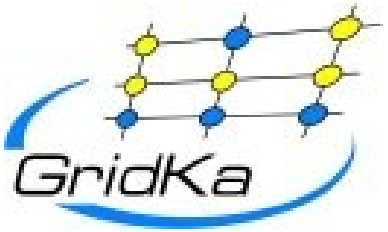
GridKa – Some feedback and suggestions…



Clemens Koerdt

Forschungszentrum Karlsruhe
in der Helmholtz-Gemeinschaft

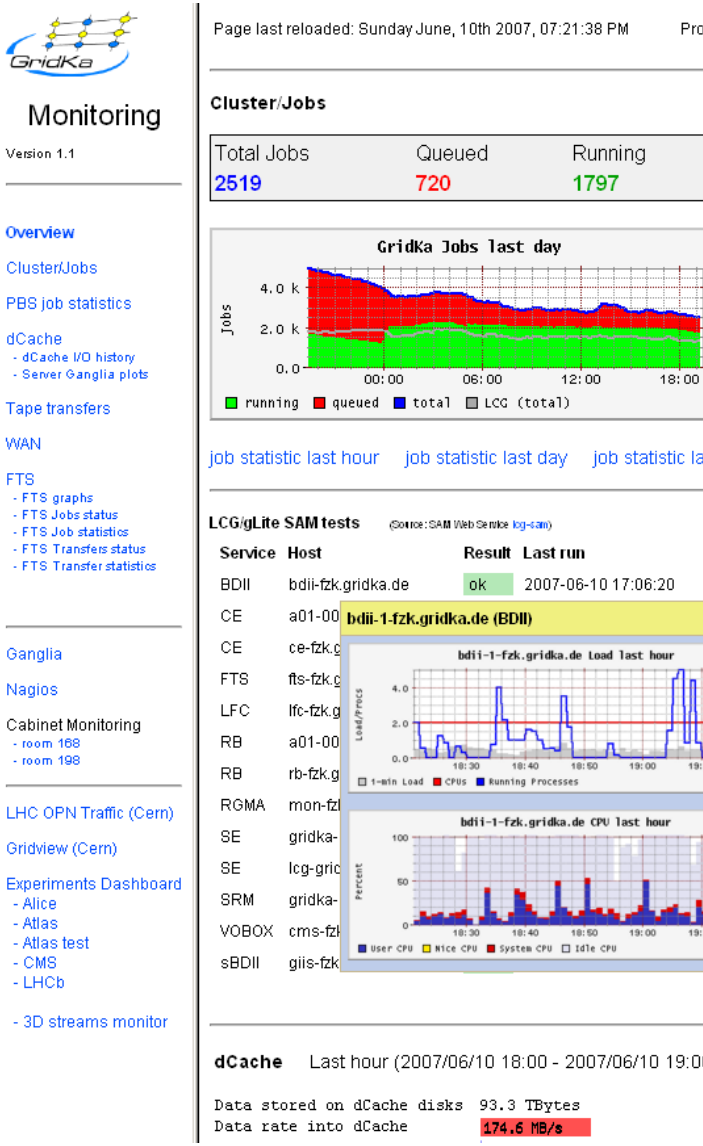
People involved in GridKa site management

- Primary Site Admin: Sven Gabriel
- Secondary Site Admin: Marc Hemberger
- dCache Admins: Doris Ressmann, Silke Halstenberg
- FTS Admins: Andreas Heiss, Silke Halstenberg
- Storage: Jos van Wezel
- PBS: Manfred Alef
- VOMS: Ingrid Schäffner
- Network: Bruno Höft
- Databases: Andreas Motzke
- Security: Ursula Epting, Bruno Höft
- ROC: Sven Hermann, Clemens Koerdt

