

Using VAPOR to improve VO administration and operations

Vo Administration and operations PORTal

EGI Community Forum 2014

Franck Michel, CNRS




- VO operation/administration
 - Manage the community of users, life-cycle, their data
 - Monitor resources supporting the VO
 - Figure out infrastructure QoS from the VO point of view
- Requires costly IT support
 - Achievable for large organized communities
 - Out of reach of smaller/younger communities
- Need to lower down VO admin/operations entry cost
 - Provide tools tailored to small communities
 - Help existing communities sustain their model
 - Facilitate the outreach of new communities

- Small to medium-size user VOs
- No or few dedicated support
 - Volunteering contribution
- Fragmented user groups
 - Scattered scientific activities
 - No control on submission tools, policies...
 - Difficult to anticipate future needs
- Dedicated and/or opportunistic usage of resources

Nagios®

myEGI

 OPERATIONS PORTAL

 dashboard

 dashboard
Virtual Organizations
admin

VOMS admin

EGI ACCOUNTING PORTAL

 GGUS

 CESGA

 GOCDB

 LCC GStat 2.0

- **Complement** existing tools: SAM, GStat2, Ops Portal...
- Provide tools **tailored** to such communities
 - Experiment independent, no install/admin required
- Help mutualise admin/operation effort
 - Between several VOs, within a VRC or beyond
- Main axes:
 - Resource status indicators, statistical reports
 - VO data management
 - *User community management*

Resources status

- Detailed report of all resources supporting the VO (BDII, GOCDB)
- Running vs. Waiting jobs
 - Queues with all jobs failed
 - Queues with no job at all

Computing Elements

Site	Hostname	Path	Status	Downtime	Total CPUs	Free CPUs
LSG-AMC	gb-ce-amc.amc.nl	cream-pbs-long	Unknown	None	131	81
LSG-AMC	gb-ce-amc.amc.nl	cream-pbs-medium	Unknown	None	131	81
LSG-AMC	gb-ce-amc.amc.nl	cream-pbs-express	Unknown	None	131	81
SARA-MATRIX	creamce.gina.sara.nl	cream-pbs-short	Production	None	2464	146
SARA-MATRIX	creamce.gina.sara.nl	cream-pbs-medium	Production	None	2464	146
SARA-MATRIX	creamce2.gina.sara.nl	cream-pbs-short	Production	None	2464	150
SARA-MATRIX	creamce3.gina.sara.nl	cream-pbs-short	Production	None	2464	150
SARA-MATRIX	creamce3.gina.sara.nl	cream-pbs-medium	Production	None	2464	150
SARA-MATRIX	creamce2.gina.sara.nl	cream-pbs-medium	Production	None	2464	150

Storage Elements

Site	ID	Status	Downtime	Impl. name	SA Total size	SA Used size	SA
LSG-AMC	gb-se-amc.amc.nl	Unknown	None	DPM	40000	8479	31
SARA-MATRIX	sm.grid.sara.nl	Production	None	dCache	27488	6	27
SARA-MATRIX	sm.grid.sara.nl	Production	None	dCache	245119	0	24

Companion of



GOCDB



GStat 2.0

Resources status

- Detailed report of all resources supporting the VO (BDII, GOCD)
- Running vs. Waiting jobs
 - Queues with all jobs failed
 - Queues with no job at all

