CNAF status and production plan

Armando Fella & Giovanni Fontana
on behalf of Italian distributed computing group

Production tools

Two Production suites under development

1) site LCG/EGEE/OSG enabled: VO enabled, Gridftp, LCG-Utils

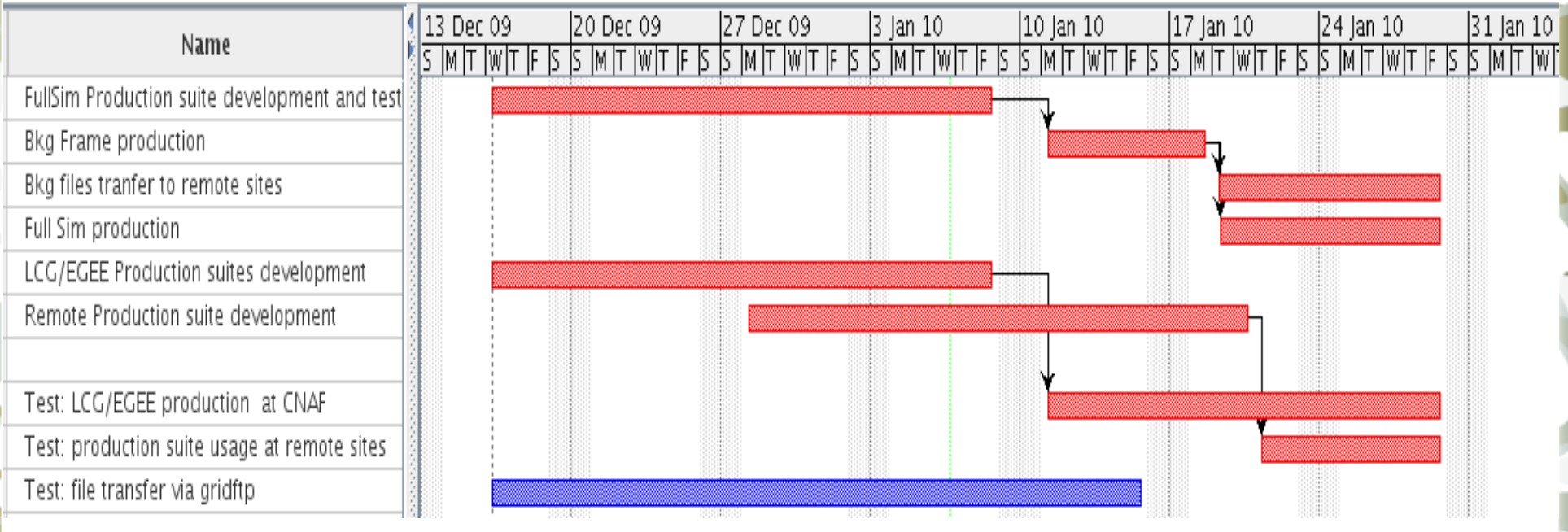
- input bkg files will be transferred to the site
- the production will be lauched from CNAF via GANGA grid job manager
- the job will transfer the output directly to CNAF via LCG-Utils (Lcg-cr)
- the DB update will be performed offline or online (under investigation)
- no need of Production suite installation on the site

2) site not LCG/EGEE enabled:

- production suite installation on the site, submission via batch system
- the transfer of input and output should be performed via gridftp
- the DB update will be performed offline

The production suite flavors are under development, **need site info**

Production time line



- Bkg Frame production started on Jan 7th
- Start of Bkg Frame transfers to remote sites: Jan 18th
- Grid submission tests of production job to remote sites start today

BkgFrame transfer to remote sites

Within the week BkgFrame production will end, my proposal for BkgFrame file transfer operation is the following:

- the transfers will be performed via gridftp command by each sites via the use of a script that will been provided during this week.

