Distributed system updates

Armando Fella

XVII SuperB Workshop and kick off meeting
La Biodola (Isola d'Elba), Italy
from 28 May 2011 to 02 June 2011

Presentation Layout

- Introduction: general view
- Distributed resources status
- Distributed frameworks evaluation
- Data distribution system update
- GANGA as user tool for Grid resources access
- CNAF status report

Check point intro

- Our current commitment is to focus on support for the SuperB detector Technical Design Report writing
  - **Concentrating on physics, detector, and background simulation studies**
  - **Development of a Production system prototype is on course**
- R&D works for computing TDR preparation are in progress
  - **Started on March '10 in occasion of first R&D Computing Workshop http://www.fe.infn.it/superb10/**
  - **Next step on July 4-7 in Ferrara http://www.fe.infn.it/superb11/**
  - **The SuperB computing TDR is expected to be released one year after the detector TDR**

Check point intro II

- **Computing TDR: the distributed computing side**
- **Transverse to all topics: collect and organize SuperB specific requirements**
  - evaluate distributed frameworks already in place in HEP scenario
  - participate into development and beta testing phase of coming up projects of interest
  - based on requirements, plan for services still not on the table

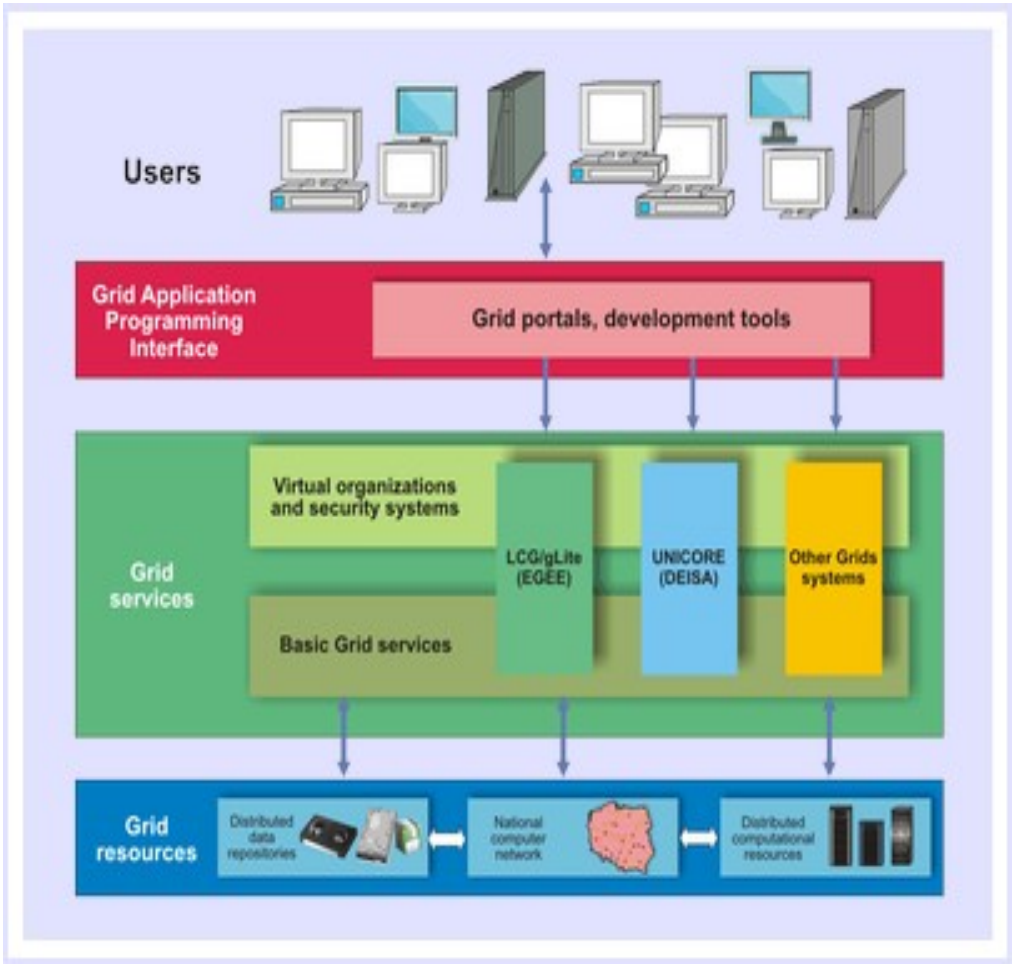
Site info summary

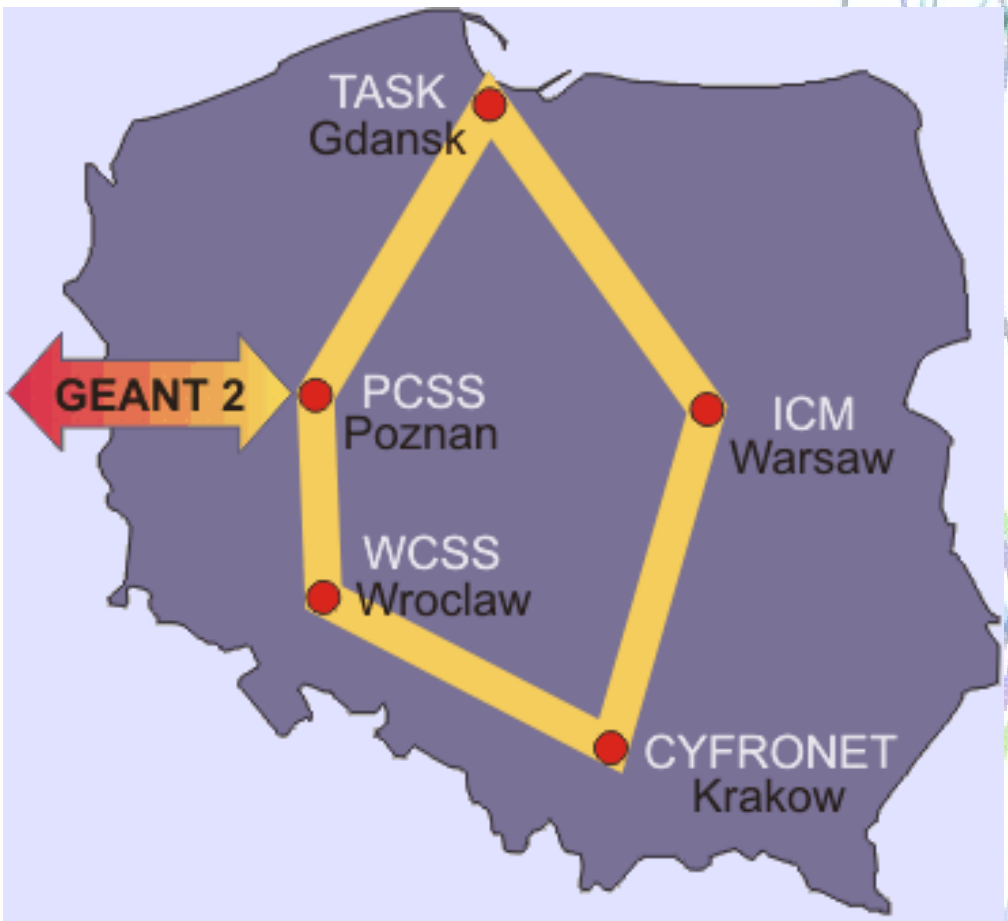
Site	Min (cores)	Max (cores)	Available (TB)	Data Access	Grid Flavor	Site contacts
CNAF	500	1000	66	StoRM	EGI	A. Fella
SLAC	400	400	10	NFS	OSG	S. Luiz, W. Yang
CALTECH	250	680	4.5	NFS	OSG	F. Porter, P. Ongmongkolkul, S. Lo
RAL	200	1000	10	Castor	EGI	F. Wilson, C. Brew, A. Martin
RALPP	50	300	5	dCache	EGI	F. Wilson, C. Brew, A. Martin
Queen Mary	300	3456	120	StoRM	EGI	A. Martin, C. Walker
Oxford Univ.	50	200	1	DPM	EGI	K. Mohammad, E. MacMahon
CCIN2P3	500	1000	10	dCache	EGI	N. Arnaud, O. Dadoun
GRIF	50	300	2	DPM	EGI	N. Arnaud, O. Dadoun
Victoria	50	100	1	dCache	EGI	A. Agarwal
Pisa	50	500	0.5	StoRM	EGI	A. Ciampa, E. Mazzoni, D. Fabiani
Legnaro	50	100	1	StoRM	EGI	G. Maron, A. Crescente, S. Fantinel
Napoli	50	100	1	DPM	EGI	S. Pardi, A. Doria
Bari	80	130	0.5	StoRM/Lustre	EGI	G. Donvito, V. Spinoso
Ferrara	10	50	0.5	StoRM	EGI	L. Tomassetti, A. Donati
Cagliari	10	50	1	StoRM	EGI	D. Mura
Perugia	10	50	1	StoRM	EGI	R. Cefala'
Torino	50	100	2	DPM	EGI	S. Bagnasco, R. Brunetti
Milano	50	100	2	StoRM	EGI	N. Neri, L. Vaccarossa, D. Rebatto
OSC	?	?	?	?	OSG	R. Andreassen, D. Johnson
Polish Grid	?	?	?	?	EGI	J. Chwastowski
Total	2710	9616	239			

