



# **ATLAS Data Challenge Production and U.S. Participation**

**Kaushik De**

**University of Texas at Arlington**

**BNL Physics & Computing Meeting**

**August 29, 2003**

# ATLAS Data Challenges



## ❄ Original Goals (Nov 15, 2001)

- ❑ Test computing model, its software, its data model, and to ensure the correctness of the technical choices to be made
- ❑ Data Challenges should be executed at the prototype Tier centres
- ❑ Data challenges will be used as input for a Computing Technical Design Report due by the end of 2003 (?) and for preparing a MoU

## ❄ Current Status

- ❑ Goals are evolving as we gain experience
- ❑ Computing TDR ~end of 2004
- ❑ DC's are ~yearly sequence of increasing scale & complexity
- ❑ DC0 and DC1 (completed)
- ❑ DC2 (2004), DC3, and DC4 planned
- ❑ Grid deployment and testing is major part of DC's

# ATLAS DC1: July 2002-April 2003



Goals : *Produce the data needed for the HLT TDR*  
*Get as many ATLAS institutes involved as possible*

**Worldwide collaborative activity**  
**Participation : 56 Institutes (39 in phase 1)**



❄ **Australia**

❄ **Austria**

❄ **Canada**

❄ **CERN**

❄ **China**

❄ **Czech Republic**

❄ **Denmark \***

❄ **France**

❄ **Germany**

❄ **Greece**

❄ **Israel**

❄ **Italy**

❄ **Japan**

❄ **Norway \***

❄ **Poland**

❄ **Russia**

❄ **Spain**

❄ **Sweden \***

❄ **Taiwan**

❄ **UK**

❄ **USA \***

**New countries or institutes**

**\* using Grid**

# DC1 Statistics (G. Poulard, July 2003)



Process	No. of events	CPU Time	CPU-days (400 SI2k)	Volume of data
		kSI2k.months		TB
Simulation Physics evt.	$10^7$	415	30000	23
Simulation Single part.	$3 \times 10^7$	125	9600	2
Lumi02 Pile-up	$4 \times 10^6$	22	1650	14
Lumi10 Pile-up	$2.8 \times 10^6$	78	6000	21
Reconstruction	$4 \times 10^6$	50	3750	
Reconstruction + Lvl1/2	$2.5 \times 10^6$	(84)	(6300)	
<b>Total</b>		<b>690 (+84)</b>	<b>51000 (+6300)</b>	<b>60</b>

# DC2: Scenario & Time scale (G. Poulard)



❄ **End-July 03: Release 7**

❄ **Mid-November 03: pre-production release**

❄ **February 1<sup>st</sup> 04: Release 8 (production)**

❄ **April 1<sup>st</sup> 04:**

❄ **June 1<sup>st</sup> 04: "DC2"**

❄ **July 15th**

➤ **Put in place, understand & validate:**

- Geant4; POOL; LCG applications
- Event Data Model
- Digitization; pile-up; byte-stream
- Conversion of DC1 data to POOL; large scale persistency tests and reconstruction

➤ **Testing and validation**

- Run test-production

➤ **Start final validation**

- **Start simulation; Pile-up & digitization**
- **Event mixing**
- **Transfer data to CERN**

➤ **Intensive Reconstruction on "Tier0"**

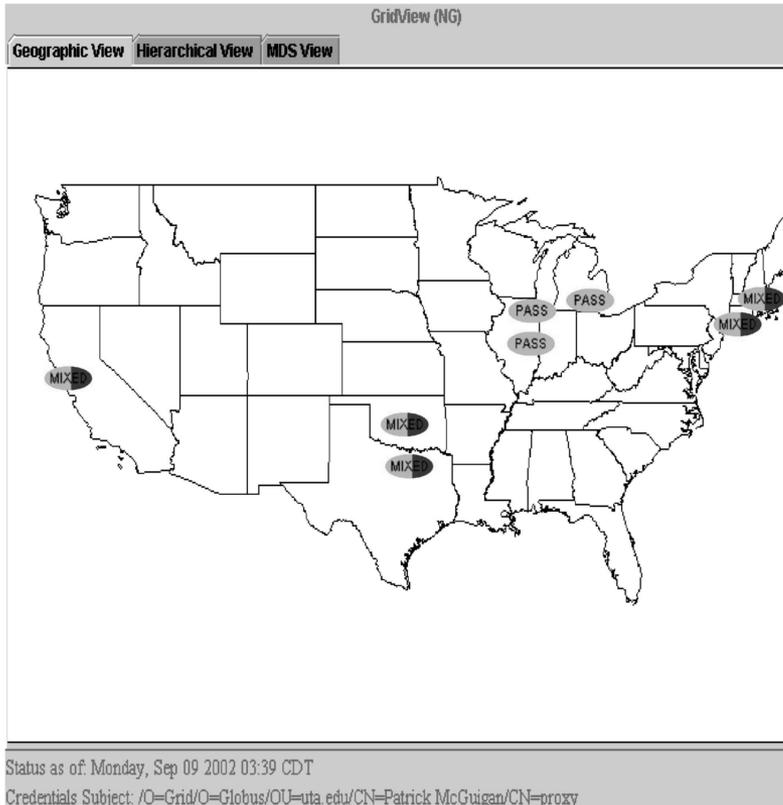
- **Distribution of ESD & AOD**
- **Calibration; alignment**
- **Start Physics analysis**
- **Reprocessing**