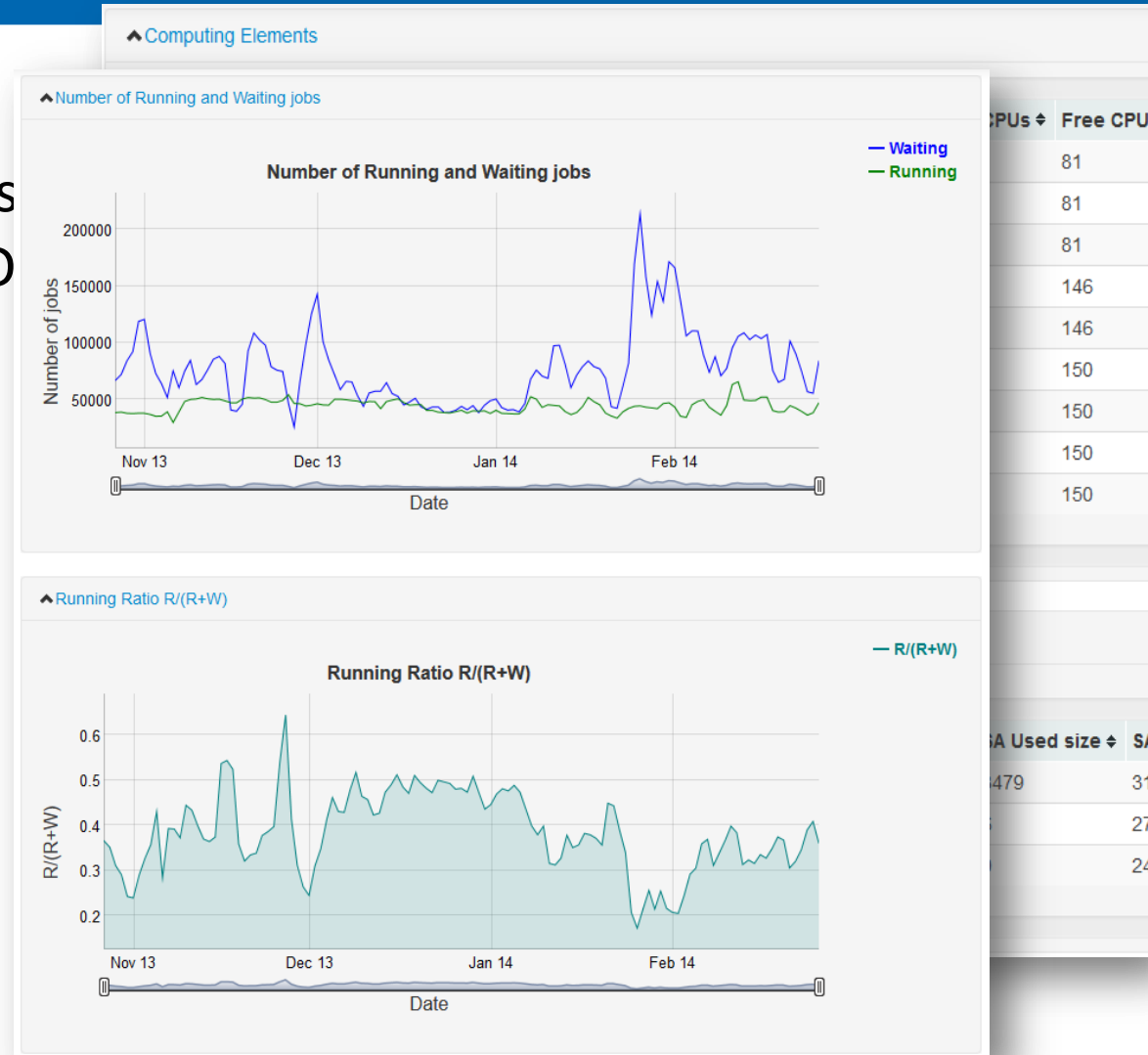
Companion of



GOCDB



GStat 2.0



Resources status

- Detailed report of all resources supporting the VO (BDII, GOCD)
- Running vs. Waiting jobs
 - Queues with all jobs failed
 - Queues with no job at all

