

# DISTRIBUTED GRID ANALYSIS USING THE GANGA JOB MANAGEMENT SYSTEM

Johannes Elmsheuser

Ludwig-Maximilians-Universität München, Germany

09 July 2007/DESY Computing Seminar



- ① DISTRIBUTED ANALYSIS MODEL IN ATLAS
- ② DISTRIBUTED ANALYSIS WITH GANGA
- ③ CONCLUSIONS

- Heterogeneous grid environment based on 3 grids:  
LCG/EGEE, OSG and Nordugrid



- Grids have different middle-ware, catalogs to store data, software to submit jobs

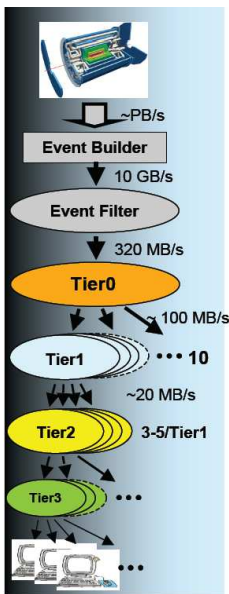
⇒ Hide differences from the ATLAS user

# DISTRIBUTED ANALYSIS MODEL I

The distributed analysis model is based the LHC computing model

- Data is distributed in Tier1/Tier-2 facilities by default  
available 24/7
- user jobs are sent to the data  
large input datasets (100 GB up to several TB)
- Results must be made available to the user  
potentially already during processing
- Data is added with meta-data and bookkeeping in catalogs

# ATLAS DATA REPLICATION AND DISTRIBUTION



- Event Filter Farm at CERN
  - Located near the Experiment, assembles data into a stream to the Tier0
- Tier0 at CERN
  - Raw data → Mass storage at CERN and to Tier 1s
  - Swift production of Event Summary Data (ESD) and Analysis Object Data (AOD)
  - Ship ESD, AOD to Tier1s → Mass storage at CERN
- Tier1s distributed worldwide (10 centers)
  - Re-reconstruction of raw data, producing new ESD, AOD
  - Scheduled, group access to full ESD and AOD
- Tier2s distributed worldwide (~ 30 centers)
  - MC Simulation, producing ESD, AOD → Tier 1s
  - On demand user physics analysis
- CERN Analysis Facility
  - Analysis
  - Heightened access to ESD and RAW/calibration data on demand
- Tier3s distributed worldwide

# DISTRIBUTED ANALYSIS MODEL II

Need for: **Distributed Data Management (DDM)**

- Managed by DDM system DQ2 (Don-Quijote 2)
- Automated file management, distribution and archiving throughout the whole grid using a Central Catalog, FTS, LFCs
- Random access needs a pre-filtering of data of interest  
e.g. Trigger or ID streams or TAGs (event-level meta data)

**Current situation and implementation**

- Data from MC Production System is currently consolidated by DDM-operations team on all Tier1 and then all Tier2 sites
- Analysis model foresees Athena analysis of AODs/ESDs and interactive use of Athena-aware-ROOT tuples
- Analysis tuple format(s) in enhancement

# DATA REPLICATION STATUS

## AOD and NTUP Replication Status (Tier-1s)

- **Data replication period Feb - 19 Jun 2007, DQ2 0.2**
- **Total data volume : 3200+ datasets, 570+Kfiles, 23+TB**

to \ from	ASGC	BNL	CERN	CNAF	FZK	LYON	NG	PIC	RAL	SARA	TRIUMF	%
ASGC												80
BNL												92
CERN												45
CNAF												21
FZK												84
LYON												85
NG												82
PIC												X
RAL												25
NIKHEF												36
TRIUMF												36

95+% 90-95% 80-90% 60-80% 25-60% <25%  
no AODs consolidation within the cloud or/and replication was stopped

Jun 2007

A.Klimentov, ATLAS SW WS

5

At DESY-HH trying to replicate all AODs as well

# ANALYSIS COMPARISONS TEVATRON/LHC

## Tevatron:

- All data (MC/real data) is available at FNAL
- Single computing center, with additional few sites for MC or data reprocessing
- use e.g. d0tools to submit jobs local batch system
- Job numbers: 10-100 per analysis cycle

