Approved projects

- EnviroGRIDS
  - Recent project conference
  - Deliverables under control
  - Fellow being renewed (second 12-month block)
  - Good synergies with the rest of the group
    - Ganga (part of the Ganga team) and Dashboard (ATLAS and CMS user monitoring)
- PARTNER
  - 4 Fellows hosted in our group (IT-ES)
    - They are PhD students (at various universities in UK and ES)
  - Deliverables under control
  - Looking for good synergies with the rest of the group
- ULICE
  - Hope to have a fellow from the last round

Submitted proposals

- Gridimob (submitted on April 13th)
  - ICT Call 6, Objective 6.2 ICT for Mobility of the Future
    - 37 M€
      - Project duration: 36 months
      - Total project budget: 7.3 M€
        - EC funding: 4.8 M€ (as a STREP)
        - EC funding for CERN: approx 400k euros (we are not WP leader)
  - Looks to be very competitive
    - 114 projects have been submitted !
    - Gridimob is ~12% !

INFRA2011-1.2.1

- The call FP7 INFRA2011-1.2.1 reserves a budget of 27M EUR for applications – eScience environments

  - The EU instrument is CP-CSA hence relatively lighter than an I3 or similar

  - The time scale is November 2010 for the submission

  - This is **not** the data management call!

- Information day: June 11th in Brussels

Call topics: e-Science environments

- 4 out of 3...

  1) Integrated service provision based on cloud, grid or hybrid cloud grid

  2) User-friendly interfaces which abstract service provision

  3) *Access to (remote) instruments*

 4) Deployment of eScience support and training centres

Ganga

User-oriented grid access

- Natural layer to use the grid

- Unified computing environment (batch, Grid, HPC, Cloud)

Stable, mature, high-quality user environment

- \> 3,000 unique users since 2007

- Originated in HEP (ATLAS and LHCb) – used by many other communities

- Integrated with other tools (Monitoring/Dashboard)

Unique users by week and experiment
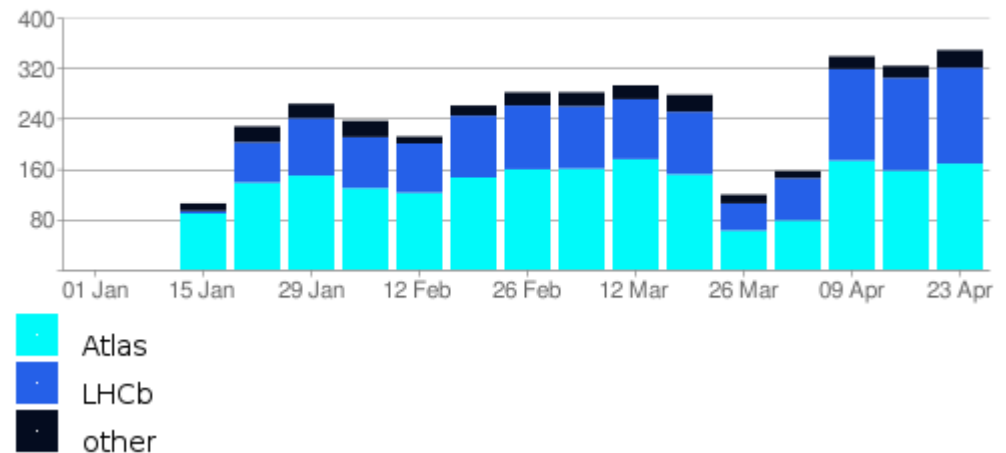
Ganga: High-Energy Physics
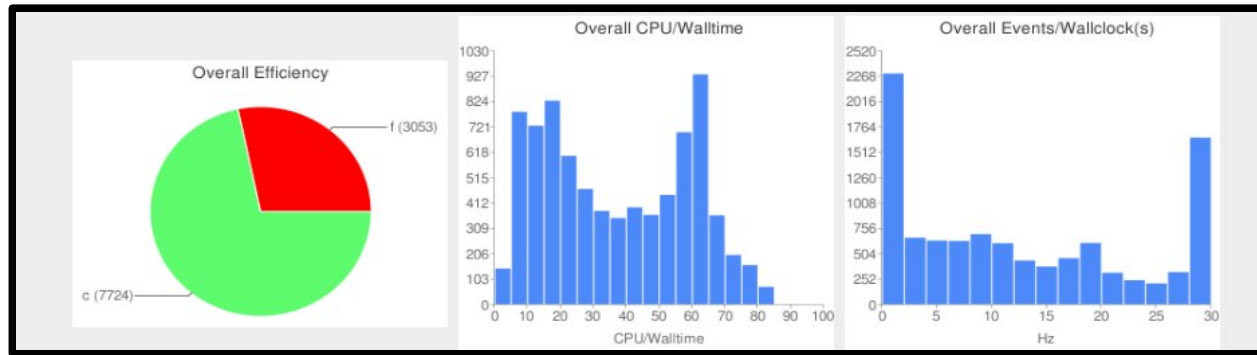
- Originated in HEP (ATLAS+ LHCb)
    - Main goal: user analysis
- Extended to site commissioning, testing, trouble shooting activities
    - GangaRobot
    - HammerCloud
- Integrated with the monitoring into a uniquely mature tool used by scientists
    - ...and makes support of large communities possible

