



ATLAS Data Challenges

19 March 2004

Dario Barberis
CERN & Genoa University



ATLAS DC1 (July 2002 - April 2003)

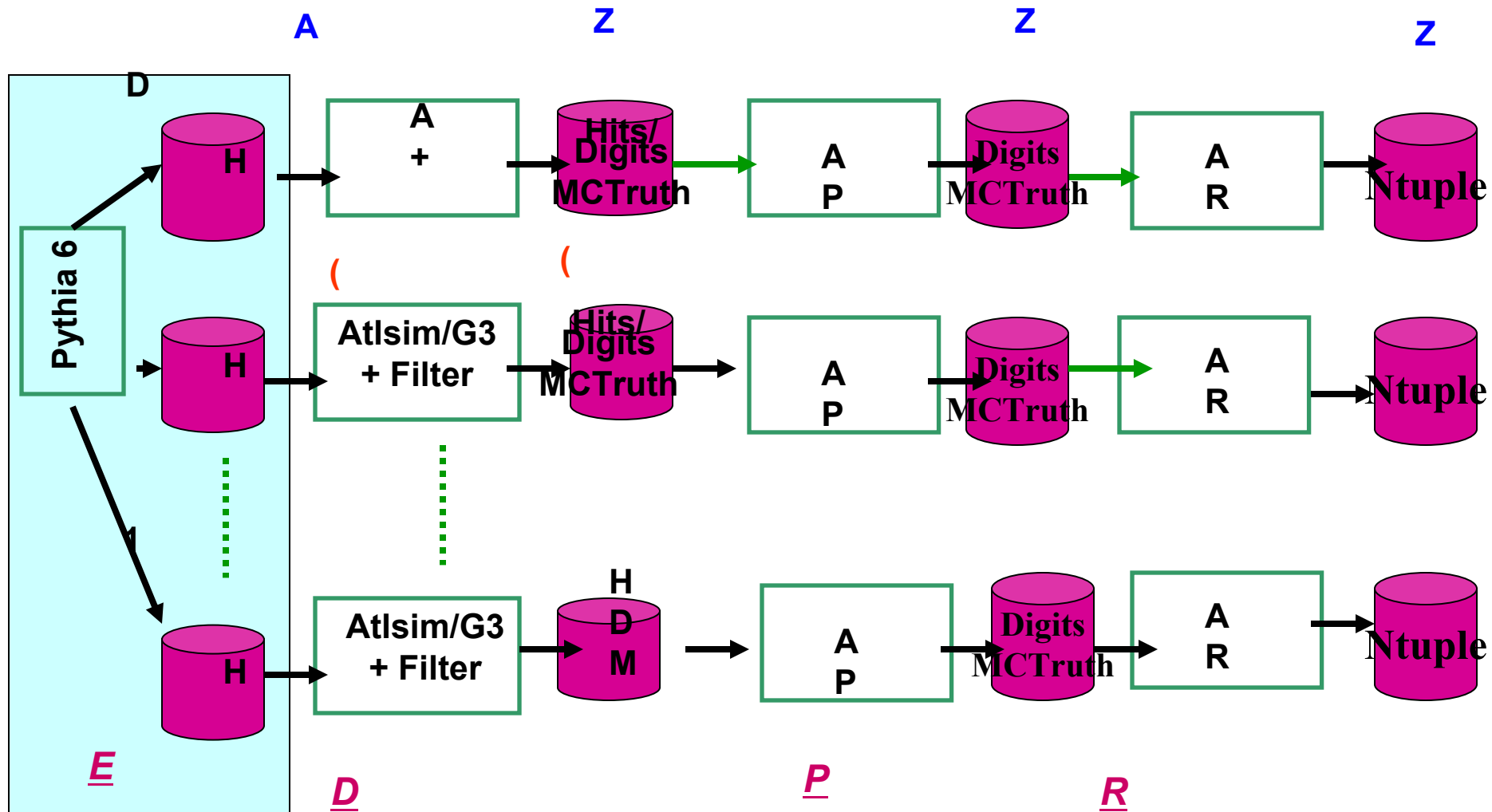
- Primary concern was delivery of events to High Level Trigger (HLT) and to Physics communities
 - HLT-TDR due by June 2003
 - Athens Physics workshop in May 2003
- Put in place the full software chain from event generation to reconstruction
 - Switch to AthenaRoot I/O (for Event generation)
 - Updated geometry
 - New Event Data Model and Detector Description
 - Reconstruction (mostly OO) moved to Athena
- Put in place the distributed production
 - "ATLAS kit" (rpm) for software distribution
 - Scripts and tools (monitoring, bookkeeping)
 - AMI database; Magda replica catalogue; VDC
 - Job production (AtCom)
 - Quality Control and Validation of the full chain
- Use as much as possible Grid tools



