INFN Roma Grid Activities

Alessandro De Salvo
Alessandro.DeSalvo@roma1.infn.it
7/5/2009

Outline

- Introduction
- Grid computing at INFN Roma
- The Tier2 Computing Center
- Activities and perspectives
- Conclusions

What is a grid?

- Relation to WWW?
  - Uniform easy access to shared information
- Relation to distributed computing?
  - Local clusters
  - WAN (super)clusters
    - Condor
- Relation to distributed file systems?
  - NFS, AFS, GPFS, Lustre, PanFS…

- **A grid gives selected user communities uniform access to distributed resources with independent administrations**
  - Computing, data storage, devices, …

Why is it called grid?

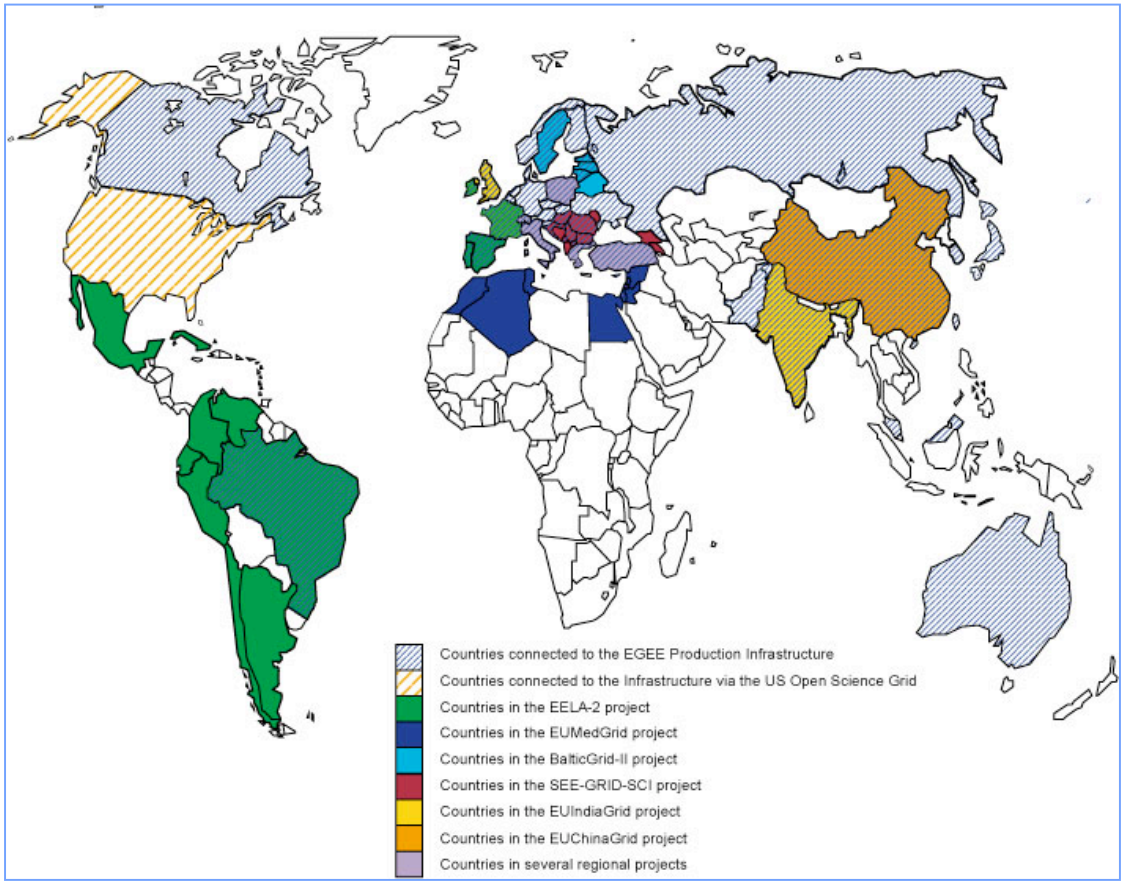
- Analogy to power grid
  - You do not need to know where your electricity comes from
  - Just plug in your devices
- You should not need to know where your computing is done
  - Just plug into the grid for your computing needs
- You should not need to know where your data is stored
  - Just plug into the grid for your storage needs

Where do we use the grids?

- Scientific collaborations
  - High-energy physics
  - Astrophysics
  - Theoretical Physics
  - Gravitational Waves
  - Computational chemistry
  - Biomed – biological and medical research
    - Health-e-Child – linking pediatric centers
    - WISDOM – "in silico" drug and vaccine discovery
  - Earth sciences
  - UNOSAT – satellite image analysis for the UN
  - Can also serve in spreading know-how to developing countries
- Industry? Commerce?
  - Research collaborations
  - Intra-company grids
- Homes? Schools?
  - E-learning

Grids in the World

- **The general INFN Roma computing infrastructure is part of the following Grids**
  - **IGI (Italian Grid Infrastructure)**
  - **WLCG (World LHC Computing Grid)**
  - **EGEE (Enabling Grid for E-sciencE)**
- **The most important grid users in Roma are**
  - **The ATLAS experiment at LHC**
  - **The CMS experiment at LHC**
  - **The VIRGO experiment**
  - **The Theoretical Physics group**



Why Grid at LHC?