# U.S. ATLAS DC1 Data Production



- ❄ Year long process, Summer 2002-2003
- ❄ Played 2nd largest role in ATLAS DC1
- ❄ Exercised both farm and grid based production
- ❄ 10 U.S. sites participating
  - ❑ Tier 1: BNL, Tier 2 prototypes: BU, IU/UC, Grid Testbed sites: ANL, LBNL, UM, OU, SMU, UTA (UNM & UTPA will join for DC2)
- ❄ Generated ~2 million fully simulated, piled-up and reconstructed events
- ❄ Largest grid-based DC1 data producer in ATLAS
- ❄ Data used for HLT TDR, Athens physics workshop, reconstruction software tests...

# U.S. ATLAS Grid Testbed



- \* BNL - U.S. Tier 1, 2000 nodes, 5% for ATLAS, 10 TB, HPSS through Magda
- \* LBNL - pdsf cluster, 400 nodes, 5% for ATLAS (more if idle ~10-15% used), 1TB
- \* Boston U. - prototype Tier 2, 64 nodes
- \* Indiana U. - prototype Tier 2, 64 nodes
- \* UT Arlington - new 200 cpu's, 50 TB
- \* Oklahoma U. - OSCER facility
- \* U. Michigan - test nodes
- \* ANL - test nodes, JAZZ cluster
- \* SMU - 6 production nodes
- \* UNM - Los Lobos cluster
- \* U. Chicago - test nodes

# U.S. Production Summary



- \* Exercised both farm and grid based production
- \* Valuable large scale grid based production experience

	Number of	Number of	CPU hours	CPU hours	CPU hours
	Files in Magda	Events	Simulation	Pile-up	Reconstruction
25 Gev di-jets	41k	1M	~60k	56k	60k+
50 Gev di-jets	10k	250k	~20k	22k	20k+
Single particles	24k	200k	17k	6k	
Higgs sample	11k	50k	8k	2k	
SUSY sample	7k	50k	13k	2k	
minbias sample	7k	?	?		

- \* Total ~30 CPU YEARS delivered to DC1 from U.S.
- \* Total produced file size ~20TB on HPSS tape system, ~10TB on disk.
- \* Black - majority grid produced, Blue - majority farm produced

# Grid Production Statistics

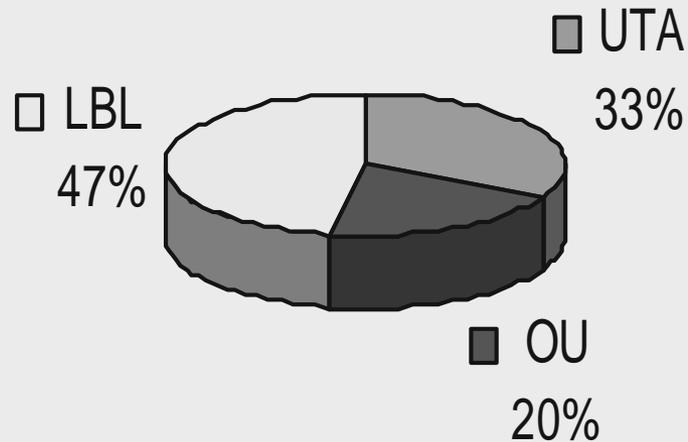


Figure : Pie chart showing the sites where DC1 single particle simulation jobs were processed. Only three grid testbed sites were used for this production in August 2002.

