Status report from ATLAS

Dietrich Liko

Overview

- DDM Development
- DDM Operation
- Production
- Distributed Analysis

DDM Development

Accomplishments
of SC4 Tier-0 Scaling Exercise

- Run a full-scale exercise, from EF, reconstruction farm, T1 export, T2 export
  - Realistic data sizes, complete flow
- Included all T1s sites in the exercise from first day (except NG)
- Included ~ 15 T2s sites on LCG by the end of the second week
- Maximum export rate (per hour) ~ 700 MB/s
  - Nominal rate should be ~ 780 MB/s (with NG)
- ATLAS regional contacts were actively participating in **some** of the T1/T2 clouds

T1 export

QuickTime™ and a
TIFF (Uncompressed) decompressor
are needed to see this picture.

On this day, missing stable CNAF, FZK, RAL for nominal rate
(excluding NGDF since not part of SC4 for ATLAS)

Lessons from SC4 Tier-0 Scaling Exercise

- From post-mortem analysis:
  - ***All 9 sites up only for a handful of hours during a whole month!!***
  - SRM and storages are not sufficiently stable and ATLAS team spent a **long** time sending notifications of errors to sites
    - Self-monitoring desirable!
  - LFC services provided by the Tier-1s largely insufficient
    - LFC down? Site not used for any Grid production or data transfer
  - Proposed GGUS improvements
    - Allowing faster ticket forwarding
  - We are willing to help:
    - What shall we monitor? Suggestions welcomed (atlas-dq2-support@cern.ch)
    - Will be including automated notifications of errors, removing repetitions

Future plans

- ATLAS SC4 Tier-0 re-run starting ~25th September
  - Ramp up already next week
- Urgent issue with LFC services ***now***
  - Connections failing at **most** Tier-1s!
  - Review **LFC service: h/w, monitoring, operations**
- Self-monitoring site services
  - **LFC, SRM, storage, FTS**
  - **All services must operate reliably**
- ATLAS regional contacts (allocated or being allocated to all sites) must help but sites have to feel a ultimately responsible for fulfilling dataset transfers
  - *… or their own physicists communities will suffer*