- 10-15 PetaByte/year
- Data analysis requires at least ~100k typical PC processors
- Scientists in tens of countries worldwide
- CERN can provide up to 20-30% of the storage and CPU
- 70-80% are provided by WLCG partners

Why Grid for VIRGO?

- Gravitational wave detector Virgo has accumulated several months of data up to now (130 TB in total). The next scientific run, VSR2, will start in July 2009 and will last ~6 months, producing additional ~130 TB of data
- Some of the analysis to be carried on the data require large computational resources. In particular the 'blind' search for continuous gravitational signals from rotating neutron stars, an activity on which the Rome group has the leadership
- VIRGO uses the grid to transparently access the computing resources needed for the analysis. The minimal amount of computing power needed to analyze 6 months of data is ~1000 dedicated CPUs for 6 months.
  The bigger the available computing power, the larger is the portion of the parameter space that can be explored (computing bounded problem)

Why Grid for Theoretical Physics?

- **Study of the QCD properties, i.e. the theory of the quark-gluon interactions**
  - Mainly based on numerical simulations
  - SU(3) Lattice QCD computations in 4 dimensions
- **Computing requirements**
  - Different requirements on RAM, depending on the lattice volume:
    - 12^4, 14^4, 18^4 <--> 250 MB, 500 MB, 1.1 GB respectively
  - Trivial parallelism, using N copies of the same program with different initial random seeds (N=200 for the biggest lattice)
    - Grid as a probability distribution generator
  - Very small, optimized executable, written in C
  - Using the SSE2 set of instructions, available in the most recent Intel CPUs
  - Very limited amount of disk space needed for the output
  - 2005-2006 (L.Giusti, S.Petrarca, B.Taglienti: POS Tucson 2006, PRD 76 094510 2007)
    - 134800 configurations ≈ 800000 CPU hours ⇒ ~1 year @ 100 CPU dedicated farm
  - Last 1.5 years
    - 70 000 configurations (run) + 40000 configurations (test) ≈ 660000 CPU hours

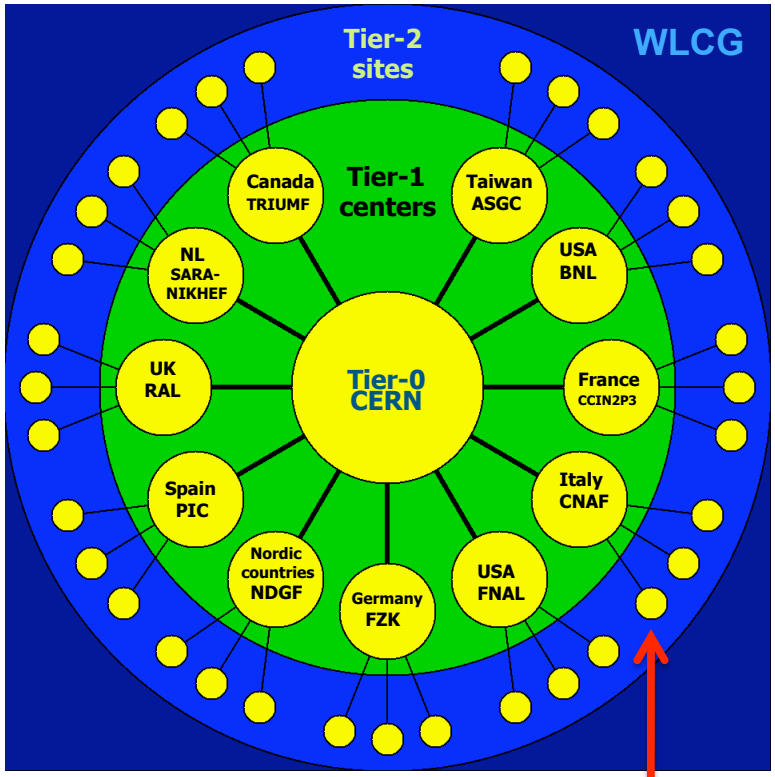
Work to be presented at *Lattice2009*

- ***Only possible via Grid Computing!***
  - New lattice dimensions will require ≥ 2 GB of RAM and dedicated run sessions on more than 100 CPUs (or ~100 dedicated CPUs for 6 months)

Challenge for the resources currently available in the grid

Organization of the Grid computing centers

- The Grid Sites are organized hierarchically
  - Generally referred as “Tiers”
- The biggest sites, in terms of resources and support level, are at the top of the hierarchy (lower levels), while the smaller ones are at the bottom (higher levels)
- Example: the WLCG Tiered structure
  - CERN (Tier-0)
  - Regional (National) Centers (Tier-1s)
  - Local Centers (Tier-2s)
- The sites located higher in the hierarchy are coordinating and supporting the sites below them

