



OSG Requirements for the Gratia External Project

Version 4.3

June 6th, 2008

Abstract

This document gives the OSG requirements on the Gratia external accounting project that have been agreed to between the Gratia project sponsor (Database Department of the Fermilab Computing Division) and the OSG management. It is recognized that the OSG, US ATLAS, and US CMS - contributors to the project - may have other requirements outside of the Gratia project deliverables. These will be discussed in a further note.

Table of Contents

1. Introduction.....	1
2. Data Collected.....	1
3. Data Collection and Storage	4
4. Reporting.....	7
5. Security	9
6. Performance and Scalability	10
7. Maintenance and Operations.....	11
8. Out of Scope	12

1. Introduction

This document outlines the requirements for the accounting system from the OSG which have been agreed to by the Gratia software development project. It is an extension and update of OSG document #130-v3. OSG-130 provides a more thorough motivation and explanation for the accounting system components, while this one (OSG-736) focuses more on concrete requirements for the external Gratia project. In addition this document starts to address the requirements for operations and data retention - the maximum retention periods that can be supported by the software developed.

2. Data Collected

Req-1.	The accounting system must be able to track the usage of the OSG CE batch system and provide the information in a format compatible with the OGF JobUsageRecord format (See Req-39)
---------------	---

Explanation	<p>Users of the service typically request the running of one or multiple batch jobs. The resources consumed for providing the service are typically the utilization of a batch system (such as LSF, PBS, Condor ...) and of an N number of computing systems (called hosts or worker nodes).</p> <p>The accounting system needs to track for each batch job on each CE instance:</p> <ul style="list-style-type: none"> • User's DN and VOMS extensions (if used) • User's VO • Batch system ID • Wall Duration • CPU time • For batch systems which monitor it, the accounting system may track memory and swap. • Hosts utilized
Note	This is an original requirement.

Req-2.	The accounting system must be able to track the usage and utilization of the OSG CE from pilot-type jobs.
Explanation	<p>A pilot job is sent to a batch system by a central service for the VO, performs checks to make sure the worker node is usable, then downloads – “pulls” - a user job from the central server. Current pilot systems include GlideInWMS and PanDA; the current user-switching system is glExec. So, in addition to recording the batch information separately, the accounting system must be able to record the information specified in requirement 1.0 for each user job that the pilot job starts.</p>
Note	This is a new requirement.

Req-3.	The accounting system must collect the usage data of the OSG file transfer service and provide the information in a format compatible with the OGF JobUsageRecord format. (See Req-39)
Explanation	<p>Users of an OSG SE transfer files between sites, within sites, and from external hosts. For both auditing and accounting purposes, we need to keep track of operations occurring on OSG SEs.</p>
	<p>The storage accounting must collect, for each transfer:</p> <ul style="list-style-type: none"> • Site name • Time • Protocol • IsNew (True for writes, False for reads) • Remote hostname (we must be able to convert from hostname to

	<p>site name, but that might require information from outside the accounting DB).</p> <ul style="list-style-type: none"> • Transfer owner's DN • VO • Local pool • Remote pool • Connection Time • Transfer Size • Exit status • Initiating host • Destination filename
Note	This is an adaptation of an original requirement.

Req-4.	The accounting system must report resource utilization per grid user and VO role.
Explanation	<p>Reporting resource usage per user and VO is the only way the accounting system can provide detailed accounting information that can be use to support the missions of all the system users.</p> <p>The VO information is necessary because the same user might have different privileges based upon the current VO role they were using.</p>
Note	This is an adaptation of an original requirement.

Req-5.	The accounting system must calculate and retain summary data. The needed summary tables may change over time.
Explanation	<p>Summary data is the main data deliverable of the system and will be retained for later querying. Any new summary tables defined will be generated only for future data – prior data will not be available in newly defined summaries.</p> <p>The following information should be summarized for the data transfers:</p> <ul style="list-style-type: none"> • Site name • Time bin • Protocol • IsNew • Remote site (if available; if not, remote hostname) • Transfer owner's DN

	<ul style="list-style-type: none"> • Transfer owner’s VO • Number of failed transfers • Number of successful transfers • Sum of connection time • Sum of transfer sizes <p>The following information should be summarized for the CE:</p> <ul style="list-style-type: none"> • Resource name • Time bin • Owner’s DN • VO Name • Exit code • Resource type • Count of number of jobs • Sum of Wall duration • Sum of CPU user duration • Sum of CPU system duration
--	--

Req-6.	The Summary information shall have a variety of granularities with different retention tines.
Explanation	<p>Decisions on the retention and summary of information is reserved total for the OSG, but in most cases will follow the accounting system’s recommendations. We envision the following level of summarization for transfer information:</p> <ul style="list-style-type: none"> • Daily – for up to 30 days. • Weekly – for up to 2 years. • Monthly – for up to 20 years.