ATLAS DC1 (July 2002-April 2003)

- DC1 was divided in 3 phases
 - Phase 1 (July-August 2002)
 - Event generation and detector simulation
 - Phase 2 (December 2002 - April 2003)
 - Pile-up production
 - Classical batch production
 - With Grid tools on NorduGrid and US-ATLAS-Grid
 - Reconstruction (April-May 2003)
 - Offline code only
- Worldwide exercise with many participating institutes

DC1 Task Flow



DC1 in numbers

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.	3×10^7	125	9600	2
Lumi02 Pile-up	4×10^6	22	1650	14
Lumi10 Pile-up	2.8×10^6	78	6000	21
Reconstruction	4×10^6	50	3750	
Reconstruction + Lvl1/2	2.5×10^6	(84)	(6300)	
Total		690 (+84)	51000 (+6300)	60

DC1 Phase 1: July-August 2002

3200 CPU's
110 kSI95
71000 CPU days

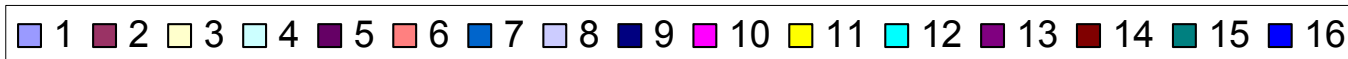
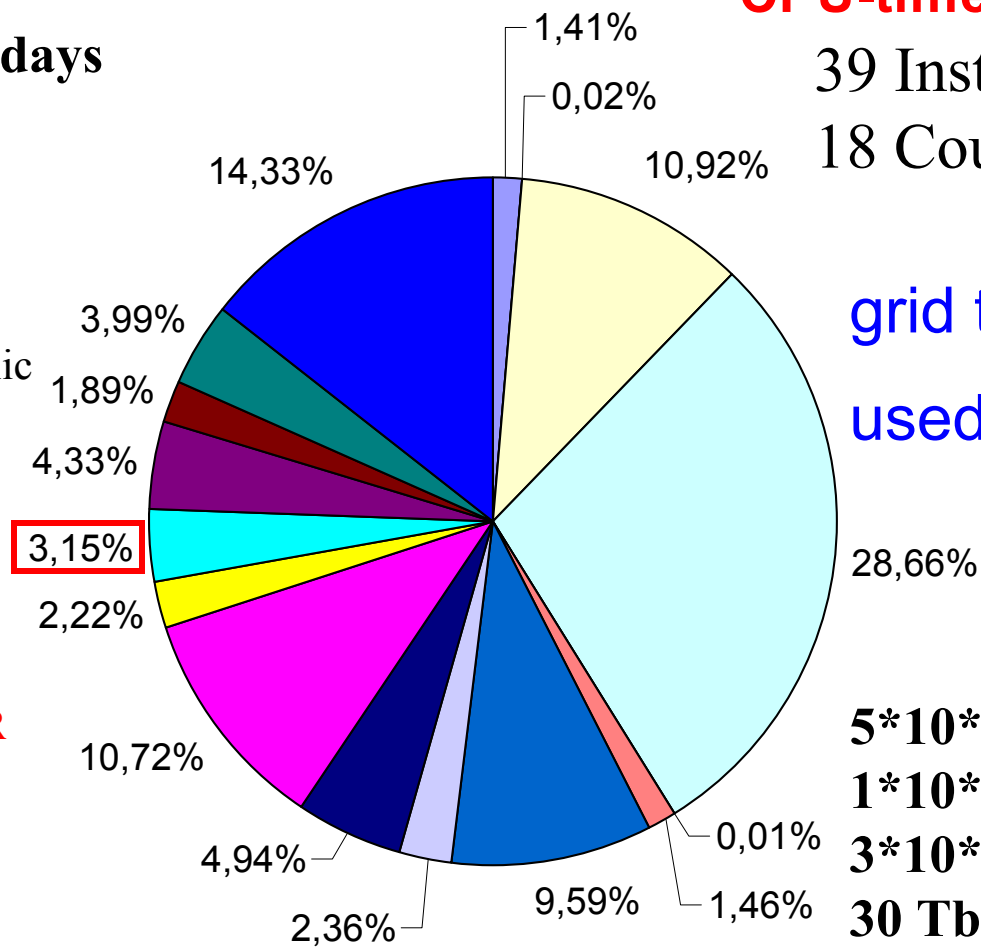
Contribution to the overall CPU-time (%) per country