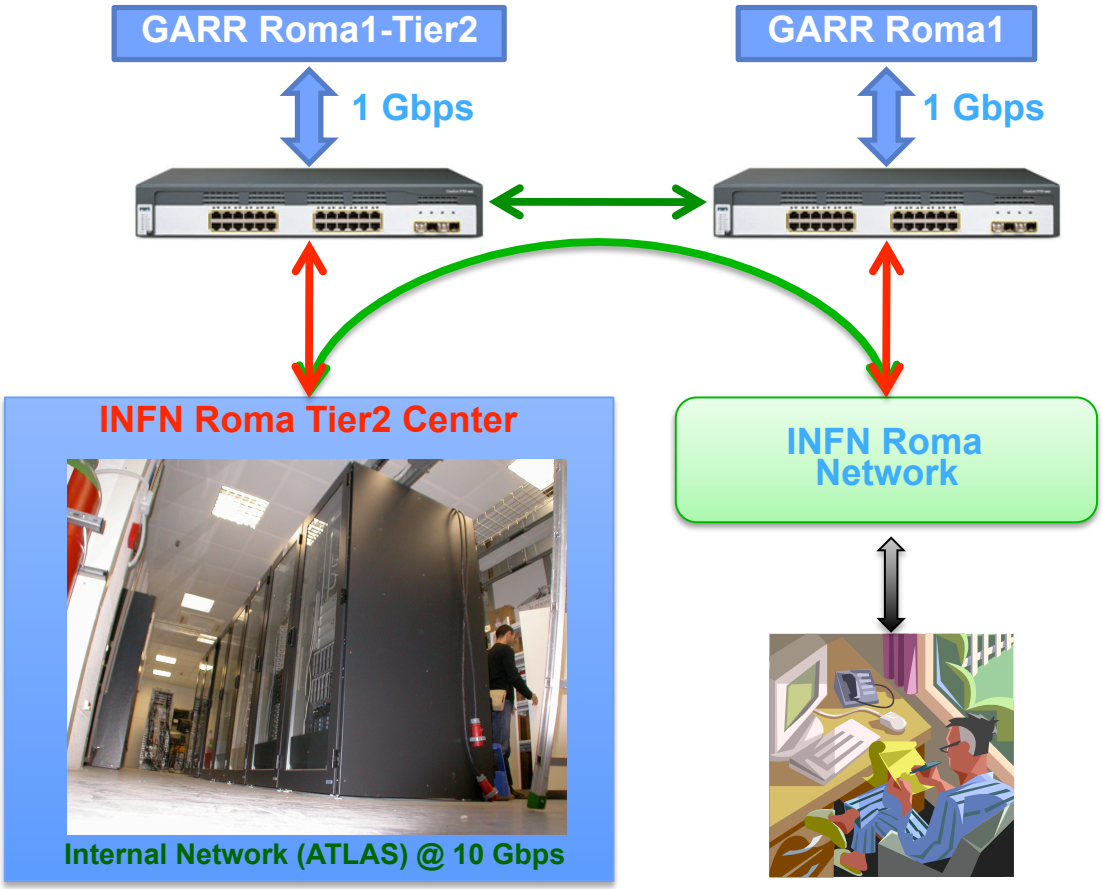


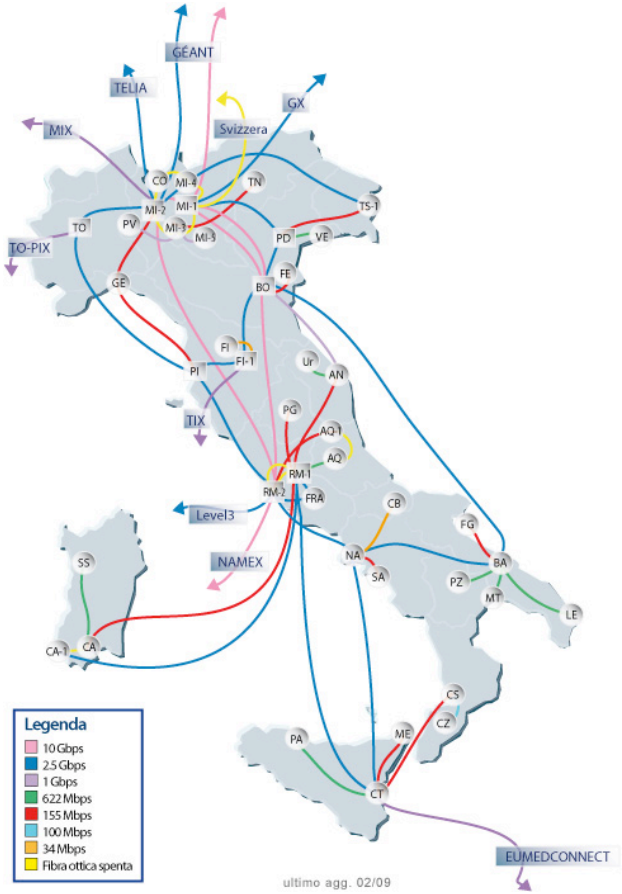
The INFN Roma Tier2 center

- 7 cooling racks, expandable to 14
  - Self-conditioned, using water
  - Remote Monitoring System
  - Up to 40 units per rack
    - With the current CPU form factor, it would be ~9000 CPU cores!

