

Grid Operations Centre

LCG Workshop

CERN 24/3/04



GOC Processes and Activities

- Coordinating Grid Operations

- Defining Service Level Parameters
- Monitoring Service Performance Levels
- First-Level Fault Analysis
- Interacting with Local Support Groups
- Coordinating Security Activities
- Operations Development
- Grid Accounting



- Within the scope of LCG we are responsible for monitoring how the grid is running – who is up, who is down, and why
- Identifying Problems, Contact the Right People, Suggest Actions
- Providing scalable solutions to allow other people to monitor resources
- Manage site Information definitive source of information
- Accounting Aggregate Job Throughput (per Site, per VO)
- Established at CLRC (RAL)
- Status of LCG2 Grid here:
- http://goc.grid-support.ac.uk/



GOC Site Database

- Develop and maintain a database to hold Site Information
- Contact Lists, Nodes, IP, URLs, Scheduled Maintenance
- Each Site has its own Administration Page where Access is Controlled through the use of X509 certificates. (GridSite)
- Monitoring Scripts read information in database and run a set of customised tools to monitor the infrastructure
- To be included in the monitoring a site must register its resources (CE,SE,RB,RC,RLS,MDS,RGMA,BDII,...)



RC



D.Kant@RL.AC.UK



CLRC EXAMPLE: RAL Site

People: Who do we notify when there are problems



Name Trevor Daniels	Description	Email address	Tel	Hours		
	Deployment Team Member	<u>t.daniels@rl.ac.uk</u>	+44 (0)1235 778093	0800 - 1700 (Mon-Wed)	Edit Delete	
Dave Kant	Deployment Team Member	<u>d.kant@rl.ac.uk</u>	+44 (0)1235 778178	0900 - 18.00 (Mon-Fri)	Edit Delete	
Martin Bly	Deployment Team Member	m.j.bly@rl.ac.uk	+44 (0)1235 446981		Edit Delete	
Steve Traylen	Deployment Team Member	s.traylen@rl.ac.uk	+44 (0)1235 446777		Edit Delete	
Andrew Sansum	Deployment Team Member	r.a.sansum@rl.ac.uk	+44 (0)1235 445863		Edit Delete	
John Gordon	Deployment Team Member	j.c.gordon@rl.ac.uk	+44 (0)1235 446574		Edit Delete	
Alistair Mills	(-)	a.b.mills@rl.ac.uk	+44 (0)1235 446084		Edit Delete	
Matt Thorpe	GOC DB Web Admin	<u>m.s.thorpe@rl.ac.uk</u>	+44 (0)1235 778178	0830 - 1700 (Mon-Fri)	Edit Delete	

[Click to add a new contact]



EXAMPLE: RAL Site

Node Information (Type, Hostname, IP Address, Group)



A	211.041404047				
	Туре	Hostname	IP Address	Group	
	LCFG	lcfg	130.246.183.111	LCG-1	Edit Delete
	MDS	lcgcs01	130.246.183.187	LCG-1	Edit Delete
	RB	lcgrb01	130.246.183.184	LCG-1	Edit Delete
	CE	lcgce01	130.246.183.182	LCG-1	Edit Delete
	SE	lcgse01	130.246.183.181	LCG-1	Edit Delete
	UI	lcgui01	130.246.183.183	LCG-1	Edit Delete
	BDII	logbdii	130.246.183.185	LCG-1	Edit Delete
	PROX	lcgrbp01	130.246.183.186	LCG-1	Edit Delete
200	RB	lcgrb02	130.246.183.189	LCG-2	Edit Delete
	CE	lcgce02	130.246.183.188	LCG-2	Edit Delete
	BDII	lcgbdii02	130.246.183.191	LCG-2	Edit Delete
	UI	logui02	130.246.183.194	LCG-2	Edit Delete
	UI	gppui04	130.246.183.172	EDG	Edit Delete
	CE	gppce05	130.246.187.142	EDG	Edit Delete
	RB	gppse05	130.246.187.140	EDG	Edit Delete
	RLS	gpprls05	130.246.187.153	EDG	Edit Delete
	MON	gpprg05	130.246.187.151	EDG	Edit Delete
	NM	gppnm06	130.246.187.145	EDG	Edit Delete

[Click to add a new node]



Monitoring Services

- There are many frameworks which can be used to monitor distributed environments
- MAPCENTRE http://mapcenter.in2p3.fr/
- GPPMON <u>http://goc.grid-support.ac.uk/</u>
- GRIDICE http://edt002.cnaf.infn.it:50080/gridice/
- NAGIOS http://www.nagios.org/
- MONALISA http://monalisa.cacr.caltech.edu/
- Example: Mapcentre 30 sites ~ 500 lines in config file (static version)
- Example: Nagios 30 sites, 12 individual config files with dependencies
- Developed Tools to Configure these services to make the job easier NAGIOS, MAPCENTER and GPPMON



