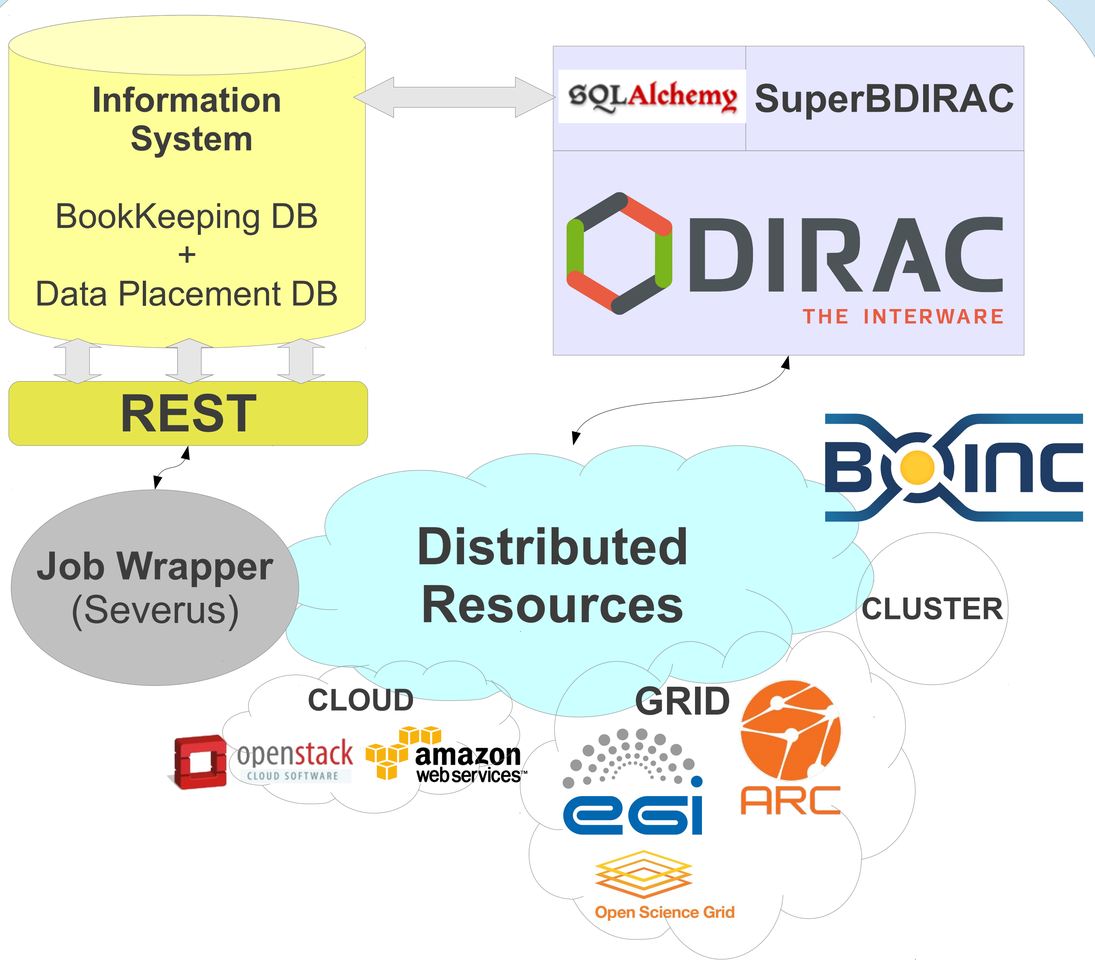
# Dirac integration with a general purpose bookkeeping DB: a complete general suite for distributed resources exploitation

Contributors: F Bianchi[a,b], M Chrzaszcz[c,d], V Ciaschini[d], M Corvo[e], C De Santis[f,g], N De Simone[f,t], D Del Prete[h], A Di Simone[f,g], G Donvito[i], A Fella[l,m], P Franchini[d], F Giacomini[d], A Gianelle[e], R Grzymkowski[d], S Longo[e], S Luitz[o], E Luppi[p,q], M Manzali[p,q], M Rama[r], G Russo[h], S Pardi[h], L Perez Perez[l], B Santeramo[i,s], R Stroili[e], L Tomassetti[l,m,q], M Zdybal[d]

a) University of Torino, Turin, Italy b) INFN – Sezione di Torino, Turin, Italy c) Physik-Institut, Universität Zürich, Zürich, Switzerland d) Henryk Niewodniczanski Institute of Nuclear Physics Polish Academy of Sciences, Krakow, Poland d) INFN CNAF, Bologna, Italy e) INFN – Sezione di Padova, Padua, Italy f) INFN, Sezione di Roma Tor Vergata, Rome, Italy – g) Department of Physics, University of Rome Tor Vergata, Rome, Italy h) INFN – Sezione di Napoli, Naples, Italy i) INFN, Sezione di Bari, Bari, Italy I) INFN - Sezione di Pisa, Pisa, Italy m) Department of Mathematics and Computer Science, University of Ferrara, Ferrara, Italy o) SLAC – Menlo Park – CA - USA p) Department of Physics, University of Ferrara, Ferrara, Italy q) INFN - Sezione di Ferrara, Ferrara, Italy r) INFN LNF – Frascati, Italy s) Department of Physics, University and Polytechnic of Bari, Bari, Italy t) Department of Electronic Engineering, University of Rome "Tor Vergata", Rome, Italy

