimated limits	
- Limit for num files	500000
- Limit for sum size	500000.0 GiB
Status	DONE (2014-01-28 17:56:59)
Approved by	AutoApprove (2014-01-28 13:09:12)
Modified by	(-)
Justification	Cannot use the existing replicas - only replica is on tape
Extra emails	
Comments	
Addon params	Request type user Is Periodic 0 Priority 0

Request actions:

For request management certificate is needed. Please follow this [link](#) to register certificate.

DaTRi

Panda info and help

- Jobs [Search](#)
- Job Types
- Jobs' Timing
- Jobs' Metrics
- Jedi Tasks
- FAX
- Quick search
- Errors
- Summaries
- Tasks [Search](#)
- Datasets [Search](#)
- Datasets Distribution
- DaTRi:
 - Data Transfer Request
 - List User Requests
 - List Pathname Requests
 - List Gauge Requests
 - Group Production
- PDZP
- AODs
- EVENTs
- Conditions DS
- DB Releases
- SIT packages
- Validation Samples
- Functional Tests
- ATLAS Data
- Reprocessed Datasets
- Logging monitor
- [Incidents](#)
- Analytics

For more information click here *(hidden text)*

Data Pattern: *

Cloud: ANY PhysGroup: ANY DDMEndpoint: PRAGUELCG2_LOCALGROUPDISK

Time period for request creation : last day last week last 30 days last 60 days

Request status: ANY Request owner (nickname/email):

Request ID: OR Subscription ID: *(Parent ID in subscription info)*

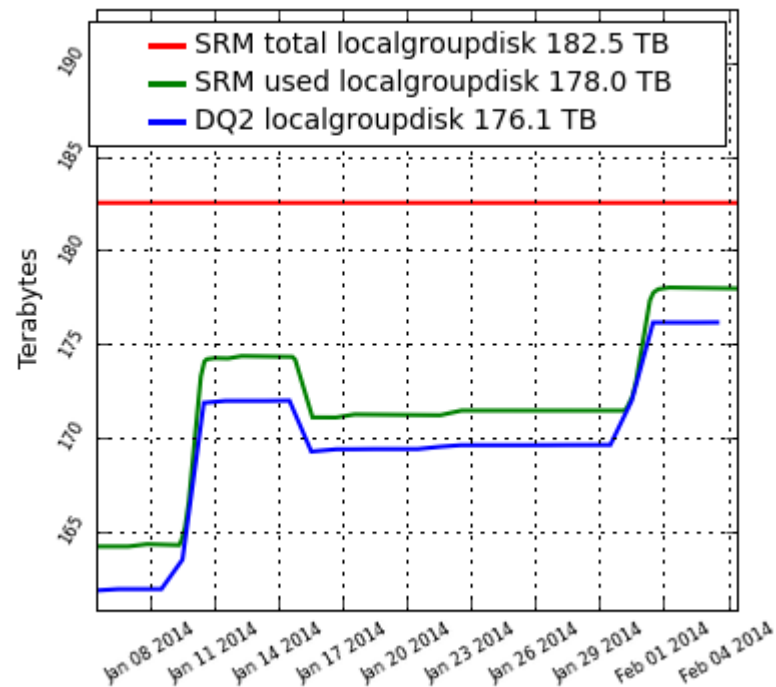
List [Actions that will use certificate from browser to identify user:](#) [User Requests](#) [Requests For Approval](#)

Num Requests 14 Num Datasets In Requests 182 Num Subscribed Datasets 182 Summary Size Of All Datasets (GiB) 10087.275628