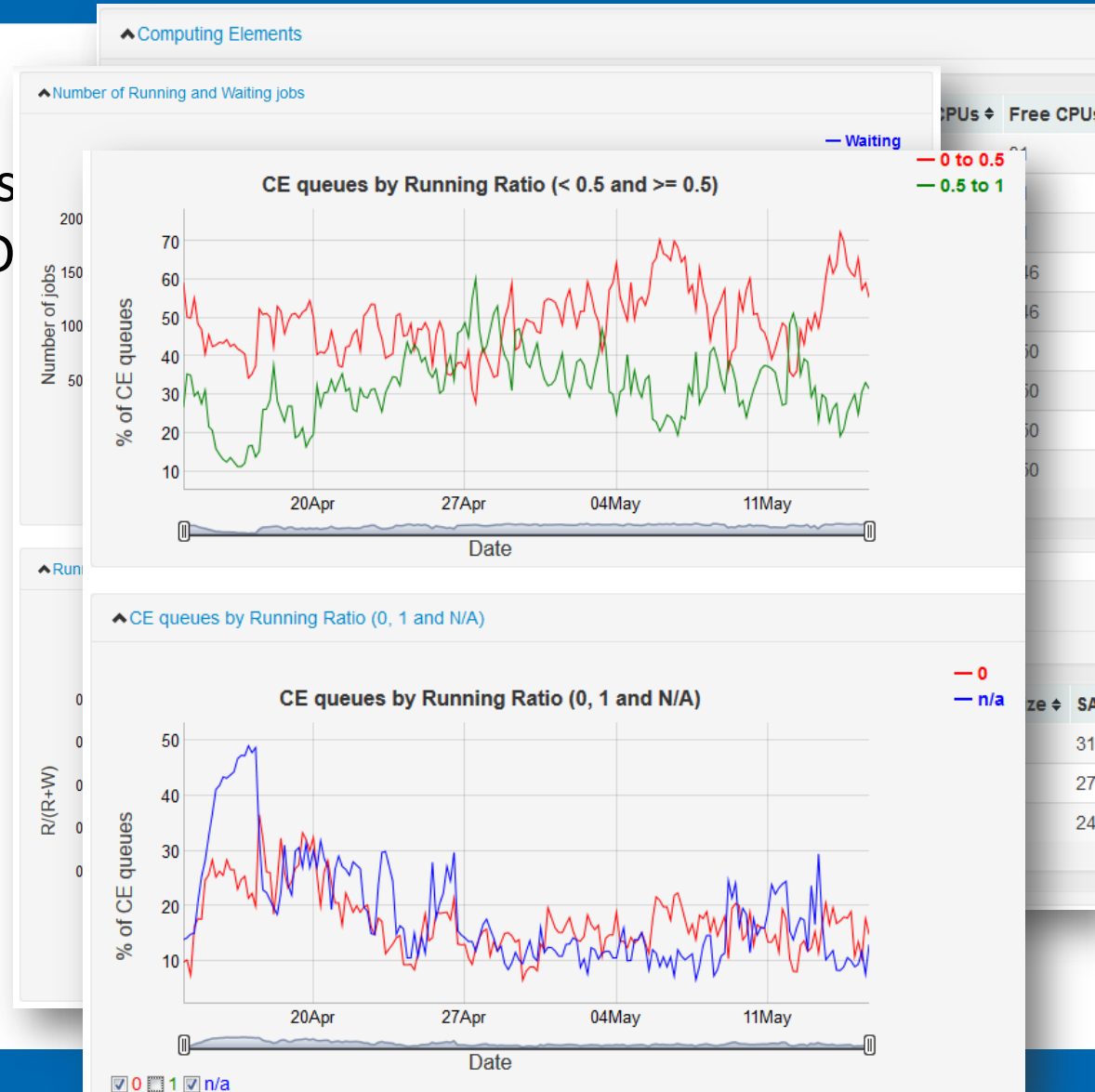
Companion of



GOCDB



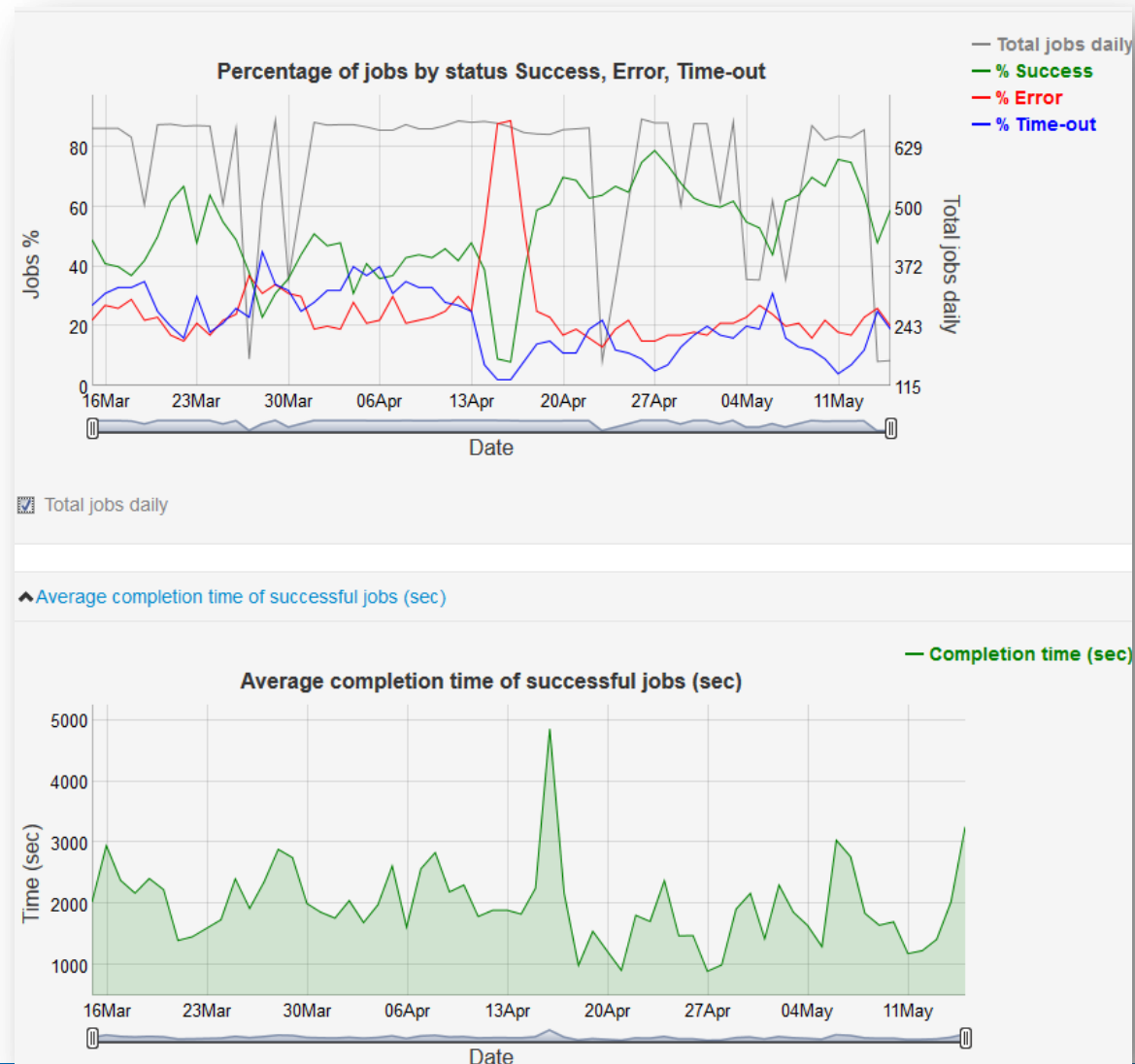
GStat 2.0



Computing Elements reports: JobMonitor

- Monitor how the infrastructure is responsive for the VO
- Success/Error/Time out rates
- Avg. jobs completion time
- Graphical view (time), or table view (by site, CE)