39 Institutes in
 18 Countries

grid tools
 used at 11 sites

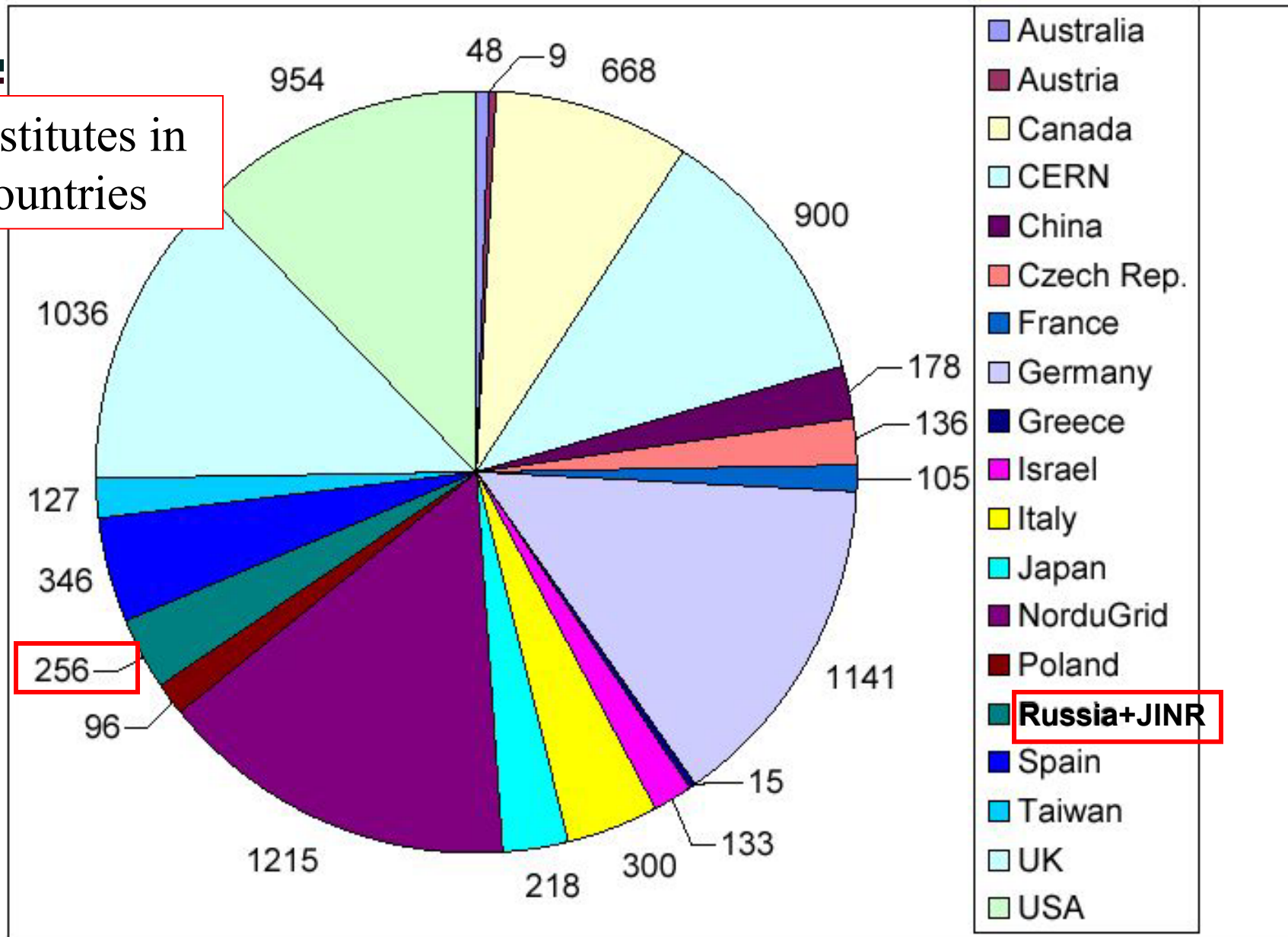
$5 \cdot 10^7$ events generated
 $1 \cdot 10^7$ events simulated
 $3 \cdot 10^7$ single particles
30 Tbytes
35 000 files