DDM Operation

DDM Ops team

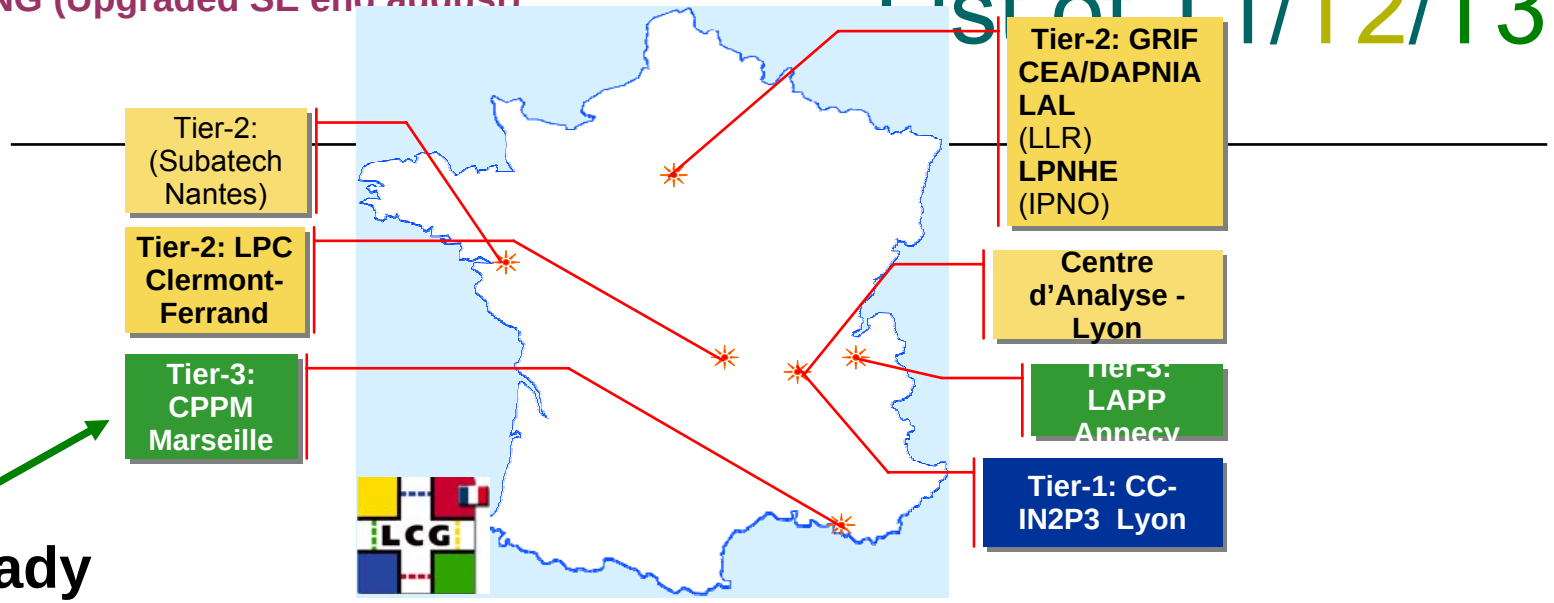
- ASGC (Jason Shih, Suijian Zhou)
- BNL (Wensheng Deng, Hironori Ito, Xin Zhao)
- CERN (MB,DC,AK)
- CNAF (Guido Negri, Paolo Veronesi and Giuseppe Lo Re)
- FZK (John Kennedy, Jiri Chudoba, Andrzej Olszewski)
- IN2P3 Lyon (Stephane Jezequel, Ghita Rahal)
- NG (Alex Read, Oxana Smirnova)
- PIC (Mireia Dossil, Xavier Espinal)
- RAL (Catalin Condurache, Frederique Brochu)
- SARA (Jiri Chudoba , grid.support@sara.nl)
- TRIUMF (Rod Walker, Denice Deatrich, Reda Tafirout)
- Central Team :
  - AK, Cunfeng Feng, Zhijun Liang, Pavel Nevski and Nectarios Benekos
  - *Stephane, Oxana, Jiri*
  - *Yuri Smirnov, Tomasz Wlodek (DB support and monitoring)*

Non french sites:

- TOKYO
- BEIJING (Upgraded SE end august)

LYON CLOUD

List of T1/T2/T3



Ready since end August

Preparation and follow-up

Meeting each 3-4 weeks of french sites

+ Tokyo (I. Ueda)

.Fast and effective interactions

with responsibles in each T2/T3

.Tests of plain FTS made before SC4

S.Jezequel, G.Rahal : Computing Operations Session

DDM Operations at SARA Cloud

- Site Issues
  - vobox failure (local hard disk)
    - restored from backup
  - FTS server
    - problems since August
    - local BDII stability – solved
    - user mappings, channels for atlas vs atlassgm
    - still not able to do transfers to Tier2's – solved today
  - incidents logged on wiki:

https://twiki.cern.ch/twiki/bin/view/Main/AtlasDDMatSARA

DDM Operations (US cloud)

- Alexei Klimentov and Wensheng Deng
- BNL : Wensheng Deng, Hironori Ito, Xin Zhao
  - GLT2, MWTier2, NETier2, SWTier2, WETier2
  - Weekly meeting (with all sites reps)
  - Coherent DDM/DQ2 installation and deployment policy
  - Testbed machines at BNL to debug and test new versions
  - Clear separation between production, users and development DDM facilities
  - OSG specifics DDM tools
  - RT problems report for each site

Tier-1/Tier-2s Data Transfer Functional Test 1/3

- The test was proposed by Zhongliang Ren within Computing Operations Group
- DDM ops was asked to define the scope of the test and to conduct it

The goal is to have ***system*** test

ASGC	Suijian Zhou	2
BNL	Wensheng Deng	3
CNAF	Guido Negri	4
GRIDKA	John Kennedy	7
LYON	Stephane Jezequel	8
PIC	Xavier Espinal	3
RAL	Frederique Brochu	7
SARA	Jiri Chudoba	3
TRIUMF	Rod Walker	5