Ganga and Supercomputers

Scientific workflows need **all** available resources:

- Local resources (batch)
- Supercomputers
- Grids

possible only if "porting" costs are minimal

✓ Optimal workflow execution (time and resources) requires the combination of different resources (for example, laptop/batch/grid or grid/HPC). Ganga is being used also in this way (several examples)

Example: Lattice QCD

CERN IT + CERN PH + ETH Zürich

NEC-SX8 – HLRS Stuttgart (**DEISA**)   **WLCG/EGEE**   **WLCG/EGEE**

"mother snapshot"

thermalization — 1 month

decorrelate — 2 months

measurements — 4 months

Up to 1,500 jobs, over 750 CPU•years

0   MC steps

supercomputer → Grid
analysis → run1 → optimization of scheduling → run2 → analysis → run3 → optimization of scheduling
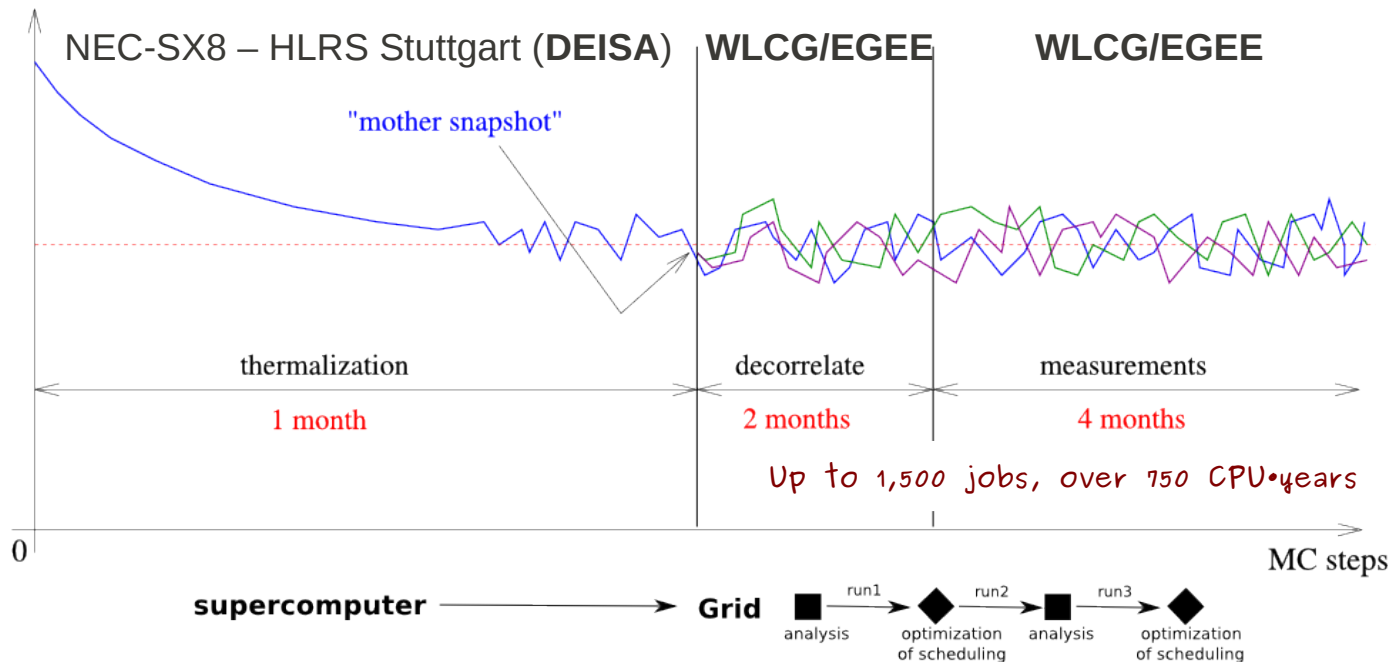
TeraGrid