1. Australia
2. Austria
3. Canada
4. CERN
5. Czech Republic
6. France
7. Germany
8. Israel
9. Italy
10. Japan
11. Nordic
- 12. Russia+JINR**
13. Spain
14. Taiwan
15. UK
16. USA



DC1 Phase 2: Nov. 2002 - Jan. 2003

56 Institutes in
21 Countries





Data Challenge 2

- DC2 operation in 2004:
 - distributed production of ($>10^7$) simulated events in May-June
 - events sent to CERN in ByteStream (raw data) format to Tier-0
 - reconstruction processes run on prototype Tier-0 in a short period of time (~ 10 days, "10% data flow test")
 - reconstruction results distributed to Tier-1s and analysed on Grid
- Main "new" software to be used (wrt DC1 in 2002/2003):
 - Geant4-based simulation, pile-up and digitization in Athena
 - complete "new" EDM and Detector Description interfaced to simulation and reconstruction
 - POOL persistency
 - LCG-2 Grid infrastructure
 - Distributed Production and Analysis environment



Phases of DC2 operation

- Consider DC2 as a three-part operation:
 - part I: production of simulated data (May-June 2004)
 - needs Geant4, digitization and pile-up in Athena, POOL persistency
 - "minimal" reconstruction just to validate simulation suite
 - will run on any computing facilities we can get access to around the world
 - part II: test of Tier-0 operation (July 2004)
 - needs full reconstruction software following RTF report design, definition of AODs and TAGs
 - (calibration/alignment and) reconstruction will run on Tier-0 prototype as if data were coming from the online system (at 10% of the rate)
 - output (ESD+AOD) will be distributed to Tier-1s in real time for analysis
 - in parallel: run distributed reconstruction on simulated data
 - this is useful for the Physics community as MC truth info is kept
 - part III: test of distributed analysis on the Grid (Aug.-Oct. 2004)
 - access to event and non-event data from anywhere in the world both in organized and chaotic ways

