

# 2011 Project proposal

## General distributed production system for medium and small VO

Armando Fella on behalf of  
SuperB Distributed Computing Group

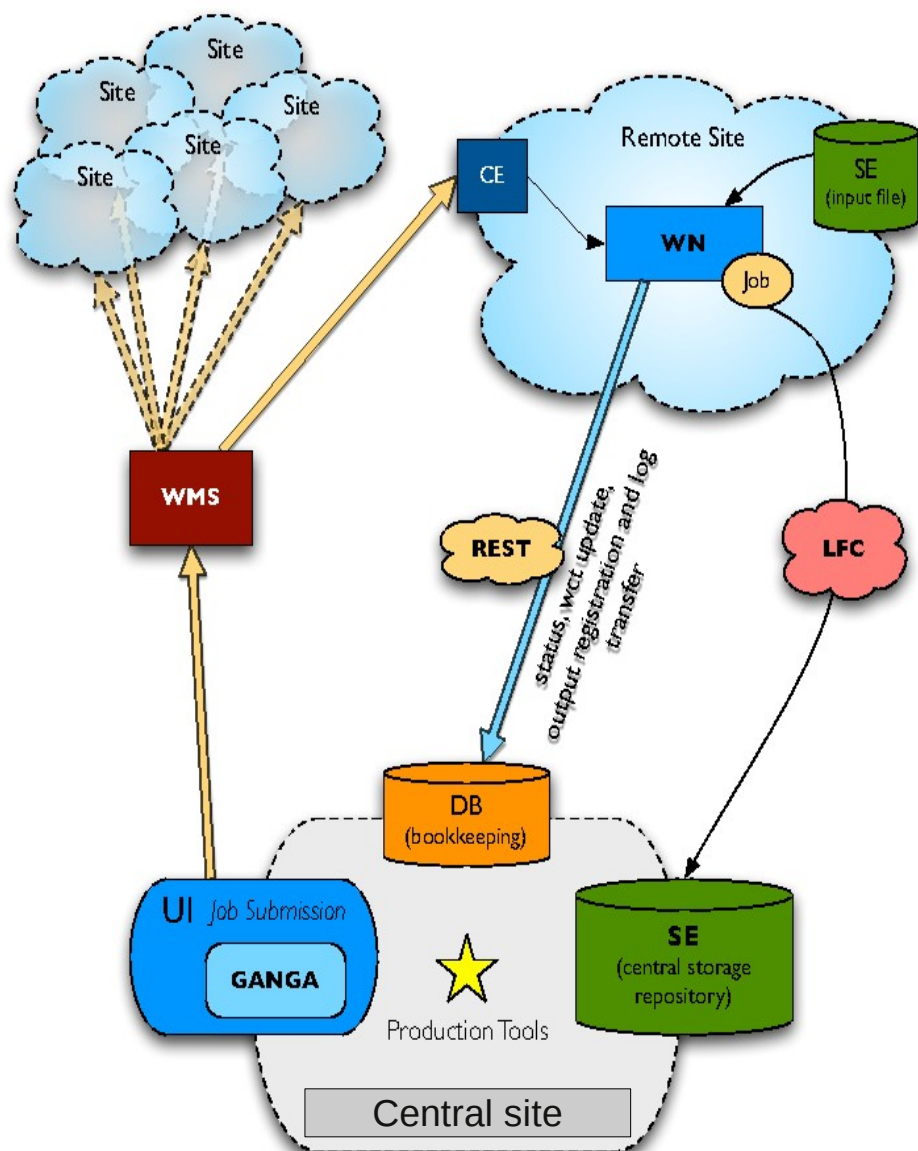
# Intro

- Capitalization of SuperB experience
  - Starting from the distributed production system developed for SB detector and machine simulation arriving to a platform capable to accomplish more general requirements
  - Target: small and medium VO size
- Minimize the platform requirements in terms of HW and human resources, custom development effort, Grid service customization and configuration
- Keywords: standard and simple

