Distributed computing status

V SuperB Collaboration Meeting
Pisa, Sept. 20th 2012

Armando Fella

Topics

- Distributed resources management
- Production system
- Distributed analysis system
- Dirac project
- Data management
- SuperB software installation and access in distributed environment: CVMFS
- CNAF resource and service status
- Grid monitoring service

Distributed resourse summary

Site	Min (cores)	Max (cores)	Disk (TB)	SRM layer	Grid Org.	Site contacts
RAL(T1)	200	1000	25	Castor	EGI	F. Wilson, C. Brew
Ralpp	50	500	5	dCache	EGI	F. Wilson, C. Brew
Queen Mary	300	2000	150	StoRM	EGI	A. Martin, C. Walker
Oxford Univ.	50	200	*1*	DPM	EGI	K. Mohammad, E. MacMahon
IN2P3-CC(T1)	*500*	*1000*	*16*	dCache	EGI	N. Arnaud, O. Dadoun
Grif	50	300	2	DPM	EGI	N. Arnaud, O. Dadoun
in2p3-lpsc	50	100	2	DPM	EGI	J.S. Real
in2p3-ires	50	100	2	DPM	EGI	Y. Patois
CNAF(T1)	500	1000	180	StoRM	EGI	A. Fella, P. Franchini
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	500	2000	15	DPM	EGI	S. Pardi, A. Doria
Bari	160	260	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	L. Fano'
Torino	50	100	2	DPM	EGI	S. Bagnasco, R. Brunetti
Frascati	30	100	2	DPM	EGI	E. Vilucchi, G. Fortugno, A. Martini
Milano	50	100	2	StoRM	EGI	N. Neri, L. Vaccarossa, D. Rebatto
Catania*	?	?	?	StoRM	EGI	G. Platania
Slac	400	400	10	NFS	OSG	S. Luiz, W. Yang
Caltech	200	400	4.5	NFS	OSG	S. Lo, F. Porter, P. Ongmongkolkul
Fnal	50	400	1	dCache	OSG	M. Slyz
OhioSC*	?	?	?	dCache	OSG	R. Andreassen, D. Johnson
Victoria	50	100	5	dCache	EGI	A. Agarwal
McGill*	100	200	1	StoRM	EGI	S. Robertson, S.K. Nderitu
Cyfronet	100	500	10	DPM	EGI	L. Flis, T. Szepienie, J. Chwastowski
Total	3570	11510	440			

* VO enabling procedure in progress

27 sites are available to the SuperB VO. From: Canada, France, Italy, Poland, UK and USA