DC2: Scenario & Time scale

September 03: **Release7**

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

March 17th 04: **Release 8
(production)**

Start final validation

May 3rd 04:

Start simulation; Pile-up & digitization

Event mixing

Transfer data to CERN

July 1st 04: “DC2”

Intensive Reconstruction on “Tier0”

Distribution of ESD & AOD

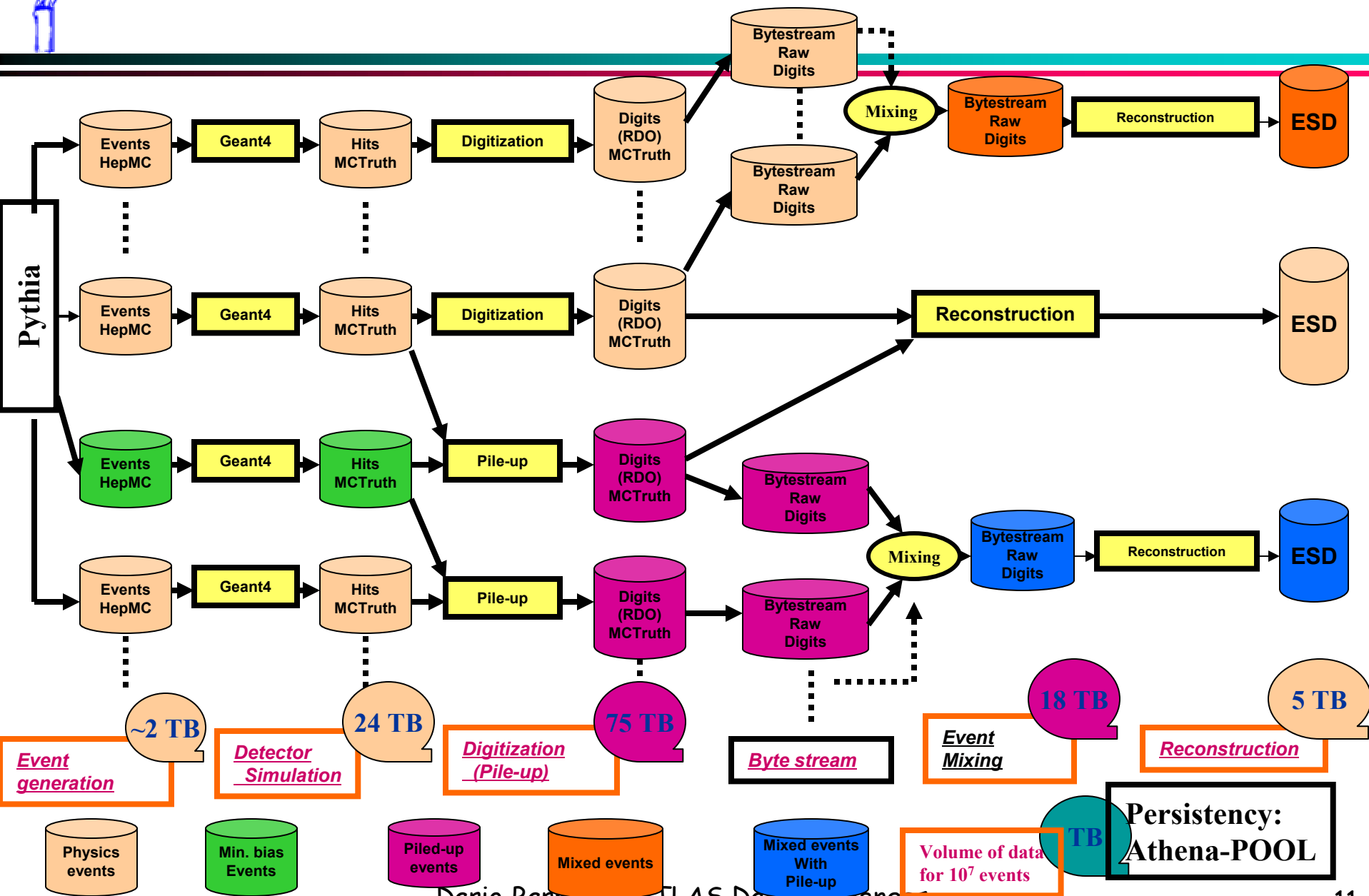
Calibration; alignment

August 1st 04:

