# Summary

- SuperB overview
- SuperB distributed computing needs
- SuperB: current development status
- What we are looking for
- Testbed DIRAC infrastructure for SuperB
- Testbed results
- DIRAC vs WebUI
- Current Production System
- Customizing DIRAC for SuperB: SuperBDIRAC
  - Porting Production System under DIRAC
  - Using Severus via DIRAC
  - SQLAlchemy class for Bookkeeping database
  - Extending Webportal
- Comments
- Request for features
- Credits

SuperB overview

- International collaboration
  - Italy, France, UK, Poland, USA, Canada, Russia, Mexico, Brazil

- HEP experiment

- B factory

- Highest integrated luminosity

- Quite similar experiments:
  - LHCb
  - Belle II
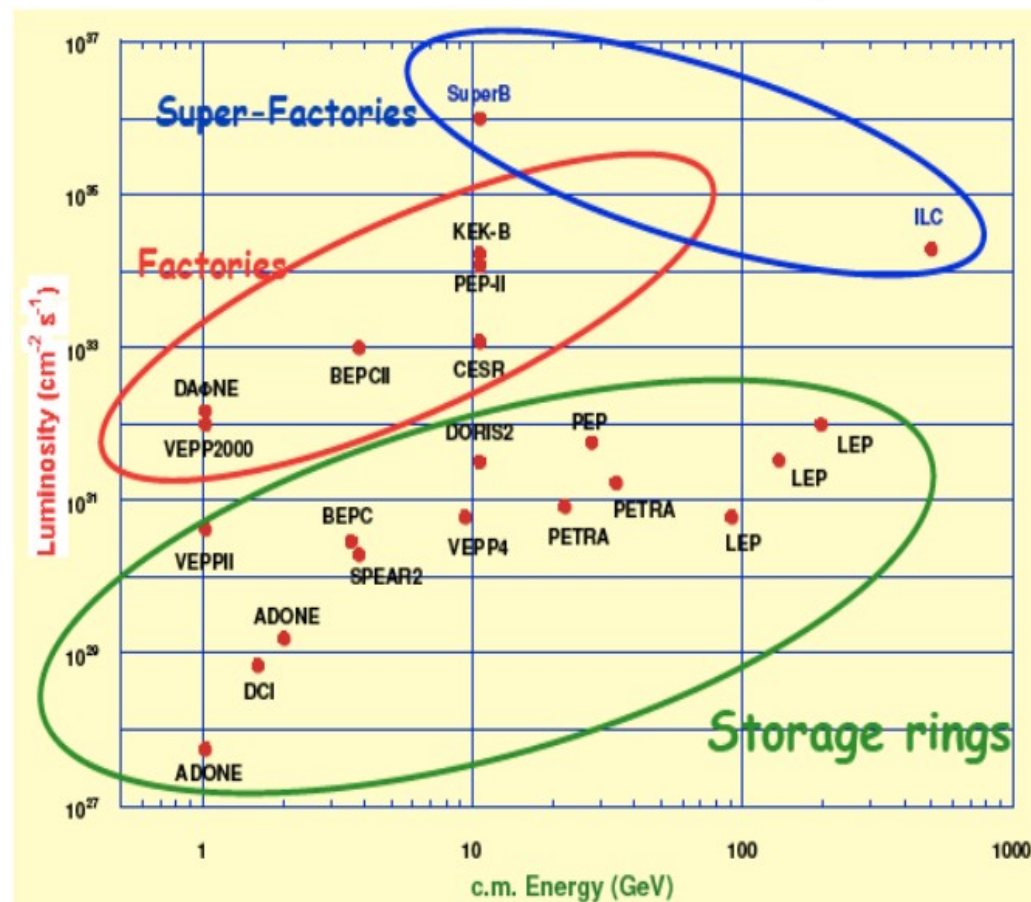
Asymmetric B-factory	detector	luminosity	Integrated luminosity
PEP-II	BaBar	10^{34} cm^{-2} s^{-1}	1 ab^{-1}
KEKB	Belle	10^{34} cm^{-2} s^{-1}	1 ab^{-1}
SuperKEKB	Belle II	8×10^{35} cm^{-2} s^{-1}	50 ab^{-1}
SuperB	BaBar (upgraded)	$3,4 \times 10^{36}$ cm^{-2} s^{-1}	75 ab^{-1}

3

SuperB distributed computing needs

- 5GByte/s of network bandwidth as experiment output
- 1000 PB of data collected in 5 years
  - Mass data transfer tool needed
- 50-60K CPU cores for first year data taking

