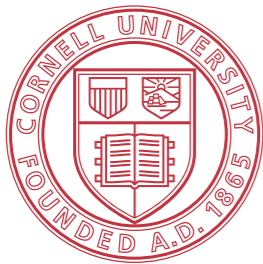


The dCacheBillingAggregator

Gregory J. Sharp
Daniel S. Riley



Cornell University
Laboratory for Elementary-Particle Physics



Overview

The dCache file system manages storage for many OSG SEs, including CMS and Atlas tier 1 & 2 sites.

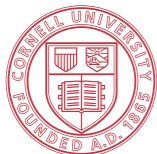
dCache records various interesting transfer statistics in its billing database that are of interest to OSG.

- Write a **dCache billing aggregator** to run at each dCache site.
 - *Read new entries from the dCache billing data base.*
 - *Send relevant statistics to a central site for analysis and display.*

Gratia summarizes and displays the accounting data.

- *Original scope: write a standalone aggregator, analysis and display system.*
- *Revised scope: write a dCache billing data probe for the Gratia system*

Work done with the dCache and Gratia teams, and Abhishek Singh Rana (UCSD).



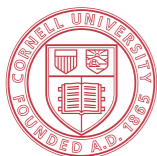
Why?

Track dCache grid resource usage.

Predicting future resource needs.

Detecting abnormal traffic patterns.

- ▶ *System faults*
- ▶ *Security violations*



How it works

The aggregator is a daemon written in python.

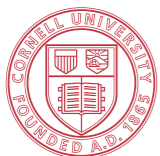
- *Different use case for the Gratia probe library.*

Typically runs on dCache DB host.

- *Accesses the dCache DB using SQLAlchemy and psycopg2.*

The steps are:

- *Check billinginfo and doorinfo tables of dCache Billing DB for new entries.*
- *Package them and send them to Gratia.*
- *Pause T seconds, then repeat.*



Status

Prototype active at UCSD since Jan 07.

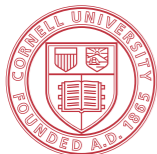
Performance is adequate, thus far.

Found some bugs.

- e.g., *GridFTP reads did not record an initiator in the database.*

Incorporated into VDT dCache distribution.

- *Awaiting deployment of some patches to dCache and Gratia.*



Future Work

Web interface to statistics.

- Negotiating responsibility.

Add any additional statistics requested.

Ensure aggregator performance is adequate for high-traffic sites.

Further code review of dCache components?