Start Physics analysis

Reprocessing

Task Flow for DC2 data



Production scenario

	Input		Output		Comments
Event generation	none		Generated events	< 2 GB files	
G4 simulation	Generated Events	"part of" < 2 GB files	Hits + MCTruth	< 2 GB files	~ 400-500 jobs/day ~ 500 GB/day ~ 6 MB/s
Detector response	Hits + MCTruth (Generated events)	1 file	Digits +MCTruth	RDO (or BS)	No MCTruth if BS
Pile-up	Hits "signal" +MCTruth Hits "min.b"	1 file Several 10 files	Digits +MCTruth	RDO (or BS)	~ 1000 jobs/day ~ 10 GB/job ~ 10 TB/day ~ 125 MB/s
Byte-stream	"pile-up" data RDO	1 (or few) files	BS		Still working on scenario
Events mixing	RDO or BS	Several files	BS		"
Reconstruction	RDO or BS		ESD		
AOD production	ESD		AOD		Streaming?

DC2 resources

Process	No. of events	Time	CPU power	Data volume	At CERN	Off site	
		months	kSI2k	TB	TB	TB	
Simulation	10^7	2	600	25	5	20	Phase I (May-June)
Pile-up & Digitization	10^7	2	400	75	15	60	
ByteStream Production	10^7	2	(small)	20	20	16	
Total Phase I	10^7	2	1000	120	40	96	
Reconstr. Tier-0	10^7	0.5	600	5	5	10	Phase II (Jul)
Reconstr. Tier-1	10^7	2	150	5	0	5	
Total	10^7			130	45	111	