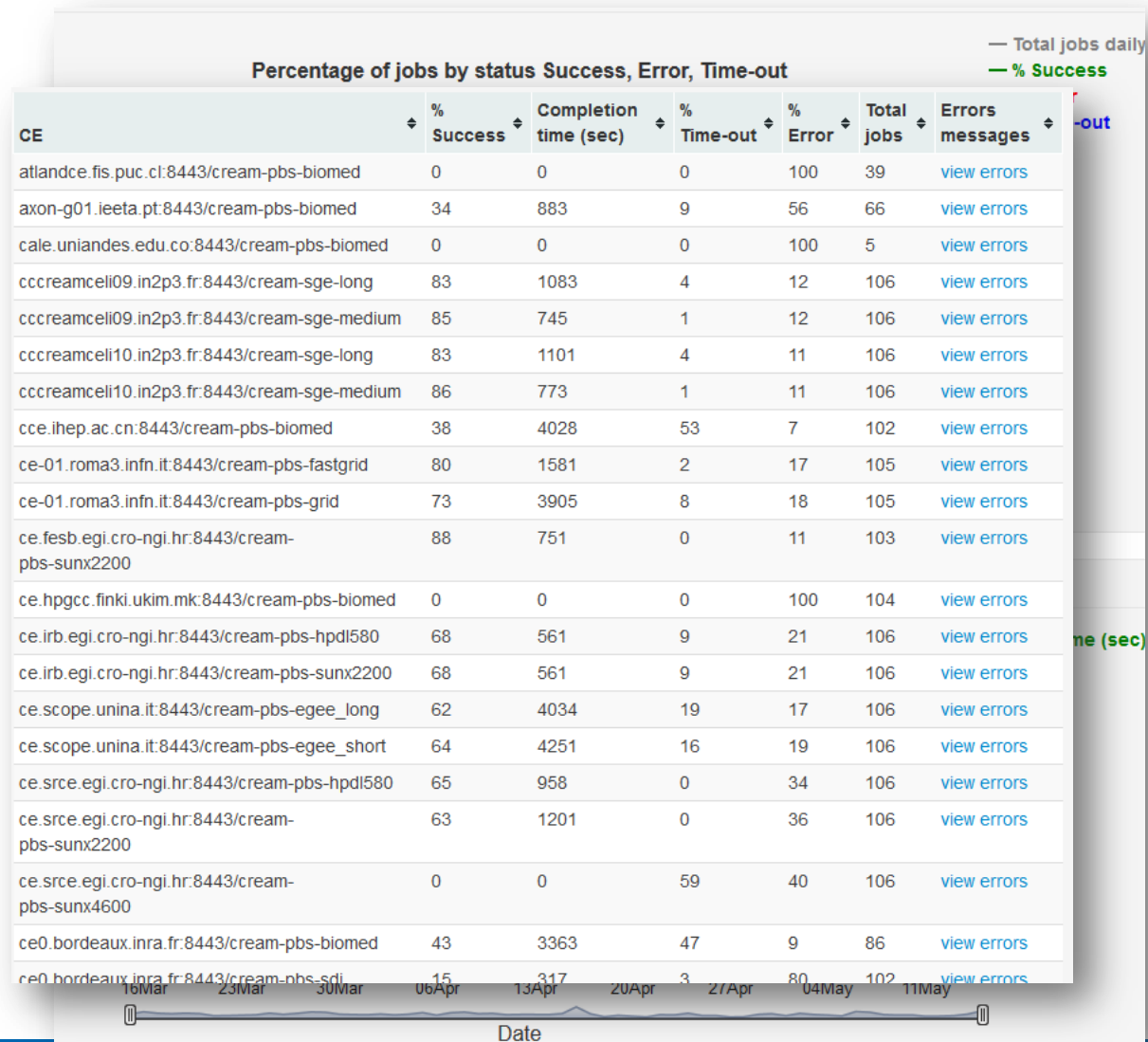
Companion of
Nagios®



Computing Elements reports: JobMonitor

- Monitor how the infrastructure is responsive for the VO
- Success/Error/Time out rates
- Avg. jobs completion time
- Graphical view (time), or table view (by site, CE)

Companion of
Nagios[®]



White list of CEs

- CEs “working fine lastly”
 - Does not replace WMS, pilot jobs
- Parametric request
- Dynamically generated XML file to feed submission systems

Minimum jobs success rate: 0.90 during the last 3 day(s).
 The last 3 consecutive jobs were successful.

[View White List](#) [View Download Url](#)

CE hostname	Queue	Job success rate	Number of jobs in success rate	Number of last consecutive successful jobs	Consecutive successful jobs reached
svr010.gla.scotgrid.ac.uk	cream-pbs-biomed	1.0000	10	3	OK
cygnus.grid.rug.nl	cream-pbs-short	1.0000	10	3	OK
fal-pygrid-44.lancs.ac.uk	cream-pbs-q	1.0000	10	3	OK
grid001.fc.up.pt	cream-pbs-biomed	1.0000	6	3	OK
svr011.gla.scotgrid.ac.uk	cream-pbs-biomed	1.0000	10	3	OK
cream-ce-grid.obspm.fr	cream-pbs-biomed	1.0000	10	3	OK
ce-01.roma3.infn.it	cream-pbs-fastgrid	1.0000	10	3	OK
ce08-lcg.cr.cnaf.infn.it	cream-lsf-biomed	1.0000	10	3	OK
ce0.m3pec.u-bordeaux1.fr	cream-pbs-biomed	1.0000	9	3	OK

Catalog-based SE scan: helps manage several issues

- Monitoring of SEs feeling up: ask heavy VO users to migrate their files
- SE decommissioning: warn users priori to decomm
- Clean-up files left behind: report by status of user (active, expired, suspended, unknown)

▲ Catalog-based SE scan: 2014-05-01 02:00:23

Hostname	% Used Space	Used Space (GB)	Free Space (GB)	Total Space (GB)	Scan Status
grid001.jet.efda.org	80	1192	282	1474	Completed
lcgse0.shef.ac.uk	81	5134	1146	6280	Completed
se.sg.grena.ge	81	1528	343	1871	Completed
prod-se-02.pd.infn.it	82	3832	817	4649	Completed
se0.m3pec.u-bordeaux1.fr	83	27797	5499	33296	Completed
epgse1.ph.bham.ac.uk	84	106335	20112	126447	Completed
dpm.cyf-kr.edu.pl	85	17087	2883	19970	Completed
eymir.grid.metu.edu.tr	86	304429	46493	350922	Completed
node12.datagrid.cea.fr	86	5624	872	6496	Completed
sbgse1.in2p3.fr	86	691218	112082	803300	Completed
hepgrid11.ph.liv.ac.uk	87	4797	677	5474	Completed
gridsm.ts.infn.it	88	3365	442	3807	Completed
gridse.ilc.cnr.it	89	390	45	435	Completed
tlapiacalli.nucleares.unam.mx	92	328192	27277	355469	Completed
se01.afroditi.hellasgrid.gr	94	7563	433	7996	Completed