Tier-1/Tier-2s Data Transfer Functional Test 2/3

- Functional test is a first phase of system check-up
- A coherent data transfer test between Tier-1 and Tier-2s for all ‘clouds’, using existing DDM and monitoring SW
  - Data generated by ATLAS MC production
  - Organized in DQ2 datasets
  - It is not this test goal to evaluate performance
  - We want to know how functional is our system is
- Two steps
  - Step 1
    - Data distribution from Tier-0 to Tier-1 and then to Tier-2s
  - Step 2
    - For Tier-2s passed step 1, consolidate data from Tier-2s on Tier-1
    - Step 2a
      - Data access from Tier-1 (BNL) to Tier-2s in “foreign cloud”

Tier-1/Tier-2s Data Transfer Functional Test 3/3 *(Preliminary status Sep 14 15:00)*

Tier-1	Rqstd / Rcvd Datasets	Success Rate %	Comments
ASGC	3/3	98%	After site problem was fixed (2-3h). T1->T2
BNL			Site problem. FTS. Fixed, resubscribed, problem to access lfc://prod-lfc-atlas-local.cern.ch
CNAF			Site problem : SE (CASTOR). Fixed. Data transfer in progress
GRIDKA	8/8	100%	From the 2nd attempt, initial efficiency 12%. T1->T2
LYON	10/8	80%	FTS errors, after Sep 11 20:00 efficiency ~80%, system was blocked by T2-T1 datatransfer, resubscribed. T1->T2 DONE
PIC	5/5	90%	From the 3rd attempt (initially 100%, then 12%). T1->T2
RAL	6/6	90%	From the 2nd attempt, initial efficiency 10%, data transfer still in progress (Request : 24h ago). T1->T2
SARA			site problem was fixed (2-3h), the problem with access to central services and FTS. FTS transfer timeout. Data transfer in progress
TRIUMF	5/5	75%	From the 1st attempt, FTS errors on Sep 11 17:00 Data transfer still in progress (Request : 48h ago). T1->T2