DC2 resources - notes

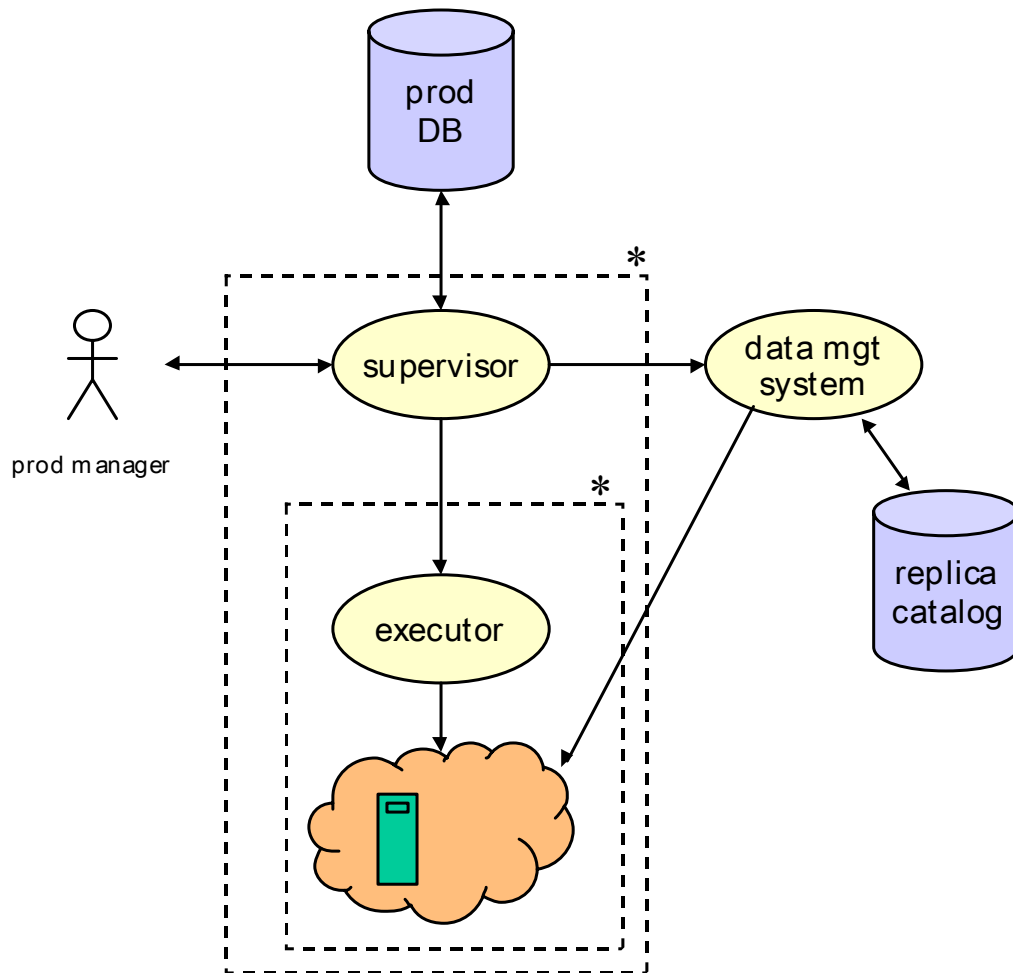
- CPU needs are now based on Geant4 processing times
- We assume 20% of simulation is done at CERN
- All data in ByteStream format are copied to CERN for the Tier-0 test (Phase II)
- Event sizes (except ByteStream format) are based on DC1/Zebra format: events in POOL are a little larger (~1.5 times for simulation)
- Reconstruction output is exported in 2 copies from CERN Tier-0
- Output of parallel reconstruction on Tier-1s, including links to MC truth, remains local and is accessed for analysis through the Grid(s)



DC2: Grid & Production tools

- We foresee to use:
 - 3 Grid flavors (LCG-2; Grid3+; NorduGrid)
 - (perhaps) "batch" systems (LSF; ...)
 - Automated production system
 - New production DB (Oracle)
 - Supervisor-executer component model
 - Windmill supervisor project
 - Executors for each Grid and LSF
 - Data management system
 - Don Quijote DMS project
 - Successor of Magda
 - ... but uses native catalogs
 - AMI for bookkeeping
 - Going to web services
 - Integrated to POOL

New Production System (1)



- DC1 production in 2002/2003 was done mostly with traditional tools (scripts)
 - Manpower intensive!
- Main features of new system:
 - Common production database for all of ATLAS
 - Common ATLAS supervisor run by all facilities/managers
 - Common data management system
 - Executors developed by middleware experts (LCG, NorduGrid, Chimera teams)
 - Final verification of data done by supervisor

New Production System (2)