## LHC:

- All data distributed to various sites
- CERN, 10 Tier1, ~100 Tier2 sites for reconstruction, MC and reprocessing
- use grid tools to submit jobs which go to the data
- Job numbers: a few 1000 per analysis cycle

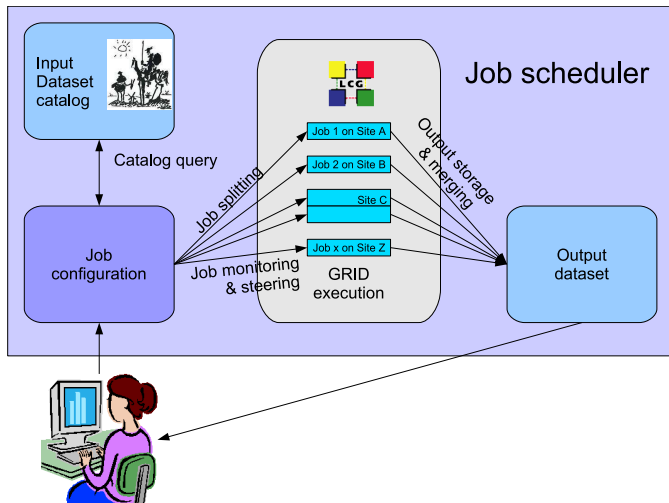
Naive assumption: Grid  $\approx$  large batch system

- Provide complicated job configuration jdl file (Job Description Language)
- Find suitable Athena software, installed as distribution kits in the Grid
- Locate the data on different storage elements
- Job splitting, monitoring and book-keeping
- etc.

$\implies$  Need for automation and integration of various different components

# DISTRIBUTED ANALYSIS

How to combine all these: Job scheduler/manager: GANGA





- 6 community projects and D-Grid integration project for a sustainable Grid infrastructure in Germany:  
DGI, AstroGrid-D, C3-Grid, HEP-Grid, InGrid, MediGrid, TextGrid
- HEP-CG, 3 working packages:
  - Data management with Job-Scheduling and Accounting, dCache
  - Job monitoring, Error identification and job steering
  - Distributed and Interactive Analysis
- Close connection to the LHC experiments and the German Tier1 and Tier2 centers
- LMU Munich:
  - Distributed and Interactive Grid Analysis
  - Gap analysis within D-Grid: Identify job scheduler candidates and review needs and existing features

# FRONT-END CLIENT: GANGA

- A **user-friendly** job definition and management tool.
- Allows simple switching between testing on a **local batch system** and large-scale data processing on distributed resources (**Grid**)
- Developed in the context of **ATLAS** and **LHCb** :
  - For **ATLAS**, have built-in support for applications based on **Athena** framework, for Production System **JobTransforms**, and for **DQ2** data-management system
- **Component** architecture readily allows extension
- Python framework

# WHO IS GANGA ?

- Ganga is an ATLAS - LHCb joint project



- Support for development work from UK (PPARC/GridPP), Germany (D-Grid) and EU (EGEE/ARDA)



- Core team

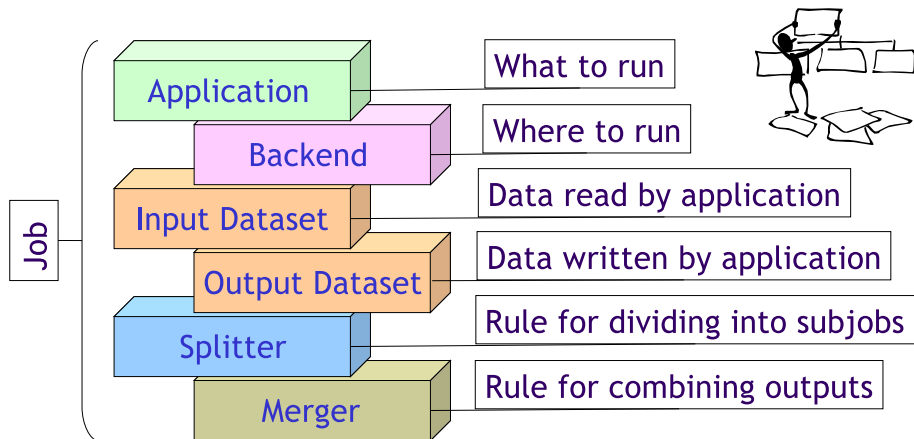
F.Brochu (Cambridge), U.Egede (Imperial), J. Elmsheuser (Munich), K.Harrison (Cambridge), H.C.Lee (ASGC Taipei), D.Liko (CERN), A.Maier (CERN), J.T.Moscicki (CERN), A.Muraru (Bucharest), W.Reece (Imperial), A.Soroko (Oxford), CL.Tan (Birmingham)