Production

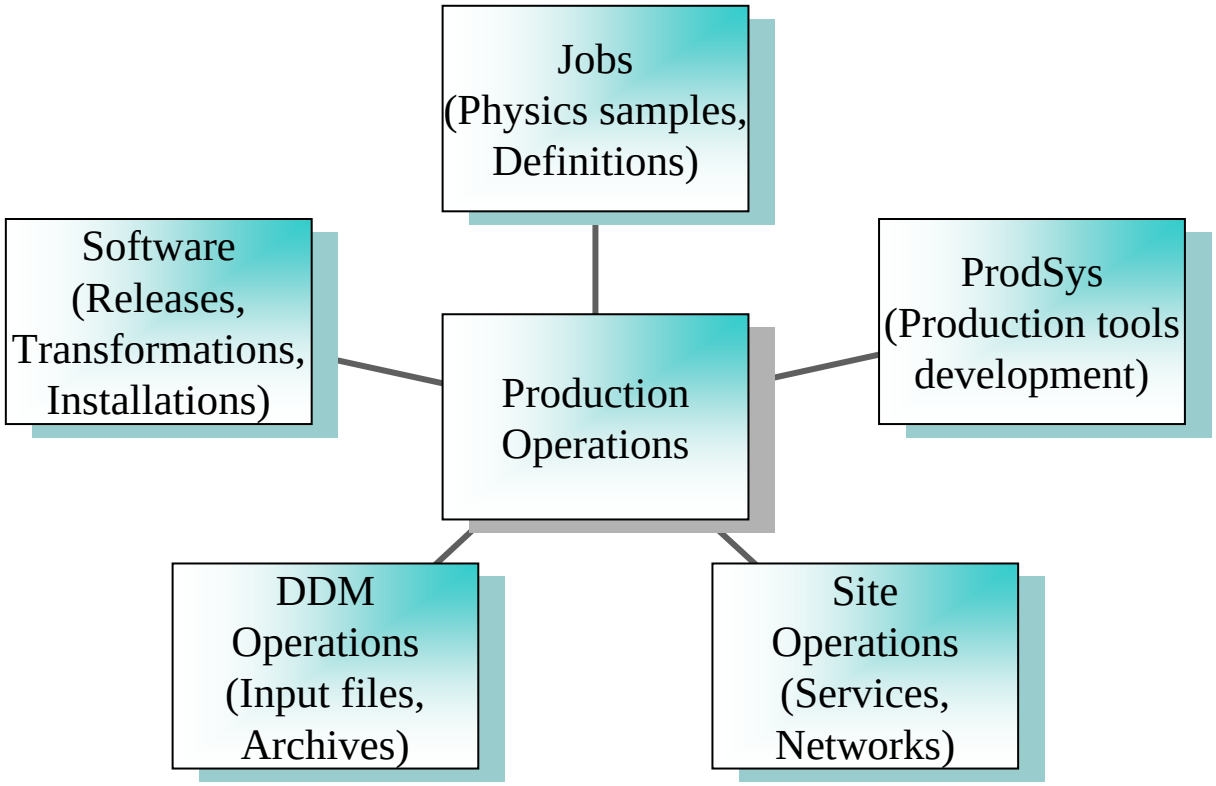
Simulation Production

- Goals of production operations
  - Run distributed production worldwide
  - Three different grids – LCG/EGEE, NorduGrid, OSG
  - Scale ~about 100 sites (10 Tier 1's, 30 T2's, T3's, opportunistic)
  - Schedule – continuous (year-round, started since late fall 2005)
- Tasks and responsibilities
  - Get production jobs done on time (average of 10 tasks have been defined per day since Dec 2005 - but rate higher with new release)
  - Scale: ~3000 tasks in past 10 months; 1.7 million jobs; 175 million events (includes x3 repeat counting of events due to many steps)
  - Provide rapid bug/error reporting to sites, services, developers (this is the most time consuming part of production – producing events is mostly automated, if nothing goes wrong!)

Organization

- Production teams (at least one per grid, also support DA)
- LCG/EGEE:
  - Four teams in operation (2 for lexor, 2 for CondorG)
  - Lexor teams led by Simone Campana/Guido Negri
  - CondorG teams led by Rod Walker
  - LCG teams merged this week into one LCG operations team
- OSG/Panda:
  - One team (2.25 FTE) running shifts 6 days/week
  - Coordinator: Kaushik De
- NorduGrid:
  - One team, led by Alex Read
- Many other independent teams in computing operations like installation, job definition, DDM operations provide support