- Hosted experiments
  - ATLAS (WLCG Tier2)
    - 286 CPU cores
      - LSF batch system
    - ~200 TB storage space
      - DPM Storage
  - CMS (WLCG Tier2)
    - 320 CPU cores
      - LSF batch system
    - 110 TB storage space
      - dCache Storage
  - VIRGO
    - 416 CPU cores
      - Torque/Maui batch system
    - 10 TB storage space
      - DPM Storage

Network connectivity for the Tier2 center

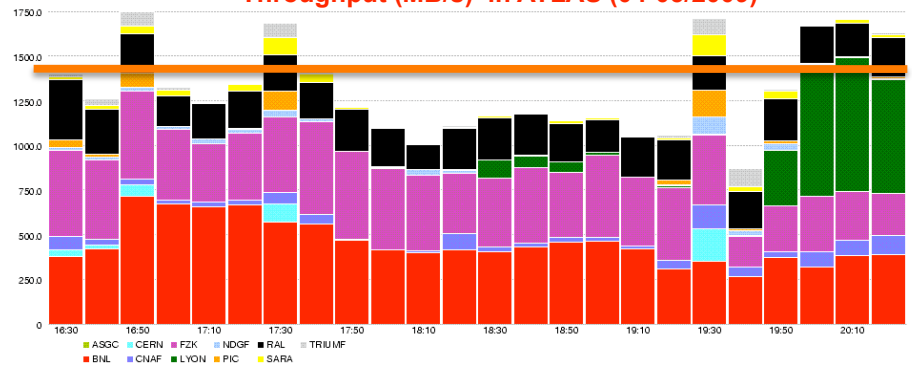




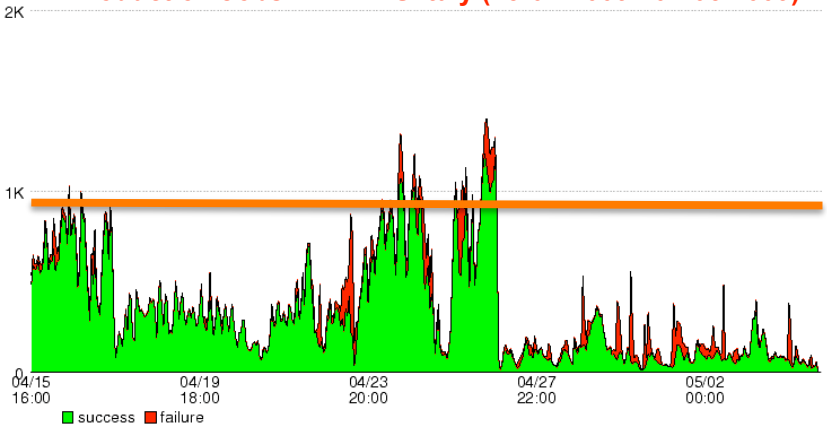
Data processing using Grid tools ATLAS

- MonteCarlo Production
  - PANDA (Production ANd Distributed Analysis)
    - http://panda.cern.ch:25880/server/pandamon/query?dash=prod
- User Analysis
  - PANDA
  - GANGA (Gaudi / Athena aNd Grid Alliance)
    - http://cern.ch/ganga
- Online Muon Calibration

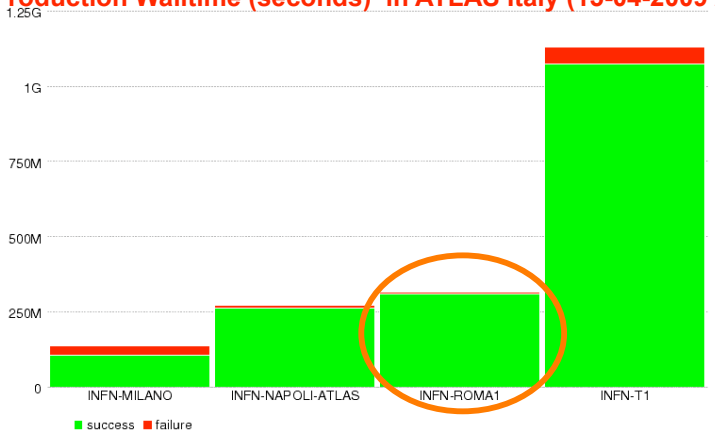
Data Transfer via the ATLAS DDM (DQ2)
Throughput (MB/s) in ATLAS (04-05/2009)



MC Production Jobs in ATLAS Italy (15-04-2009 / 04-05/2009)



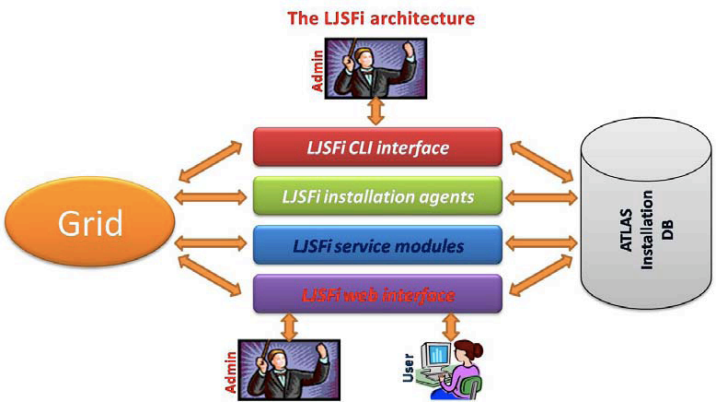
MC Production Walltime (seconds) in ATLAS Italy (15-04-2009 / 04-05/2009)