GOC Features – GPPMon

Status of Grid, based on the success of job submission to resources, displayed as a world map, with sites represented by coloured dots

- SQL Query of Database -> List of Resources (CE, RB)
- Job Submission to each Site in Two Ways:

Direct to CE= globus-job-runIndirect to CE via Resource Brokers= edg-job-submit

- Responses Collected and Translated into a Site Status Colour Index Success via RB = Green, Globus Only = Orange, Fail = Red
- Geographical View Presented Against World Map



LCG2 CORE SITES Status: 23 March 2004 17.00



Site	History	via Globus	via RAL_lcgrb02
CERN	Graph	11:00:38 - 27/02/2004	11:10:34 - 27/02/2004
CNAF	Graph	11:02:14 - 27/02/2004	11:03:41 - 27/02/2004
FNAL	Graph	11:01:12 - 27/02/2004	11:01:30 - 27/02/2004
FZK	Graph	11:03:18 - 27/02/2004	11:04:30 - 27/02/2004
PIC	Graph	11:02:46 - 27/02/2004	11:04:19 - 27/02/2004
RAL	Graph	11:01:42 - 27/02/2004	11:03:11 - 27/02/2004
Taiwan	Graph	11:03:56 - 27/02/2004	11:05:42 - 27/02/2004





GOC Features – CERTIFICATE Monitoring

Status of host certificates

- SQL Query of Database -> List of Resources (CE, SE)
- Download Certificate using OPENSSL
- Responses Collected and Translated into a Site Status Colour Index

Lifetime > 1 Month	OK	- Green
1 Week < Lifetime < 1 Month	Warning	- Yellow
Lifetime < 1 Week	Critical	- Red

Geographical View Presented Against World Map



BNL

FNAL

TRIUME

Site	Label	Node	State	Lifetime (wks:days:hrs:mins:sec)
Torino	То	t2-ce-01.to.infn.it	ok	29 weeks, 1 days, 2:13:33
UAM	UAM	grid003.ft.uam.es	critical	0 weeks, 5 days, 3:13:32
UB	UB	lcg-ce.ecm.ub.es	ok	35 weeks, 0 days, 5:36:5
USC	USC	lcg-ce.usc.cesga.es	ok	32 weeks, 2 days, 5:41:23



GOC Features – Nagios Monitoring

Nagios is a powerfull monitoring service that supports notifications, and the execution of remote agents to correct problems when faults are discovered.

- Advantages => proactively monitor grid (NRPE daemon)
- Automatic Configuration of Nagios based on Database
- Developed a set of plugins which focus on service behaviour and data consistency
 - Do RBs find resources?
 - Does Site GIIS's publish correct hostname?
 - Is the site running the latest stable software release?
 - Does the Gatekeeper authentication service work?
 - Are the host certificates valid e.g Issued by Trusted CA
 - Are essential services running e.g GridFTP
- Further plugins are being developed (e.g certification)



Nagios Screen Shot

Service Summary for Nodes:

Certificate Lifetime Check, GridFTP, GRAM Authentication

Site Attributes via GIIS (siteName, Tag, ...)

HOST	PLUGIN	STAT	TUS			STATUS INFORMATION
Host ↑↓	Service ↑ ↓	Status 1	Last Check ↑↓	Duration ${\bf \uparrow}ullet$	Attempt ↑↓	Status Information
adc0015.cern.ch	Certificate Lifetime	ок	03-02-2004 13:27:59	7d 21h 49m 45s	1/3	Certificate expires: {36}wks,{0}dys,{0}hrs,{21}min,{34}sec
	<u>GIIS attr</u> <u>GlueCEInfoHostName</u>	ок	03-02-2004 12:41:39	7d 21h 11m 39s	1/3	GlueCEInfoHostName attribute is adc0015.cern.ch
	<u>GIS attrisiteName</u>	ок	03-02-2004 12:54:40	7d 20h 58m 7s	1/3	siteName is CERN-LCG1 dataGridVersion is LCG1-1_1_3
	GateKeeper Authentication Test	ок	03-02-2004 13:07:49	7d 21h 49m 29s	1/3	GRAM Authentication test successful
	GridFTP Service	ок	03-02-2004 13:28:00	7d 21h 31m 7s	1/3	GRIDFTP Test Pass
itlasgrid04.usatlas.bnl.gov	Certificate Lifetime	ОК	03-02-2004 13:28:00	7d 21h 30m 25s	1/3	Certificate expires: {46}wks;{0}dys;{0}hrs;{42}min;{29}sec
	<u>GIIS attr</u> <u>GlueCEInfoHostName</u>	CRITICAL	03-02-2004 12:43:00	7d 21h 48m 20s	1/3	IO::Socket::INET: connect: Connection refused
	<u>GIIS attr siteName</u>	CRITICAL	03-02-2004 12:55:32	7d 20h 57m 13s	1/3	O::Socket::INET: connect: Connection refused
	GateKeeper Authentication Test	ок	03-02-2004 13:08:30	7d 20h 43m 43s	1/3	GRAM Authentication test successful
	GridFTP Service	ок	03-02-2004 13:28:20	7d 21h 48m 9s	1/3	GRIDFTP Test Pass
	RRDTool	ОК	03-02-2004 13:34:17	7d 21h 9m 3s	1/1	GRAM Authentication test successful
						Dava Kant