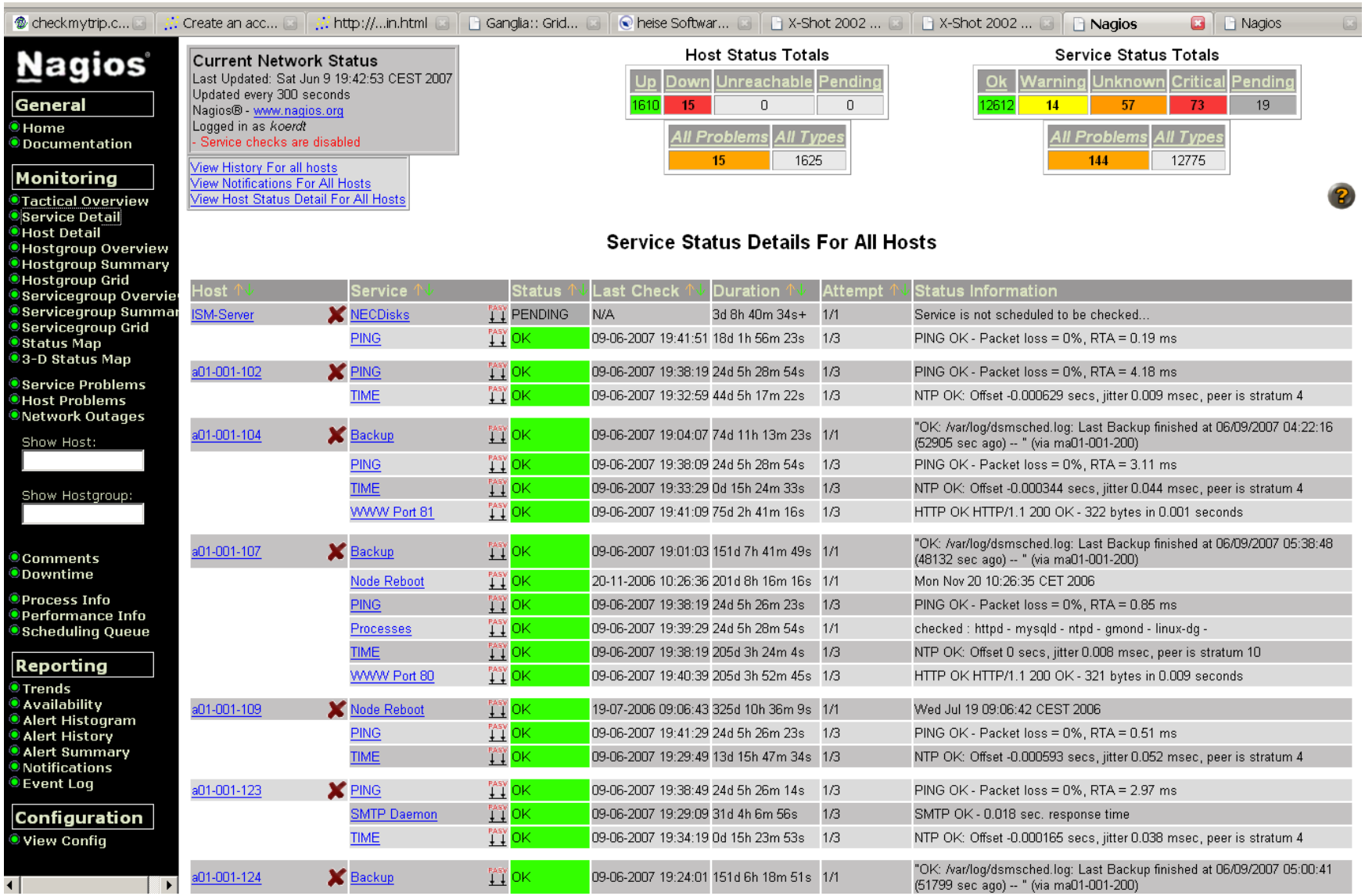
Overall responsible for GridKa: Holger Marten

Tools we use at GridKa

- Nagios (monitoring and alarms)
- Ganglia
- Cacti, LCG-OPN monitoring
- SAM/Gridview
- Gstat
- regional Site Functional Test server
- Experiment dashboards
- Collection of scripts for testing and parsing logfiles