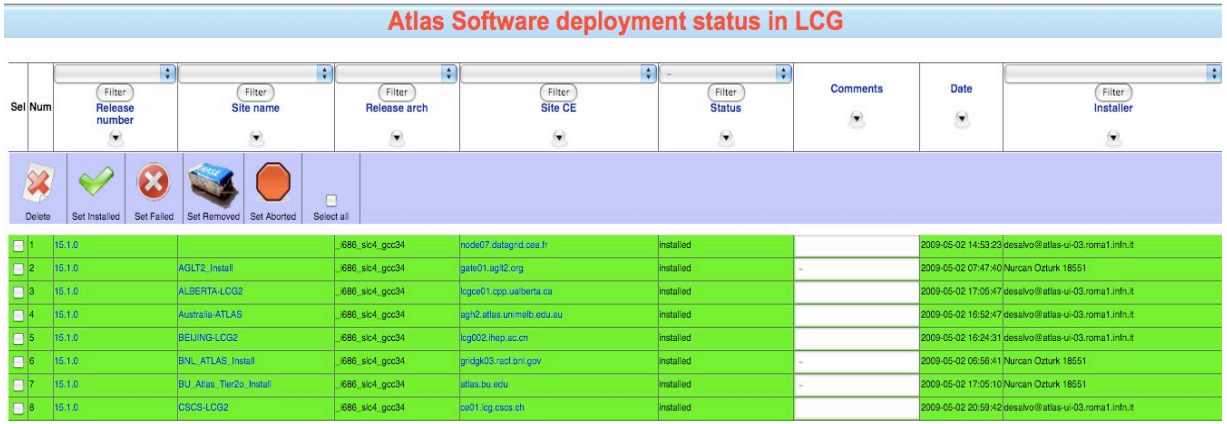


Grid services in Roma: the ATLAS Software Installation System

- **Core service for the ATLAS Software Installations in LCG/EGEE and OSG Grids**
  - Serving more than 200 sites in unattended or semi-automatic (assisted) mode
  - Single instance running in Rome and serving all the ATLAS collaboration since 2003
  - Online view of the availability of the software deployment in all the LCG/EGEE (Europe) and OSG (Open Science Grid, US) sites
  - On-demand installation, testing and management of the releases
  - Based on the LJSFi (Light Job Submission Framework for Installation) engine, entirely developed in Rome for LCG in 2003, exanded to OSG n 2009 and planning to include NorduGrid (Nordic Countries) in the next months, unifying the Installation System for full WLCG ATLAS Collaboration
    - Work to be done here!
  - Designed for multi-Grid and multi-VO operations
  - https://atlas-install.roma1.infn.it/atlas_install