On going site setup works

- Thanks to all the site contacts for the valuable work
  - **Ohio Supercomputer Center (OSC), http://www.osc.edu/**
    - Contact: R. Andreassen, D. Johnson
    - Status: Under Testing
  - **INFN-Milano**
    - Contact: N. Neri, L. Vaccarossa
    - Status: VO enabling completed --> exp SW installation
  - **Polish Grid (PL-Grid)**
    - Contact: J. Chwastowski
    - Status: VO enabling in progress

Polish Grid


Users
Grid Application Programming Interface
Grid portals, development tools
Grid services
Virtual organizations and security systems
Basic Grid services
LCG/gLite (EGEE)
UNICORE (DEISA)
Other Grids systems
Grid resources
Distributed data repositories
National computer network
Distributed computational resources


TASK
Gdansk
GEANT 2
PCSS
Poznan
ICM
Warsaw
WCSS
Wroclaw
CYFRONET
Krakow

Zeus site in PL-Grid (snapshot)

- HP BL2x220c blade servers - ca. **900 nodes**
- **9840** Intel Xeon cores
- **17TB** of RAM
- Interconnect:
  - 1GbE per node, 10GbE switch-switch
  - Infiniband 4x QDR (**40Gbps**) per node
- Disk resources: **1.5PB**
- Scientific Linux 5
- 105Tflops, **#84** on Top500, **#1** in Poland

http://www.plgrid.pl/en

Thanks to
T.Szepieniec

Zeus site GPU resources (snapshot)

- HP SL390z servers - **24 nodes**
- **288** Intel Xeon cores
- **1,7TB** RAM
- **48** nVidia M2050 'Tesla' GPGPU cards
- Interconnect:
  - 1GbE per node
  - Infiniband: 4x QDR (**40Gbps**)
- Scientific Linux 5
- 28Tflops, **#1** GPGPU cluster in Poland

http://www.plgrid.pl/en

Thanks to
T.Szepieniec

SuperB VO enabling in OSG

- **Steffen Luitz** is the **OSG** responsible for SuperB VO (superbvo.org)
  - **The VO has been formally defined in OSG (pending approval)**
  - **The OSG VO support center (SUPERB-SC) has been created**