Nagios



checkmytrip.c... | Create an acc... | http://...in.html | Ganglia:: Grid... | heise Softwar... | X-Shot 2002 ... | X-Shot 2002 ... | Nagios | Nagios

Nagios

General
- Home
- Documentation

Monitoring
- Tactical Overview
- Service Detail
- Host Detail
- Hostgroup Overview
- Hostgroup Summary
- Hostgroup Grid
- Servicegroup Overvie
- Servicegroup Summar
- Servicegroup Grid
- Status Map
- 3-D Status Map
- Service Problems
- Host Problems
- Network Outages

Show Host:

Show Hostgroup:

- Comments
- Downtime
- Process Info
- Performance Info
- Scheduling Queue

Reporting
- Trends
- Availability
- Alert Histogram
- Alert History
- Alert Summary
- Notifications
- Event Log

Configuration
- View Config

Current Network Status
Last Updated: Sat Jun 9 19:42:53 CEST 2007
Updated every 300 seconds
Nagios® - www.nagios.org
Logged in as *koerdt*
- Service checks are disabled

View History For all hosts
View Notifications For All Hosts
View Host Status Detail For All Hosts

Host Status Totals

Up	Down	Unreachable	Pending
1610	15	0	0

All Problems	All Types
15	1625

Service Status Totals

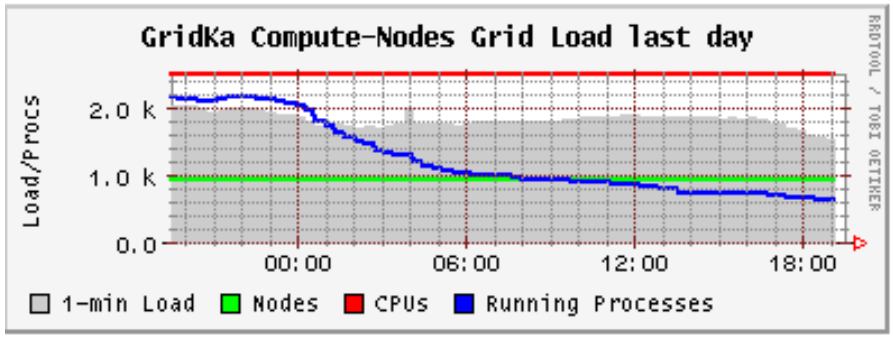
Ok	Warning	Unknown	Critical	Pending
12612	14	57	73	19