Ganga – a unified approach for accessing diverse computing environments

... to reach out for **new resources**

... to achieve **application-level interoperability**

Examples of systems supported by Ganga

✔ **via SAGA → TeraGrid**

- Collaboration with Lousiana State Univ.
- SAGA back-end for Ganga contributed by LSU

✔ via gLite →(EGEE)/WLCG, OSG

✔ via Globus/GridWay → PRAGMA (Asia/Pacific)

✔ *Experiment-specific system (ATLAS-PanDA and DIRAC-LHCb)*

High Performance Systems						
Name	Institution	System	Peak TFlops	Memory TBytes	Status	Load
Kraken	NICS	Cray XT5	1030.00	129.00	Up	
Ranger	TACC	Sun Constellation	579.40	123.00	Up	
Abe	NCSA	Dell Intel 64 Linux Cluster	89.47	9.38	Up	
Lonestar	TACC	Dell PowerEdge Linux Cluster	62.16	11.60	Up	
Steele	Purdue	Dell Intel 64 Linux Cluster	60.00	12.40	Up	
Queen Bee	LONI	Dell Intel 64 Linux Cluster	50.70	5.31	Up	
Lincoln	NCSA	Dell/Intel PowerEdge 1950	47.50	3.00	Up	
Big Red	IU	IBM e1350	30.60	6.00	Up	
Frost	NCAR	IBM BlueGene/L	22.90	2.00	Up	
Cobalt	NCSA	SGI Altix	6.55	3.00	Up	
Pople	PSC	SGI Altix 4700	5.00	1.54	Up	
NSTG	ORNL	IBM IA-32 Cluster	0.34	0.07	Up	
		Total:	1984.62	306.3		

Storage Systems				
Name	Institution	System	Online Storage TB	Offline Storage TB
IU Archival Storage	IU	HPSS	N/A	28000
SDSC Tape Storage	SDSC	HPSS	N/A	25000
Ranch	TACC	Sun StorageTek Mass Storage Facility	57	10000
MSS	NCAR	Mass Storage System	N/A	2000
Data Capacitor	IU	Lustre	535	N/A
SDSC GPFS-WAN	SDSC	Global Parallel File System-Wide Area Network	700	N/A
		Total:	1292	65000

Clouds

Google Summer of Code 2009:

