# *Distributed production managers meeting*

Armando Fella on behalf of Italian distributed computing group

# 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**

# Site setup procedure

Wiki reference page:
http://mailman.fe.infn.it/superbwiki/index.php/How_to_Grid/Site_setup

**Three main setup procedure steps**:

A) the grid site should enable the VO on each Grid element
B) data handling solution validation: LCG-Utils
C) software installation procedure

For each point, the wiki section for simple test exists:
http://mailman.fe.infn.it/superbwiki/index.php/How_to_Grid/Site_setup/Tests

and job based test:
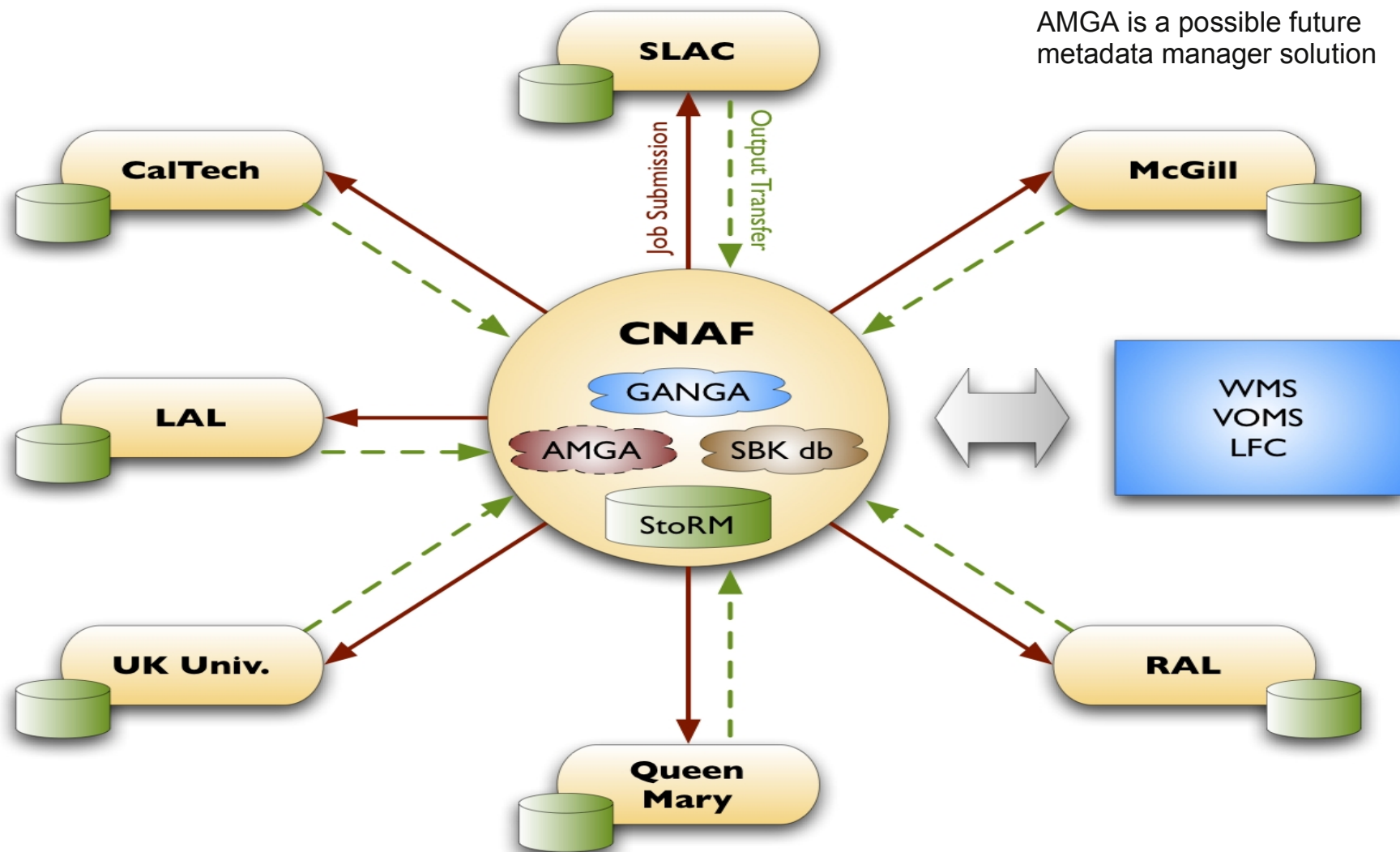http://mailman.fe.infn.it/superbwiki/index.php/How_to_Grid/Site_setup/Job