- week 11-18 Jan: site contacts perform transfer tests as they are described into test transfer requirement page (number 4 excluded)
http://mailman.fe.infn.it/superbwiki/index.php/How_to_Grid/Site_setup/Requirements/Checks/transfer-test

- week 18-25 Jan: the validated sites will start the files transfer from CNAF, two parallel transfer streams should be a good ratio resources load/expected time of transfer completion.

Site status, enabled CE

- The next V0.1.4 FastSim release will be the official software suite to be used for February production

- The set of executables used in production will be provided with the release. No need of compilation into test release

- The Grid job (LCG/EGEE + OSG) will bring with it (via Input Sandbox mechanism) the software test release and will setup an autonomous environment for execution of applications located into the official Software Installation path ($VO_SUPERBVO_ORG_SW_DIR)

- The Job will be submitted by CNAF UI via GANGA tool and will transfer the .root output files at its completion. The DB info will be retrieved offline at CNAF via log file parsing.

Site status, enabled CE

CE/queue VO enabled (lcg-infosites command)	Good-ce	Ganga job submission
ce.scope.unina.it:2119/jobmanager-lcgpbs-egee_long		ok
ce.scope.unina.it:2119/jobmanager-lcgpbs-egee_short		ok
ce02.esc.qmul.ac.uk:2119/jobmanager-lcgsge-lcg_long	ok *	ok
ce03.esc.qmul.ac.uk:2119/jobmanager-lcgsge-lcg_long2_x86	*	ok
ce03.esc.qmul.ac.uk:2119/jobmanager-lcgsge-lcg_long_x86	*	ok
ce04-lcg.cr.cnaf.infn.it:2119/jobmanager-lcglsf-debug	ok	ok
ce04-lcg.cr.cnaf.infn.it:2119/jobmanager-lcglsf-superb	ok	ok
ce05-lcg.cr.cnaf.infn.it:2119/jobmanager-lcglsf-debug	ok	ok
ce05-lcg.cr.cnaf.infn.it:2119/jobmanager-lcglsf-superb	ok	ok
ce06-lcg.cr.cnaf.infn.it:2119/jobmanager-lcglsf-debug	ok	ok
ce06-lcg.cr.cnaf.infn.it:2119/jobmanager-lcglsf-superb	ok	ok
grid-ce-01.ba.infn.it:2119/jobmanager-lcgpbs-infinite	ok	
grid-ce-01.ba.infn.it:2119/jobmanager-lcgpbs-long	ok	
grid-ce-01.ba.infn.it:2119/jobmanager-lcgpbs-short	ok	
grid0.fe.infn.it:2119/jobmanager-lcgpbs-superb	ok	ok
grid10.lal.in2p3.fr:2119/jobmanager-pbs-superbvo.org	ok *	
gridce.pg.infn.it:2119/jobmanager-lcgpbs-superbvo	ok	ok
gridce1.pi.infn.it:2119/jobmanager-lcglsf-superb	ok	ok
grisuce.scope.unina.it:2119/jobmanager-lcgpbs-grisu_long		ok
grisuce.scope.unina.it:2119/jobmanager-lcgpbs-grisu_short		ok
heplnx206.pp.rl.ac.uk:2119/jobmanager-lcgpbs-grid1000	ok	ok
heplnx206.pp.rl.ac.uk:2119/jobmanager-lcgpbs-grid2000	ok	ok
heplnx206.pp.rl.ac.uk:2119/jobmanager-lcgpbs-grid500	ok	ok
heplnx206.pp.rl.ac.uk:2119/jobmanager-lcgpbs-short	ok	
heplnx207.pp.rl.ac.uk:2119/jobmanager-lcgpbs-grid1000	ok	ok
heplnx207.pp.rl.ac.uk:2119/jobmanager-lcgpbs-grid2000	ok	ok
heplnx207.pp.rl.ac.uk:2119/jobmanager-lcgpbs-grid500	ok	ok
heplnx207.pp.rl.ac.uk:2119/jobmanager-lcgpbs-short	ok	
t3-ce-01.pd.infn.it:2119/jobmanager-lcglsf-superbvo	ok	ok