Ganga is now under GPL

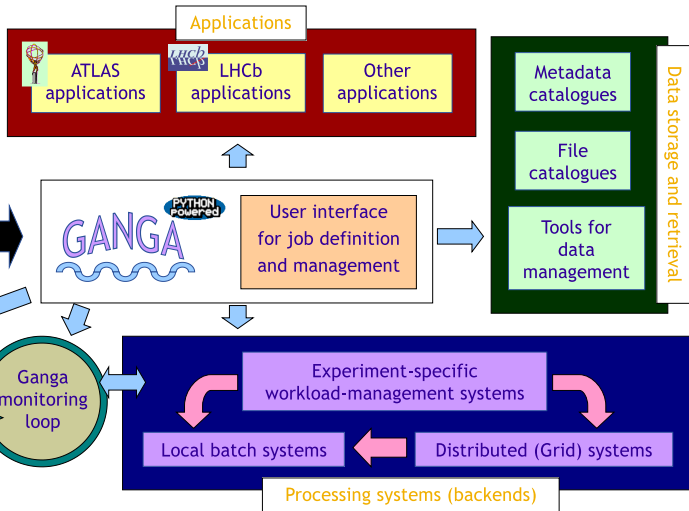
# GANGA JOB

- Ganga is based on a simple, but flexible, job abstraction
- A job is constructed from a set of building blocks, not all required for every job

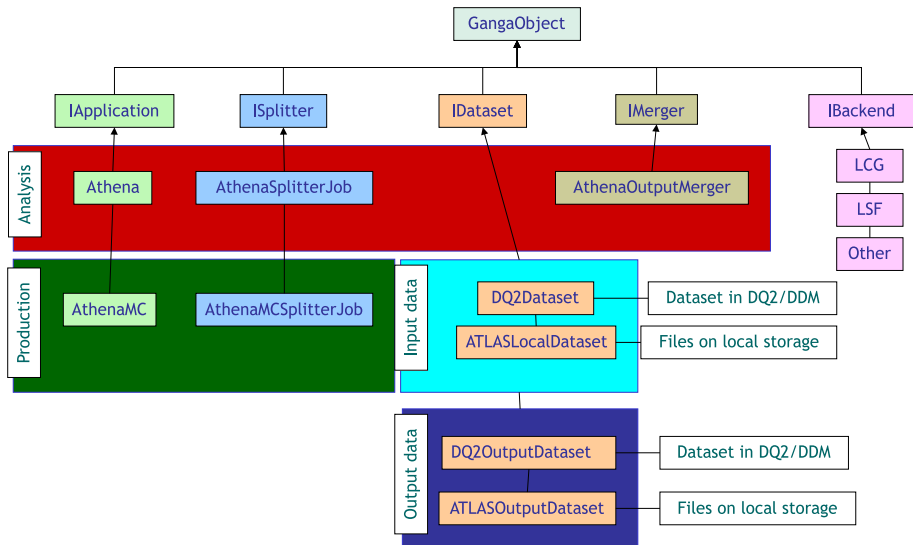


# GANGA BUILDINGBLOCKS

- Ganga has built-in support for ATLAS and LHCb
- Component architecture allows customization for other user groups