Please send an email to list superb-grid-mng@lists.infn.it reporting the
site setup status A, B or C to keep track of progress and permit the planning of
next step distributed model validation.

# 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

SLAC

Job Submission

Output Transfer

CalTech

McGill

CNAF

GANGA

AMGA    SBK db

StoRM

WMS
VOMS
LFC

LAL

UK Univ.

RAL

Queen
Mary

# 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
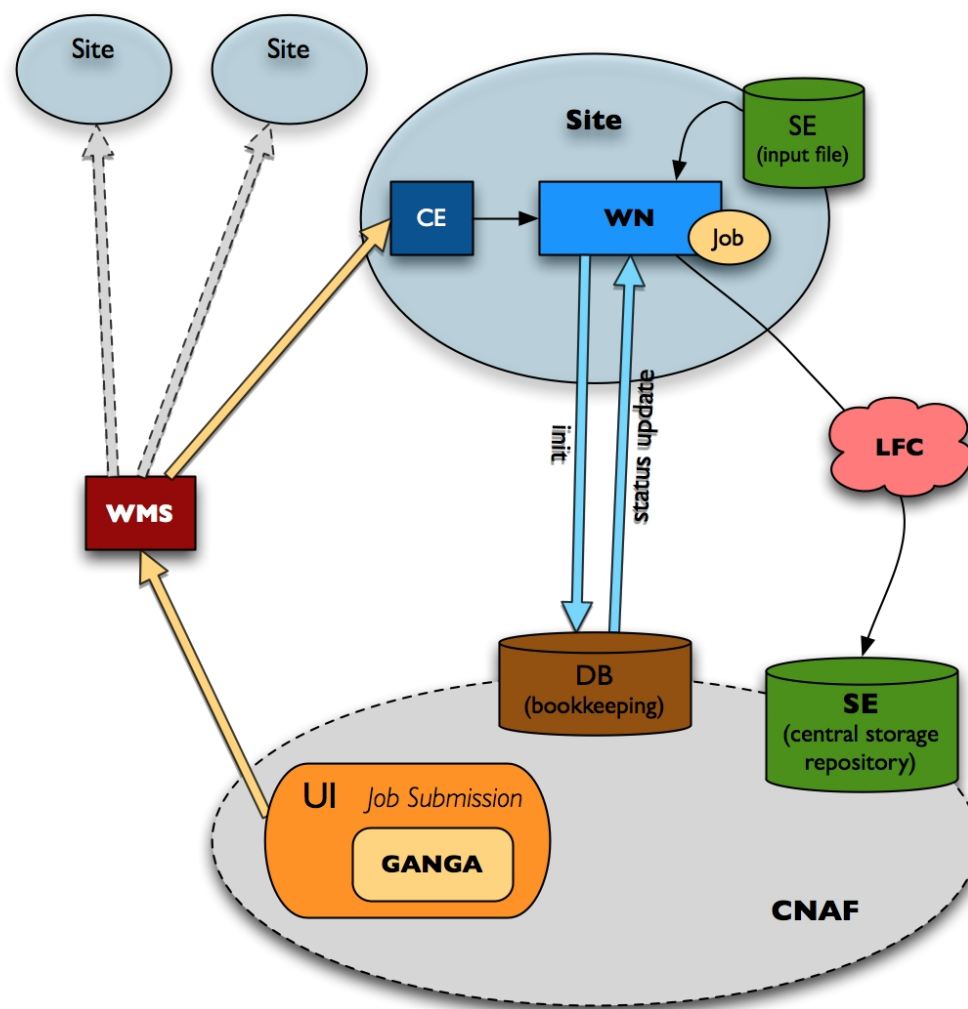
# Full Grid integrated production timeline
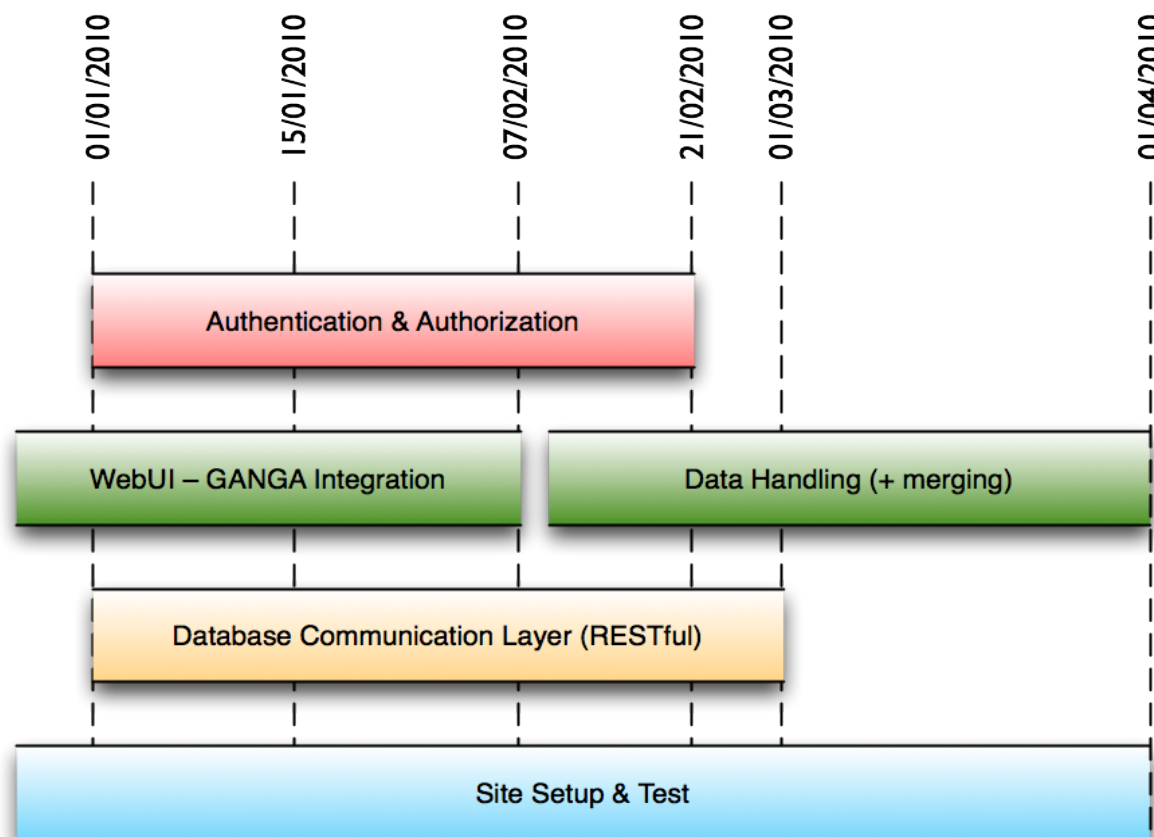
# February production design description