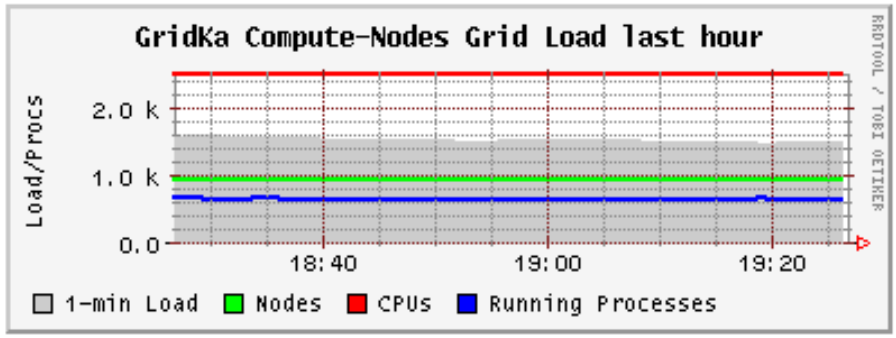
All Problems	All Types
144	12775

Service Status Details For All Hosts

Host	Service	Status	Last Check	Duration	Attempt	Status Information
ISM-Server	NECDisks	PENDING	N/A	3d 8h 40m 34s+	1/1	Service is not scheduled to be checked...
	PING	OK	09-06-2007 19:41:51	18d 1h 56m 23s	1/3	PING OK - Packet loss = 0%, RTA = 0.19 ms
a01-001-102	PING	OK	09-06-2007 19:38:19	24d 5h 28m 54s	1/3	PING OK - Packet loss = 0%, RTA = 4.18 ms
	TIME	OK	09-06-2007 19:32:59	44d 5h 17m 22s	1/3	NTP OK: Offset -0.000629 secs, jitter 0.009 msec, peer is stratum 4
a01-001-104	Backup	OK	09-06-2007 19:04:07	74d 11h 13m 23s	1/1	"OK: /var/log/dsmsched.log: Last Backup finished at 06/09/2007 04:22:16 (52905 sec ago) -- " (via ma01-001-200)
	PING	OK	09-06-2007 19:38:09	24d 5h 28m 54s	1/3	PING OK - Packet loss = 0%, RTA = 3.11 ms
	TIME	OK	09-06-2007 19:33:29	0d 15h 24m 33s	1/3	NTP OK: Offset -0.000344 secs, jitter 0.044 msec, peer is stratum 4
	WWW Port 81	OK	09-06-2007 19:41:09	75d 2h 41m 16s	1/3	HTTP OK HTTP/1.1 200 OK - 322 bytes in 0.001 seconds
a01-001-107	Backup	OK	09-06-2007 19:01:03	151d 7h 41m 49s	1/1	"OK: /var/log/dsmsched.log: Last Backup finished at 06/09/2007 05:38:48 (48132 sec ago) -- " (via ma01-001-200)
	Node Reboot	OK	20-11-2006 10:26:36	201d 8h 16m 16s	1/1	Mon Nov 20 10:26:35 CET 2006
	PING	OK	09-06-2007 19:38:19	24d 5h 26m 23s	1/3	PING OK - Packet loss = 0%, RTA = 0.85 ms
	Processes	OK	09-06-2007 19:39:29	24d 5h 28m 54s	1/1	checked : httpd - mysqld - ntpd - gmond - linux-dg -
	TIME	OK	09-06-2007 19:38:19	205d 3h 24m 4s	1/3	NTP OK: Offset 0 secs, jitter 0.008 msec, peer is stratum 10
	WWW Port 80	OK	09-06-2007 19:40:39	205d 3h 52m 45s	1/3	HTTP OK HTTP/1.1 200 OK - 321 bytes in 0.009 seconds
a01-001-109	Node Reboot	OK	19-07-2006 09:06:43	325d 10h 36m 9s	1/1	Wed Jul 19 09:06:42 CEST 2006
	PING	OK	09-06-2007 19:41:29	24d 5h 26m 23s	1/3	PING OK - Packet loss = 0%, RTA = 0.51 ms
	TIME	OK	09-06-2007 19:29:49	13d 15h 47m 34s	1/3	NTP OK: Offset -0.000593 secs, jitter 0.052 msec, peer is stratum 4
a01-001-123	PING	OK	09-06-2007 19:38:49	24d 5h 26m 14s	1/3	PING OK - Packet loss = 0%, RTA = 2.97 ms
	SMTP Daemon	OK	09-06-2007 19:29:09	31d 4h 6m 56s	1/3	SMTP OK - 0.018 sec. response time
	TIME	OK	09-06-2007 19:34:19	0d 15h 23m 53s	1/3	NTP OK: Offset -0.000165 secs, jitter 0.038 msec, peer is stratum 4
a01-001-124	Backup	OK	09-06-2007 19:24:01	151d 6h 18m 51s	1/1	"OK: /var/log/dsmsched.log: Last Backup finished at 06/09/2007 05:00:41 (51799 sec ago) -- " (via ma01-001-200)