Ganga + CernVM + EC2 + ATLAS
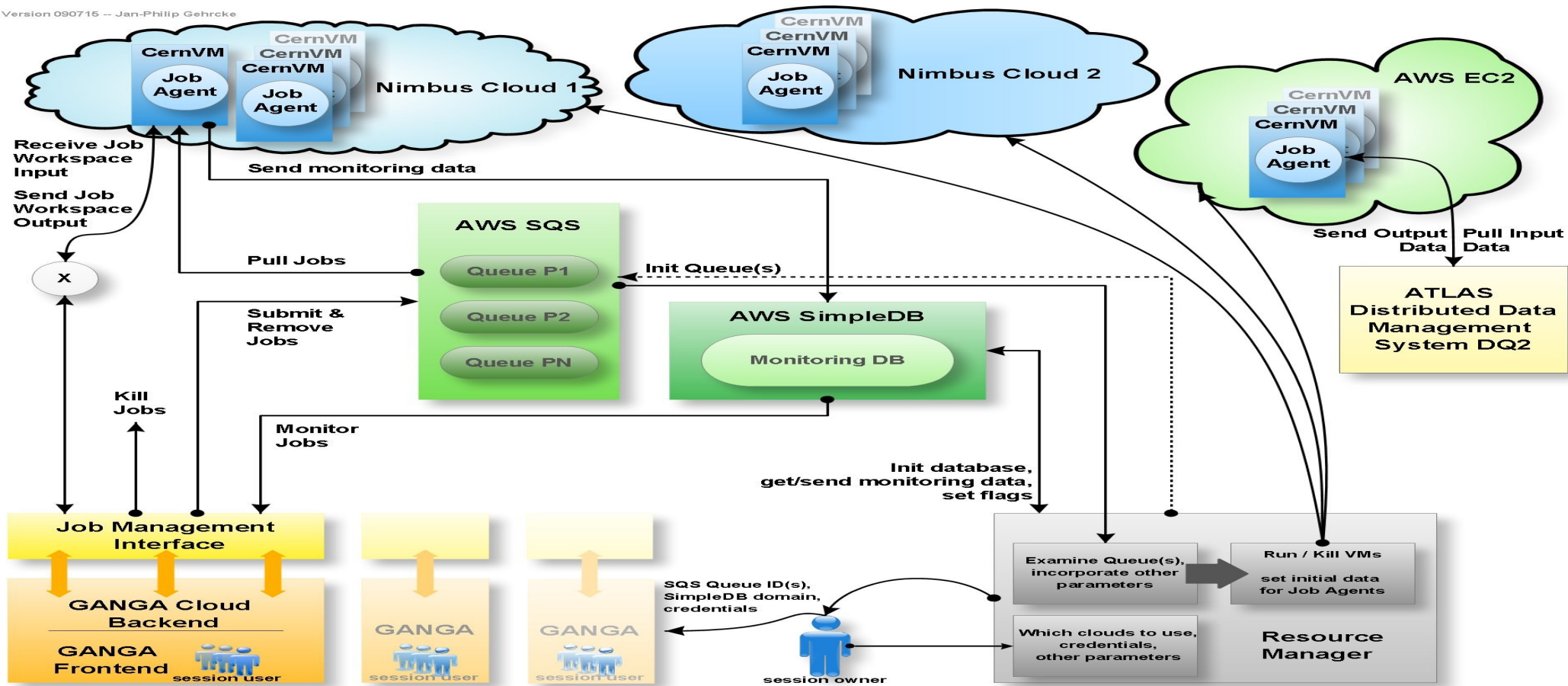
Possible partners (material for internal discussion)

- Imperial College London
  - Contributing 1-2 FTE to Ganga for several years + the project leader (U. Egede)
- An handful from HEP
  - Instituto Nazionale Fisica Nucleare
    - CRAB and CMS user analysis support
  - Munich LMU
    - Contributing 1-2 FTE to Ganga for several years + the ATLAS distributed analysis coordinator (J. Elmsheuser)
  - CERN
    - Core contributions to analysis tools and support (notably Ganga), monitoring (dashboard), solid links with all the 4 LHC experiments...

- Non-HEP "application portfolio"
  - WMO
  - WHO?
  - ITU??
  - ...

INFRA2011-1.2.1

- Elements of continuity with ROSCOE:

  - The CERN allocation for the "matching FTEs" foreseen for ROSCOE can be reused for this project

  - WLCG is our main goal. This time it should be even clearer: concentrate on LHC rather than water it down with other initiatives like FAIR or teaming up with too many applications/partners (like in the full ROSCOE proposal)

- Notable differences:

  - Scaled-down collaboration: focused only on LHC or on tools we want to support for LHC

  - Prepare LCG to support users on different (emerging) infrastructures (HPC and Clouds)

  - Focusing on ATLAS and CMS means Ganga and CRAB. We should build on the recent successful collaboration (HammerCloud) and interest (user support, error reporting, job peeking, etc...). We need concrete starting points for a coherent synergy across participants.

  - Due to its success as a community tool (in EGEE and beyond), Ganga is a credible candidate to be the backbone of a working e-Science environment:

    – Ganga should also be the glue for selected collaborations between us CERN and other communities.

    – We need the freedom to choose trusted partners which will effectively contribute to both the proposal preparation and then deliver convincing results in their domain

    – Both the quality and the research area of these communities should be aligned with our laboratory standards and priorities

- A refocused project looks very attractive because at the same time it reduces all the overheads (reduced project complexity, clear global scope) and increases the direct benefit for our current and future activities