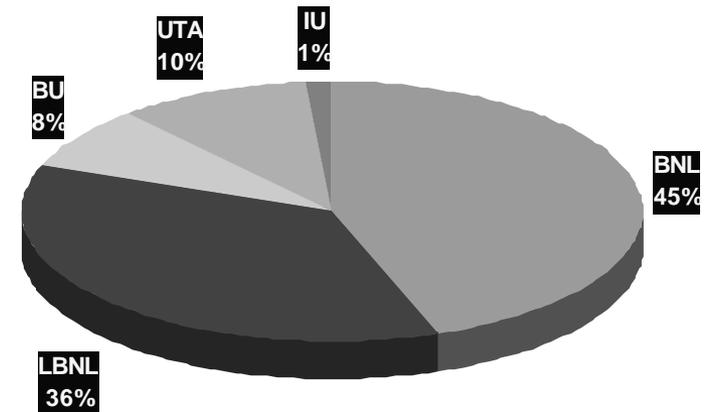


Figure : Pie chart showing the number of pile-up jobs successfully completed at various U.S. grid sites for dataset 2001 (25 GeV dijets). A total of 6000 partitions were generated.

These are examples of some datasets produced on the Grid. Many other large samples were produced, especially at BNL using batch.

# DC1 Production Systems



- \* Local batch systems - bulk of production
- \* GRAT - grid scripts, generated ~50k files produced in U.S.
- \* NorduGrid - grid system, ~10k files in Nordic countries
- \* AtCom - GUI, ~10k files at CERN (mostly batch)
- \* GCE - Chimera based, ~1k files produced
- \* GRAPPA - interactive GUI for individual user
- \* EDG - test files only
- \* + systems I forgot...
- \* More systems coming for DC2
  - LCG
  - GANGA
  - DIAL

# Databases used in GRAT



## ❄ Production database

- ❑ define logical job parameters & filenames
- ❑ track job status, updated periodically by scripts

## ❄ Data management (Magda)

- ❑ file registration/catalogue
- ❑ grid based file transfers

## ❄ Virtual Data Catalogue

- ❑ simulation job definition
- ❑ job parameters, random numbers

## ❄ Metadata catalogue (AMI)

- ❑ post-production summary information
- ❑ data provenance

# U.S. Middleware Evolution



**Globus**

Used for 95% of DC1 production



**Condor-G**

Used successfully for simulation



**DAGMan**

Used successfully for simulation  
(complex pile-up workflow not yet)



**Chimera**

Tested for simulation, used for  
all grid-based reconstruction



**LCG**

# U.S. Experience with DC1



- ❄ ATLAS software distribution worked well for DC1 farm production, but not well suited for grid production
- ❄ No integration of databases - caused many problems
- ❄ Magda & AMI very useful - but we are missing data management tool for truly distributed production
- ❄ Required a lot of people to run production in the U.S., especially with so many sites on both grid and farm
- ❄ Startup of grid production slow - but learned useful lessons
- ❄ Software releases were often late - leading to chaotic last minute rush to finish production

# U.S. Plans for DC2



- ❄ Computing organization in the U.S. has been restructured to reflect growing importance of grid in DC2 (we hope to use only grid based production for DC2 in the U.S.)
- ❄ R. Gardner leading effort to develop grid tools and services, K. De & P. Nevski leading production
- ❄ New tools being developed for DC2, based on Chimera see: <http://www.usatlas.bnl.gov/computing/grid/gce>
- ❄ Joint CMS/ATLAS preDC2 exercise underway - called Grid3, for next 6 months
- ❄ Need to develop plans and have software ready and tested before real DC2 production/user analysis starts