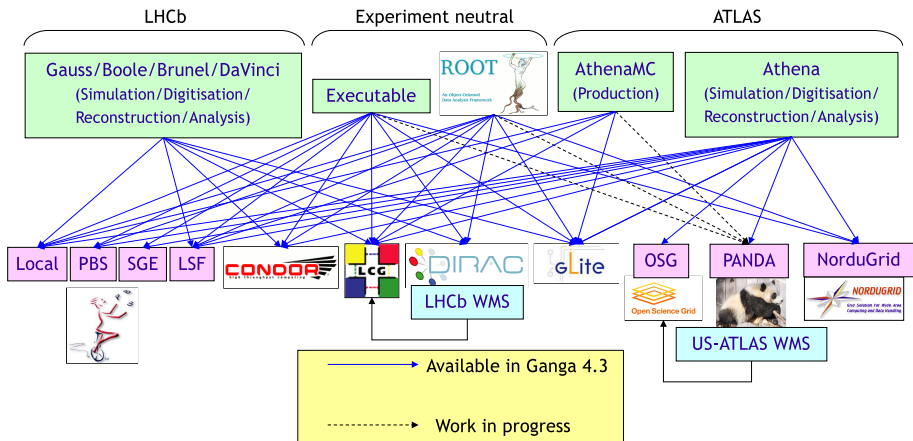


# GANGA ATLAS OBJECTS



# GANGA BACKENDS AND APPLICATIONS

- Ganga simplifies running of ATLAS (and LHCb) applications on a variety of Grid and non-Grid back-ends



# JOB DEFINITION USING ATLAS SOFTWARE I

- Job definition at command line on local desktop:

```
athena AnalysisSkeleton_topOptions.py
```

- Job definition at command line for GRID submission:

```
ganga athena  
  --inDS csc11.005320.PythiaH170ww11.recon.A0D.v11004107  
  --outputdata AnalysisSkeleton.aan.root  
  --split 3  
  --lcg  
  --ce ce-fzk.gridka.de:2119/jobmanager-pbspro-atlasS  
  AnalysisSkeleton_topOptions.py
```

No need to use option `-ce`, job can also be automatically send to data

## JOB DEFINITION USING ATLAS SOFTWARE II

Job definition within GANGA IPython shell:

```
j = Job()
j.application=Athena()
j.application.prepare(athena_compile=False)
j.application.option_file='$HOME/athena/12.0.5/InstallArea/jobOptions
j.splitter=AthenaSplitterJob()
j.splitter.numsubjobs = 3
j.merger=AthenaOutputMerger()
j.inputdata=DQ2Dataset()
j.inputdata.dataset='csc11.005145.PythiaZmumu.recon.AOD.v11004103'
j.inputdata.match_ce=True
j.outputdata=DQ2OutputDataset()
j.outputdata.outputdata=['AnalysisSkeleton.aan.root']
j.backend=LCG()
j.submit()
```