3. Data Collection and Storage

Req-7.	The accounting system must publish an interface so that a service or resource can report its usage across the network (push model) using a common XML format.
Explanation	<p>In the OSG model, each service should be responsible to push it usage information to the accounting system.</p> <p>The services are expected to adhere to the JobUsageRecord format as specified in Req-39.</p>

Req-8.	The accounting system must provide the ability to have a persistent data store for the accounting records, allowing configuration of the retention period.
Explanation	The raw data may be stored in a separate format than summary data. Access to the raw data may be as original data files. Summary data should be queryable.

Req-9.	The software infrastructure for the probes, collection and transport of the accounting records shall be reusable components in other OSG services. In particular, this is supported for the OSG Site Functional Testing infrastructure.
Explanation	The Gratia repository shall be able to accept new records that follow the OGF UsageRecord format. Extensions to the repository schema, and reporting of information when the software infrastructure is reused in this way is outside of the responsibility of the Gratia project. The project shall accept such changes only after a review of the technical and implementation details.

Req-10.	The accounting system must be able to recover from temporary errors without loss of data, or if data is lost be able to record the length of time data loss occurred.
Explanation	Expected reasons for the central OSG components include both downtimes due to unexpected database outage, upgrades, etc. The system should be able to recover from 4 days of downtime and require only 1 day to return to normal operations.

Req-11.	The accounting system software must be built to allow the retention of raw data and to retain summary data for up to 20years, some of which will be retained in long term storage and some directly accessible.
Explanation	Retention policies must be flexible enough for the OSG to set whatever retention policy they desire up to a maximum dictated by operational constraints. Typical retention times are: <i>In Long Term Storage:</i> Individual Job Records-1 year Individual Transfer Records-1 month <i>Accessible Online:</i>

	<p>Daily Job Summaries-2 month</p> <p>Weekly Job Summaries-2 years</p> <p>Monthly Job Summaries-Up to 20 years</p> <p>Hourly Transfer Summaries-168 hours</p> <p>Daily Transfer Summaries-30 days</p> <p>Weekly Transfer Summaries-Up to 20 years</p> <p>Monthly Transfer Summaries-Up to 20 years</p>
Note	This is a new requirement.

Req-12.	The accounting system must be able to accommodate future name changes for sites, DNs, and VOs.
Explanation	The user DN, site name, and even VO name is not static. User certificates may change DN formats, sites might change their names to reflect local organizational changes, and VO capitalization is often inconsistent between sites. The accounting system must be able to correct all of this information.

Req-13.	The system must allow data collection site to choose to retain, summarize, store and provide access to its local data, and provide means to forward filtered and/or summarize data to other repositories (VO or OSG wide).
Explanation	<p>Sites may want to run a collector if they have their own accounting needs; this requirement allows them to reuse the existing infrastructure.</p> <p>This requirement also allows them to store locally more information than is required by OSG and still not overload the infrastructure when forwarding the information (by sending it in summary form).</p>

Req-14.	The accounting system must provide a mechanism through which each resource provider can specify which accounting records should be reported and which should not.
Explanation	<p>The batch systems at the grid site may also be shared with local users. This means that sites must be able to filter out the records originating from local, not grid, use.</p> <p>Storage element transfers may also need to be filtered to remove non-grid movement.</p>

Req-15.	The accounting system must provide a mechanism to link an accounting record with the usage data (records of actions and events related to the
----------------	---

	user's resource usage) for up to 90 days from collection of the data, such that from a tabular view of the accounting record it is possible to see its detailed usage data.
Explanation	<p>Because the accounting information will be used by many users to evaluate contracts and resource allocation models, it is paramount that the users have a certain confidence in the produced accounting information.</p> <p>The accounting system will provide a mechanism to provide to external interfaces the accounting information (usage data, authentication information....) that is used to create the final accounting records.</p> <p>This requirement applies to both SE and CE data.</p>
Note	This is an original requirement.

Req-16.	The accounting system must provide a mechanism to store usage and accounting records for a certain time, which is configurable, and must have the ability to destroy the records after that time.
Explanation	Because auditing of the records will happen later than the data collection, the records must be kept and archived for the time duration agreed by the OSG, stakeholders and the site owners. Once the time period expires, the records must be destroyed.