Good-ce test command: globus-job-run $ce /bin/hostname

* Need certificate upgrade, see http://mailman.fe.infn.it/superbwiki/index.php/VO_technical_info

Site status, enabled SE

lcg-infosites se --vo superbvo.org	lcg-cr	globus from local
grid2.fe.infn.it		
se.scope.unina.it	ok	ok
gridse.pg.infn.it	ok	
gridse3.pg.infn.it		
storm02.cr.cnaf.infn.it	ok	ok
se03.esc.qmul.ac.uk	ok	
grisuse.scope.unina.it	ok	ok
t3-srm-01.pd.infn.it	ok	ok
t3-srm-02.pd.infn.it		
grid2.fe.infn.it		
gridse.pg.infn.it		
storm02.cr.cnaf.infn.it		
se03.esc.qmul.ac.uk		ok
heplnx204.pp.rl.ac.uk	ok	
grid05.lal.in2p3.fr	ok	ok

Lcg-cr command (%s= SE hostname):

```
lcg-cr -v --vo superbvo.org -d ${s} -l lfn:/grid/superbvo.org/test1/test_${s}_${randstr} file:$PWD/test.txt
```

Globus-url-copy command:

```
globus-url-copy -vb file:$PWD/test.txt gsiftp://${s}/tmp/${s}_gsiftp.test
```

Discussion

- The authoritative information about site setup stage it is the wiki page
http://mailman.fe.infn.it/superbwiki/index.php/How_to_Grid/Site_setup/Site_setup_status_report

- Please report any problem you meet to the email list and update the status page as soon as each step is completed

- Is it necessary a specific session dedicated to SW installation?

- Need to discuss in specific the OSG site status and plan

BACKUP

Site setup procedure notes

- Transfer tool Gridftp is the only one supported by CNAF center,
  to use the tool you need to be a member of superbvo.org.
  A new wiki requirement was added: “Transfer test page”

- A production service test procedure is under development,
  please report on “Site status” wiki page the requested info, in particular
  the Storage Element hostname for LCG-Utils/Gridftp test

- The Grid way (OSG/LCG) is the preferred, please try to solve the
  VO site enabling “bureaucratic” steps. To achieve the goal is necessary
  to have a profitable communication with Grid people in each site.

Distributed Computing human network

CNAF	Italian group
Caltech	Frank Porter, Piti Ongmongkolkul
SLAC	Steffen Luiz, Wei Yang
McGill	Steven Robertson
Queen Mary	Adrian Bevan
RAL	Fergus Wilson
LAL and Lyon	Nicolas Arnaud
Bari	Giacinto Donvito, Vincenzo Spinoso
Legnaro - Padova	Gaetano Maron, Alberto Crescente
Napoli	Silvio Pardi
Ferrara	Giovanni Fontana, Marco Ronzano
Pisa	Alberto Ciampa, Enrico Mazzoni, Dario Fabiani

Email list: **superb-grid-mng@lists.infn.it**

Production design proposals

Next MC productions involve the exploitation of distributed computing resources, two different distributed design systems are proposed:

the February production distributed system proposal include the installation at remote sites of tools used at CNAF during the Nov '09 test production.

the web production interface, local Bookkeeping DB communication, Data handling via direct access to file system