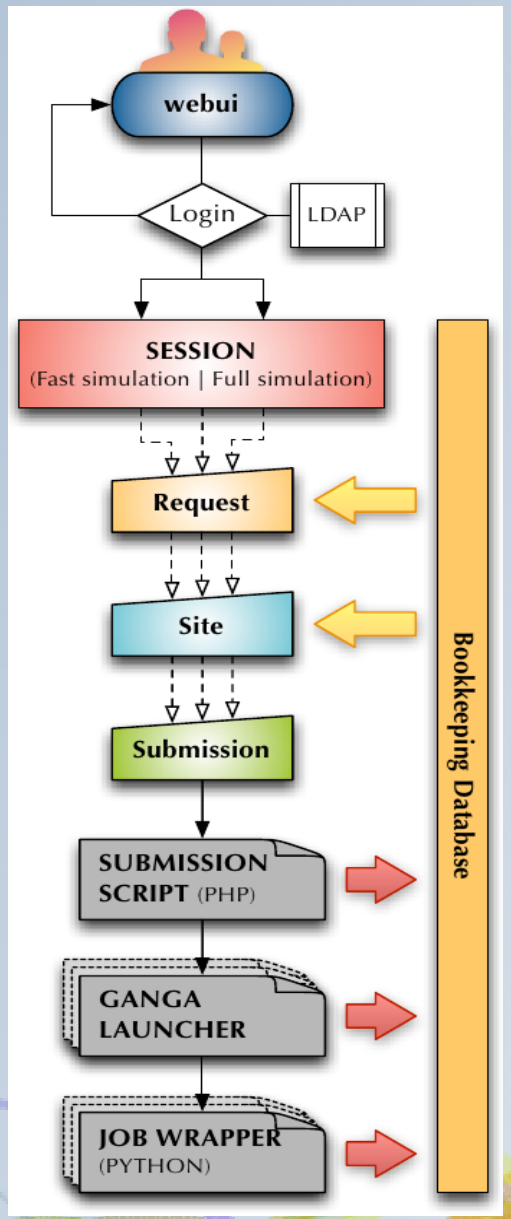
OSG collaboration

- **OSG council - SuperB meeting** on July 10th
  - Meeting goal is broader the collaboration about crucial subjects in interoperability field
  - SuperB presents its distributed computing related comments and requirements
    - **OSG support to distributed software installation CVMFS**
    - **Use of GlideinWMS as unique submission system**
    - **New OSG information service publishing on WLCG/EGI BDIIs is under discussion**
      - **We are invited to participate to the work**
    - **Enabling VOMS Role in OSG is a big issue**
    - **Nagios per VO enabled in OSG Grid**
      - **Request for installation of required packages in OSG WN release has been accepted**

Distributed resources status

- **Full distributed resource testing phase** is mostly in stand-by
  - ~1/3 of the sites results misconfigured
    - Testing, fixing, reinstalling, contacting is slowly going on
    - Need manpower
- VO enabling operations at remote sites
  - **McGill**: testing services in progress
  - **SLAC**: ok
  - **Caltech**: ok *(involved in 2012_Fullsim_summer prod)*
  - **Fermilab**: ok *(involved in 2012_Fullsim_summer prod)*
  - **Ohio Supercomputing Center**: enabling process is stopped

Production system



- **FullSim summer production**
  - **The first on distributed resources**
  - **The first using the new generation production system**
    - Production goals have been accomplished
    - Outcome: list of improvements to be applied and bugs to be solved

- M.Manzali, the primary developer left (not replaced)
- C.De Santis is moving to production system field (see next presentation by C.De Santis)
- Two students from Ferrara (three years degree):
  - REST interface refactoring
  - Production system debugging

Analysis system prototype

- The main developer left the group in June, need a replacement
- Minor progress in testing functionality and debugging
- Collaboration with Ganga developer team
  - Evaluation of Dirac backend adoption
    - Dirac setup requested to Ganga team; the task is in progress
  - We are looking for new developers

- Tutorial page:
  - **http://mailman.fe.infn.it/superbwiki/index.php/Tutorial_%28draft%29**

Dirac project

- **A fruitful collaboration with Polish computing group started on June '12**
  - Goal: setup and configure a Dirac system to fulfil the SuperB requirements
  - **General work plan in priority order:**
    - Simulation production use case (in progress)
      - Porting of SuperB specific environment
      - Bookkeeping DB integration
    - Workload Monitor system
    - Analysis use case integrated with Ganga system
    - Mass data transfer system

- Bi-weekly meeting http://superb.infn.it/restricted-distributed-computing
- **See next talks by B.Santeramo and M.Zdybal**

Data management

- **Storage system evaluation**
  - HadoopFS on WAN: testbed on Bari and Napoli
- **Data access framework library development**
  - Data access optimization on local and WAN scenario
  - Mask the low level data access layer at the sites
  - Useful support from ROOT development team
- **Mass data transfer system**
  - FTS3 evaluation
  - PhEDEx evaluation process (stand-by)
- **File catalog ng** (dynamic LFC ng by EMI R&D, stand-by)
- **Data model definition** (stand-by)
- **Geographically distributed data center study** (stand-by)
- **See the presentation by P.Franchini on Computing R&D session**

SuperB software management system in distributed environment: CVMFS

- CernVM is a baseline Virtual Software Appliance. The Appliance represents a complete, portable and easy to configure user environment for developing and running HEP data analysis locally and on the Grid, independently of Operating System software and hardware platform. The goal is to remove a need for the installation of the experiment software and to minimize the number of platforms (compiler-OS combinations) on which experiment software needs to be supported and tested.
- **CVMFS server-side has been installed and tested at CNAF**
  - IGI + SuperB (P. Veronesi + P.Franchini)
- **CNAF farming group setup the client side components**
  - Squid proxy service is shared with CMS and LHCb
- **The setup of a geographical redundancy of the service is a constraint to move to production state**
  - INFN Torino will host the secondary service, IGI is responsable for installation, configuration and maintenance (within October)
- **The CVMFS installation campaign on distributed resources may be able to start on November 1st**
- A detailed installation and configuration report has been produced:
  - http://mailman.fe.infn.it/superbwiki/index.php/CNAF_portal/CVMFS