Jobs monitoring

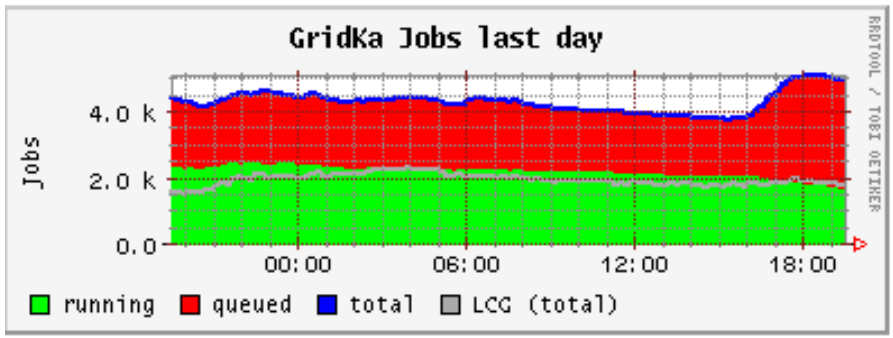


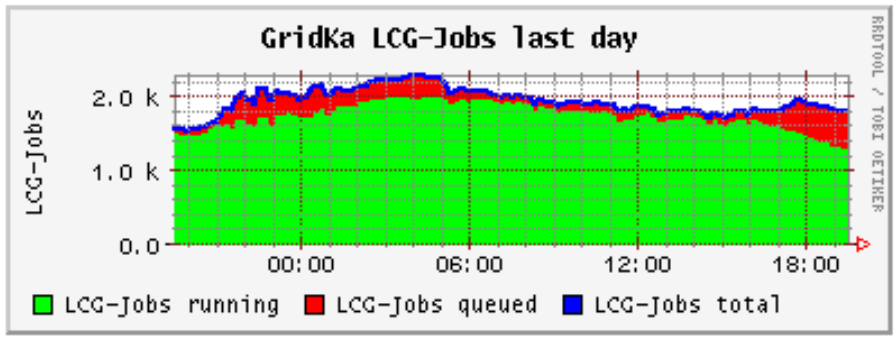


Ganglia main page

Total Jobs	Queued	Running
4925	3205	1718

LCG Jobs	Queued	Running
1 (0.0%)	1	0 (0.0%)





dCache – gridftp doors

Proxy cerificate used for gridftp and srm tests: valid

SRM door Update: every 15mins

Answer on port 8443: OK srmput/advisory-delete test: OK Last test finished: 09/06/07 19:47:10 (test history)

gridftp Server Update: every 5mins (write/read test every approx. 15mins).

Server	Port 2811 answer	write/read test	last test finished	last 20 tests (latest left)
f01-015-101-e.gridka.de	OK	OK	09-06-2007 19:45	
f01-015-102-e.gridka.de	OK	OK	09-06-2007 19:45	
f01-015-105-e.gridka.de	OK	OK	09-06-2007 19:45	
f01-015-106-e.gridka.de	OK	OK	09-06-2007 19:45	
f01-015-107-e.gridka.de	slow response	OK	09-06-2007 19:47	
f01-015-108-e.gridka.de	OK	OK	09-06-2007 19:45	
f01-015-109-e.gridka.de	OK	OK	09-06-2007 19:45	
f01-015-110-e.gridka.de	slow response	OK	09-06-2007 19:45	
f01-015-112-e.gridka.de	OK	OK	09-06-2007 19:45	
f01-015-113-e.gridka.de	DOWN	OK	09-06-2007 19:49	
f01-015-114-e.gridka.de	OK	OK	09-06-2007 19:45	
f01-015-115-e.gridka.de	OK	OK	09-06-2007 19:46	
f01-101-110-e.gridka.de	OK	OK	09-06-2007 19:45	
f01-101-111-e.gridka.de	OK	OK	09-06-2007 19:45	
f01-101-113-e.gridka.de	OK	OK	09-06-2007 19:46	

dCache statistics monitoring: OK

This inidcator is updated once per hour, approx. 15mins after each full hour. If no dCache stats raw file is sent by the dCache head node shortly after the full hour, an error is indicated. Possible reasons are: broken scp connection from the headnode to the webserver, very high load on the headnode or failure of the headnode or the dCache system.