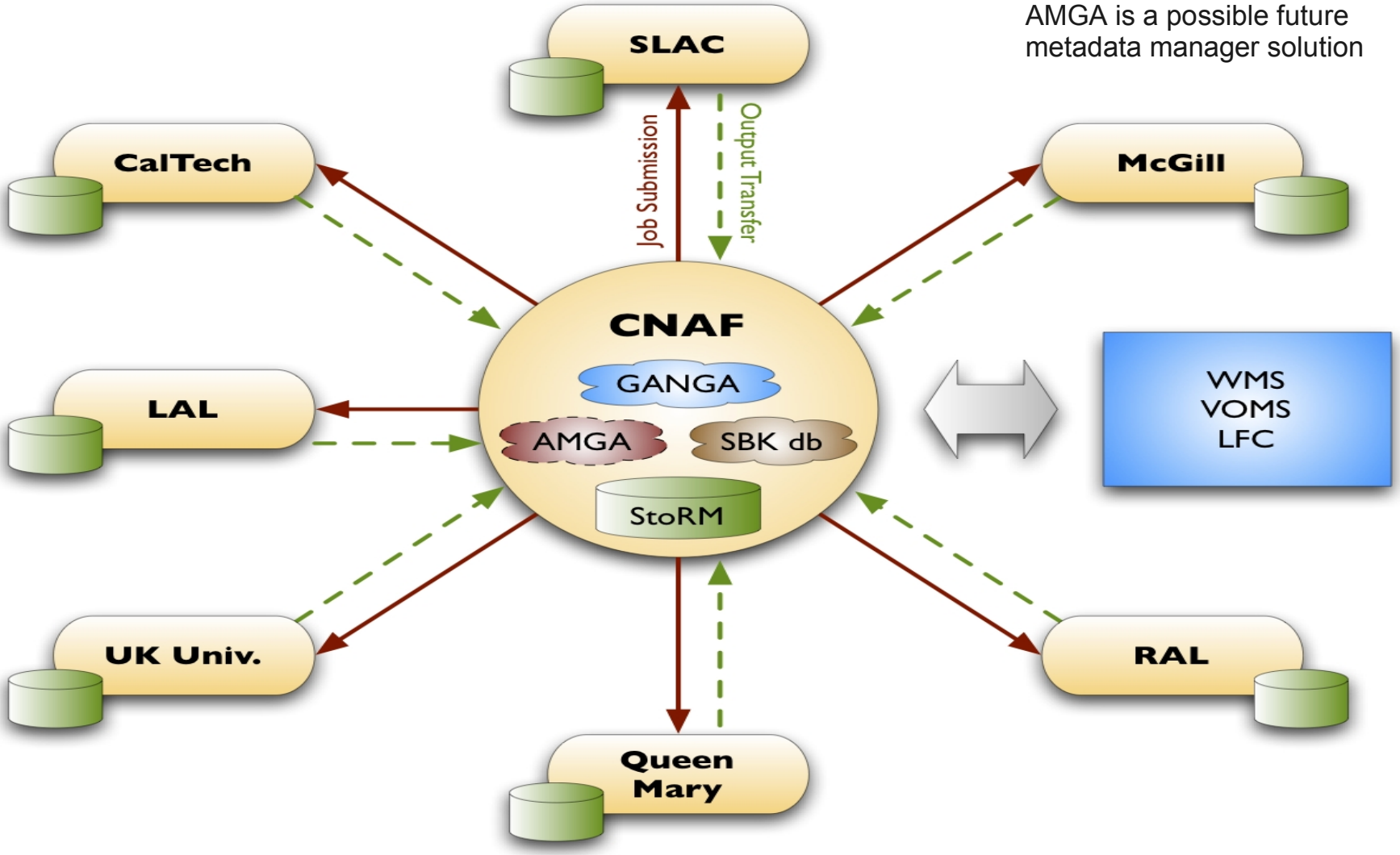
the setup of a full Grid compliant system will be proposed as solution for April '10 production

Full GANGA based job management, Grid Security Interface authentication, central bookkeeping DB communication, data handling via lcg-utils.

the site resources exploitation during official production should be ruled:
- the production managers will have priority on systems utilization
- an agreement on submission policy with user community should be defined