Catalog-based SE scan: helps manage several issues

- Monitoring of SEs feeling up: ask heavy VO users to migrate their files
- SE decommissioning: warn users priori to decomm.
- Clean-up files left behind: report by status of user (active, expired, suspended, unknown)

Catalog-based SE scan: 2014-05		Active users	Used Space (GB)
Hostname	% S	/C=IT/O=INFN/OU=Personal Certificate/L=CNR-ITB/CN=Ivan Merelli	1.00
grid001.jet.efda.org	8	/O=GRID-FR/C=FR/O=CNRS/OU=CREATIS/CN=Sorina Camarasu	1.00
lcgse0.shef.ac.uk	8	/DC=es/DC=irisgrid/O=upw/CN=abel.carrion	0.41
se.sg.grena.ge	8	/O=GRID-FR/C=FR/O=CNRS/OU=CREATIS/CN=Tristan Glatard	8.00
prod-se-02.pd.infn.it	8	/C=IT/O=INFN/OU=Personal Certificate/L=Bari/CN=giacinto donvito	1.00
se0.m3pec.u-bordeaux1.fr	8	/O=GRID-FR/C=FR/O=CNRS/OU=CREATIS/CN=Rafael Silva	2.00
epgse1.ph.bham.ac.uk	8	/O=GRID-FR/C=FR/O=CNRS/OU=LPC/CN=The Quang Bui	1.00
dpm.cyf-kr.edu.pl	8	/O=GRID-FR/C=FR/O=CNRS/OU=CREATIS/CN=Robot: vip.creatis.insa-lyon.fr - Tristan Glatard	89.00
eymir.grid.metu.edu.tr	8		
node12.datagrid.cea.fr	8	Suspended or expired users	Used Space (GB)
sbgse1.in2p3.fr	8	/O=GRID-FR/C=FR/O=CNRS/OU=I3S/CN=Ketan Maheshwari	3.00
hepgrid11.ph.liv.ac.uk	8	/C=KR/O=KISTI/O=GRID/O=KISTI/CN=97968351 Sehoon Lee	0.10
gridsm.ts.infn.it	8	/O=GRID-FR/C=FR/O=CNRS/OU=CREATIS/CN=Pierre Gueth	1.00
gridse.ilc.cnr.it	8	/O=GRID-FR/C=FR/O=U-PSUD/CN=Dawei Feng	0.32
tlapiacalli.nucleares.unam.mx	9	/O=GRID-FR/C=FR/O=CNRS/OU=CREATIS/CN=Carole Lartizien	0.23
se01.afroditi.hellasgrid.gr	9		
		Unknown users	Used Space (GB)
		/O=GRID-FR/C=FR/O=CNRS/OU=IBCP/CN=Jaipheessa Ibecepe	31.00
		/O=GRID-FR/C=FR/O=CNRS/OU=LPC/CN=Jean Salzemann	0.18
		/O=GRID-FR/C=FR/O=CNRS/OU=LPC/CN=Vincent Bloch	9.00

Cleanup of Dark data/lost files

- Track SE-LFC consistency issues
- Delete dark data
- Report lost files
- Supports DPM, dCache, Castor SEs that allow GridFTP as access protocol