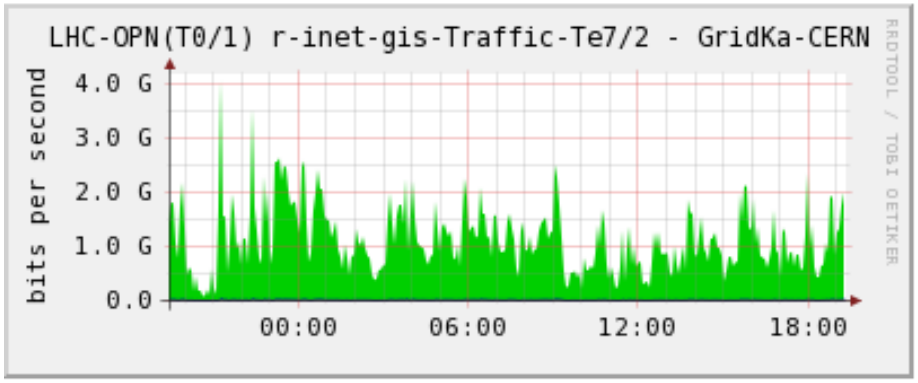
Data stored on dCache disks	97 TBytes	Data in disk-only pools	48.2 TBytes
No. of files on dCache disks	487848	Files in disk-only pools	393616

Last hour (2007/06/09 18:00 - 2007/06/09 19:00)

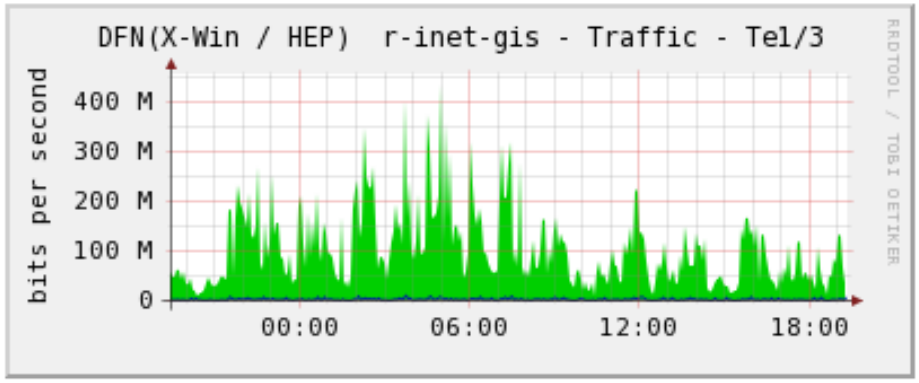
Filetransfers	839		
Errors	32		
Data into dCache [MBytes]	472467.1	Data rate into dCache	131.2 MB/s
Data out of dCache [MBytes]	48512	Data rate out of dCache	13.5 MB/s
Data written to tape[MBytes]	215025.4	Average tape write speed	59.7 MB/s
Data read from tape[MBytes]	13323.1	Average tape read speed	3.7 MB/s

LCG Optical Private Network Monitoring

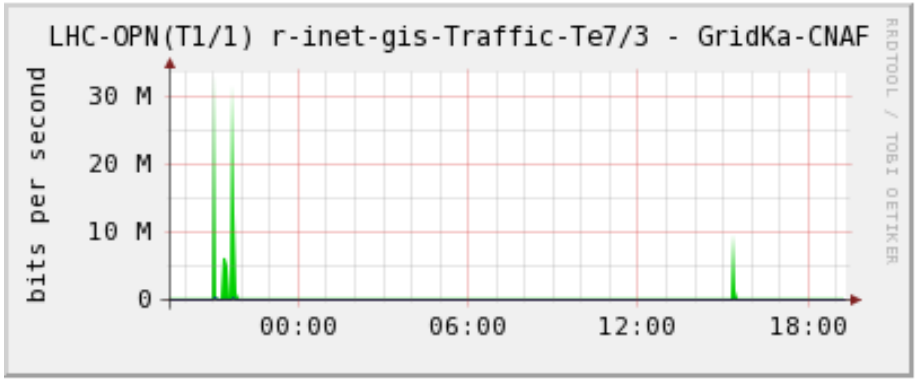
WAN traffic (last 24h) incoming, outgoing



LHC **O**ptical **P**rivat **N**etwork 10 Gbit/s link to CERN



10 Gbit/s Internet (DFN/X-WIN) uplink.



T1-T1 LHC **O**ptical **P**rivat **N**etwork 10 Gbit/s link to CNAF

currently not available

T1-T1 LHC **O**ptical **P**rivat **N**etwork 10 Gbit/s link to SARA/NIKHEF

Missing features to make work easier..