Full Grid integrated production design

AMGA is a possible future metadata manager solution



Full Grid integrated production workflow

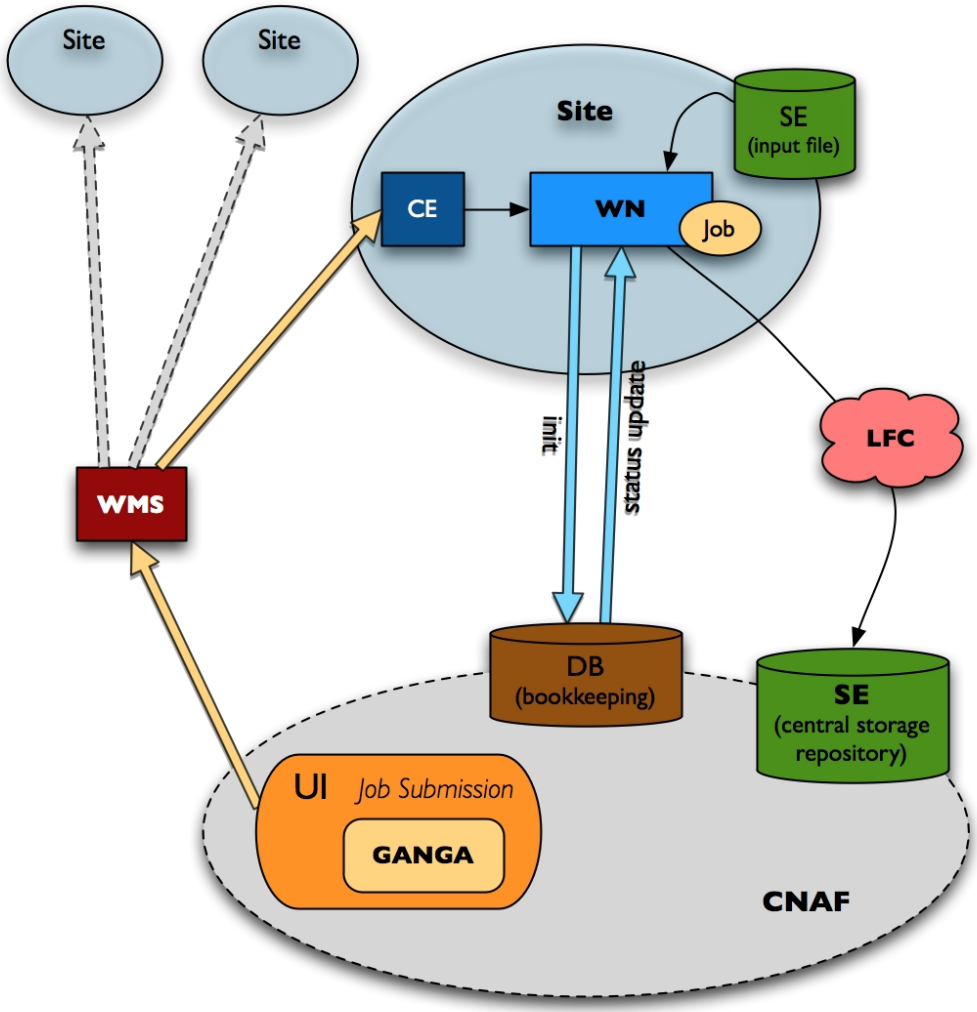
The job input files are transferred via
- LCG-Utils to the involved sites
- Storage Elements