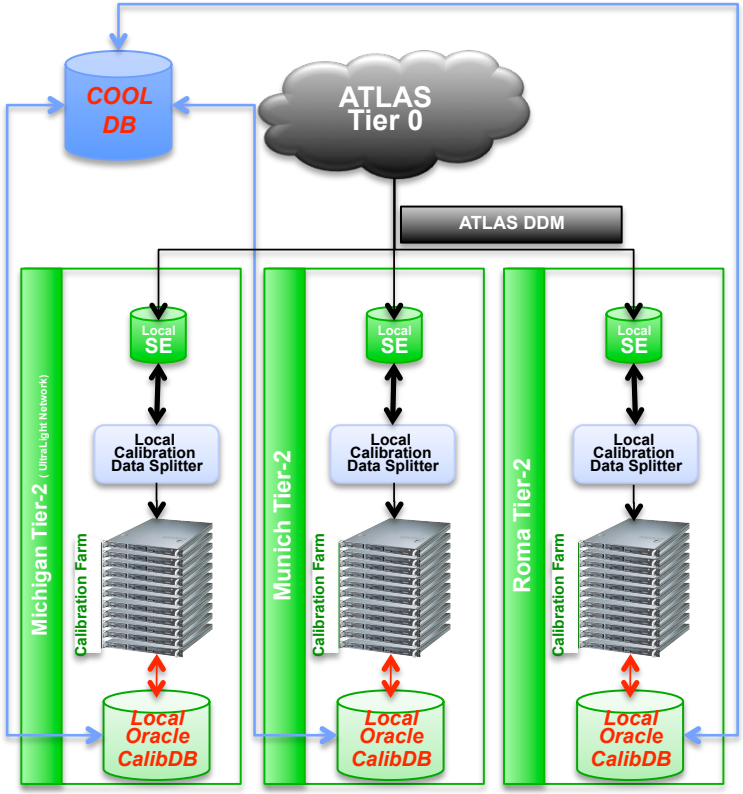
Presented at CHEP 2007



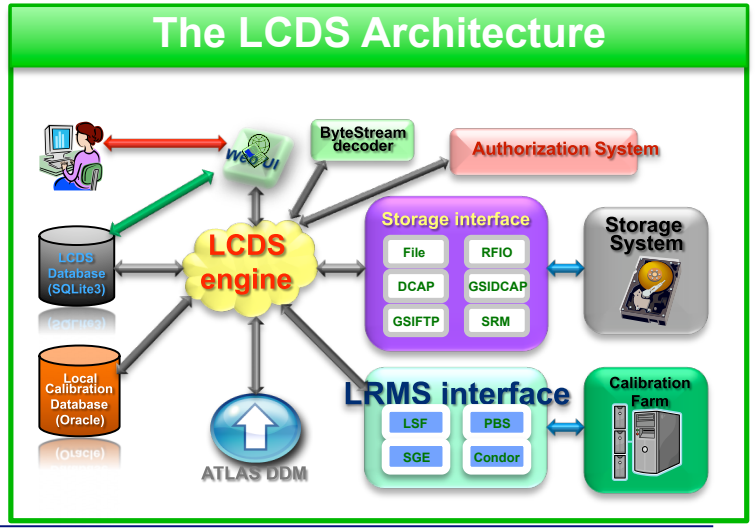


The ATLAS MDT calibration center

- **Online Calibration of the Monitored Drift Tubes (MDT) of the ATLAS Muon Spectrometer**
  - Data coming from CERN via the ATLAS Distributed Data Management system (aka DQ2) must be processed in quasi-streaming mode to calculate the calibration constants in 24/48 hours
  - 3 calibration centers (Michigan, Munich and Rome) share the calibration tasks in different regions of the detector
  - The full calibration procedure is supervised in each center by an agent called Local Calibration Data Splitter (LCDS), fully integrated with the Grid and local tools needed
  - The architecture design and the control tools are direct responsability of the INFN ATLAS Roma group

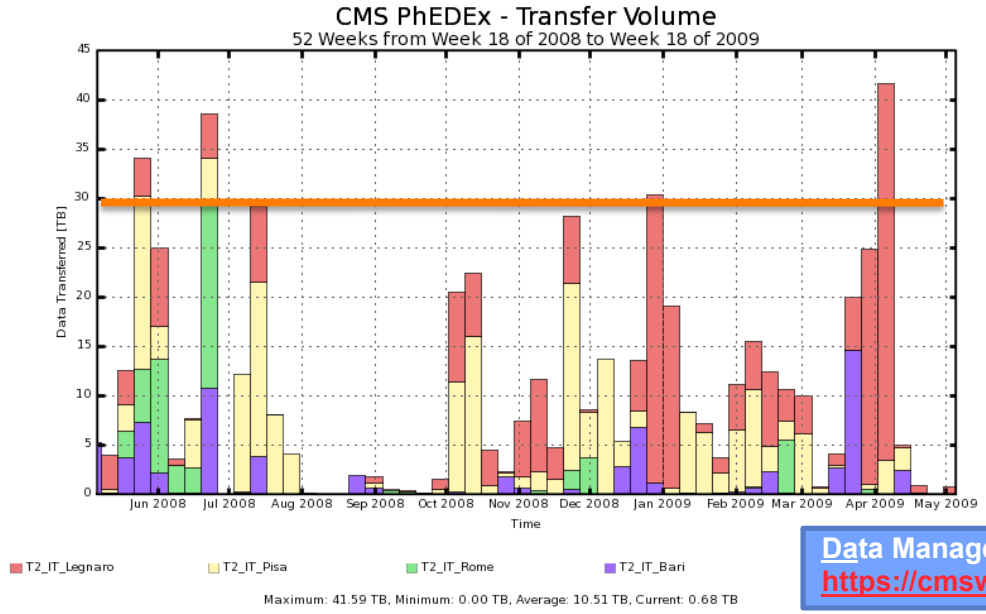


Presented at CHEP 2009



Data processing using Grid tools CMS

- MonteCarlo Production
- User Analysis
  - CRAB (CMS Remote Analysis Builder)
    - https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideCrab
- Offline Calibration



Data Management from PhedEx
https://cmsweb.cern.ch/phedex/prod/Info::Main

CMS ECAL calibration

Strategy	Time	Precision
Mean energy deposited by jet triggers independent of ϕ at fixed η (after correction for Tracker material)	few hours	~2-3%
Neutral pion mass peak: @ $2x10^{33}$ cm^{-2} s^{-1}	few days	≤1%
Z→ ee: absolute calibration	100 pb^{-1}	< 1%
W→ eν: E/p measurement	5-10 fb^{-1}	0.5%

"on-line": short term calibration (+laser monitoring correction) at CAF (CERN)