- An OSG site contacts group will be formed to coordinate the efforts, a meeting on regular basis will be setup.
- Short term topics proposal:
  - **VO enabling/integrating operations**
  - **EGI/OSG interoperability**
  - **OSG VOMS replica setup**

Production system prototype

- The development of a distributed production system prototype is in progress since end of 2009
- A lot of work has to be done, join the project!
- See next presentation by **L. Tomassetti**

	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 x 10^8	1.6 x 10^9	8.6 x 10^9
Involved site	1	9	15
WallClockTime	6 years	19 years	150 years
Disk occupancy (TB)	0.5	5	25
Peak job running	500	2500	7000

Projects evaluation

- Understanding the status of the projects already in place in the field of distributed computing for HEP. Very important
  - **Distributed system frameworks: job management, data handling, resource monitor. LHC Computing Grid FIRST**
    - Dirac, Panda, PanDA Dynamic Data Placement , PhEDEx
  - **ATLAS framework for software installation in distributed environment, see S. Pardi presentation in this session**
  - **HTTP access layer to StoRM Storage Element: test phase**
  - **Distributed data access systems**
    - HadoopFS, Gluster, xRootD, EOS

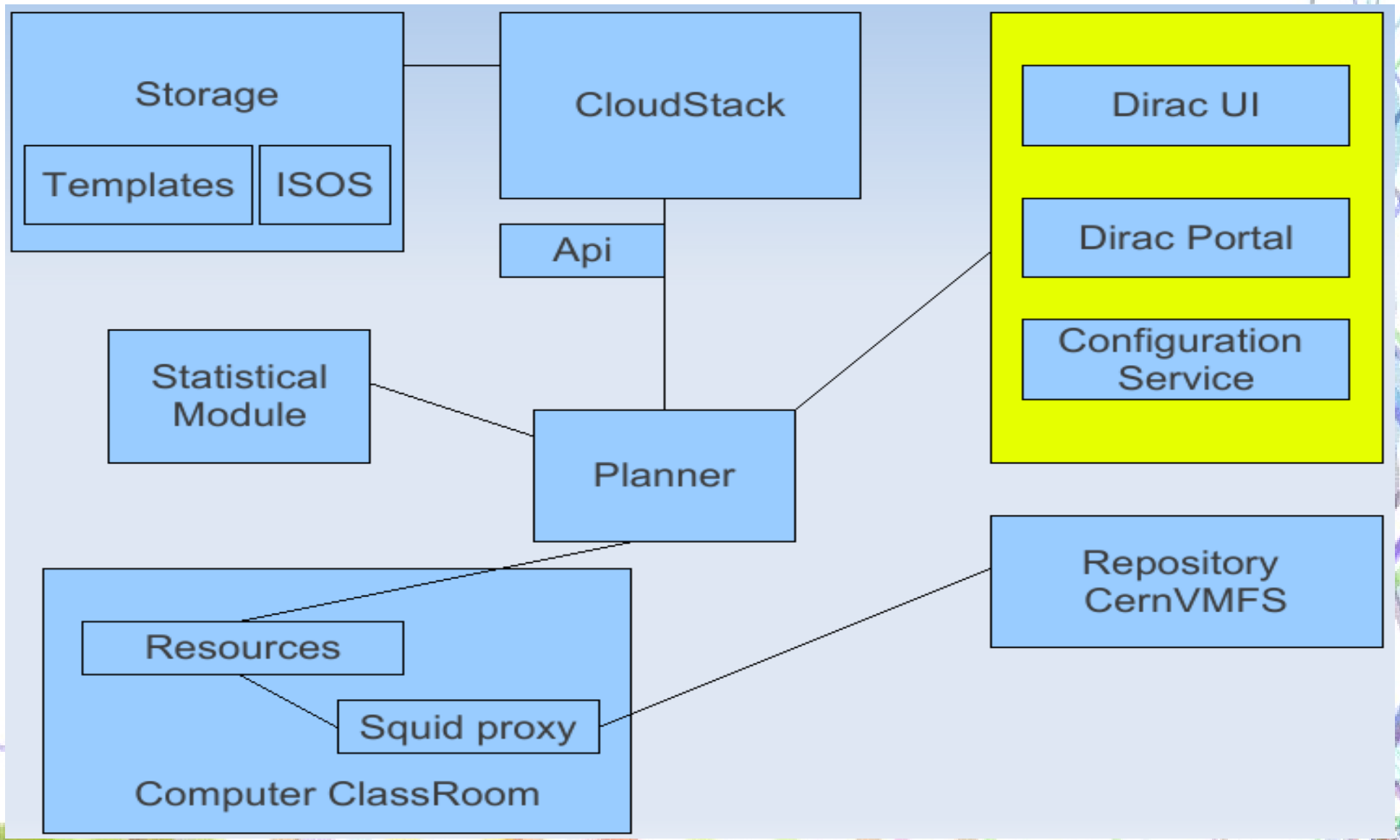
Dirac evaluation

- **DIRAC (Distributed Infrastructure with Remote Agent Control)**
  - **Dirac a complete Grid solution for a community of users needing access to distributed computing resources. DIRAC forms a layer between a particular community and various compute resources to allow optimized, transparent and reliable usage.**
- **Adopted/in use by 5 communities and South American Grid**