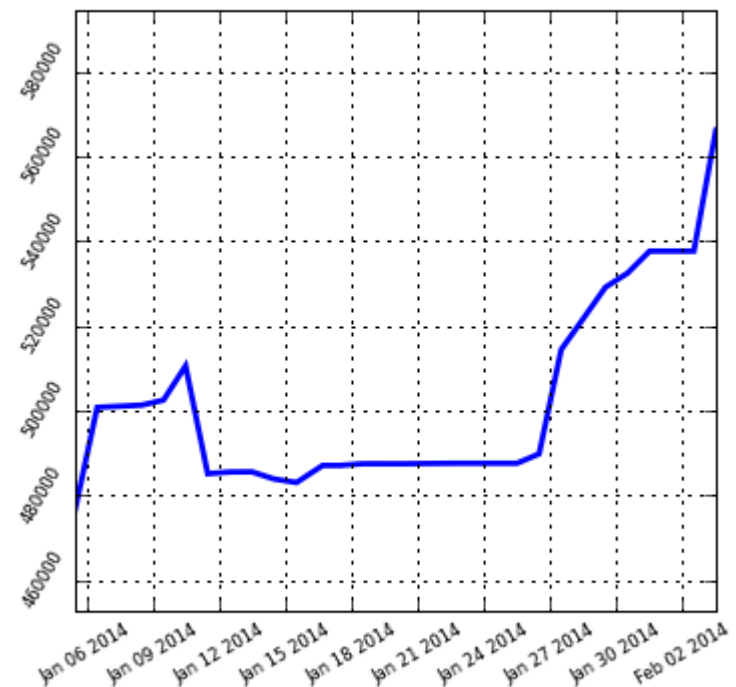
Expand information about requests

Req ID	Data Pattern Status Creation time Modification time	DDM Endpoint	Num Datasets All/Subscribed	Num Files	Size (GiB)
91727	data12_8TeV.00207113.physics_Egamma.merge.NTUP_COMMON.r4065_p1278_p1562_tid01316235_00 status: DONE (2014-01-28 12:52:30 2014-01-28 17:56:59)	PRAGUELCG2 LOCALGROUPDISK	1 / 1	2	0.0030
90525	user.okepka*Pythia8*sv03/ status: DONE (2014-01-17 22:30:17 2014-01-18 10:21:40)	PRAGUELCG2 LOCALGROUPDISK	5 / 5	1596	39.2789
90316	user.okepka.MC12_14TeV.998003.Herwigpp_DPE_H12007_jetjet_JZ2.HITS.e03.sv03/ status: DONE (2014-01-16 08:19:43 2014-01-16 20:45:42)	PRAGUELCG2 LOCALGROUPDISK	1 / 1	997	32.5181
90315	user.okepka.MC12_14TeV.998002.Herwigpp_DPE_H12007_jetjet_JZ1.HITS.e03.sv03/ status: DONE (2014-01-16 08:18:36 2014-01-16 20:45:36)	PRAGUELCG2 LOCALGROUPDISK	1 / 1	997	33.4966
90314	user.okepka*MC12_14*EVNT*e04/ status: DONE (2014-01-16 08:17:03 2014-01-16 20:45:36)	PRAGUELCG2 LOCALGROUPDISK	2 / 2	80	2.2664
90313	user.okepka.mc12_14TeV.11999*HITS*.e1133_der0.sv03/ status: DONE (2014-01-16 08:15:05 2014-01-16 20:45:33)	PRAGUELCG2 LOCALGROUPDISK	3 / 3	1196	29.3202
90312	user.okepka*MC12_14*EVNT*e03/ status: DONE (2014-01-16 08:13:13 2014-01-16 20:45:30)	PRAGUELCG2 LOCALGROUPDISK	18 / 18	358	15.2244