## Abstract

In the context of High Energy Physics computing field the R&D studies aimed to the definition of the data and workload models have been carried on and completed by the SuperB community beyond the experiment life itself. The work resulted of great interest for a generic mid- and small size VO to fulfill Grid exploiting requirements involving CPU-intensive tasks. We present the R&D line achievements in the design, developments and test of a distributed resource exploitation suite based on DIRAC. The main components of such a suite are the information system, the job wrapper and the new generation DIRAC framework. The DB schema and the SQL logic have been designed to be able to be adaptive with respect to the VO requirements in terms of physics application, job environment and bookkeeping parameters. A deep and flexible integration with DIRAC features has been obtained using SQLAlchemy technology allowing mapping and interaction with the information system. A new DIRAC extension has been developed to include this functionality along with a new set of DIRAC portal interfaces aimed to the job, distributed resources, and metadata management. The results of the first functionality and efficiency tests will be reported.




**Simple, standard and long term solution**

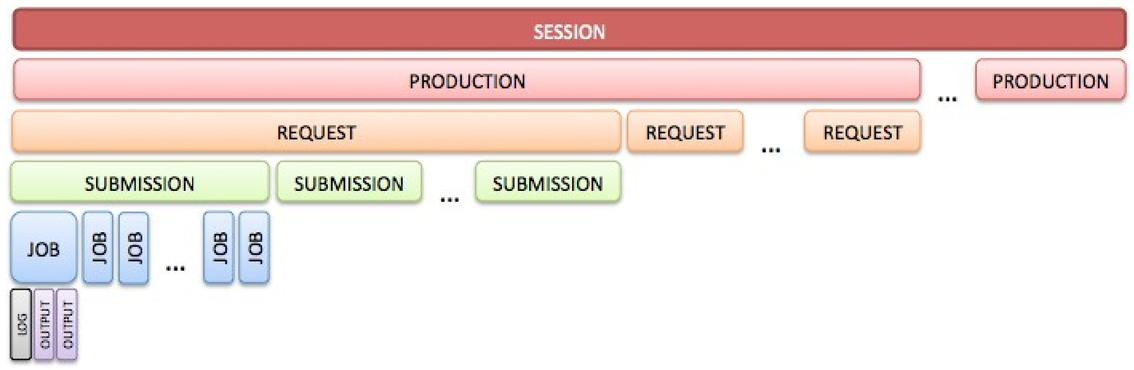
**Flexible enough to fulfill needs of a generic VO**

## Information System – SBK database

- **Relies on PostgresSQL 9.1**
- **Exploits hstore fields, trigger and PL/pgSQL language**
- **Normalized structure: NF1, NF2 and NF3 compliant**
- **Capable to sustain high transaction load (900 operations/sec)**

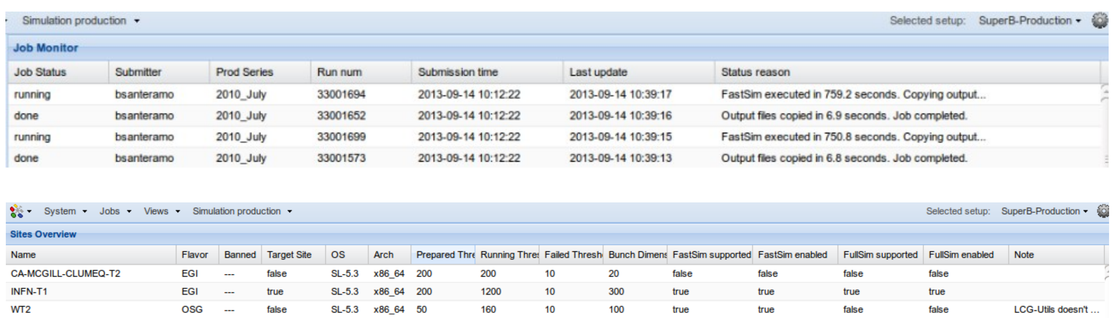
## SBK entities

- **Session**: defines a simulation (parameter and software)
- **Production**: Session subset to produce all needed to simulate a scenario (eg. Background in detector)
- **Request**: Production subset – set required events to simulate
- **Submission**: final step is job submission
  - Scheduling
  - Re-submission
  - Monitoring