CCLRC http://grid-ice.esc.rl.ac.uk/gridice

GOC is monitoring	;	Gric	es of the Grid				
Site view	VO view	Geo view	Gris view			Help	about
		Select Site	🗾 and/or I	Role 🔽	Show		

	Computing Resources							Storage Resources			
Site	Slot#	SlotFree	SlotLoad	RunJob	WaitJob	Power	CPU#	CPULoad	Available	Total	96
cern.ch	408	180	55 <mark>%</mark>	0	0	-	-	-	67.5 Gb	69.1 Gb	2%
cnaf.infn.it	-	-	-	-	-	-	-	-		-	-
cr.cnaf.infn.it	2154	1086	4 9 %	253	0	762647	387	51×	868.0 Gb	999.7 Gb	13%
fnal.gov	12	12	0%	0	0	-	-	-	-	-	-
fzk.de	-	-	-	-	-	-	-	-		-	-
gridka.de	-	-	-	-	-	-	-	-	-	-	-
gridpp.rl.ac.uk	438	273	37%	55	0	-	-	-	59.8 Gb	69.0 Gb	13%
grid.sinica.edu.tw	294	294	0%	0	0	-	-	-	-	-	-
hep.ph.ic.ac.uk	126	126	0%	0	0	-	-	-	9.2 Gb	16.8 Gb	<mark>4</mark> 5%
ifae.es	480	480	0%	0	0	433978	160	0%	5.6 Tb	22.4 Tb	25%
nikhef.nl	500	230	54%	137	13	-	-	-	1.4 Tb	1.7 Tb	20%
triumf.ca	4490	30	99%	0	0	-	-	-	729.1 Gb	731.1 Gb	0%
TOTAL	8902	2711	33%	445	13	1196625	547	26%	8.6 Tb	25.9 Tb	17%

← • → • ⑧ ◙ ゐ | ◎, ☜ ③ | ₽- ∌ 🖬 🗄



gridpp.rl.ac.uk

cgrb02.gridpp.rl.ac.uk	RB	UpTime: 1-	2:11	Reg.Op	enFiles.: 31	41	Socket	: TCP(29)) UDP(17) ^{FS NA}	PA Full
	Process	5		Instance	:S	С	PU	Merr	iory	Tir	ne
	Process Name	Status	Inst#	First	Last	1Max	All	1Max	Ávg	1Max	All
	condorg-schedul	er S	1	1-3:15	1-3:15	0	0	0	0	0-0:0	0-0:0
	condor-master	S	1	1-3:15	1-3:15	0	0	0	0	0-0:0	0-0:0
	fmon-agent	S	1	1-3:16	1-3:16	0	0	0	0	0-0:0	0-0:0
	ftp-server	S	1	1-2:16	1-2:16	0	0	0	0	0-0:0	0-0:0
	job-controller	S	1	1-3:15	1-3:15	0	0	0	0	0-0:2	0-0:2
brokor	local-logger	S	1	1-2:1	1-2:1	0	0	0	0	0-0:0	0-0:0
broker	local-logger- interlog	s	6	1-4:1	0-3:11	o	0	ο	O	0-0:3	0-0:3
	logging-and- bookkeeping	s	11	1-2:55	0-0:7	o	0	O	O	0-0:0	0-0:0
	log-monitor	S	1	1-3:11	1-3:11	0	0	1	1	0-0:2	0-0:2
	network-server	S	21	1-3:13	1-3:13	0	0	1	1	0-0:3	0-0:4
	proxy-renewal	S	4	1-3:16	1-3:14	0	0	0	0	0-0:0	0-0:0
	workload-manag	er S	4	1-3:16	1-3:16	0	0	1	1	0-0:4	0-0:8
cgce02.gridpp.rl.ac.uk	CE	UpTime: 7-	22:3	Reg.Op	enFiles.: 28	22	Socket	: TCP(38)) UDP(18) ^{FS NA}	PA Full
cgse01.gridpp.rl.ac.uk	SE	UpTime: 7-	22:34	Reg.Op	enFiles.: 12	87	Socket	: TCP(16)) UDP(19) ^{FS NA}	PA Full

Generated: Tue, 23 Mar 2004 16:12:58 +0000

GridICE Homepage

-

顫 – 🗗 ×

CCLRC A Second GOC Site

Second Centre to Provide Monitoring (Taipei = GMT + 8 hours)

GOC does a remote Database Query to collect a list of resources and monitors those resources

GOC tools packaged for distribution











LCG Accounting – How it Works



PBS filter to extract data from the event log records.