Req-17.	The accounting system must have a way to identify and easily eliminate "bad" data from probes for 90 days
Explanation	As an example, the current Gratia probes sometimes report the unix timestamp as the Wall Duration on LSF batch systems, causing usage to be over-reported by hundreds of thousands of hours for that job. A system administrator should feel comfortable in removing these entries when necessary.
Note	This is a new requirement.

4. Reporting

Specific reporting requirements will change as the data provided by Gratia is used and reviewed. Evolving reporting requirements will be discussed and agreed to between OSG and the project.

<https://twiki.grid.iu.edu/twiki/bin/view/Accounting/ReportRequirements>

Req-18.	The accounting system must provide a web interface that OSG staff and Consortium members can use to query the accounting information. This interface should include controlled access so that some data is restricted to
----------------	--

	specific groups.
Explanation	The OSG and Gratia project collaborate on what pre-defined reports and data a user can access. A mix of statically generated and dynamically queryable reports should be provided.

Req-19.	Daily and weekly text based reports shall be provided and automatically sent to specified email lists.
Explanation	The OSG and Gratia project collaborate on what pre-defined reports and data are automatically generated and mailed.

Req-20.	The accounting system must provide the OSG with direct database access for summary data so the OSG can reuse the data in its own web interfaces.
Explanation	The OSG has a need to present its users (VO managers, grid users, and site admins) with a variety of data; only a part of this is accounting data. In order to provide a single webpage, the OSG will need to query the database directly.

Req-21.	The accounting system must provide an export function that allows exporting a selected number of accounting summary records in Windows Comma Separated format (for importing in Excel) and XML format.
Explanation	This feature will allow users to use various tools for accounting records processing and viewing.
Note	This is an original requirement.

Req-22.	The accounting system must allow for collections of sites to be defined for reporting purposes of summary information
Explanation	Reporting for simple sets of resources or sites shall be supported by the system, for example defining all sites that are ATLAS or CMS Tier-2s.

Req-23.	The system must report US ATLAS and US CMS accounting information to the WLCG.
Explanation	The WLCG has accounting requirements are currently: <ul style="list-style-type: none"> • Monthly reports of summary processing accounting data from Tier 1 and Tier 2 facilities • Retention of data until the report has been accepted by WLCG management (approximately 90 days). This should be well within

	<p>the retention policy set by the OSG in section Error! Reference source not found.</p> <ul style="list-style-type: none"> • Normalization of the accounting information based on information provided by each site.
Note	This requirement is new, but it clarifies and expands a specific portion of 16.0 in OSG-136.

5. Security

Req-24.	The accounting system must provide security to ensure the authenticity and reliability of the collected data. The accounting system must have a complete and accurate list of resources that it collects data from. The resource list not only list resource IP addresses, but also digital certificates in order to protect against attacks such as IP spoofing.
Explanation	If service probes are not authenticated and authorized: 1) a malicious user can corrupt entire accounting system. The attacker can upload corrupt data on behalf of legitimate OSG services. 2) service probes may mistakenly overwrite one another, leading to data corruption.

Req-25.	The accounting system must have the ability to protect accounting information from unauthorized read access.
Explanation	Accounting information contains detailed information about users and systems. Due to privacy laws in various nations, it is important to have authentication for information which reveals a users' identity.
Note	This is an original requirement.

Req-26.	The accounting system communication channels must be secure (they will support confidentiality and data integrity)
Explanation	In addition, international agreements require that user's accounting data must be encrypted when crossing national borders.
Note	This is an original requirement.

Req-27.	The accounting system communication channels from the probe to the collector must be authenticated.
Explanation	In light of the auditing requirements from the OSG Security Officer, it must be possible to verify the validity of the accounting data.

Note	This is a new requirement.
------	----------------------------

6. Performance and Scalability

Req-28.	The accounting system must minimize its interference with the services of which it measures the usage.
Explanation	<p>The accounting system may not significantly impact the performance of the services it measures. As a guideline:</p> <ul style="list-style-type: none"> • Periodically running tasks should only be running 5% of the time; a probe which normally takes 1 minute to run should only be run once every 20 minutes. • Probes which run as daemons should average no more than 5% of the CPU. • Probes should not wastefully utilize common resources. For example, if the probe queries a database, the queries ought to be reasonably optimized. If a periodic probe queries the batch scheduler for status information, it should only do this once or twice per run.

Req-29.	The accounting system must be able to track at least 1,000 services instances.
Explanation	Currently deployed grids have no more than 100 CE services running. However, with the addition of more services (glExec, SE, RSV), 1,000 reporting services is not difficult to imagine.
Note	This an original requirement.