# Requirement

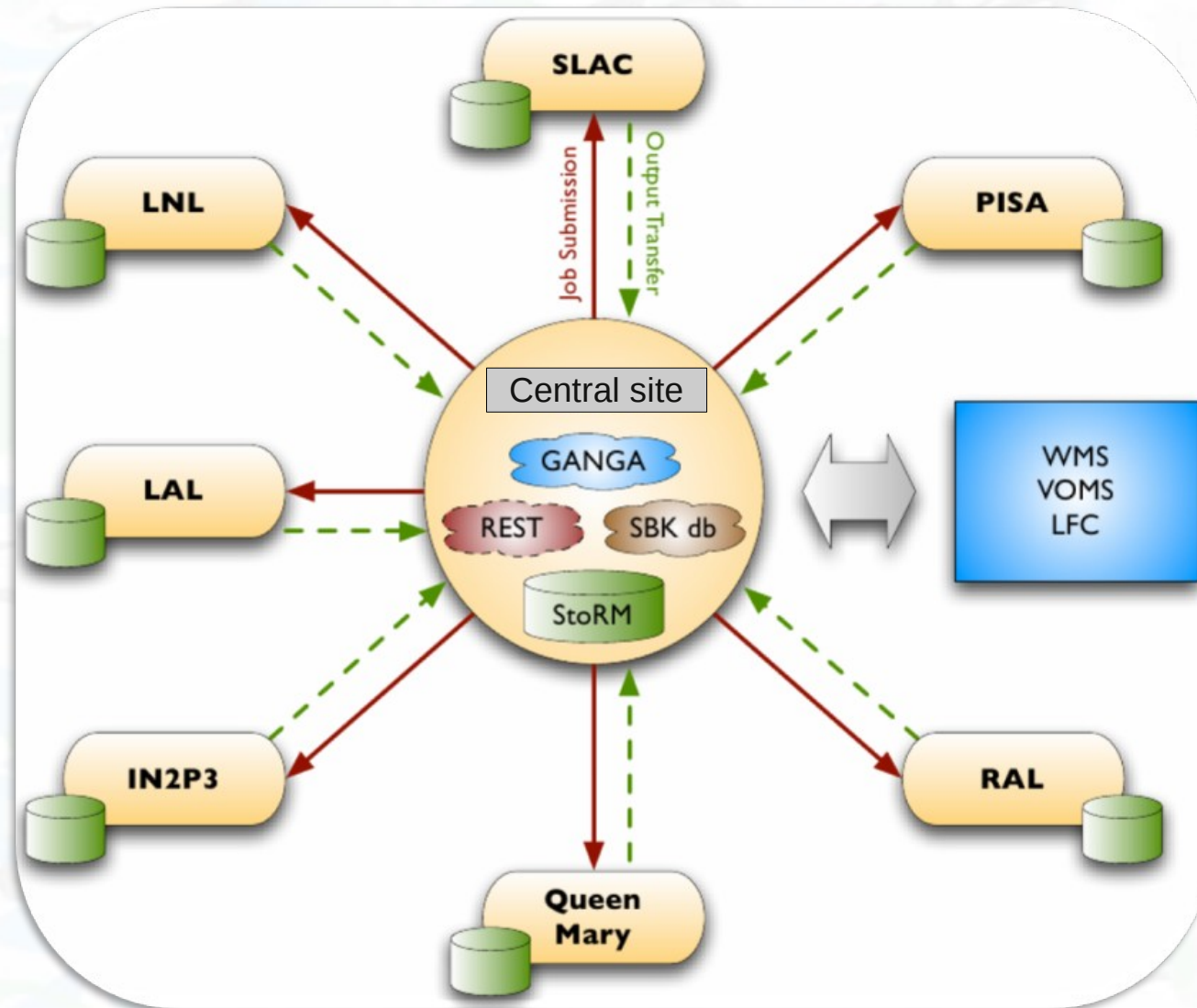
- One head node:
  - gLite UI, GANGA, MySQL, Apache, ProdTools (web interface)
- LFC server, VO shared, standard, per site, name space
- VOMS, default configuration
- WMS, per site CE:queue filter
- LCG-Utills:
  - Job input transfer from site SE to WN
  - Job output transfer to “Central site” (CNAF for SuperB)