Task = [job]*

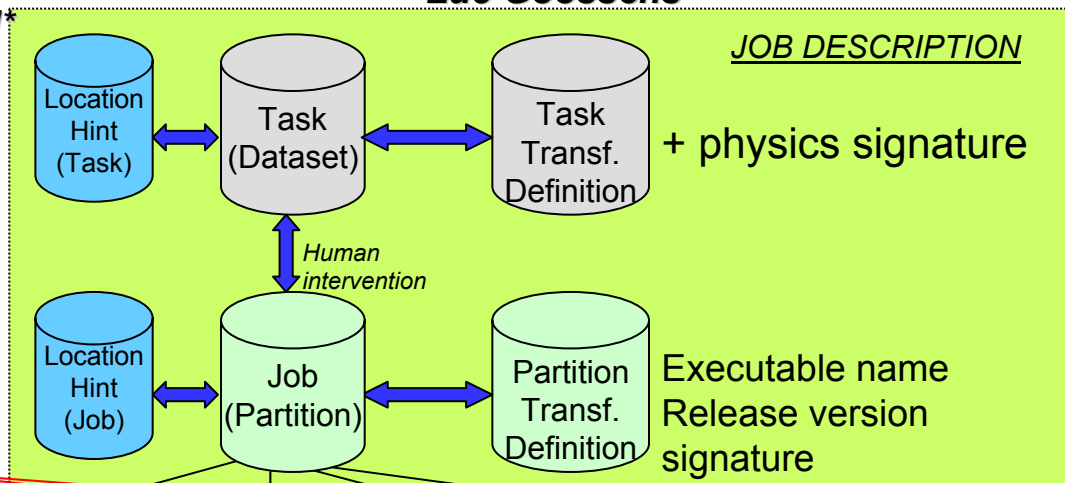
Dataset = [partition]*

Luc Goossens

Data Management System

Luc Goossens

Job Run Info



Kaushik De

Supervisor 1

Supervisor 2

Supervisor 3

Supervisor 4

Rob Gardner

Alessandro De Salvo

Oxana Smirnova

Luc Goossens

US Grid Executer Chimera

LCG Executer

NG Executer RB

LSF Executer

US Grid

LCG

NG

Local Batch



Tiers in DC2

- Tier-0
 - 20% of simulation will be done at CERN
 - All data in ByteStream format (~16 TB) will be copied to CERN
 - Reconstruction will be done at CERN (in ~10 days).
 - Reconstruction output (ESD) will be exported in 2 copies from Tier-0 (2 x ~5 TB).

Tiers in DC2

- Tier-1s will have to
 - Host simulated data produced by them or coming from Tier-2s; plus ESD (& AOD) coming from Tier-0
 - Run reconstruction in parallel to Tier-0 exercise (~2 months)
 - This will include links to MCTruth
 - Produce and host ESD and AOD
 - Provide access to the ATLAS V.O. members
- Tier-2s
 - Run simulation (and other components if they wish to)
 - Copy (replicate) their data to Tier-1
 - ATLAS is committed to LCG
- All information should be entered into the relevant database and catalog

Available (pledged) resources

TIERs	Site	kSI2k	TB (Disc)	TB (Tape)	kSI2k-extra
Australia	Melbourne	12	1		
Canada	TRIUMF	331	5		
Czech Republic	Prague	25	3		
CERN	CERN	200	45		
France	CCIN2P3	200	3		
Germany	Gridka	90	8	10	
Israel		23			
Italy	CNAF	200	10		
Japan	Tokyo	200			
Netherlands	NIKHEF	75	2	10	
NorduGrid	NG	380			
Poland	Cracow	80	3.5		
Spain	PIC	50	4		
Russia+JINR		140	18	20-30	
Switzerland	Manno	18	4.6	20	
Taiwan	Taipei	78	3.5		
UK-Batch			20		675
UK	RAL	1450	15		
US	BNL	200	50		
		3650	180.6		
Data being collected and evolving quickly					



Russian contribution to DC2

- As for DC1, we foresee 4 production centres:
 - Dubna
 - Moscow State University
 - ITEP
 - Protvino
- They act collectively as a Tier-2
- Still open discussions:
 - which Tier-1 will act as technical support for middleware and ATLAS software installation?
 - in which Tier-1 storage facility will data produced in Russia be saved?
 - not necessarily the same as above!



Russian contribution to S&C manpower

- Russian contribution to software & computing manpower (FTEs, current estimate for 2004 average):

Funding Agency	Inner Detector	LAr	TileCal	DC/Grid/Prod.	Total
Russia	2.6	0.3	1.0	0.1	4.0
JINR	-	-	1.0	0.1	1.1

- No contribution to core software (Framework, Core Services, Data Base, Infrastructure)
- Several infrastructure tasks are still not properly covered
 - a lot of "best effort" support for developers and users alike
- Please contact us for details and discussions!