- Interaction with EGI and OSG sites
- 30 sites currently involved in 7 countries

- Official productions yet done (not updated list)

	September 2009	February 2010	July 2010
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
Wallclock time	6 years	19 years	150 years
Disk occupancy (TB)	0,5	5	25
Peek job running	500	2500	7000

4

SuperB: current development status

- Detector based on BaBar, but need upgrades
  - final design not yet ready
  - Detector TDR just terminated, will be published asap
- Will be built in new founded Cabibbo Lab (Tor Vergata - Italy)

- Simulation and Analasys software based on BaBar framework
  - R&D activities for improvements
- Final design of computing model not yet ready
  - Data model to be finally established
  - Fully distributed environemnt (grid based)
  - A working production system already exists, but...
    - We have time to evaluate if better solutions available

What we are looking for

- MC production tool

- Analysis tool

- Mass data management tool

- User management tool

- Easy solution for "final users"

- Best choice shoul be an instrument able to manage all these aspects
  - Minimizing human effort

- Can DIRAC accomplish all these tasks ?
  - Preliminary evaluation says YES
- Now we are in the second evaluation phase:
  - Porting MC production under DIRAC

Testbed DIRAC infrastructure for SuperB

- 2 sites involved: INFN-T1 (aka CNAF) and INFN-BARI

host	OS	UI	DIRAC	HW	purpose
bbrbuild01.cr.cnaf.infn.it	SL 5.4	gLlite 3.2.0	v6r5p5	64bit VM, 2GB ram, 47GB disk space	production
sb-serv04.cr.cnaf.infn.it	SL 5.4	gLlite 3.2.0	v6r5p6	64bit VM, 2GB ram, 17GB disk space	SuperBDIRAC development
gridtest-05.ba.infnf.it	SL 5.8	EMI 2.1.0-1	v6r5p7	64bit VM, 2GB ram, 20GB disk space	SuperBDIRAC installation procedure

- New physical server at INFN-T1 expected in next weeks
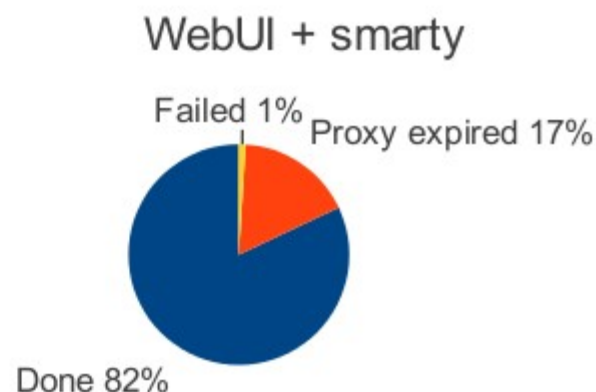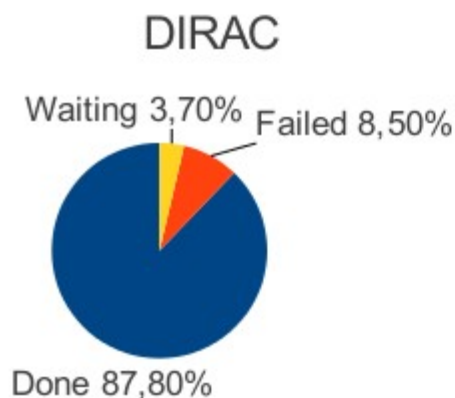 – Will host a DIRAC production instance

- Job submission
  - Normal and filling mode
  - Parametric jobs
  - Stagein and stageout (Sandbox, Ses)
  - DIRAC API

- Data management (DIRAC File Catalog)
  - Upload and register, replicate and delete files
 - Metadata
  - FTS data transfers

DIRAC vs WebUI

- In a comparative test between DIRAC and WebUI (current SuperB production system), DIRAC performance a bit better (despite no optimization nor customization) – [November 2011]