# SCREENSHOT OF THE IPYTHON SHELL

```
elmsheus@lxplus014:~ - Befehlsfenster - Konsole
Sitzung Bearbeiten Ansicht Lesezeichen Einstellungen Hilfe

Ganga.GPIDEV.Lib.Job : INFO killing job 34
Ganga.GPIDEV.Lib.Job : INFO killing job 34.0
Ganga.GPIDEV.Lib.Job : INFO job 34.0 status changed to "killed"
Ganga.GPIDEV.Lib.Job : INFO killing job 34.1
Ganga.GPIDEV.Lib.Job : INFO job 34.1 status changed to "killed"
Ganga.GPIDEV.Lib.Job : INFO killing job 34.2
Ganga.GPIDEV.Lib.Job : INFO job 34.2 status changed to "killed"
Ganga.GPIDEV.Lib.Job : INFO job 34 status changed to "killed"
Out [2]: 1

In [3]:

In [3]:execfile('/afs/cern.ch/user/e/elmsheus/split/5320_cern.py')
Ganga.GPIDEV.Lib.Job : INFO submitting job 35
Ganga.GPIDEV.Adapters : INFO submitting job 35.0 to LOG backend
Ganga.GPIDEV.Adapters : INFO submitting job 35.1 to LOG backend
Ganga.GPIDEV.Adapters : INFO submitting job 35.2 to LOG backend

In [4]:

In [4]:jobs
Out [4]: Statistics: 25 jobs
-----
#  id  status      name      subjobs  application      backend      backend.actualCE
#  ---  ---
#  6   completed   5320      3        Executable      LCG          ce-fzk.gridka.de:2119/jobmanager-pbspro-atlas
#  7   new         5320      10       Executable      LCG
#  8   new         5320      2        Executable      LCG
#  9   completed   5300      3        Executable      LCG          grid10.lal.in2p3.fr:2119/jobmanager-pbs-atlas
#  10  completing  5300      3        Executable      LCG          ce-fzk.gridka.de:2119/jobmanager-pbspro-atlas
#  12  completed   Athens    1        Athens          LSF
#  13  new         Executable  Local
#  14  completed   Athens    1        Athens          LSF
#  15  completed   2         Athens    1        Athens          LCG
#  16  completed   Athens    1        Athens          LSF
#  17  completed   athens_log_dq2_test1,  GRIDKA, DQ2_LOCAL, 7233  Athens      LCG          ce-fzk.gridka.de:2119/jobmanager-pbspro-atlas
#  18  new         athens_log_dq2_test2,  LYON, DQ2_LOCAL, 5300  Athens      LCG
#  19  completed   athens_log_dq2_test2,  LYON, DQ2_LOCAL, 5300  Athens      LCG          cclgce1102.in2p3.fr:2119/jobmanager-bqg-shor
#  20  completed   Athens
#  22  submitted   5320      3        Athens          LCG
#  24  submitted   5320      10       Athens          LCG
#  26  completed   5145_short  2        Athens          LCG
#  27  completed   5300      3        Athens          LCG
#  28  new         5300      1        Athens          Local
#  29  completed   5300      3        Athens          LCG
#  30  killed     5145      10       Athens          LCG
#  31  completed   5145_short  5        Athens          LCG
#  32  running    5300      3        Athens          LCG
#  34  killed     5320_cern  3        Athens          LCG
#  35  submitted   5320_cern  3        Athens          LCG

In [5]:
```

# JOB DEFINITIONS WITH THE GUI

The image shows the GANGA GUI interface with several callouts pointing to different components:

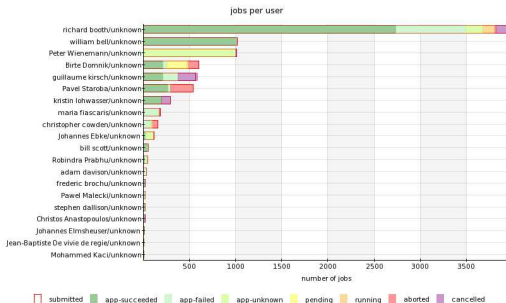
- Logical Folders:** A tree view on the left side of the main window.
- Job Monitoring:** A table in the center showing a list of jobs with columns for ID, status, application, backend, and LCG reason.
- Job details:** A window on the right showing the configuration for a specific job, including input/output paths and application settings.
- Job builder:** A window at the bottom right showing the configuration for a job's attributes, such as application, args, env, and backend.
- Log window:** A window at the bottom left showing the log output for a job, including messages like "logging Gangas in INFO mode".
- Scriptor:** A window on the top right showing a script editor with Python code for job execution, including a "Script" tab and an "Execute" button.

# GANGA JOBS MONITORED BY THE DASHBOARD

<http://dashb-atlas-job.cern.ch/dashboard/request.py/jobsummary>

JOB SUMMARY

any user ▾  
 any site ▾  
 any vosite ▾  
 any cloud ▾  
 any tier ▾  
 any ce ▾  
 ganga ▾  
 any dataset ▾  
 any application ▾  
 any rb ▾  
 any activity ▾  
 any grid ▾  
 unk  pend  run  term  
 done  canc  abort  g-unk  
 succ  fail  a-unk  
 donesuccess  
 2007-04-30 00:00:00 to  
 2007-05-06 23:00:00  
 sort by user ▾  
 bars in the plot

