

Distributed computing work status

A. Fella on behalf of Distributed Computing group

Outline

- Distributed storage works overview
- Dirac evaluation
- PhEDEx evaluation
- Distributed resources status
- Ganga use case analysis
- High level model
- CNAF services

Distributed storage works overview

● **R&D Data access:**

- HTTP and WebDAV in Wide Area Network scenarios
- collaboration setup with EMI group and CNAF StoRM team

● **Distributed Tier1 model study:**

- functional evaluation, data distribution policy, resource allocation

● **Data access and storage system evaluation:**

- gLaster, NFS4, Hadoop, EOS
 - See next talk “Data access R&D status”
 - See talk on Computing planning session on Thu 15, 15:00 – 16:30

Dirac evaluation

- Use case production:
 - **Comparative test performed between official production system and Dirac system**
- Use case analysis:
 - **Large test performed using SuperB analysis task**
- Evaluation of Dirac new release
- Intense collaboration with developer
 - **See next talk by B. Santeramo**

PhEDEx generalization project

- Oracle test instance ready at CNAF
- All the HW/SW configuration requirement have been provided
 - **Support by CMS, thanks to:
Nicolo' Magini, Jacek Wojcieszuk, Claudio Grandi**
- CNAF offers one machine for Service installation
- D.Andreotti and A.Sartirana can help sporadically (again thanks)
- A collaboration CMS-SuperB can be activated on Phedex generalization project
 - **This action is in standby cause of lack of people:**
 - Need a full time person coordinating the task SuperB side

Distributed resources

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	3456	120	StoRM	EGI	A. Martin, C. Walker
Oxford Univ.	50	200	1	DPM	EGI	K. Mohammad, E. MacMahon
IN2P3-CC(T1)	500	1000	10	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	127	StoRM	EGI	A. Fella
Pisa	50	500	0.5	StoRM	EGI	A. Ciampa, E. Mazzone, D. Fabiani
Legnaro	50	100	1	StoRM	EGI	G. Maron, A. Crescente, S. Fantinel
Napoli-infn	50	100	5	DPM	EGI	S. Pardi, A. Doria
Napoli-grisu	50	300	5	DPM	EGI	S. Pardi, A. Doria
Napoli-unina	50	300	5	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	R. Cefala'
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	F. Porter, P. Ongmongkolkul, S. Lo
OhioSC	?	?	?	dCache	OSG	R. Andreassen, D. Johnson
Victoria	50	100	5	dCache	EGI	A. Agarwal
Cyfronet	100	500	10	DPM	EGI	L. Flis, t.Szepienie, J. Chwastowski
Total	3070	11066	349			

Storage resources

- Data driven model constraints on analysis tasks:
 - **E.g.: all the datasets a job needs to analyze in the same execution should be present on the same site.**
- => need to plan next SP output dataset placement
- Disk availability estimation on 2q 2012 for production output purpose: 340TBi
 - **RAL T1: 50TBi**
 - **QMUL: 120TBi**
 - **IN2P3-CC: 50TBi**
 - **CNAF: 120TBi**

Distributed resource status

● Work in progress:

- Ohio SC: OSG site setup in progress
- SLAC: OSG enabling in progress
 - OSG porting plan for SLAC site and SuperB requirement have been discussed in a meeting among OSG responsible, scientific communities and SLAC management on Nov 30th
- INFN-Catania: to be planned
- RAL-PPD: to be planned