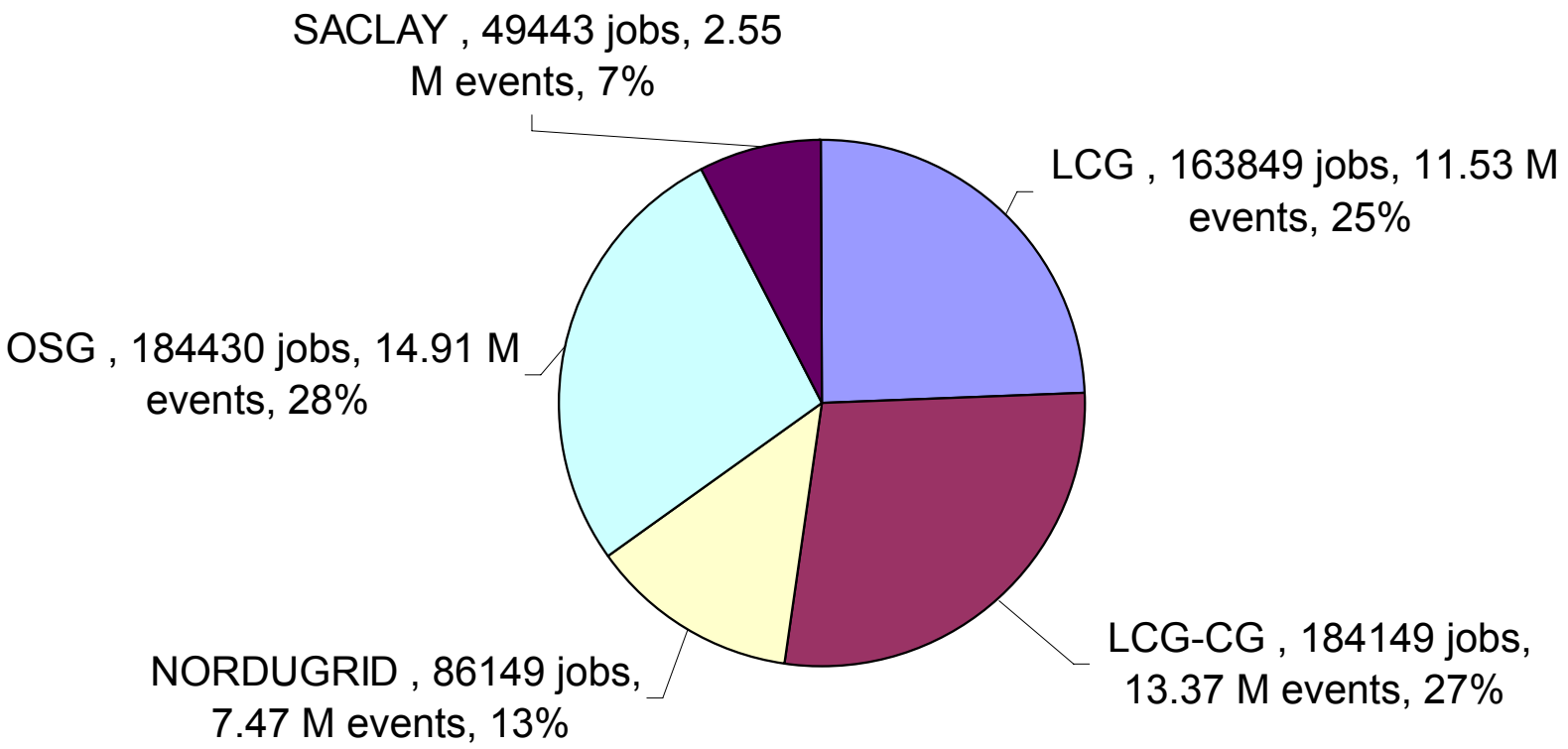
Production Dependencies



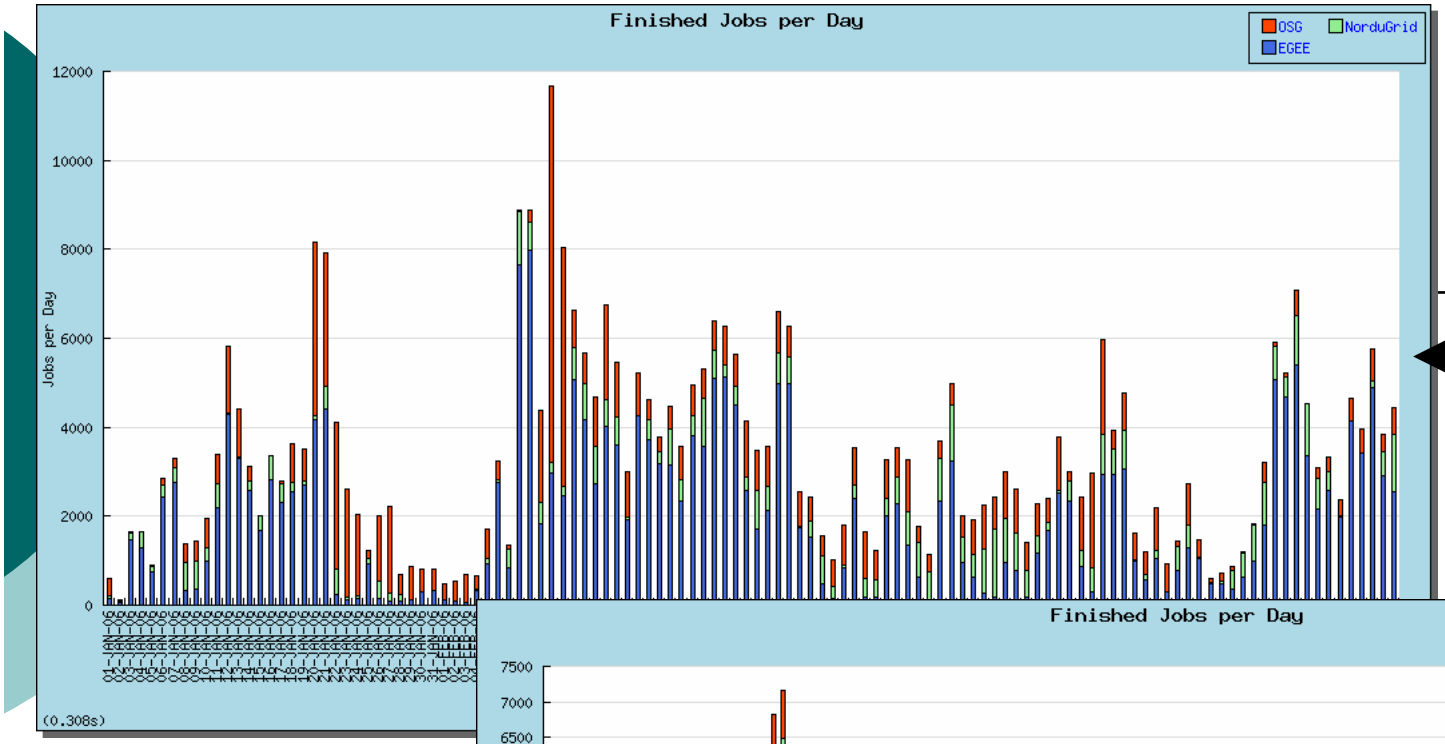
Production Status

- Run continuous CSC production since Dec. 2005
- Completed Release 11 validation & CSC Production
- Intensive Release 12 validation this summer
- Overall – highly successful
- But – need to scale up urgently