CNAF resource and service status

- The assigned disk space for 2012 is 200TB, currently 127TB have been allocated
  - **73TB will be added within November**
  - **AMS will use SuperB disk space (~8TB) until end of October**

- **Four new blade machines** will be added to the SuperB machine set in one month, courtesy of CNAF.
- SuperB will use the new machines for the following services:
  - **DB clusterization**
  - **Dirac system**
  - **Production head node replacement**

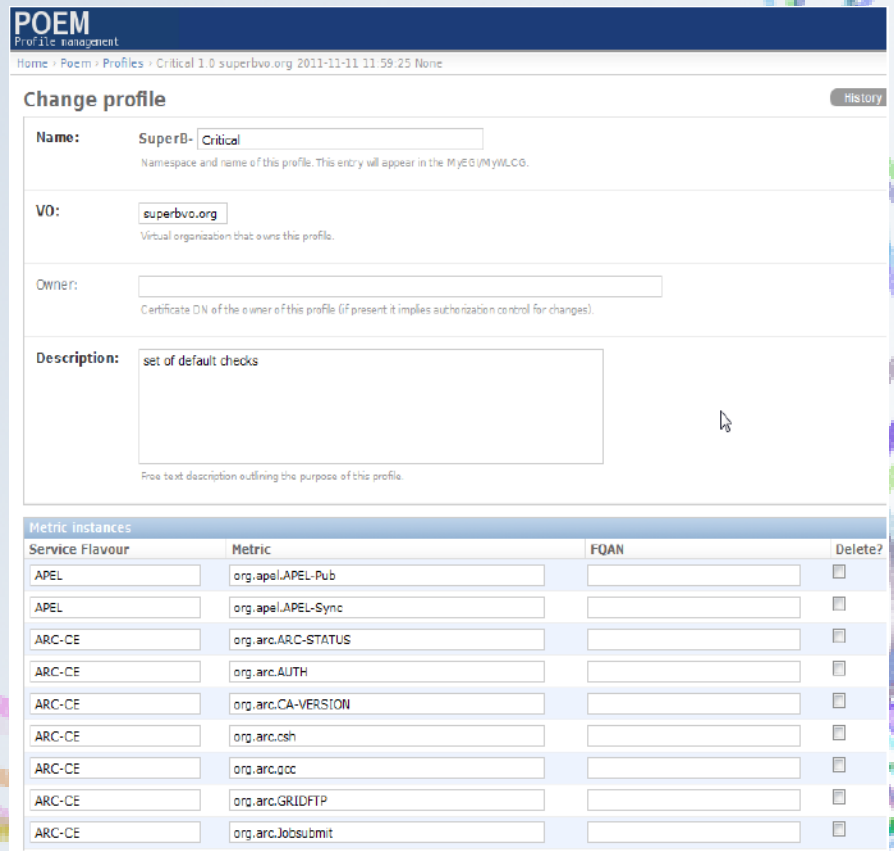
Grid monitoring service

- Nagios system at CNAF has been upgraded to last version
  - Solve configuration problems and add new services
  - Active collaboration with developers
- **POEM (Profile Management)**, feature released with EGI SAM* update 17.1

Improved metric management
- Creation of new profiles via web portal
- Adding of new probes via web portal

Two specific SuperB probes have been added (need test)
- Simulation production validation (FastSim and FullSim)

Nagios portal https://sb-serv01.cr.cnaf.infn.it/nagios/
- Standard probes: lcg-CE, CREAM-CE and SRM
- *Service Availability Monitoring (SAM)



Work plan

- **Distributed resources**
  - Complete the few on going enabling processes (within 2013)
  - OSG integration process: start CVMFS setup, GlideinWMS setup evaluation (December)
- **Simulation production use case**
  - Complete the bug fixing and usability issues updates on WebUI (December)
  - Dirac SP use case: deployment of beta release (December)
- **Distributed analysis use case**
  - Apply the bug fixing of priority (December)
- **Distributed storage R&D works** --> see P.Franchini presentation in R&D session
- **CVMFS delpoyment**
  - Start the CVMFS setup campaign on EGI distributed environment (within December)
- **CNAF related works** (within 2012):
  - Service setup/upgrade on new HW (Dirac service, SP Head node, DB clustering)
  - EMI2 WN validation test
- **Grid monitoring service**
  - Add new SuperB probes and complete the enabling process (within November)