Gratia End-of-Development Impact Report

Author: Brian Bockelman.
Delivered to the August 21, 2008 OSG Executive Team meeting.

Summary:

The development of Gratia, the accounting software used by the OSG, will transition from active development into maintenance and support as of the delivery of V1.0 (on or shortly after) October 1. The project will remain to deliver the remaining requirements for the local Fermilab site, and is likely to be staffed by Fermilab at the level of 0.5 FTE. This report explores impact this transition will have on the OSG. This note talks only about the software development issues, not operations issues. The OSG Operations area is responsible for operations of the OSG accounting repository. This is temporarily delegated to FermiGrid – there are discussions between Rob Quick and Keith Chadwick on plans for the longer term.

OSG documents 736 lists the agreed upon requirements for OSG expectations of the Gratia software.

Post-transition support of Gratia:

The transport layer and repository software are also used by the OSG RSV project. However, RSV does not foresee any support needs.

The switchover of support from Gratia development to FermiGrid happened officially in July 2008 after the release of the Gratia v0.36. The upgrade, which required an extensive DB migration, took a very long time to finish (about 1 month). The Gratia team **does not foresee another long database upgrade** for Gratia 1.0.

The OSG is funding ¼ of an FTE for Gratia operations and ¼ of an FTE for Gratia development during Year 3.

Remaining technical requirements:

If Gratia management chooses to make the October 1 deadline for halting of Gratia development inflexible, it is likely that some requirements will not be met. Requirements in danger of not making this cutoff date are:

- **Security of probes**. The security system is designed to allow probes to self-register with the Gratia collector; this prevents an attacker from uploading data under the name of a different probe. Unfortunately, the implementation suffers from a long-present flaw (registering any new probe requires the collector to restart). Because this portion of the code was contributed by outside consultants and it's unclear how extensive the problem is, the timeline and **manpower needed to solve is unknown**. The worst-case estimate is 1 month of developer time.

- o Mine Altuney or Brian Bockelman may be in the position to help technically on this task.
- **Summary tables for transfer information**. Summary information is critical for the ability to do transfer reporting, and it will be difficult to drive adoption of the transfer reporting if sites do not have any reports based on this data. **The Gratia team believes this will be in for 1.0.**
- A "handful" of specific reports. There are some reports OSG has previously requested, but has not made high-priority. These can be finished if requested by stakeholders. Each individual report takes a small amount of time, and could be done as necessary throughout Year 3 by the remaining ¼ FTE.
- **Record recording rate**. The OSG requirements state that a rate of 100 records / second must be achievable. The current hardware can hit around 25 records / second under completely optimal conditions, and the current load is 12k records an hour.
  - o The current rate will increase dramatically if sites start participating heavily in transfer reporting.
  - o The **target rate might not be achievable**, even considering future improved hardware deployments.
  - o One way to mitigate this performance problem would be a heavier emphasis on local collectors, and the ability to send only summary records to the central collector. This would allow the processing rate to stay the same even if the number of transfers reported significantly increases.

Going concerns:

There are a few issues which were not reflected strongly in the Gratia requirements document which the OSG will need to see resolved in Year 3.

- The requirements for **storage accounting information were not clearly defined** in the document. We did not specify what data needs to be gathered and if it needs to be summarized. Currently, dCache reports space reservations and the total amount of space available and used. This point needs to be revisited with stakeholders, as there is overlap with the GIP.
  - o If the storage accounting needs stay similar to the current design, there are minor probe bugfixes needed and collector work needed for summary tables. **Gratia development estimates this at 20 developer hours for minimal implementations**.
  - o OSG should update the requirements document with storage accounting information.
- We may need the ability to **upload only summarized data from a local collector to the central collector**. At the cost of being able to audit individual jobs or transfers, this would be a way to significantly reduce the number of records processed by the central collector. For transfer information, this may be an acceptable – or eventually necessary – compromise.

- OSG should resolve whether **central auditing requirements are necessary for transfers**.
- OSG or its stakeholders may have OSG-specific visualization needs (i.e., a certain graph type or appearance that Gratia might not provide). Those which can be implemented easily in the Gratia framework and are generic accounting graphs can be done within Gratia. **For specialized data views, the work may have to be done by OSG Metrics and Measurements**.
- The level of **integration between Gratia and OIM is ill defined**. OSG will need to examine what, if any, integration needs to exist, and will have use development time during Year 3. The reports that currently require VORS will be transitioned over to OIM before Gratia 1.0.