- Current rate average ~ 1M events (all steps) per week
- Sufficient to keep up with physics needs so far
- Need to double rate with Release 12.0.3 (starting in few weeks)
- Need to double again every 3-4 months – target 10M events/wk by Fall 2007

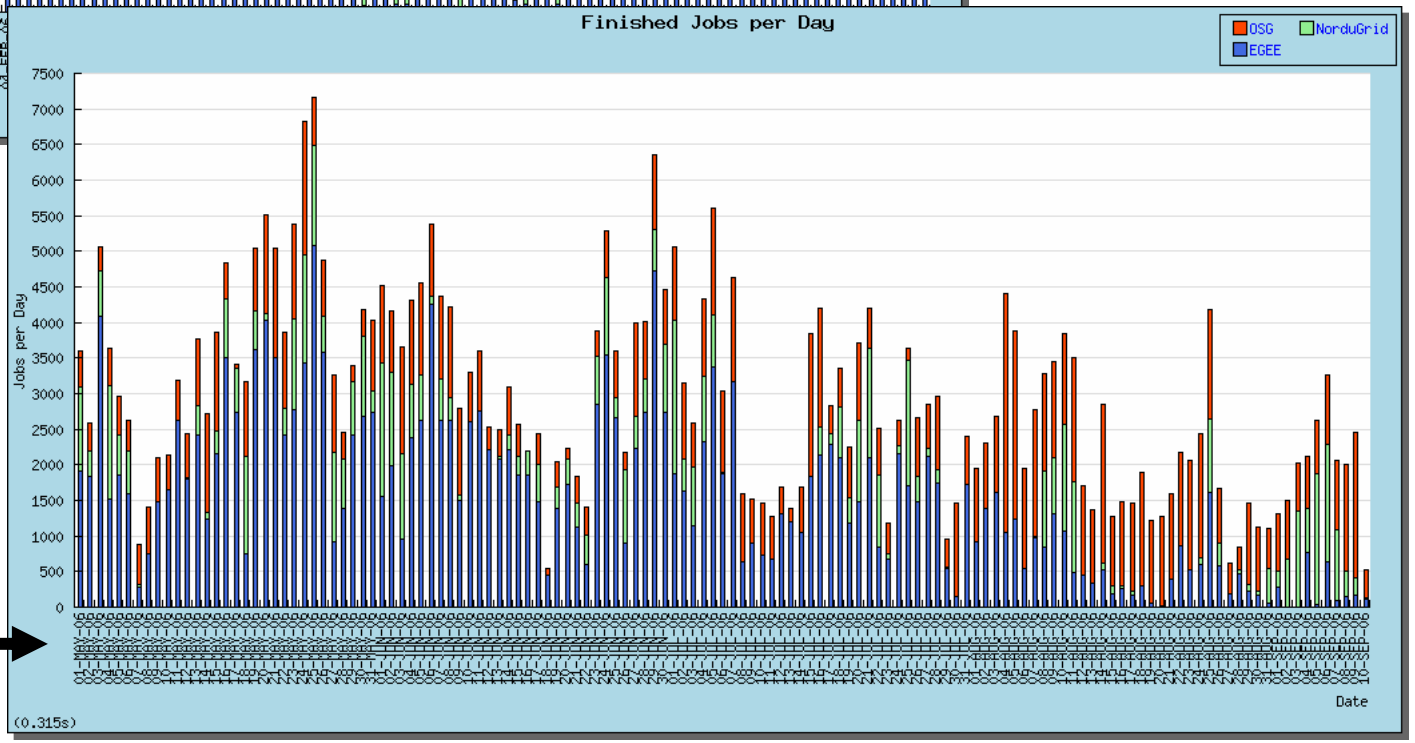
Number of successful jobs (Release 11.0.x)