LOCALGROUPDISK



Number of files for PRAGUELCG2_LOCALGROUPDISK

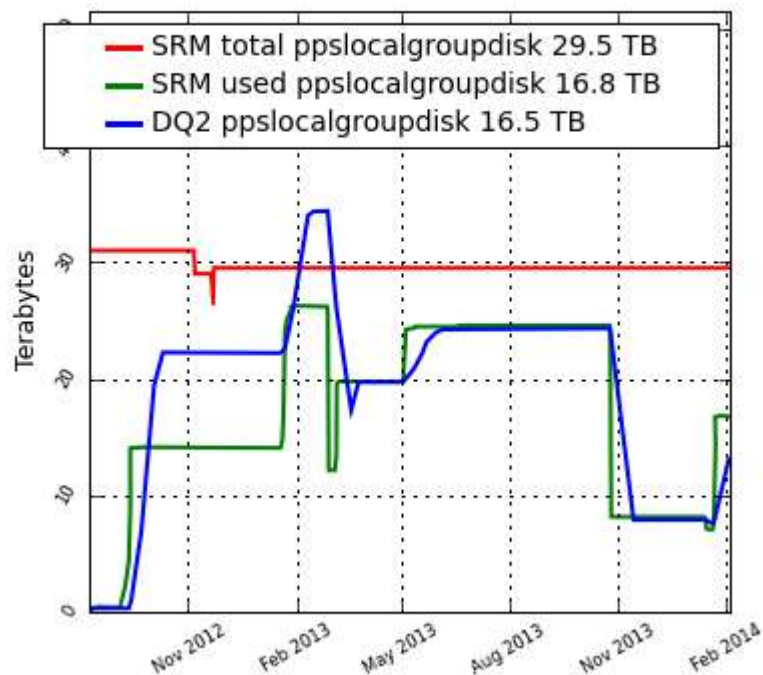


DQ2 Accounting: Storage Usage Plots

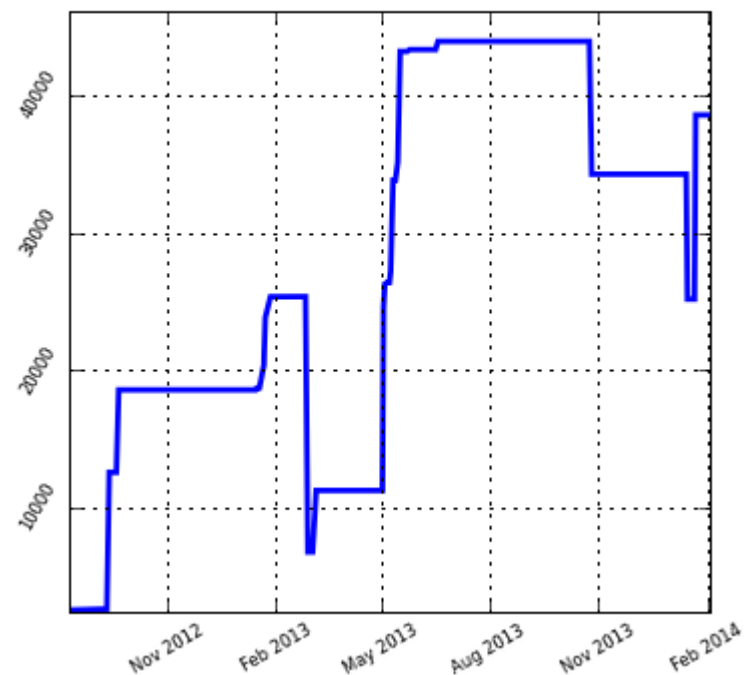
http://bourricot.cern.ch/dq2/accounting/site_view/PRAGUELCG2_LOCALGROUPDISK/30/

PPSLOCALGROUPDISK

Used disk space for PRAGUELCG2_PPSLOCALGROUPDISK



Number of files for PRAGUELCG2_PPSLOCALGROUPDISK



Datová úložiště CESNET v Plzni
dCache
publikováno přes naše Tier2

Přístup

- Pomocí specializovaných serverů – User Interface
 - ui5 – virtuální stroj se SL5
 - ui6, ui7 – virtuální stroje se SL6
 - problémy s výkonem při využití CVMFS
 - ui8 – fyzický stroj (iberiso2) se SL6
- Nové servery pro virtualizaci budou používat KVM
 - předpoklad, že budou lépe fungovat s CVMFS
 - ui6 od dnešního dne k dispozici na novém hw

Podpora

- Správci
 - obecná podpora (výpadky, opravy hardware, instalace OS, spouštění úloh, ...)
 - fzu@rt4.cesnet.cz
 - kritický nedostatek pracovníků, brzké vyřešení není zaručené
- ATLAS specifické problémy
 - postupy při vytváření úloh, nastavení software, výběr dat, ...
 - globální ATLAS skupiny
 - lokálně jen **svépomocí**
 - Twiki

Další dostupné zdroje

- EGI
 - zdroje i mimo VO ATLAS
- Metacentrum + CERIT-SC
 - <http://www.metacentrum.cz/>
 - každý akademický pracovník

Výpočetní stroje

Zkuste **osobní pohled na výpočetní stroje.**

Celkem CPU v MetaCentru: **9772**

Celkem čeká **1 943 úloh** ve frontách na spuštění.

adresář	využití - využito volné	velikost
/storage/bmo1/home	93%	85TB
/storage/bmo2/home	86%	110TB
/storage/bmo3-cerit/home	91%	260TB
/storage/budejovice1/home	39%	44TB
/storage/jihlava1-cerit/home	6%	374TB
/storage/ostrava1/home	7%	88TB
/storage/plzen1/home	45%	44TB
/storage/praha1/home	43%	51TB
/storage/praha2-elixir/home	19%	95TB
Celkem úložného místa v MetaCentru		1151TB
Celkem využito		517TB
Celkem volného místa		633TB