- **Advanced logging system**
  - like syslog (get it adopted by the software developers)
- **Adding of a stamp from local batch system to a job's ID to improve traceability of jobs from UI to WN**
- **Tool do to correlation statistics for difficult problems related to load etc.**
  - strange behavior often in connection with high load was detected from time to time.
  - Reasons are often difficult to find
  - Tool would help in identifying correlations and thus possible causes of malfunctioning
  - Probably useful in many environments….

examples of most frequent interventions

- **Storage Interventions (dCache) have been frequent**
  - Update to 1.7
  - Problems with SRM part
  - constant crisis management
- **Frequent interventions had also been necessary in connection with network**
  - but those have been short most of the time
- **Some interventions necessary for adding new hardware**
  - But mostly limited to particular services

- **(not yet seen, but at risk..): VOMS mapping if not planned with care**
  - 4 dimensions (VO, group, role, account) need to be mapped to one dimensional UID
  - number of sgm accounts?
  - later changes will lead to troubles (interrelations between CE, WN, PBS)
  - Owner of allocated memory is UID not user!

Most frequent unscheduled interventions

- **SRM instabilities**
  - Known fragility of component
  - Changes in load
  - For unknown reasons
- **Recently: problems with high load on CE**
  - Scaling problem?
  - Under investigation
- **Hardware Problems**
  - Limited in scope
- **Electric power cuts / cooling**
  - Human errors
  - Gathered some experience with unexpected behavior

Deployment planning

- **General deployment considerations**
  - Trying to setup core services redundantly,..
  - …gives fallback solution in case of instabilities after m/w service upgrades
  - gather as much experience as possible in PPS (if not in crisis management mode)
  - Still expect surprises especially when under higher load then usual
- **Ahead of deployment**
  - Planning is discussed weekly in local site admin meetings
  - Recommendation/need of experiment representatives (often embedded in local meetings) is very important for decisions taken
  - Study release notes thoroughly
  - Follow discussions on rollout list
  - Decide to deploy an update once you think you can properly estimate and handle the risk
  - Prioritize in consultation with user community and operations meeting recommendations
- **During and after the deployment**
  - sequential update
  - Rocks for WNs, yaim for grid server
  - Monitor extensively
  - Be ready to rollback
  - Central distribution of important software components, configuration

cross site issues

- operations meetings (wide overview and expertise)
- regional operations meetings (with very frank discussions)
- rollout list for time critical things and discussion with broader audience extremely helpful (a lot of details)
- GGUS search engine very useful (again details)
- workshops of course
- bilateral contacts with peers (indispensable)
- Specialized contacts, i.e. dcache user forum, specialized workshops
- …

COD – interactions and usefulness

- Alarm and ticket notifications are still crucial
- Correlation to central problems still unsatisfying (is being addressed already..)

Suggestions put forward by Gridka Site admins:

- spent more time on detecting/analysing central failures (also addressed..)
- One could think about COD people filtering information and sending broadcasts of detected cross site issues (if additional workload can be handled..)
- Interaction: from site perspective it might be better if teams would not change that often

Communications with the users

- **asking users to use GGUS ticketing system**
  - advantage: everybody can see what's going on
  - possible to get a quick overview of troubles
  - people taking more care of describing problem
  - responsibilities are clear
- **Some still prefer to use phone and other direct contacts**
  - expect this behaviour to increase with the coming LHC
  - can increase responsiveness..
  - , but hat disadvantages (see above)
  - in very urgent cases use at least a combination of both
- **High level meetings: GridKa Technical Advisory Board**
  - contact to user groups and very experienced people
- **embedded contact people represented in weekly GridKa operations meetings**
  - so far made very good experiences with that
- **Communication between Sites – Users could still be improved**
  - is going in the right direction (recent changes to ops meeting, etc..)

GridKa Computing Centre