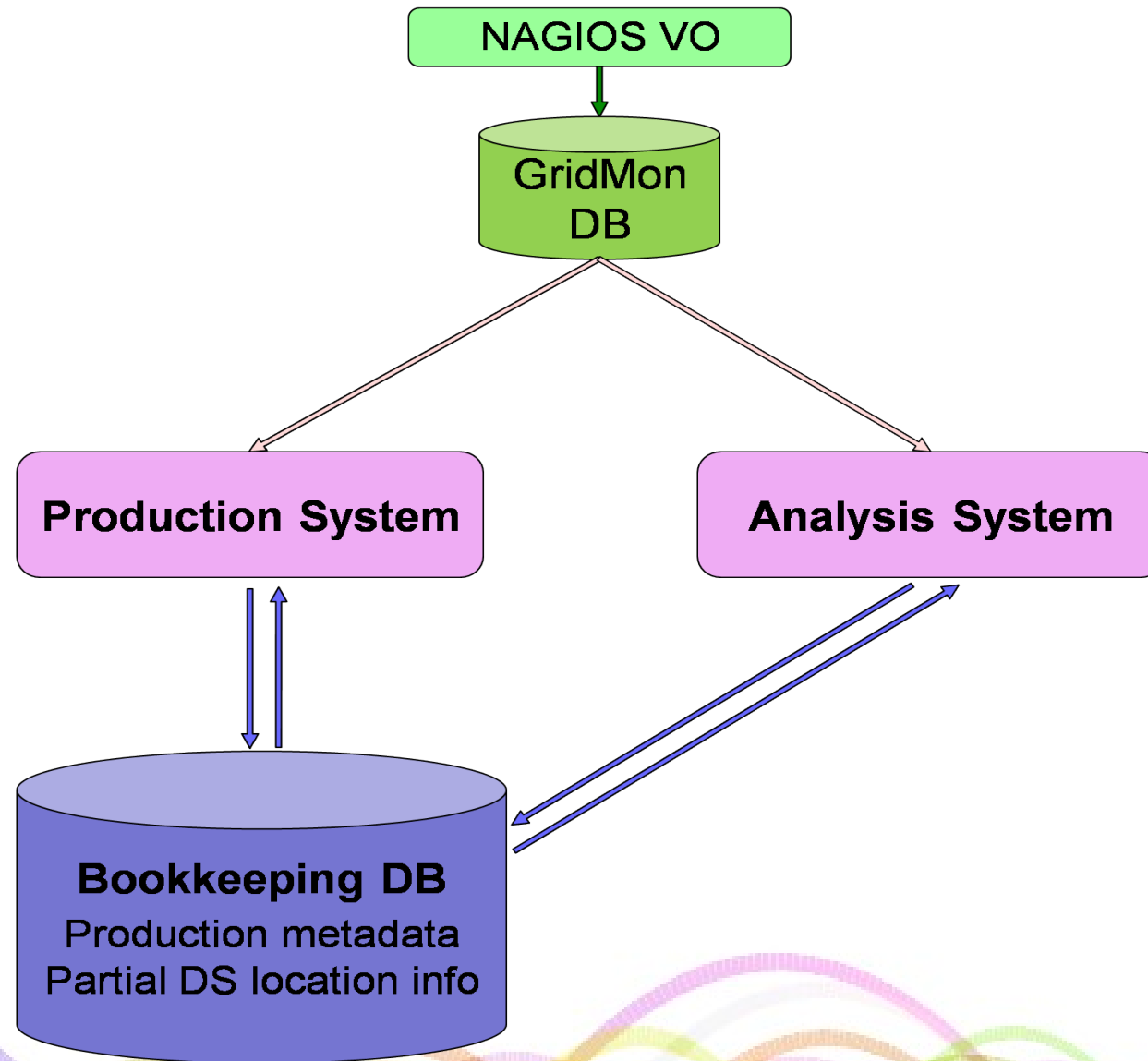
End user analysis

- Ganga plugin development is in progress

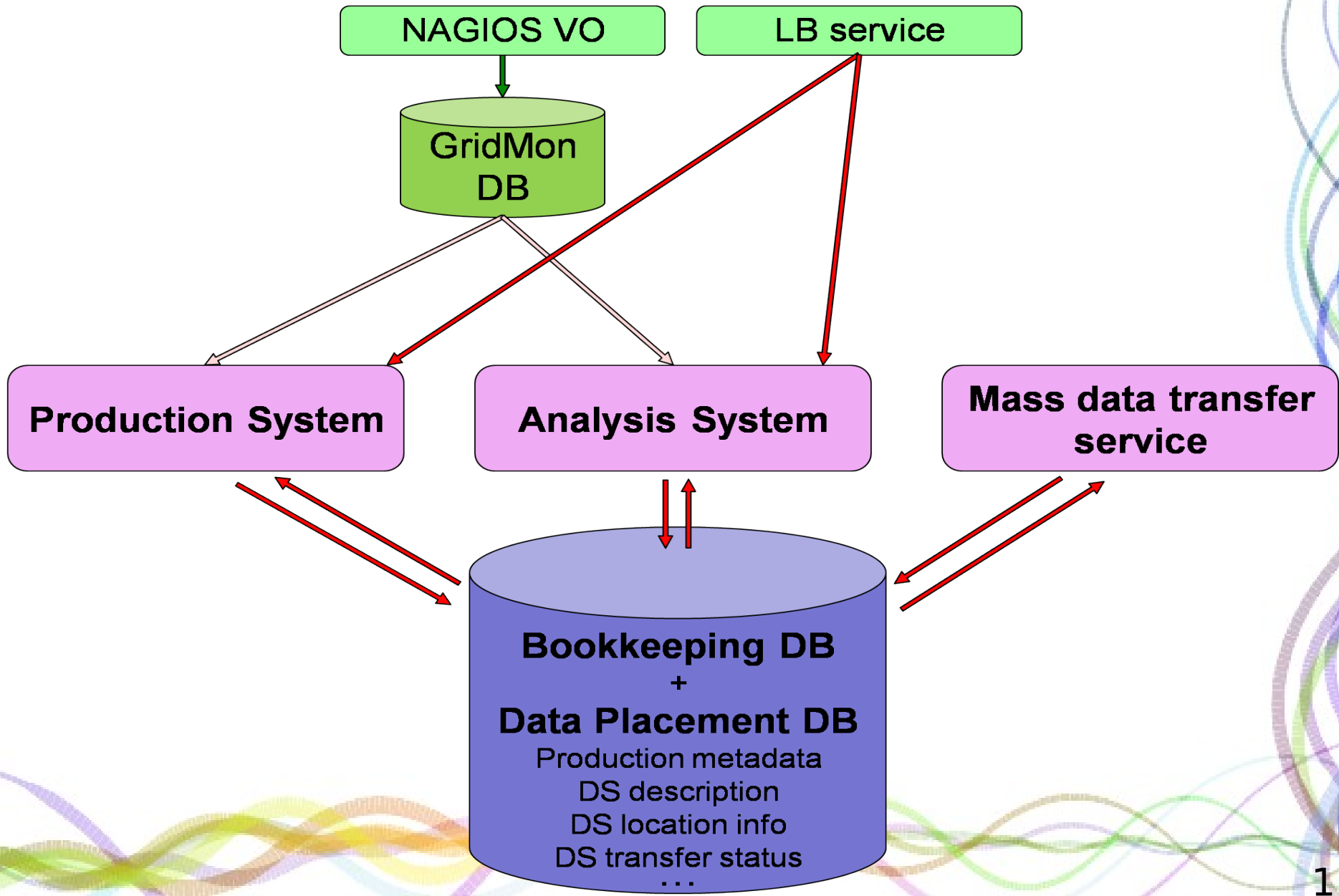
- A prototype is ready to be tested
- Join next Ganga hands on session
 - Jobs input data automatic management via dataset selection
 - Running sites check and selection
 - Grid service availability automatic check
 - On line job monitor (waiting for new Ganga mon feature)
 - New storage area access management for job output
 - authentication is driven by VOMS Role “Analysis”
 - 2010_September Generics at RAL T1 and IN2P3 can be already accessed

- **GANGA ref. <http://ganga.web.cern.ch/ganga/>**

High level model, now



High level model, next



CNAF status

- StoRM http/webdav data access under test in collaboration with StoRMers
- CVMFS setup for non LHC exp is a CNAF goal for first quarter 2012
- Disk space jump to 127TBi --> 180TBi in 2012
 - 15TB user space quota (12TB occupied, cleaning asap)
 - To be ported to new SE area management
 - 33TB production output
 - 79TB available
- HW upgrade plan for 1q 2012 under definition:
 - DB clustering --> HA, CVMFS server, UI (on demand?)

Near future plan

- Complete the production and analysis frameworks
 - **Data placement DB in production (Jan)**
 - **Ganga analysis framework development**
 - Setup of first group of expert users (Jan)
 - **GridMon DB sql logic layer development**
 - **Production and analysis system integration with new DB backend (within Feb)**
- Complete Dirac evaluation work
 - **Write a detailed report (within Feb)**
- US SuperB site integration in OSG env. The process started, completion is expected for 1q 2012
- FastSim production preparation:
 - **Large prod system test stress, planning, etc.. (start on Jan)**

People news

- Key people at Ferrara will leave shortly
 - **M. Manzali, 2 years fellowship ends on March 2012**
 - **E. Vianello degree on Dec. 15Th 2011**
 - **A. Galvani degree on March 2012**
- Collaborations
 - **S. Makarychev (ITEP, 50%): from January 2012**
 - **C. De Santis (INFN, 30%)**

Thank you, questions?

High level model, next