  - BelleII, CTA, CREATIS, ILC, LHCb and GISELA (successor of EELA2)
- **Include interface for Cloud resource exploitation**
  - CESGA: Dirac open cloud interface development
    - OpenNebula, Eucalyptus and CloudStack
- **Include framework for MPI job management on Grid resource**
  - GISELA Grid report
- **See: Dirac User Community meeting, Barcellona, 12-13 May 2011**
  Http://indico.in2p3.fr/conferenceTimeTable.py?confId=5271#all

Dirac evaluation II

- Dirac testbed has been installed at CNAF: first submission test
- Goal: evaluate system capabilities concentrating on Simulation Production use case
  - **Data access model evaluation**
  - **Cloud job and data handling management evaluation**
  - **MPI job management evaluation**
- Thanks to Dirac/LHCb community for the extensive support and enthusiasm:
  - **Andrei Tsaregorodtsev, Ricardo Graciani and Vanessa Hamar**
    - **Http://diracgrid.org**

CESGA project: Dirac-cloud interface



Data distribution, off line

- Mass data movement management via FTS
- CNAF – IN2P3 bidirectional channel has been configured
- First transfer tests are successful
  - **http://egee-jra1-dm.web.cern.ch/egee-jra1-dm/FTS/**

[Home] FTS Monitor v1.5.0 - last updated: 23/05/2011 17:29:47 (updated every 600 seconds)

Job identifier: 5213b44c-8547-11e0-a9b2-fc7d6341a4cb Submit Query

Job information

SUBMIT TIME	JOB STATE	CHANNEL NAME	SPACE TOKEN	STORAGE CLASS
23/05/2011 16:16:58 +02:00	Finished	IN2P3-CNAF		

File transfers

FILE ID	SOURCE FILE	FILESIZE	THROUGHPUT	LAST TRANSFER TIME	LOG
35746657	/testfts_in2p3-cnaf	14.57 KiB	4.86 KiB/s	23/05/2011 16:17:03 +02:00	/var/tmp/glite-url-copy-edguser/IN2P3-CNAFcompleted/IN2P3-CNAF__2011-05-23-1417_qFRL40.log