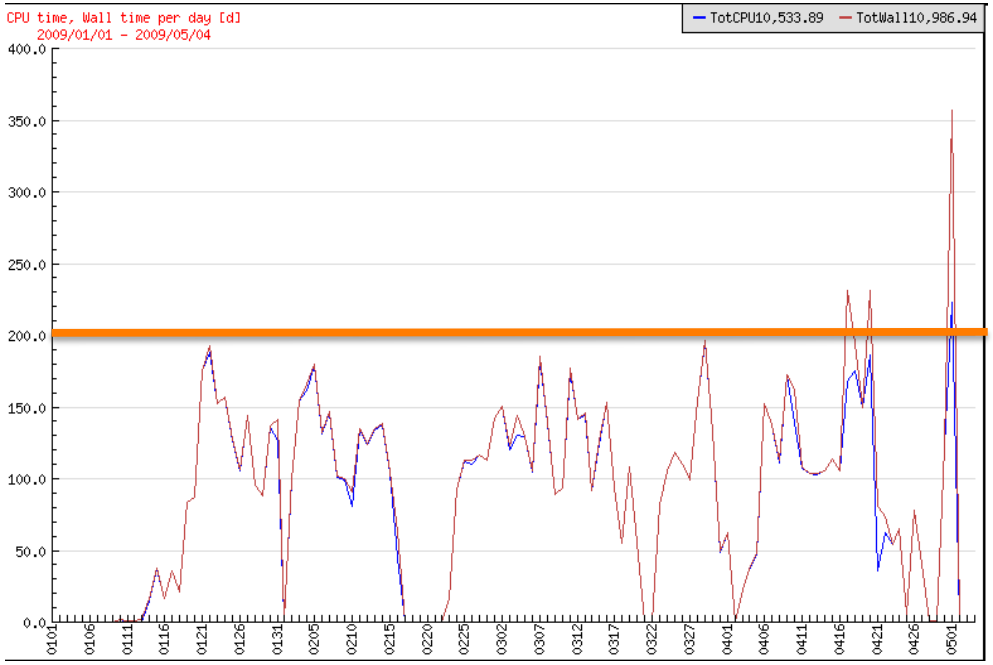
"off-line": long term calibration at Tier2

- **Copy of single and double electron dataset in RECO and ALCARECO formats**
- **Iterative calibration algorithms run every n pb^{-1}**
- **Crystal calibration constants sent back to CERN**

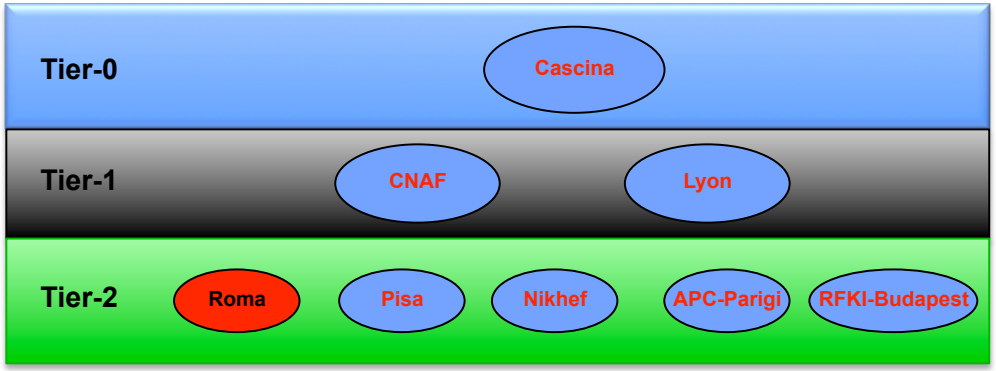
Data processing using the Grid tools VIRGO

- Mainly used for Data Analysis
- Computing resources accessible by
  - VIRGO (higher priority)
  - THEOPHYS
  - ARGO

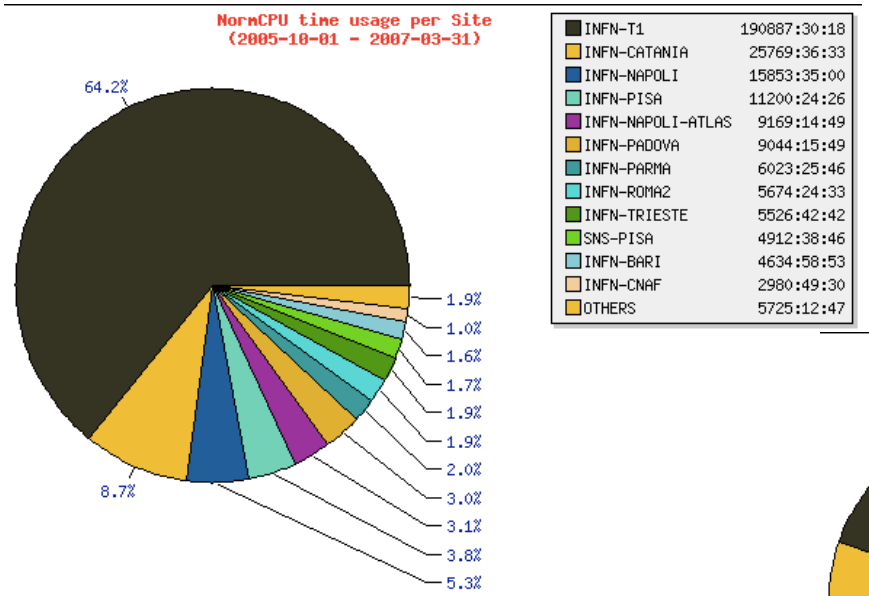
- Farm utilization in Roma since January 2009
  - 43000 jobs
  - 252800 h total CPU time