RGMA-API publishes data to a PbsRecords database table on the MON box and records the names of the processed logs for book-Dave Kant D.Kant@RL.AC.UK



"END" EVENT RECORDS CONTAIN THE FOLLOWING INFORMATION





The actual table schema contains more information that is shown here.



LCG Accounting – How it Works



Extract data from globus-gatekeeper and system messages logs

Record a list of files processed to reduce network traffic/load







The actual table schema contains more information that is shown here.



In order to match the authenticated user DN's to the corresponding jobs we need to process the system message logs.

Record ID : [GK] =/= Record ID [PBS]



Gatekeeper log

Messages log

PBS Event log



LCG Accounting – How it Works



GIIS filter to collect CPU performance benchmarks for the worker nodes from the subclusters attached to the CE.

RGMA-API publishes data to SpecRecords database table on the MON box







CPU Performance benchmarks for the worker nodes in the subclusters attached to the CE The actual table schema contains more information that is shown here.





3-Way join matches records and writes them to the LcgRecords Table. LcgRecords records are unique

Site now has a copy of its own accounting data.



LCG Accounting – How it Works



Data processed at each site is streamed to the GOC server

GOC has then aggregated information for all sites



CCLRC LCG Accounting – How it Works

GOC provides an interface to produce accounting plots "on-demand"

Total Number of Jobs per VO per Site

Total Number of Jobs per VO aggregated over all sites

Tailor plots according to the requirements of the user community







LCG Accounting – Summary

Local Testing Cycle
RPMS Packaging
Documentation (Technical / User) [DONE
Trial Deployment Released last week to two sites
Final Testing Cycle Two weeks
LCG2 Release April 2004



- PBS log processed daily on site CE to extract required data, filter acts as R-GMA DBProducer -> PbsRecords table
- 2. Gatekeeper log processed daily on site CE to extract required data, filter acts as R-GMA DBProducer -> GkRecords table
- Site GIIS interrogated daily on site CE to obtain SpecInt and SpecFloat values for CE, acts as DBProducer -> SpecRecords table, one dated record per day
- These three tables joined daily on MON to produce LcgRecords table. As each record is produced program acts as StreamProducer to send the entries to the LcgRecords table on the GOC site.
- 5. Site now has table containing its own accounting data; GOC has aggregated table over whole of LCG.
- 6. Interactive and regular reports produced by site or at GOC site as required.

CCLRC Accounting Issues

- 1. There is no R-GMA infrastructure LCG-wide, so most sites are not able to install and run the accounting suite at present. It is expected that R-GMA and the MON boxes will be rolled out in LCG2 soon after the storage problems are resolved. Until this happens the complete batch and gatekeeper logs will have to be copied to the GOC site for processing.
- 2. The VO associated with a user's DN is not available in the batch or gatekeeper logs. It will be assumed that the group ID used to execute user jobs, which is available, is the same as the VO name. This needs to be acknowledged as an LCG requirement.
- 3. The global jobID assigned by the Resource Broker is not available in the batch or gatekeeper logs. This global jobID cannot therefore appear in the accounting reports. The RB Events Database contains this, but that is not accessible nor is it designed to be easily processed.
- 4. At present the logs provide no means of distinguishing sub-clusters of a CE which have nodes of differing processing power. Changes to the information logged by the batch system will be required before such heterogeneous sites can be accounted properly. At present it is believed all sites are homogeneous.



- Getting site information in database
 When, who
- Monitor Multiple Grids (LCG1/2 -> Production, Development – Should we monitor each VO's RB/BDII?



Future of GOC

GOC needs to deliver a production quality service to the community round the clock so a third GOC should be setup somewhere in Canada or the US. Build real 24x7 operations support.

Extend the range of monitoring jobs (SRM, Registering Files verifying the information in the RLS,...) further develop the range of nagios plugins.

Automate notifications via nagios; proactively monitor grid.

Interact with experts to develop/implement more through trouble-shooting tools and to monitor/track problems in the deployment. A good place to start would be regular operations meetings – we want to do more than raise problems.

Working towards a site-status and diagnostic page to allow site administrators to change site information, get latest accounting statistics for that site, run a series of tests in real time against the site [viz certification testing on demand] and automatic report generation that is customised to the site.

Tailor GOC database to allow EGEE ROCs to monitor resources in a region.