The job submission is performed by GANGA on User Interface at CNAF

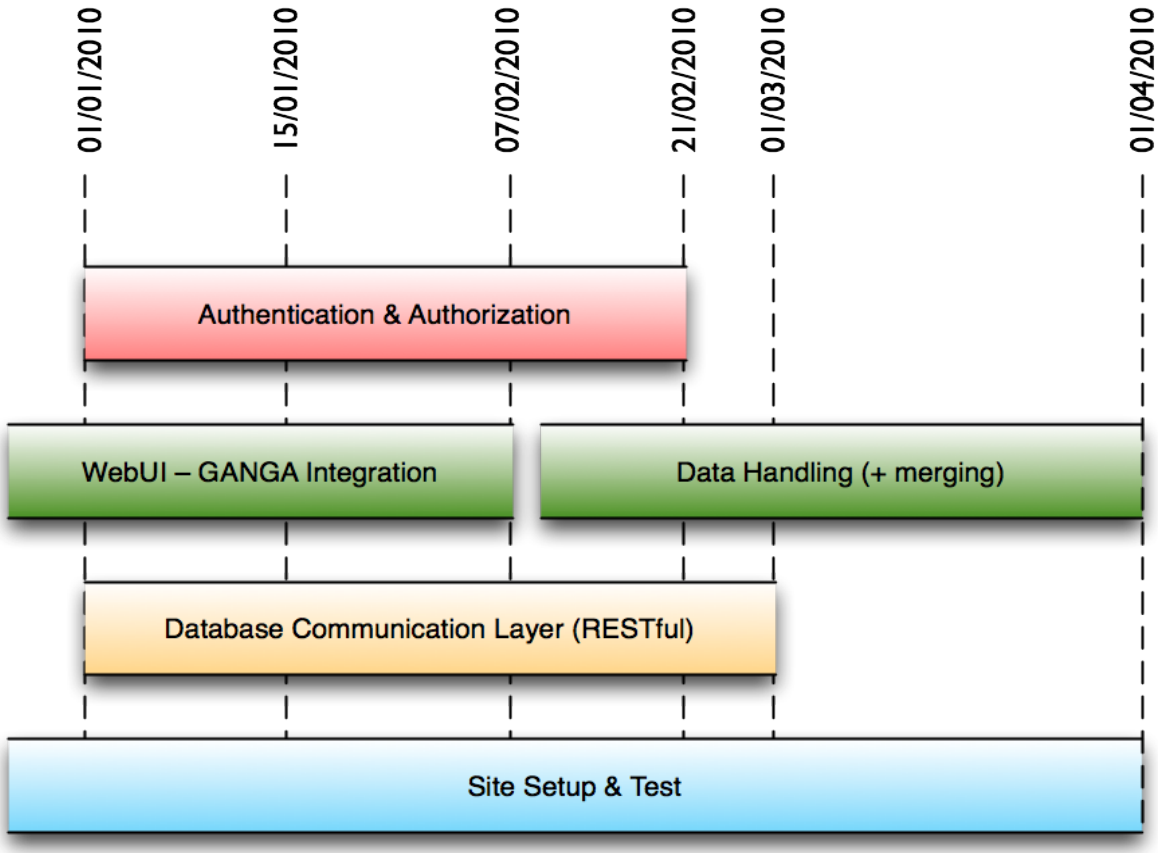
The WMS routes the jobs to the matched sites

The job is scheduled by the site Computing Element to a Worker Node

The job during running time
- accesses the DB for initialization and status update
- retrieves input files by local Storage Element
- transfers the output to the CNAF Storage Element



Full Grid integrated production timeline



February production design description

Production system components to be installed at sites:

Bookkeeping MySQL DB, phpmyadmin as management web interface (optional)
Apache web server
PHP 5 scripting language + apache php module
Web production and monitor interface software (submission to Grid and to batch)

Requirements:

One server host accessible from worker nodes where the above components should be installed (prod.server@your.site)
Disk space to contain input production files (local Storage Element)
Disk space accessible from WN, the output files should be stored in a system five like file system (Eg.: ext3, xfs, jfs, gpfs....). The output files should be copied back to CNAF central storage at end production time.

February production procedure proposal

1) Background files production at CNAF :

The background files to be used as input for Fast Simulation are produced at CNAF

2) Background files transfer to sites :