**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), in the Grid case it should be a User Interface
- Disk space to contain input/output production files
- The disk space should be accessible r/w from WN,
- 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
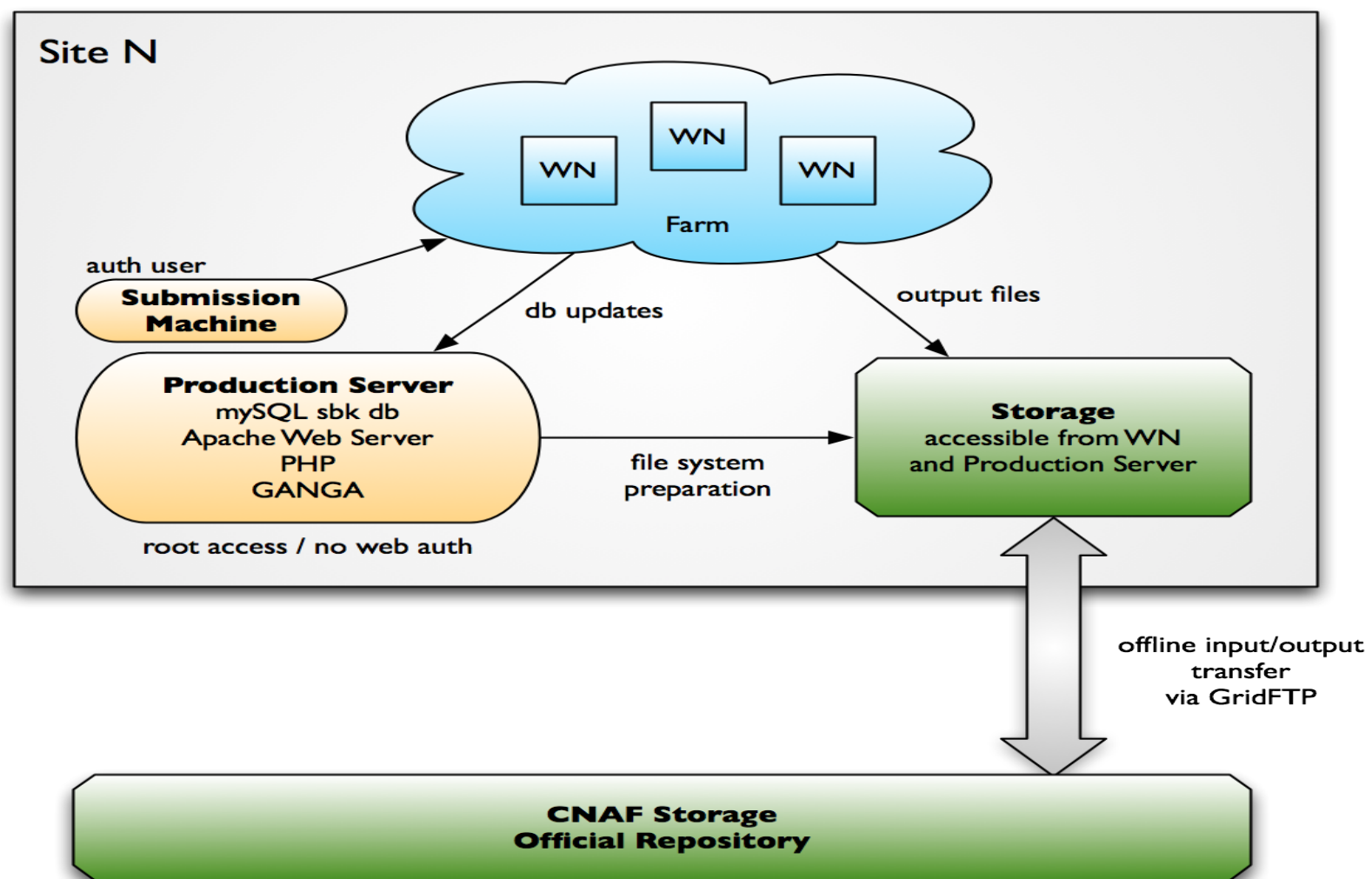
**3) Perform the production at sites :**

- 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

# Backup solution

**1) The site contacts transfer the job input files
   from CNAF to site storage**

**2) CNAF submission scripts creation and DB update**

- The production system will be customized with specific site setup (storage paths, submission launch, etc.)