# Job workflow

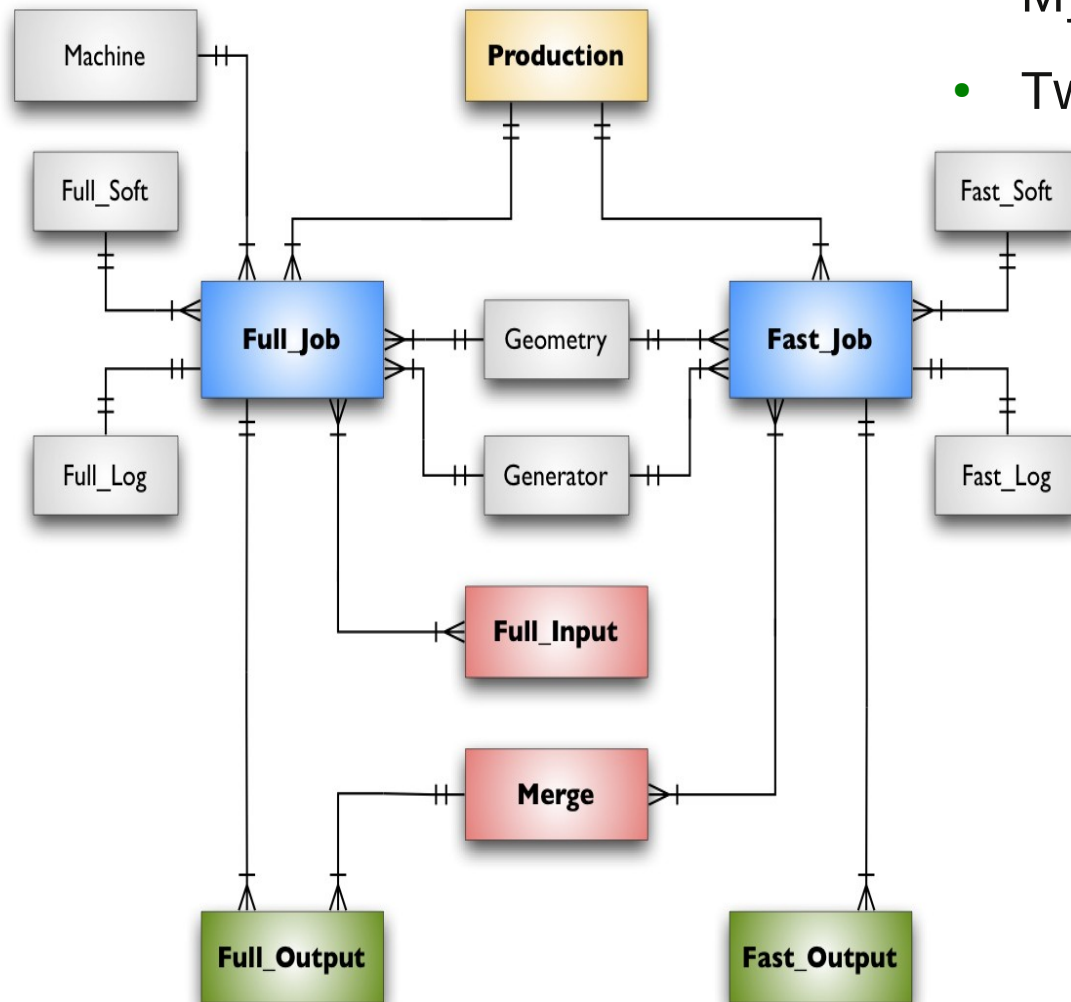


- Pre production operations:
  - Transfer job input files and proto VO sw release (if necessary) to site SEs (LCG-Utils)
- Job submission via prodtools (web interface on head node):
  - An authorized user should launch the submission script from UI --> **need automation**
  - GANGA bulk submission via WMS
- Jobs communicate status and log info to Bookkeeping DB via REST
- Jobs transfer output files to central repository (LCG-Utils)