Raw data are produced in Cascina and copied to Tier-1s
Calibrated data are replicated in (some of the) Tier-2 (Roma, Pisa, RFKI)

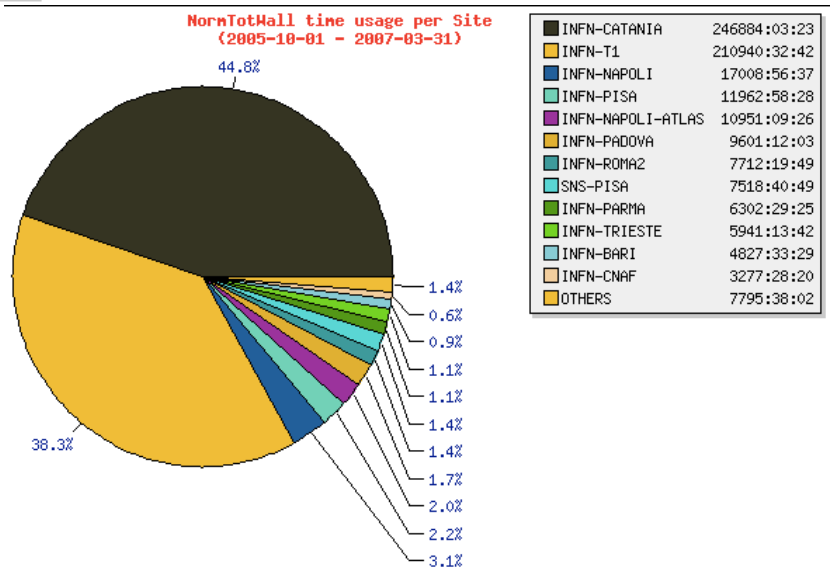


Data processing using the Grid tools Theoretical Physics



INFN-T1	190887:30:18
INFN-CATANIA	25769:36:33
INFN-NAPOLI	15853:35:00
INFN-PISA	11200:24:26
INFN-NAPOLI-ATLAS	9169:14:49
INFN-PADOVA	9044:15:49
INFN-PARMA	6023:25:46
INFN-ROMA2	5674:24:33
INFN-TRIESTE	5526:42:42
SNS-PISA	4912:38:46
INFN-BARI	4634:58:53
INFN-CNAF	2980:49:30
OTHERS	5725:12:47

2005 - 2007



INFN-CATANIA	246884:03:23
INFN-T1	210940:32:42
INFN-NAPOLI	17008:56:37
INFN-PISA	11962:58:28
INFN-NAPOLI-ATLAS	10951:09:26
INFN-PADOVA	9601:12:03
INFN-ROMA2	7712:19:49
SNS-PISA	7518:40:49
INFN-PARMA	6302:29:25
INFN-TRIESTE	5941:13:42
INFN-BARI	4827:33:29
INFN-CNAF	3277:28:20
OTHERS	7795:38:02

Plots from the Grid Monitoring Systems
http://gridice4.cnaf.infn.it:50080/gridice/site/site.php
https://dgas.cnaf.infn.it

Summary of the Grid-related activities in Roma

- **LHC experiments**
  - ATLAS and CMS calibration activities
  - ATLAS Software Installation in the Grid sites
- **European projects**
  - WLCG (World LHC Computing Grid) and EGEE (Enabling Grid for E-sciencE)
    - High Energy Computing Clusters [ATLAS]
      - High-level Job Scheduling system (Job Priority)
    - LCG 3D Distributed Database System
      - Mostly related to the Oracle installation for the ATLAS calibrations
      - Active backup of the ATLAS LCG File Catalog at CNAF
        - Work presented at CHEP 2009
- **Italian Grid Infrastructure (IGI)**
  - High-performance storage systems
    - dCache [CMS]
    - DPM [ATLAS, VIRGO]
    - StoRM [ATLAS]
  - Workload Management System interaction with HEP Clusters [ATLAS]

Conclusions

- Roma is fully integrated with the Italian Grid Infrastructure and the WLCG systems
- The ATLAS and CMS collaborations are actively using the resources located in Roma to produce Monte Carlo data, perform user analysis and computation of calibration constants
  - The system is robust and powerful enough to cope with the incoming LHC data taking
- Other experiments or groups (VIRGO and the Theoretical Physics) are already actively using the grid services to cope with their computing requirements
- Several Grid-related activities are present in Rome
  - Included in Italian, European and/or Worldwide projects
  - Room for further development and improvements
- Always trying to improve the levels of reliability and stability of the systems
- All the users in Roma are invited to take advantage of the experience of the people involved in the Grid projects and join the efforts

The Grid is waiting for you!

More information

http://www.cern.ch/lcg

http://igi.cnaf.infn.it

http://www.eu-egee.org

http://www.gridcafe.org

ATLAS

http://atlas.web.cern.ch

http://cms.web.cern.ch

http://www.virgo.infn.it