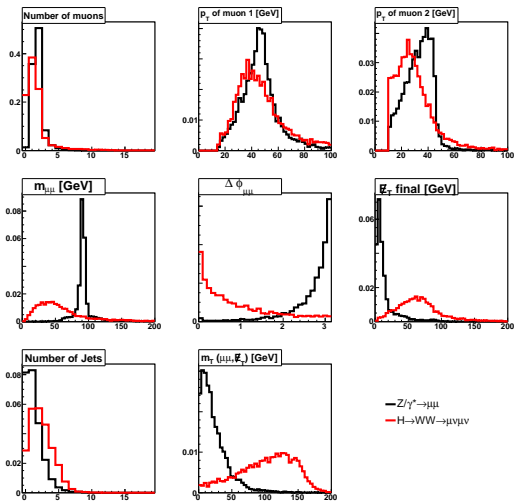


submit

user	current status					grid exit status				application exit status				overall		
	Sub	Unk	Pend	Run	Term	Done	Canc	Abort	Unk	Grid%	Succ	Fail	Unk	App%	D/S	Overall%
Birte Dornik/unknown	604	0	209	21	374	267	2	105	0	71.77	213	38	126	85.54	211	56.42
Brigitte Epp/unknown	10	0	0	0	10	10	0	0	0	100	2	8	0	20	2	20
Carlos Escobar Ibanez/unknown	4	0	0	0	4	4	0	0	0	100	4	0	0	100	4	100
Christos Anastopoulos/unknown	21	0	0	0	21	17	4	0	0	100	9	9	3	50	9	42.86
Elena Oliver Garcia/unknown	3	0	0	0	3	3	0	0	0	100	3	0	0	100	3	100
Esteban Fullana Torregrosa/unknown	6	0	0	0	6	6	0	0	0	100	0	6	0	0	0	0
Farida Fassi/unknown	6	0	0	0	6	6	0	0	0	100	6	0	0	100	6	100
Folkert Koetsveld/unknown	3	0	0	0	3	0	0	0	3	0	2	1	0	66.67	0	0
Grant Gorfine 9015/unknown	4	0	0	0	4	4	0	0	0	100	1	3	0	25	1	25
Jean-Baptiste De vivie de regie/unknown	13	0	0	0	13	13	0	0	0	100	10	3	0	76.92	10	76.92
Johannes Ebke/unknown	119	0	0	0	119	116	0	3	0	97.48	21	15	83	58.33	21	17.65
Johannes Elmsheuser/unknown	14	0	0	0	14	14	0	0	0	100	11	3	0	78.57	11	78.57
total	6924	0	209	176	6539	7566	502	434	37	94.56	4927	1403	2209	77.84	4666	55.99

# RESULTS

Download the processed events (root-tuples) to the local desktop:



- Production of Signal MC sample with AthenaMC
  - 50000 additional events:
  - event generation: 5 jobs or use already validated evgen samples where only a small fraction has officially been simulated/reconstructed
  - simulation: 1000 jobs
  - reconstruction: 50 jobs
  - From this exercise a prototype for automatic job submission has evolved which will eventually will be part of Ganga
  - Statistics from the dashboard: 11/4-11/6: 79% Grid eff. \* 77% Application eff. = 53% overall eff.
- Distributed Analysis as part of the CSC homework:
  - process signal and background MC samples at well known and maintained sites: GridKa, Lyon and LRZ
  - This involves: SUSY signal, ttbar (5200), Di-boson backgrounds, etc.
  - Using SUSYView

- Data availability and incomplete datasets
  - Newest AOD data has been replicated to many sites by DDM-ops team
- Repeated direct input files access problems (Posix I/O) on Castor, dCache and DPM storage elements due to various reasons:
  - broken ROOT or middleware plugins
  - Operational issues, like TURL resolving not working
- gLite WMS is now used for analysis