The background files should be transferred to file system accessible from WN

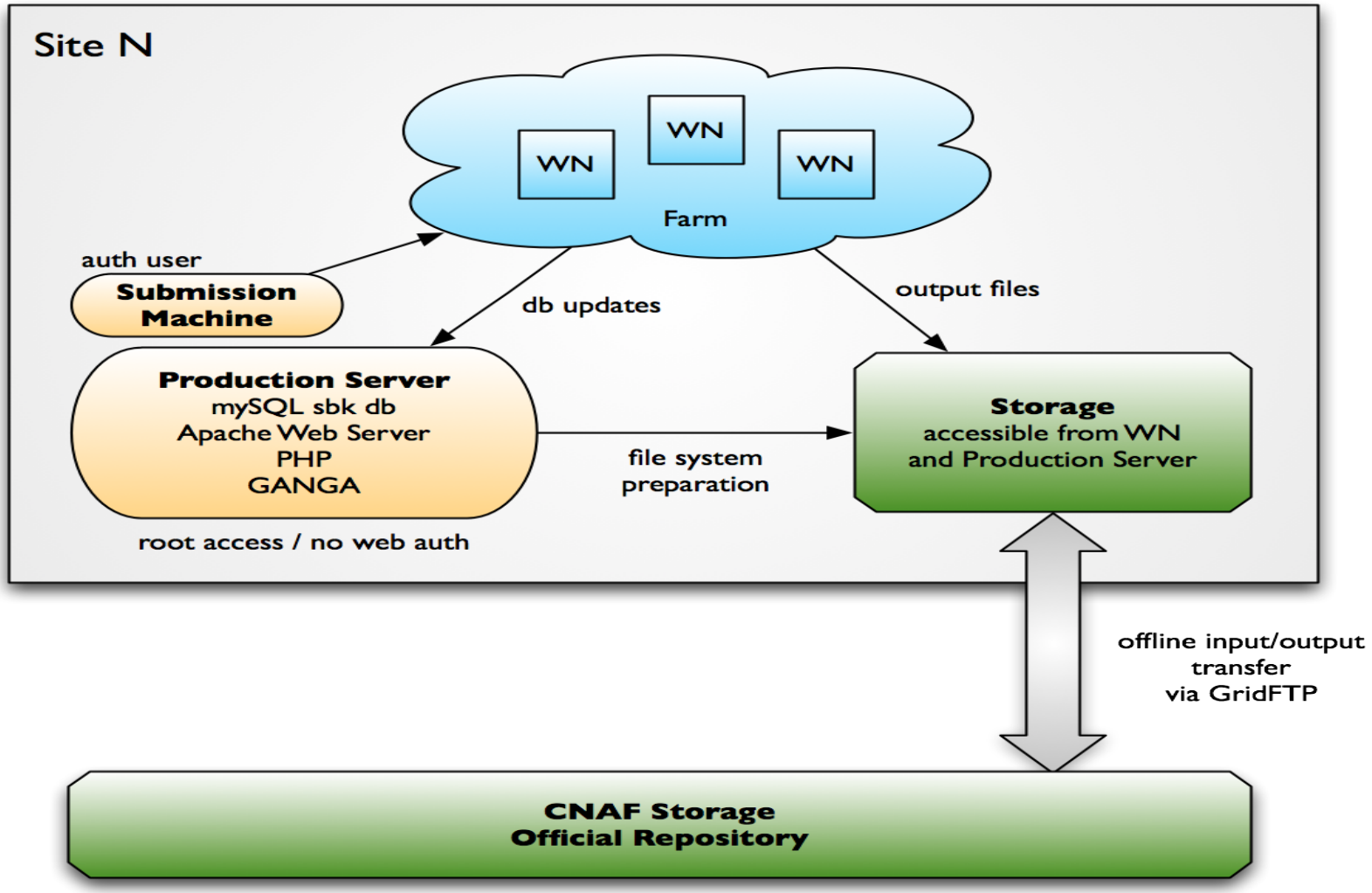
3) Perform the production at sites :

Use the web User Interface to produce the simulated data
Store the data into an accessible disk space

4) Output files transfer to CNAF central storage :

The job output files should be transferred to CNAF

February production procedure proposal



Conclusion

The February distributed solution is an hybrid solution looking toward the full Distributed (Grid compliant) solution.

The coordination/cooperation with involved site contacts is a key element

The time needed for the February production completion depends strictly by the number of sites ready at the production start time