DIRAC

Waiting 3,70% Failed 8,50%

Done 87,80%

WebUI + smarty

Failed 1% Proxy expired 17%

Done 82%

site	WebUI Smarty				DIRAC			
	done	proxy expired	failed	effieciency	done	waiting	failed	efficiency
CYFRONET-LCG2	100	0	0	100,00%	100	0	0	100,00%
IN2P3-CC	100	0	0	100,00%	20	0	80	20,00%
IN2P3-LPSC	33	67	0	33,00%	100	0	0	100,00%
INFN-BARI	97	0	3	97,00%	100	0	0	100,00%
INFN-FRASCATI	99	0	1	99,00%	59	37	4	59,00%
INFN-LNL-2	99	0	1	99,00%	100	0	0	100,00%
INFN-MILANO-ATLASC	100	0	0	100,00%	100	0	0	100,00%
INFN-PISA	100	0	0	100,00%	100	0	0	100,00%
INFN-T1	97	3	0	97,00%	100	0	0	100,00%
UKI-LT2-QMUL	0	100	0	0,00%	99	0	1	99,00%
TOTAL	825	170	5	82,50%	878	37	85	87,80%

9

Current Production System

- ## WebUI (Custom portal)
  - https://bbr-serv09.cr.cnaf.infn.it:8443/~webui_sbk5
  - access distributed resources for fastsim e fullsim
- ## based on sbk5 PostgreSQL bookkeeping database
- ## Ganga to perform job submission
  - Severus Wrapper
- ## Severus – python wrapper to perform all production related task
  - Stagein and stageout via lcg commands
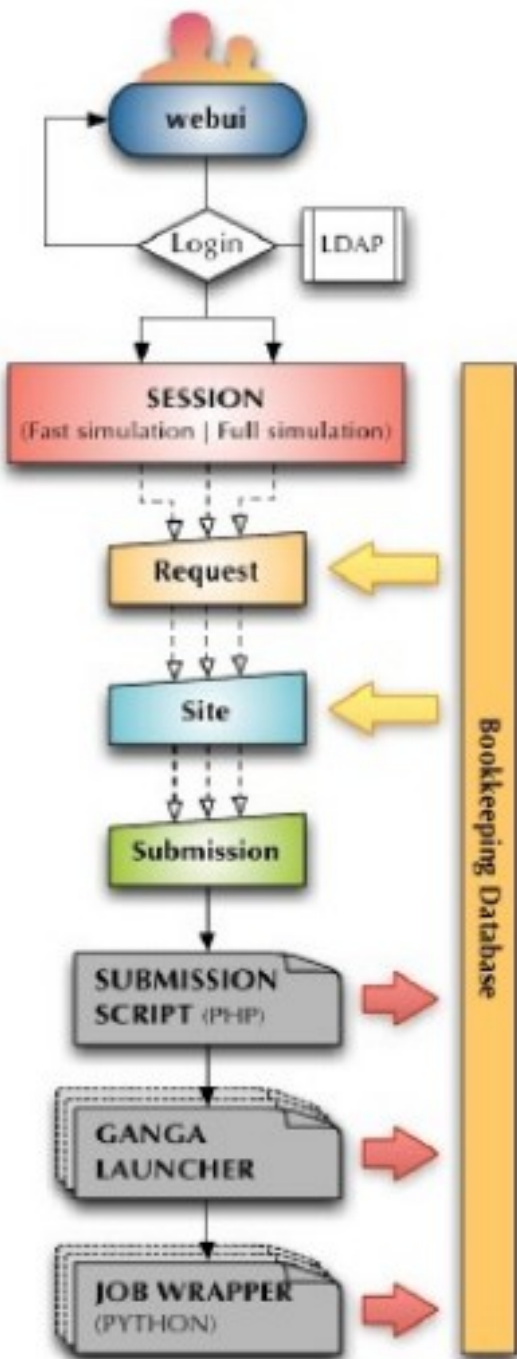
SuperBDIRAC

- Final goal:
  - Production System
  - User Analyis
  - Mass Data Transfer (via FTS)
  - Resource Management

- Current goal
  - Production System
  - Resource Management

- We want to use latest available DIRAC version until possible (freeze when production will start)

Porting Production System under DIRAC

- **Goal: manage all aspects related to Simulation Production (FastSim and FullSim) via DIRAC**

  - User management (role and permissions)

  - Site Management per Session

  - Production creation and monitoring

  - Requests creation and monitoring

  - Bunch jobs submission and monitoring

  DIRAC need to interact with Bookeeping Database (SBK5)

Using Severus via DIRAC

- Severus: python wrapper to perform SuperB MonteCarlo productions
  - Severus take care about stagein, stageout, Simulation program execution, SBK5 update
- **Severus must be launched via DIRAC**

Launching Severus via DIRAC

- Status: proof of concept
- mc_production.py
  - A python script powered by DIRAC API
    - Get php file generated by WebUI as input
    - Parse php and severus conf file
    - Retrieve necessary parameters
    - Launch DIRAC jobs using a modified version of Severus wrapper