# DISTRIBUTED ANALYSIS TUTORIALS AND SUPPORT

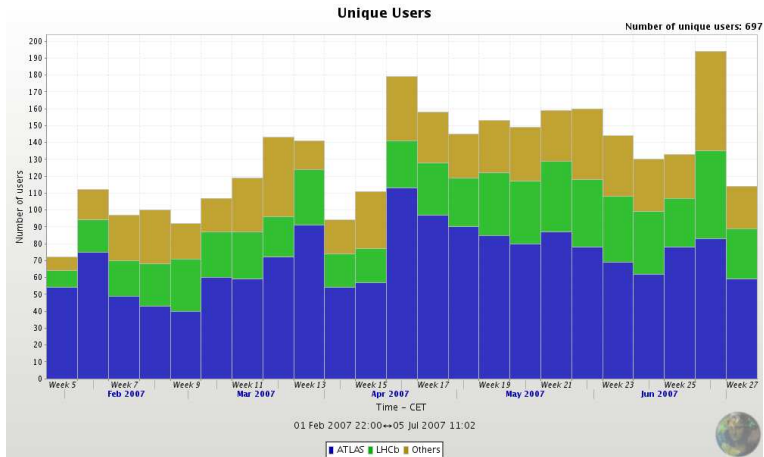
<https://twiki.cern.ch/twiki/bin/view/Atlas/GangaTutorial43>

- Edinburgh (February 1st-2nd )
- Milan (February 5th-6th )
- Lyon (March 5th-7th)
- Munich (March 29th-30th )
- Toronto (April 18th)
- Bergen (April 27th)
- Valencia (May 3rd-4th)
- **Ganga User support and Feedback:**  
[hn-atlas-GANGAUserDeveloper@cern.ch](mailto:hn-atlas-GANGAUserDeveloper@cern.ch)



# USAGE STATISTICS

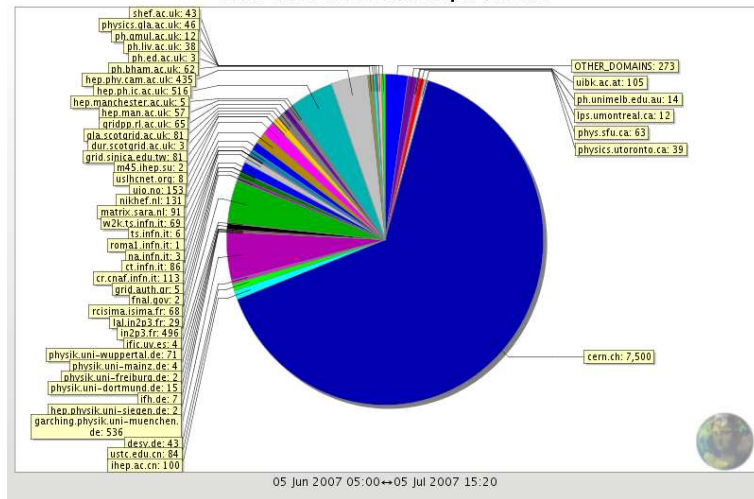
- Over 750 unique users since beginning of the year
- Over 440 ATLAS users have tried Ganga at least once
- About 50 ATLAS Ganga users per week



# DOMAIN STATISTICS

- Over 50 different domains in the last month

Total number of sessions per domain



# GANGA USERS I

- Main Users



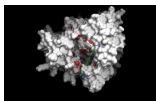
- Other activities
- Geant 4**



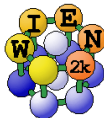
HARP



Garfield



Academia Sinica  
Genomics Research Center



med  
austron



Ganga is used as Grid abstraction layer, also in conjunction with DIANE

- Bioinformatics: Avian flu drug search 2006
- Geant4 regression testing
- Cambridge Ontology: content-based image retrieval

## The Grid improves efficiency of Geant4 production cycles

The Ganga/DIANE framework greatly improved the performance of the Geant4 production. In addition, the run completion time is more predictable and allows a better planning of the production and less intervention from the production operator.



For the distributed analysis it is vital to have:

- Easy interface that does not scare off physicists
- A reliable and robust service of many components

## Conclusions

- Growing number of users are using the Grid for analysis,  
Still some room to grow
- Have to push analysis to higher Tiers
- Data Management is a central issue for Distributed Analysis
- It might interesting to see how more interactive use of resources will evolve

Homepage:

<http://cern.ch/ganga>

Links to Documentation, Tutorials, etc.:

<http://ganga.web.cern.ch/ganga/workbook/>