**3) The site contacts launch the scripts (aka submit jobs)**

**4) The site contacts performs the transfers of output files to CNAF**

# Discussion layout

**Site resource availability/required :**
CPU, Disk, Production host

**Software related issues:**
installation procedure, Bruno compatibility

**Requirement step-by-step revision**:
moving through wiki tree looking for open issues

# Site resource availability/required

**Computational available resources per access system:**

Via Grid Computing: CPU Grid LCG, CPU Grid OSG, CPU Grid WestGrid
Via direct batch system: CPU Condor, CPU LSF, CPU via PBS, CPU batch x

- Total requirement number of cores used in two weeks: 2000/2500, 500/700 from CNAF.

**Disk space availability per access system:**

NFS like, direct Linux FS access: NFS, GPFS, GFS, AFS, Luster
Data Handling system managed: dCache, xRootd, SAM

- NFS like, direct Linux FS access, recommended solution
- For February prod a SRM interface is not needed

- Background files (job input) occupancy: ~410GB --> greater then 500GB
- Output files occupancy: more is better

| CPU | | | | | Disk space for data | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| | | | | | (2 GB per day-core) | | 9 days of prod |
| | minimum | maximum | "average" (2*min+max)/3 | | | available | needed |
| | (cores) | (cores) | (cores) | | | (TB) | (TB) |
| CNAF | **500** | **1000** | 667 | | CNAF | **50** | 12 |
| SLAC | **200** | **1000** | 467 | | SLAC | **few*10** | 8 |
| CALTEC | **250** | **250** | 250 | | CALTEC | **20** | 5 |
| McGill | **60** | **60** | 60 | | McGill | - | 1 |
| RAL | **200** | **1000** | 467 | | RAL | **10** | 8 |
| CCIN2P3 | *500* | *1000* | 667 | | CCIN2P3 | - | 12 |
| LAL | ? | ? | | | LAL | - | - |
| Pisa | **50** | **200** | 100 | | Pisa | - | 2 |
| Legnaro-Pd | **50** | **100** | 67 | | Legnaro-Pd | - | 1 |
| Napoli | **50** | **100** | 67 | | Napoli | - | 1 |
| Bari | **50** | **100** | 67 | | Bari | - | 1 |
| Ferrara | **10** | **50** | 23 | | Ferrara | - | 0 |
| Total | 1910 | 4810 | 2876.67 | | Total | 80 | |
| | **may need to be updated** | | | | | **may need to be updated** | |
| | *our guess, to be verified* | | | | | *our guess, to be verified* | |

# Software related issues

- Bruno software compatibility, the CNAF example
    - ---> Frontend machine configuration bbr-serv08

Wiki tour.

# Requirement step-by-step revision

Wiki tour

# 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

# BACKUP Slides

# Production tool status

## Grid submission: GANGA based job submission test

- job data handling tasks successfully tested
  - lfc name space creation, output transfer via lcg-util
- job submission on remote site with data handling successfully tested
- job output merger solution successfully tested

## Production tools development (FastSim)

- web production manager tool:
  - used in successful Nov '09 production
- web monitor tool:
  - the jobs update the SBK DB during run time state

See Luca Tomassetti presentation

# What is missing to the full Grid DCM

**Grid job communication with bookkeeping DB**
in a WAN scenario the job should be able to

- Authenticate it self to the DB service
  - GSI compliant authentication via proxy validation
- Creation of a RESTfull based communication service:
  - Frontier and AMGA systems under study

**Grid job management**

- Retry policy
  - resubmission procedures and error recovery in Grid env
- Grid based monitor system