## SuperBDIRAC

- **DIRAC extension**
  - **SQLALchemy layer**
  - **Integration with Information System**
  - **Severus compatible**
- **DIRAC webportal extension**
  - **Site monitoring**
  - **Simulation job monitoring**



Simulation production — Selected setup: SuperB-Production

Job Monitor

| Job Status | Submitter | Prod Series | Run num | Submission time | Last update | Status reason |
|---|---|---|---|---|---|---|
| running | bsanteramo | 2010_July | 33001694 | 2013-09-14 10:12:22 | 2013-09-14 10:39:17 | FastSim executed in 759.2 seconds. Copying output... |
| done | bsanteramo | 2010_July | 33001652 | 2013-09-14 10:12:22 | 2013-09-14 10:39:16 | Output files copied in 6.9 seconds. Job completed. |
| running | bsanteramo | 2010_July | 33001699 | 2013-09-14 10:12:22 | 2013-09-14 10:39:15 | FastSim executed in 750.8 seconds. Copying output... |
| done | bsanteramo | 2010_July | 33001573 | 2013-09-14 10:12:22 | 2013-09-14 10:39:13 | Output files copied in 6.8 seconds. Job completed. |



System — Jobs — Views — Simulation production — Selected setup: SuperB-Production

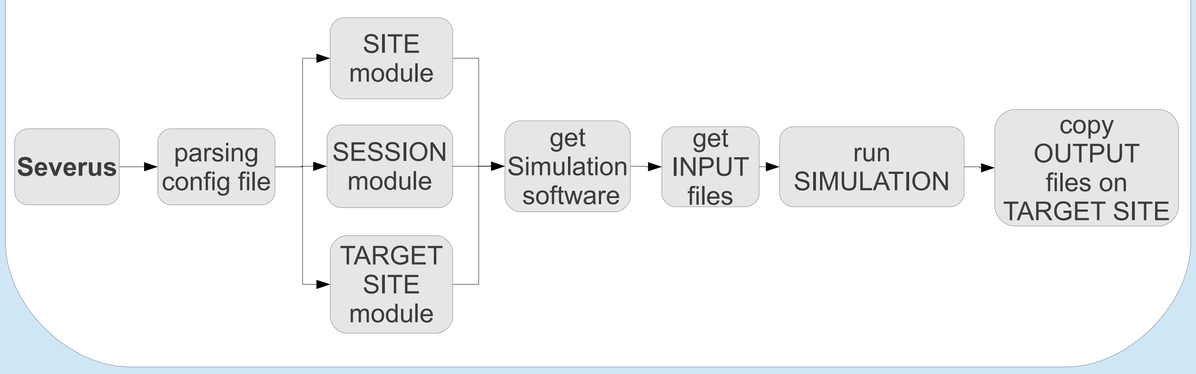
Sites Overview

| Name | Flavor | Banned | Target Site | OS | Arch | Prepared Thr | Running Thres | Failed Thresh | Bunch Dimens | FastSim supported | FastSim enabled | FullSim supported | FullSim enabled | Note |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CA-MCGILL-CLUMEQ-T2 | EGI | --- | false | SL-5.3 | x86_64 | 200 | 200 | 10 | 20 | false | false | false | false | |
| INFN-T1 | EGI | --- | true | SL-5.3 | x86_64 | 200 | 1200 | 10 | 300 | true | true | true | true | |
| WT2 | OSG | --- | false | SL-5.3 | x86_64 | 50 | 160 | 10 | 100 | true | true | false | false | LCG-Utils doesn't ... |

## Severus Job Wrapper

**Configuration file define:**

- **OPTIONS**
- **SOFTWARE**
- **REST**
- **SITE**
- **TARGET SITE**
- **INPUT**
- **OUTPUT**
- **EXPORT VARS**
- **SESSION NAME**




**CHEP2013: October 14 - 18, 2013 - Amsterdam, The Netherlands**
**contact email: bruno.santeramo@ba.infn.it**

## References:

[1] http://diracgrid.org - [2] http://www.postgresql.org/ - [3] http://www.sqlalchemy.org/ - [4] http://boinc.berkeley.edu/ - [5] SuperB Technical Design Report, http://arxiv.org/abs/1306.5655
[6] Fielding R T 2000 Architectural Styles and The Design of Network-based Software Architectures , PhD Thesis,University of California Irvine
[7] A.Fella, E.Luppi, L.Tomassetti A General Purpose Suite for Job Management, Bookkeeping and Grid Submission. International Journal of Grid Computing & Applications (IJGCA) Vol.2, No.2, June 2011. DOI: 10.5121/ijgca.2011.2202.