details	Site	Hostname	Access Protocol	Free Space Before (GB)	Free Space after (GB)	Total Space (GB)	Check Status
+	BEIJING-LCG2	ccsrn.ihep.ac.cn	gsiftp	556	555	586	Completed
+	BG03-NGCC	se02.ngcc.acad.bg	gsiftp	771	771	888	Completed
+	BG05-SUGrid	se001.grid.uni-sofia.bg	gsiftp	1525	N/A	1927	Completed with errors
+	CBPF	se.cat.cbpf.br	gsiftp	6027	6026	6027	Completed
+	CBPF	se02.cat.cbpf.br	gsiftp	1565	1565	43975	Completed
+	CESGA	se2.egi.cesga.es	gsiftp	1509	1510	6015	Completed
+	CIEMAT-LCG2	srm.ciemat.es	gsiftp	578	N/A	1328	Completed with errors
+	CYFRONET-LCG2	dpm.cyf-kr.edu.pl	gsiftp	2888	N/A	19971	Completed with errors
+	DESY-HH	dcache-se-desy.desy.de	gsiftp	9611	9610	38520	Completed
+	GARR-01-DIR	gridsv3-4.dir.garr.it	gsiftp	2989	2991	4288	Completed
+	GE-01-GRENA	se.sg.grena.ge	gsiftp	343	343	1872	Completed
+	GR-01-AUTH	se01.grid.auth.gr	gsiftp	651	695	2998	Completed
+	GR-06-IASA	se02.marie.hellasgrid.gr	gsiftp	371	372	644	Completed
+	GRISU-UNINA	grisuse.scope.unina.it	gsiftp	3605	3601	4794	Completed
+	HG-01-GRNET	se01.isabella.grnet.gr	gsiftp	2029	2030	3510	Completed
+	HG-02-IASA	se01.marie.hellasgrid.gr	gsiftp	3408	3374	9998	Completed
+	HG-04-CTI-CEID	se01.kallisto.hellasgrid.gr	gsiftp	2627	N/A	5284	Completed with errors

Cleanup of Dark data/lost files

- Track SE-LFC consistency issues
- Delete dark data
- Report lost files
- Supports DPM, dCache, Castor SEs that allow GridFTP as access protocol

details	Site	Hostname	Access Protocol	Free Space Before (GB)	Free Space after (GB)	Total Space (GB)	Check Status	
+	BEIJING-LCG2	ccsrn.ihep.ac.cn	gsiftp	556	555	586	Completed	
+	BG03-NGCC	se02.ngcc.acad.bg	gsiftp	771	771	888	Completed	
+	BG05-SUGrid	se001.grid.uni-sofia.bg	gsiftp	1525	N/A	1927	Completed with errors	
+	CBPF	se.cat.cbpf.br	gsiftp	6027	6026	6027	Completed	
+	CBPF	se02.cat.cbpf.br	gsiftp	1565	1565	43975	Completed	
+	+	CBPF	se02.cat.cbpf.br	gsiftp	1565	1565	43975	Completed
+	+	CESGA	se2.egi.cesga.es	gsiftp	1509	1510	6015	Completed
+	Total number of files on the SE: 84550		Download Dark Data dump: ⓘ					
+	Number of Dark Data files older than 12 months: 37198 (82 %)		Download Lost Files dump: ⓘ					
+	Number of Lost Files: 47 (0 %)							
+	+	CIEMAT-LCG2	srm.ciemat.es	gsiftp	578	N/A	1328	Completed with errors
+	+	CYFRONET-LCG2	dpm.cyf-kr.edu.pl	gsiftp	2888	N/A	19971	Completed with errors
+	Total number of files on the SE: 261329		Download Dark Data dump: ⓘ					
+	Number of Dark Data files older than 12 months: 177867 (80 %)		Download Lost Files dump: ⓘ					
+	Number of Lost Files: 33498 (43 %)		Show errors					
+	+	DESY-HH	dcache-se-desy.desy.de	gsiftp	9611	9610	38520	Completed
+	+	GARR-01-DIR	gridsv3-4.dir.garr.it	gsiftp	2989	2991	4288	Completed
+	+	HG-04-CTI-CEID	se01.kallisto.hellasgrid.gr	gsiftp	2627	N/A	5284	Completed with errors

- Up and running for 6 VOs:
 - Biomed, CompChem, Enmr.eu, France Grilles VO, Shiwa-workflow.eu, Vlemed
- Integrated with Operations Portal



- Exploit and benefit from VAPOR!
 - Analyse job monitoring results
 - Queues with no jobs at all?
 - Detect bottleneck or underperforming resources?
 - Analyse data management results
 - Config issues: GridFTP
- Consider extensions based on feed-back of VOs
- Support more VOs

- Migrate to Glue2
- Community users management:
 - Users database,
 - User life-cycle workflow: LFC, VOMS, EGI Apps Db
 - Apprehend the diversity of users
 - Identify users "behind" robot certificates
 - Foster resource usage acknowledgment
 - Track scientific production => *acknowledgement*
 - VT Scientific Publications Repository

A yellow sticky note is pinned to the top with two red pushpins. The text is written in a black, cursive font.

