

Date: 10/17/2008

Present: Gabriele & Andrew interviewing Burt

VO Questionnaire for the Metrics Correlation and Analysis Service (MCAS) project

- What monitoring tools do you use?

Doing different things depending on the roles (1) Tier 1 admin; (2) LPC support

1) Tier 1 admin: making sure that Grid jobs are going through.

1.1) We have home grown condor plots. Checking the occupancy of the cluster, rate of job flow at the job scheduling level. At this level, it is not easy to see if jobs succeed. This gives general information on the status by looking on whether the resource is empty or not.

1.2) Using Ganglia plots. Using Ganglia as a repository of information. For example, I'm adding the status of OSG BDII to gauge its health. However, I also keep the log file because ganglia keeps data in RRD. This has 2 problems:

(a) old data gets rotated out AND I want to keep data around for about 1 month for 2 reasons

(a.1) I want the data around for 2 weeks, to report status at the weekly operation meeting.

(a.2) If OSG BDII is failing, then this can get reflected in the monthly service availability numbers for USCMS: I want to be able to check on that at the end of the month. In this case, the granularity of ganglia (1 point every 5 mins) is enough.

(b) RRD does an average with the granularity of 5 minutes and I may want to have more granularity from time to time.

2) LPC support: these are the non-grid users. LPC cluster has a size $> 1/2$ grid cluster + bluearc. 2 monitoring tools:

2.1) Monitoring of bluearc is poor: if file access is slow, I don't know how to do troubleshooting. I've written ad hoc monitoring script, measuring latency of BlueArc writing 10GB file. I do not keep the logs of this (as opposed to 1.2) because keeping it over time is not critical.

2.2) Using standard condor monitoring,

- Can you walk me through a troubleshooting session?

1) LPC troubleshooting for BlueArc slow.

Problem: Shared file system is slow e.g. emacs takes 10sec instead of 2ms. Problem is noticed by interactive users only.

Realization of the problem: I look at plots of bluearc latencies OR I get an alarm from Zebics (check spelling).

Troubleshooting:

a) I try to identify if there is something that generates a high load from the nodes e.g. I look at the ganglia plots of the 2000 [...check or 200?] nodes.

b) Look at condor plots to see if anomalies show up... this rarely leads to results.

c) Log at the worker nodes and try to track it down with system level tools (ps, top, ...).

Solution: 2 cases analyzed

a) In one case, I could not find the problem.

b) In another case, 200 jobs were doing too many reads in BlueArcs. User logged on multiple nodes doing parallel read/write (exposes a problem in the robustness of BlueArc).

2) Troubleshooting Grid Jobs.

Problem: CMS analysis and production jobs have info there. Jobs fail.

Realization: looking at the dashboards.

Troubleshooting: Looking at Dashboard, RVS SAM site availability.

Potential benefits from MCAS: Integration between Dashboard w/ SAM may be beneficial (some of it may be already integrated). Not much integration on storage, if monitored at all.

- Any special needs on data representation? Do you like plots, tables...?

simple 1d plots may be useful most of the time e.g availability vs. time. Can check if jobs are failed at sites and site is not available.

Cannot think of useful correlation plot (e.g. bluearc latency vs. number of failed jobs ... not very useful)

- How much in real time should the information be?

Order of minutes. Typically, for networking info, the latency is 1 hour (from transfer db).

Critical systems within 2 minutes. For correlation with other data, 5-10mins could be sufficient: this is not clear.

Bluearc: getting info every 2 mins to prevent service failure interactively.

Grid jobs have longer time scales.

Ganglia plots are critical to have quickly e.g. load building on gatekeepers should have measurement with a frequency of the order of minutes.

NOTES:

- Burt will try to send the questionnaire to other members of USCMS and might have more info in 2 weeks

- Burt preferred the format of one-on-one interview, rather than filling in the questionnaire