# General design

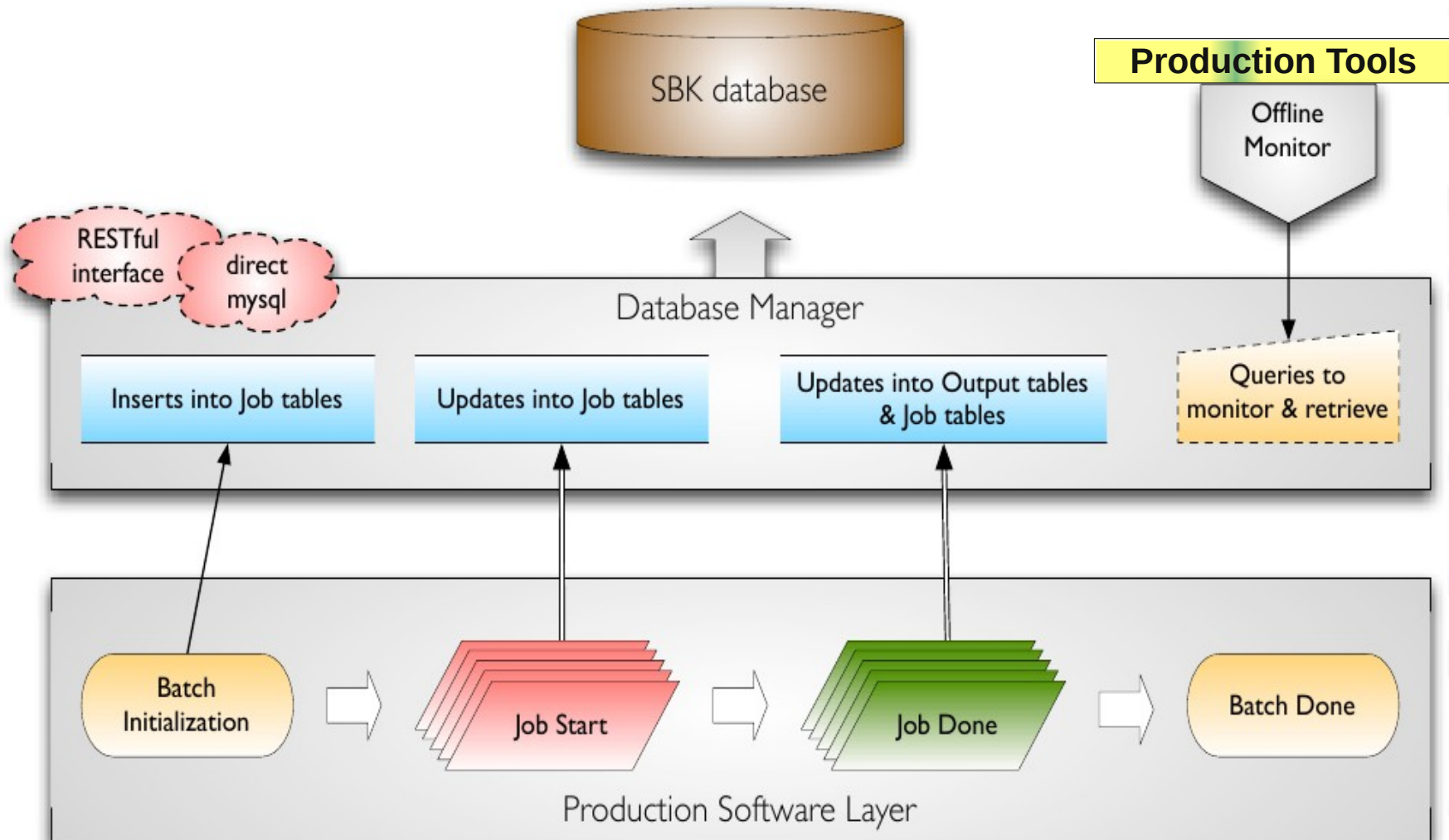


# Bookkeeping DB



- MySQL/RESTfull arch
- Two possible schema
  - Simple: only one production flavour, no merging flux metadata
  - Two productions: first output merging as input for second

# Bookkeeping-Job interaction



# System capability

	Sept. '09	Feb. '10	Jul. '10
Analysis stream	2	5	6
job done, failure rate	5K, 10%	20K, 8%	160K, 10%
Number of event	$2.25 \times 10^8$	$1.6 \times 10^9$	$8.6 \times 10^9$
Involved site	1	9	15
WallClockTime	6 years	19 years	195 years
Disk occupancy (TB)	0.5	5	25
Peak job running	500	2500	7000

- 99% of failure rate due to Grid services or sites malfunction



# July '10 production results

## Tier-1 sites:

INFN-T1 - CNAF Bologna, Italy  
IN2P3-CC - Lyon, France  
RAL-LCG2 - Oxford, UK

## Tier-2 sites:

UKI-LT2-QMUL - London, UK  
UKI-SOUTHGRID-RALPP - London, UK

GRIF - Orsay, Paris, France

SLAC - Stanford, CA, USA  
CIT\_CMS\_T2B - Caltech, Los Angeles, CA, USA  
VICTORIA-LCG2 - UVIC, Victoria, CA

INFN-PISA - Pisa, Italy  
INFN-LNL-2 - Legnaro, Italy  
INFN-BARI - Bari, Italy  
INFN-NAPOLI-ATLAS - Napoli, Italy  
INFN-CAGLIARI - Cagliari, Italy  
INFN-TORINO - Torino, Italy

- Site usage:
- **15 sites on 3 Grid flavours:**
  - **EGI, OSG, WestGrid**

ALL SITES		
Status	# of jobs	events
done	161 671	10 292 680 000
failed	4 171	285 730 000
running	3	210 000
sys-failed	11 391	790 780 000
timeout	136	9 500 000
Total	177 372	11 378 900 000



# Production tools

- **production creation:**
  - bookkeeping DB initialization for a new production;
- **monitor system:**
  - include a parametric search engine on job metadata, a job log file analysis subsystem, a production report generator, a general set of production result graphs per site and job status and a job analysis tool etc.
- **submission interface for shift takers interaction:**
  - permits the automatic submission to all the available sites.
- **submission interface for expert interaction:**
  - permits a fine grain selection of job submission parameter.
- **production request form:**
  - parametric production requests insertion interface.
- **Integrated elog system:**
  - semi automatic elog system collecting information about shift actions.

# To do list and ideas

- Collect information about generic VO requirements in term of job WCT, job I/O, Storage Element disk occupancy, bookkeeping metadata
- General design review in the optic of suppressing SuperB specifics
- Add configuration interfaces (Admin portal)
- Add automatic submission by production tools layer: authentication bridge layer permitting apache to use Grid resources
- Discuss the inclusion of the multi output repository concept into the design:
  - Jobs can be submitted to the sites where output data should be stored
  - Could we add the parameter “site” per physics channel simulation requests?
- In opposite, discuss the inclusion of Mass Transfer Service (FTS) into the design permitting the data transfer in post production phase
- Bookkeeping DB adaptation, ER schema dynamical modification?