Severus test results

- Test results:
  - Severus works properly via DIRAC
    - SBK5 updated
    - Submission properly visible from WebUI
    - Stagein and stageout via lcg commands
    - Stagein works properly
    - Stageout works properly
    - LFC properly updated
    - Tested jobs running up to 80.000 events
  - MC production must be performed by sb_productionmanager users
    - sb_productionmanager DIRAC group maps user with VOMSRole = ProductionManager

15

Severus: next steps

- Stagein and Stageout via DIRAC instaead of using lcg commands

- Submission via DIRAC webportal

- Retrieve resources data from DIRAC

- Final goal
  - Remove from Severus all functionalities that can be taken "gratis" by DIRAC

SQLAlchemy class for Bookkeeping database

- PostgreSQL Bookkeeping DB (named SBK5)

- SBK5 defines:
  - Sessions (FastSim and FullSim)
  - Productions
  - Physical data (detector geometry, background, etc..)
  - Sites enabled and validated for each Session

- Interaction with SBK5 and DIRAC mandatory

- See Milosz Zdybal presentation

Extending Webportal

- Current goal:
  - Replicate WebUI portal structure under DIRAC webportal

- Work status
  - Just started
  - We are acquiring necessary skills
    - But Webportal will change...
      - How we can be sure to not waste our time ?

- Since we plan to use latest DIRAC version until production starting (some years), we are really interested in new webportal development

- Complete SuperBDIRAC
- Evaluate all functionalities offered by DIRAC to use them
  - Especially interested in
    - Transformation System
- Sync DFC and LFC
  - LFC just used, but DFC offers very interesting features

- Documentation, tutorial, step by step guides
  - We are not sure to know all DIRAC capabilities

Credits

- Marcin Chrzaszcz - Kracow

- Christian De Santis - Rome

- Giacinto Donvito - Bari

- Armando Fella - Pisa

- Rafał Grzymkowski - Kracow

- Bruno Santeramo - Bari

- Miłosz Zdybał - Kracow

Thanks to DIRAC developers, expecially to

Adrian Casajus, Krzysztof Daniel Ciba, Ricardo Graciani, Matvey Sapunov, Federico Stagni, Andrei Tsaregorodtsev

BACKUP SLIDES

Groups

Groups and Role in VO and DIRAC

Group/Role	Type	Description	VO share	DIRAC mapping
/superbvo.org	Standard User	Generic user group	0	**sb_user**
/superbvo.org/Role=ProductionManager	Production Manager	User role related to MC production management	20	**sb_productionmanager**
/superbvo.org/Role=SoftwareManager	Software Manager	VO Software administrator	0	sb_softwaremanager
/superbvo.org/Role=Analysis	Standard User	User role related to analysis task	0	sb_analysis
	DIRAC administrator	DIRAC administrator		dirac_admin
	DIRAC pilot job	Useful for DIRAC		sb_pilot

DIRAC API

- DIRAC offers a powerful set of API
  - We need a good quality documentation to speed up development
- API documentation
  - http://diracgrid.org/files/docs/CodeDocumentation/API/index.html
  - Epydoc generated – updated – maintened by developers

Useful links

- DIRAC official site
  - http://diracgrid.org/
- DIRAC official forum
  - https://groups.google.com/forum/?hl=en&fromgroups#!forum/diracgrid-forum
- DIRAC repository
  - https://github.com/DIRACGrid
- DIRAC API documentation
  - http://diracgrid.org/files/docs/CodeDocumentation/API/index.html
- DIRAC testbed for SuperB wiki page
  - http://mailman.fe.infn.it/superbwiki/index.php/Distributed_Computing/Dirac_testbed
- DIRAC testbed webportals:
  - bbrbuild01.cr.cnaf.infn.it
    - https://bbrbuild01.cr.cnaf.infn.it:8443/DIRAC/
  - sb-serv04.cr.cnaf.infn.it
    - https://sb-serv04.cr.cnaf.infn.it:8443/DIRAC/
- Production WebUI
  - https://bbr-serv09.cr.cnaf.infn.it:8443/~webui_sbk5/
- SuperB resources (google Doc)
  - https://docs.google.com/spreadsheet/ccc?key=0AsjxRpEZ2zEIdHU2cjh4YTdrajdYdlNmajdBVGZWLWc#gid=0
- SuperB Production Shifter Guide wiki page
  - http://mailman.fe.infn.it/superbwiki/index.php/Distributed_Computing/Production_Shift_Guide