- **FILE_ID:** 35746657
- **JOB_ID:** 5213b44c-8547-11e0-a9b2-fc7d6341a4cb
- **FILE_STATE:** Finished
- **SOURCE_SURL:** srm://ccsrm02.in2p3.fr/pnfs/in2p3.fr/data/superb/prod/IN2P3-CC/test_release/testfts_in2p3-cnaf
- **DEST_SURL:** srm://storm02.cr.cnaf.infn.it/superb/test/testfts7_from_in2p3
- **FILESIZE:** 14921
- **FINISH_TIME:** 23/05/2011 16:17:24 +02:00
- **JOB_FINISHED:** 23/05/2011 16:17:25 +02:00
- Transfer
  - **TRANSFER_STATE:** Completed
  - **SOURCE_TURL:** gsiftp://ccdcacsn217.in2p3.fr:2811//pnfs/in2p3.fr/data/superb/prod/IN2P3-CC/test_release/testfts_in2p3-cnaf
  - **DEST_TURL:** gsiftp://diskserv-san-76.cr.cnaf.infn.it:2811//storage/gpfs_superb/prod/test/testfts7_from_in2p3
  - **SRC_PREP_DURATION:** 6
  - **DEST_PREP_DURATION:** 6
  - **TX_DURATION:** 3
  - **SRC_FINAL_DURATION:** 0
  - **DEST_FINAL_DURATION:** 1
  - **BYTES_WRITTEN:** 14921

[Home | Channels | Agents | Channels audit | About]

Data distribution, on line

- Production job transfers the output where it is expected to be and a replica at CNAF
  - **Production system includes now the “target site” selection feature per simulation request**
- Proposal: start discussing dataset definition
  - **CMS and ATLAS solutions have been shown at Frascati XVI workshop (slide uploaded in this slot)**
  - **Dataset definition affects the distributed data model definition**
  - **Computing group need to outcome with a proposal to the community to start converging on agreed solution**

GANGA SuperB plugin II

- Physics group requests a tool suite permitting “simple user” Grid resources exploitation during XV Caltech SuperB workshop
- GANGA is officially the adopted tool
- CNAF centralized GANGA installation and configuration is ready
- The development of a specific SuperB plugin has started
  - **Work in progress on the following functionality:**
    - **SuperB sites modeling: integration in submission tasks**
    - **Automatic output merging features**
    - **Bulk submission extending methods**
    - **Simple analysis use case, local posix access only**
- **Thanks to Ganga developers for valuable help**

CNAF site status

- CNAF Tier1 site hosts central simulated data repository, Production framework and complete user workspace

- On going works:
  - **Test new StoRM layer for file access via HTTP protocol**
  - **SAM service installation and setup for SuperB VO. Alarm/monitor system for Grid service at VO enabled sites**

Involved people

- G. Donvito INFN-Bari
- A. Fella INFN-Pisa
- S. Luitz SLAC
- E. Luppi INFN-Ferrara
- M. Manzali INFN-Ferrara
- A. Paolini INFN-CNAF
- S. Pardi INFN-Napoli
- B. Santeramo INFN-Bari
- L. Tomassetti INFN-Ferrara
- + two Ferrara students

Info, docs, email list

- SuperB wiki Distributed Computing root page:
  - **http://mailman.fe.infn.it/superbwiki/index.php/Main_Page**
- Distributed Computing mailing list:
  - **superb-distrcomp [at] lists.infn.it**
- VO enabling procedure in EGI/OSG, list of requirements in terms of services, packages and functionality SuperB requires to be present on sites
  - **http://mailman.fe.infn.it/superbwiki/index.php/How_to_Grid/Site_setup**
  - **http://mailman.fe.infn.it/superbwiki/index.php/VO_enabling_procedure_information**
  - **Site contact mailing list: superb-grid-mng [at] lists.infn.it**

A lot of work to do

- **Detector TDR definition**
  - **Several improvements are required to production system prototype (next talk)**
  - **Grid environment related tasks**
    - **Service evaluation, testing, integration**
    - **Multi-Grid environment related subjects**
- **Computing TDR definition (see next session, Comp. R&D)**
  - **Cloud resource access, virtualization**
  - **Distributed computing and parallel tasks**
  - **Projects evaluation and profiling**
  - **Distributed storage (See S. Pardi talk at Comp. R&D)**
  - **More..**