Thank
you

A yellow sticky note is pinned to the top with two red pushpins. The text is written in a black, cursive font.

Demo?
Come and
see me!

A yellow sticky note is pinned to the top with two red pushpins. The text is written in a black, cursive font.

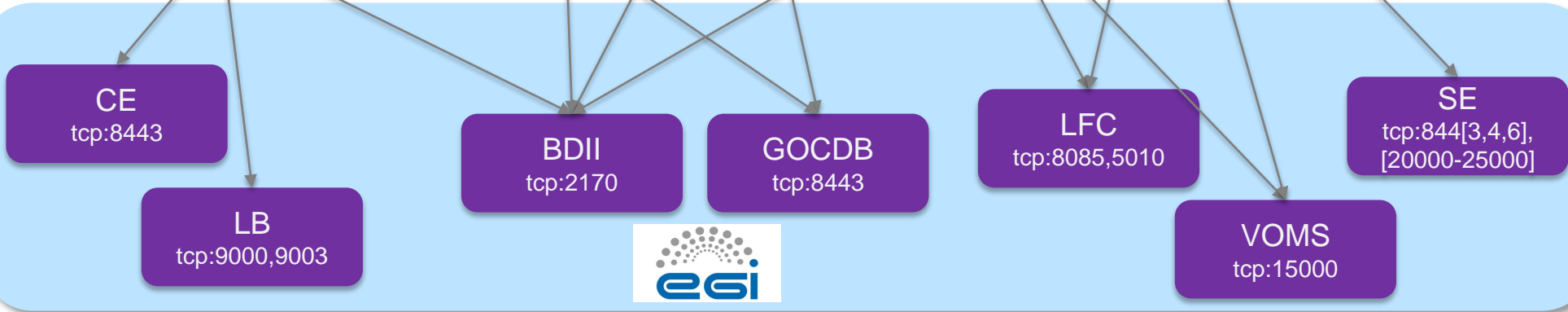
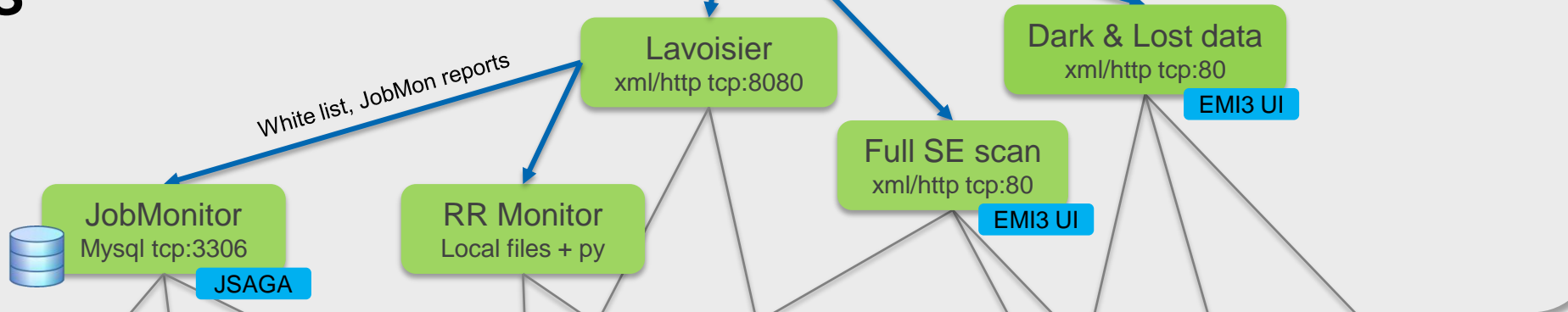
Questions
?

Backup slides

IN2P3 CC



I3S



Project duration	12 months, started April 2013
Budget	Total: 125 440€, requested: 52 200€ Engineer position recruited for 12 month at I3S
Project leader	Franck Michel, CNRS I3S
EGI contact	Diego Scardaci
Project wiki	https://wiki.egi.eu/wiki/VT_VAPOR
Mailing List	vt-vapor [at] mailman [dot] egi [dot] eu
Repository and tracker	https://redmine.i3s.unice.fr/projects/vapor

- Technical contributors:
 - CNRS I3S (FR)
 - IPHC (FR)
 - EGI Operations Portal team
 - CNRS Creatis (FR)
 - GRyCAP, Universitat Politècnica de València (SP)
- User communities: scope definition, test
 - biomed, France Grille VO, COMPCHEM, WeNMR, VLEMED