Req-30.	The accounting system must be able to handle the input data of 100 usage records a second.
Explanation	<p>We believe this to be a realistic rate of data input into they system.</p> <p>The term “usage records” applies to both job usage and storage transfers.</p>
Note	This is a new requirement.

Req-31.	An accounting system repository must support up to 50 users querying the summary tables.
Explanation	The 50 simultaneous users might be from either the built-in web interface or querying the system via another web application.

	The users will not query the database directly.
--	---

Req-32.	The accounting system must respond to most queries in a reasonable amount of time (under 30 seconds). The accounting system database should respond to queries based solely upon summary information within 5s.
Explanation	Queries based on non-summary data are not supported. Access to non summary data is provided in collaboration with the OSG team on a best effort basis within the effort and time available. Collaboration with the OSG teams in further development may be required.

7. Maintenance and Operations

Req-33.	The accounting system must provide an operations administrative and monitoring interface.
Explanation	Because the accounting information will be used to evaluate VO-resource providers contracts and resource allocation models, it is paramount that users have a certain confidence in the produced accounting information. Lack of accounting information due to an internal failure must be distinguished from the lack of accounting information due to a service inactivity or failure.
Note	This is an original requirement.

Req-34.	The accounting system (that is the system deployable as a whole with multiple servers) must be structured so that a it is possible to build a system that does not have a single point of failure
Explanation	As a complete (hardware+software) the accounting system should be deployable in such a way that the system would not have a single point of failure.
Note	This is an original requirement.

Req-35.	The Gratia project must keep its planned feature list and schedules public and discuss them with the OSG, the operations teams and other stakeholders before final decisions are made.
Explanation	This allows the stakeholders to input their requirements and schedule needs to the Gratia project.

Req-36.	The accounting system must be well documented
Explanation	In order for the accounting system to be reused by others and widely accepted by site admins, running probes and collectors must be a well documented process. Documentation on how to add probes, store the information in the job records and access this information for probe specific summary tables must be included. Documentation should be well organized and easy to find.

Req-37.	Regression tests to validate new versions of the Gratia software shall be provided.
Explanation	It is expected that there will be many deployments of the Gratia software and repositories. New versions must be testable in the Integration Testbed and in site deployments.

Req-38.	New versions and updates of the software shall support in place upgrade and updates without loss of data in the repository
Explanation	It is expected that there will be many deployments of the Gratia software and repositories.

8. Others

Req-39.	The information stored and exchanged should be compatible with the OGF JobUsageRecord but is allowed to deviate from the standard when it is missing features that are necessary to support one of this documents' other requirements.
Explanation	For example the current standard does not have a proper way to express the VO to which the usage belongs and does not have a proper way to express summary data. These deviations should be fully documented.

9. Out of Scope for Gratia Project

Not currently a requirement:

	The accounting system must provide a mechanism to allow the removal of personal information
--	---

Explanation	Other collaborations (such as the EGEE) work in countries with very stringent data retention laws. It may be necessary to remove private information, such as user DN, from Gratia data if the probe is run at non-US sites.
Note:	This is not yet in the implementation plan pending a definite request and an understanding of the requirements from the WLCG. Implementing encryption of data transferred will be available. We need to check the requirements of existing OSG off-shore sites.

Addition to Req 6:	The Gratia project has not agreed to the OSG request to have hourly data transfer summaries stored and accessible
Explanation	We envision the following level of summarization for transfer information: <ul style="list-style-type: none"> • Hourly (data transfers only) – for up to 168 hours.

Revisions:

Version number	Date	Initials of author making revisions	Summary of changes
4.0	March 28 th , 2008 April 10th	RP, IM	For agreement with project. Security concerns remaining to be addressed: How do we ensure that the collected accounting data is authentic? Do we have any protection over the probe code ? How can we prevent if a site wants to change their accounting records ? 1)site can try to submit fake jobs but this is not easy and can be caught because site owner do not have VO certificates. Or 2) site owner try to tamper with gratia probes. To prevent this for example, do we do any cross-examination: checking the jobs submitted from the acclaimed VOs ? I understand not all VOs only submit jobs from portals, but pilots do so cross-check is easy there. And some VO indeed use submission portals
3.0	January 22, 2008	BB	Initial Draft
3.0.1	February 5, 2008	BB	Incorporate suggestions from RP
3.0.2	February 6, 2008	BB	Incorporate comments from PC; updated TOC.
3.0.3	Feb 8, 2008	BB	Cleanup comments and minor clarification about summaries.
4.3	June 6, 2008	PhC	Clarify that the collector to collector connection must support sending only summary data. Add Clarification of “compatible with JobUsageRecord” (Req-39)