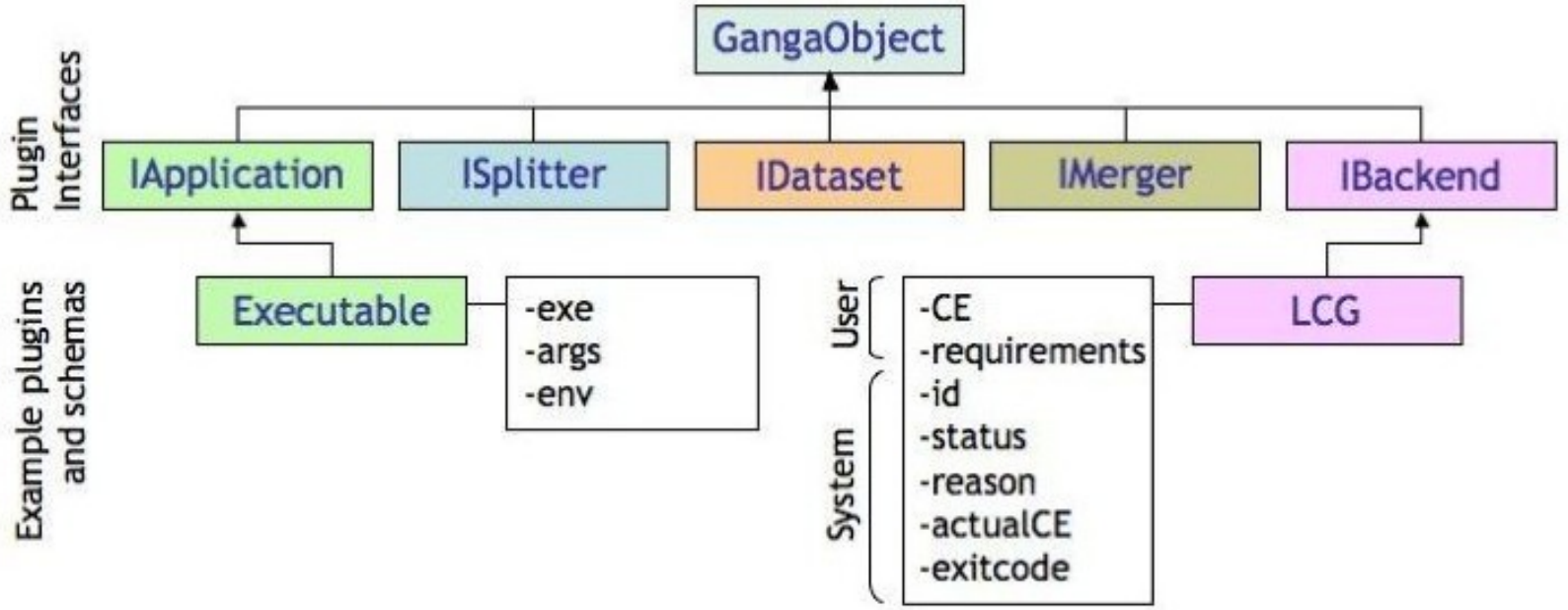
Questions, comments ?

backup

GANGA work plan

- Works follow up proposal
  - **Basic plugin feature releasing**
  - **Complete GANGA instruction wiki page**
  - **“expert” user recruitment from Fast and FullSim communities for tool suite testing period**
  - **Hands on user meeting**

GANGA SuperB plugin



"Plugins for different types of application, backend, dataset, splitter and merger inherit from interface classes, which have a common base class. Schemas for the Executable application and for the LCG backend are shown as examples."

www.gridpp.ac.uk/papers/chep06_harrison.pdf