Already 50M physics events produced in 2006!



~~Jan 1-April 30~~
(max 12k jobs
Reached in Feb)



May 1-now

Distributed Analysis

Overview

- Tools
  - Pathena on OSG/PANDA
  - GANGA on EGEE
- Data distribution
- Job Priorities

Pathena on OSG/PANDA

- PANDA is the Production and Analysis Tool used on OSG
  - A pilot job based system similar to DIRAC & Alien
- Pathena is a lightweight tool to submit analysis jobs to PANDA
- In use since spring and has been optimised for user response

GANGA on EGEE

- UI being developed togther with LHCb
- Resource Broker is used to submit jobs
- It has been integrated with the new ATLAS DDM system DQ2 over the summer
- Offers similar functionality and interface as pathena

Future

- We plan to test new features
  - TAG based analysis
- Interoperability
  - On the level of the application
  - On the level of the middleware

Dataset distribution

- In principle data should be everywhere
  - AOD & ESD during this year ~ 30 TB max
- Three steps
  - Not all data can be consolidated
  - Other grids, Tier-2
  - Distribution between Tier-1 not yet perfect
  - Distribution to Tier-2's can only be the next step

CSC11 AOD Data at Tier-1

	Datasets	Complete	Files	Size
ASGC	96	14	5634	520
BNL	226	131	17053	1736
CERN	253	106	16610	1712
CNAF	1	0	217	39
FZK	16	3	1510	172
LYON	72	13	6518	786
RAL	10	1	589	93
SARA	2	0	251	38
PIC	7	3	916	105
TRIUMF	7	3	510	77

Dataset conclusion

- AOD Analysis at BNL, CERN, LYON
- ESD Analysis only at BNL

- We have still to work hard to complete the “collection” of data
- We have to push hard to achieve equal distribution between sites

- Nevertheless:
  Its big progress to some month ago!

Job Priorities

- EGEE Job priorities WG
  - Shares mapped to VOMS attributes
  - Short queue/CE
    - In the future it could be one gLite CE
- Data <and> Queues
  - On LCG we can offer today DA at CERN and in Lyon

Summary

- We have got first operational experience with our new DDM
  - Some goals have already been achieved, other still need some work
- Operation needs personnel
- Ambitious plans to ramp up production
- Distributed Analysis is being provided as a tool to our users