# Plans for DC2 Production System



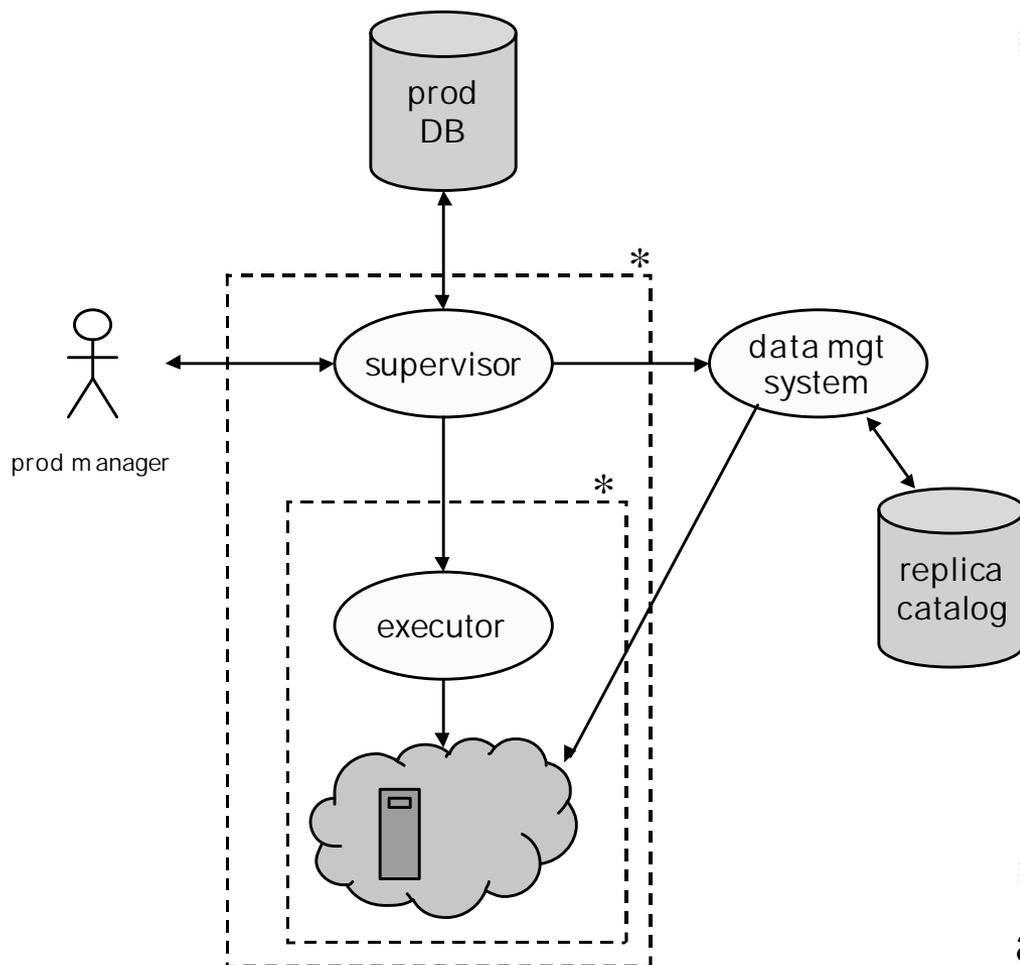
## ❄ Need unified system for ATLAS

- ❑ for efficient usage of facilities, improved scheduling, better QC
- ❑ should support all varieties of grid middleware (& batch?)

## ❄ First “technical” meeting at CERN August 11-12, 2003

- ❑ attended by Luc Goosens\*, KD, Rich Baker, Rob Gardner, Alessandro De Salvo, Jiri Chudoba, Oxana Smirnova
- ❑ design document is being prepared
- ❑ planning a Supervisor/Executor model (see fig. next slide)
- ❑ first prototype software should be released ~6 months
- ❑ U.S. well represented in this common ATLAS effort
- ❑ Still unresolved - Data Management System
- ❑ Need strong coordination with database group (Luc & Kaushik attended Database meeting at Oxford in July)

# Schematic of New DC2 System



## \*Main features

- ❑ Common production database for all of ATLAS
- ❑ Common ATLAS supervisor run by all facilities/managers
- ❑ Common data management system a la Magda
- ❑ Executors developed by middleware experts (LCG, NorduGrid, Chimera teams)
- ❑ Final verification of data done by supervisor

\*U.S. involved in almost all aspects - could use more help

# Conclusion



- ❄ Data Challenges are important for ATLAS software and computing infrastructure readiness
- ❄ U.S. playing a major role in DC planning & production
- ❄ 12 U.S. sites ready to participate in DC2, more welcome
- ❄ Production software development needs help
- ❄ Physics analysis major emphasis of DC2
- ❄ Involvement by more U.S. physicists is needed in DC2
  - ❑ to verify quality of data
  - ❑ to tune physics algorithms
  - ❑ to test scalability of physics analysis model