Lessons from ATLAS DC2 and Rome Production on Grid3

Jim Shank

Boston University

The ATLAS Experiment

CHEP06, TIFR

February 15, 2006

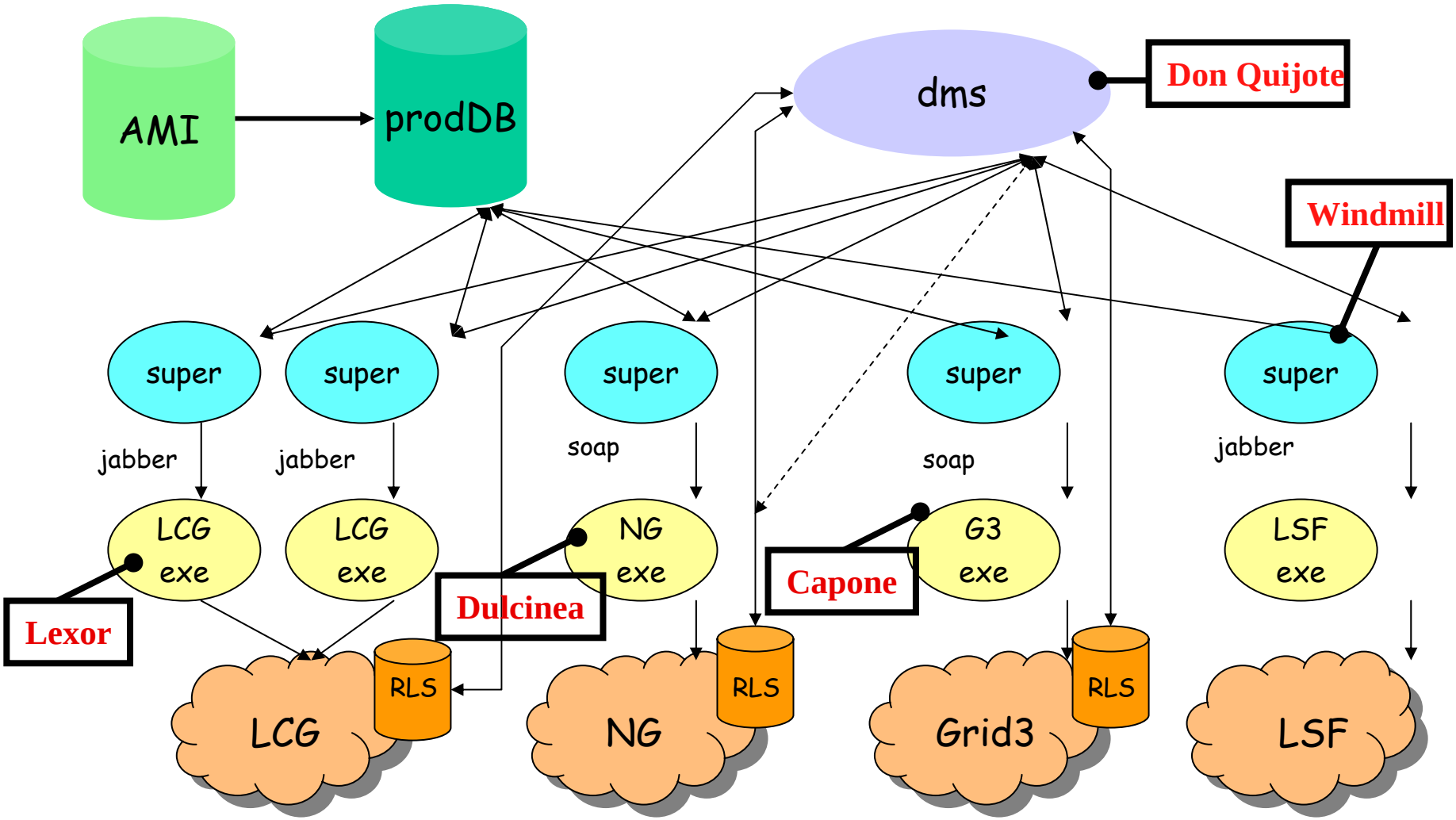
ATLAS Data Production

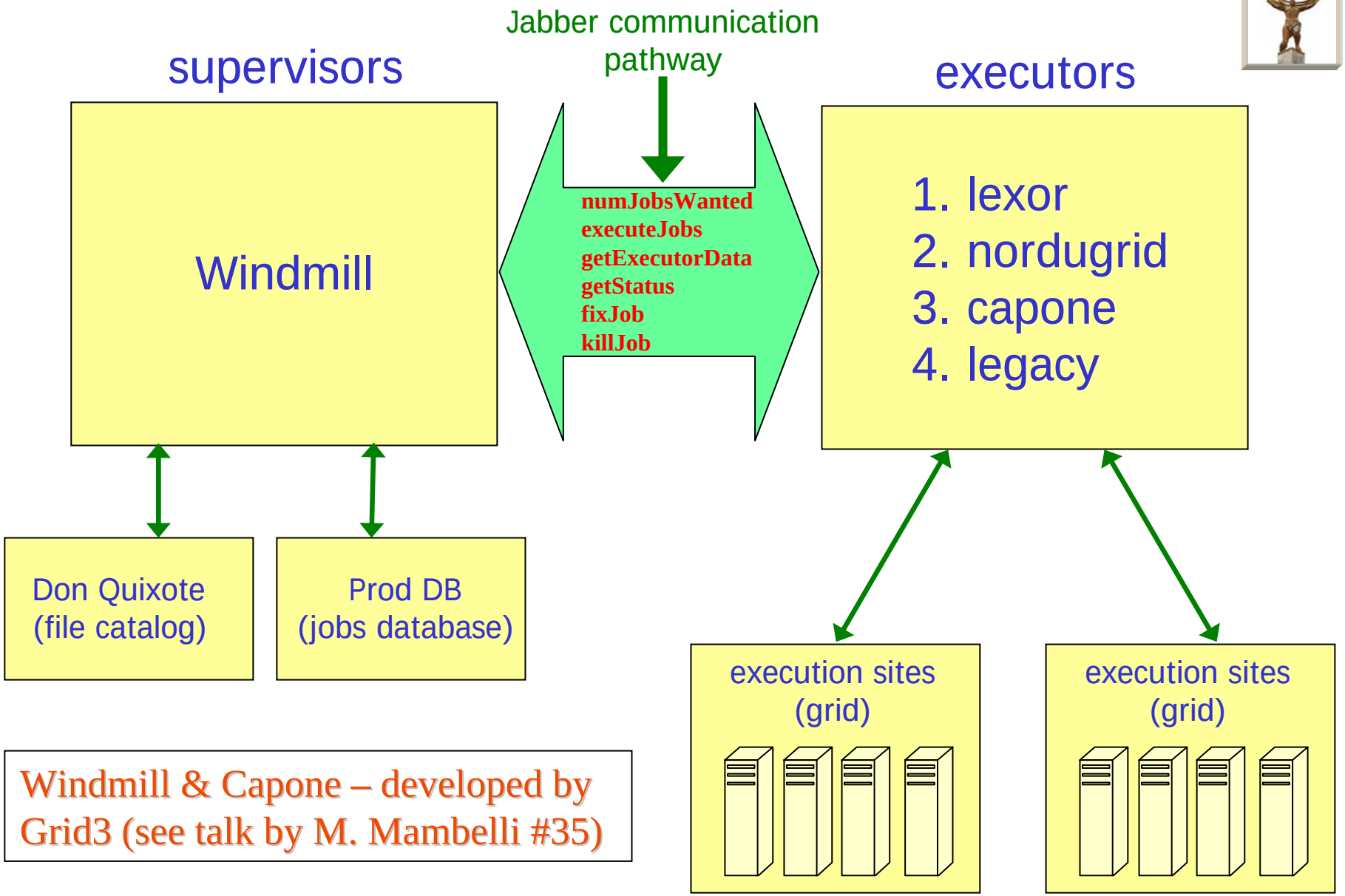
- Data Challenge 2
  - Goal: to validate ATLAS computing model
  - All production done on grid facilities ~Jul-Dec, 2004
- Rome Production
  - Goal: produce data for Rome Physics Workshop
  - Grid production ~Jan-May, 2005 (over 200 physics samples)
- See CHEP06 talk by G. Poulard #111 for more details
- This talk will focus on experience with Grid3 in U.S.
  - 250k jobs successfully finished in U.S. (~25% of ATLAS)
  - 60M SpecInt 2000 CPU days used in U.S.
  - 20 universities/national labs produced data in the U.S.
  - Over 100 papers presented in Rome from data produced on all grids

Data Challenge 2 (DC2)

- DC2 Phase I (other phases not reported here)
  - Part 1: event generation
  - Part 2: GEANT 4 detector simulation (CPU intensive)
  - Part 3: pile-up and digitization (disk intensive)
    - Pile-up: superposition of “background” events with the “signal” event
    - Digitization: response of the sensitive elements of the detector
- DC2 Scale
  - Physics channels ~10 million events
  - Calibration samples ~few million events

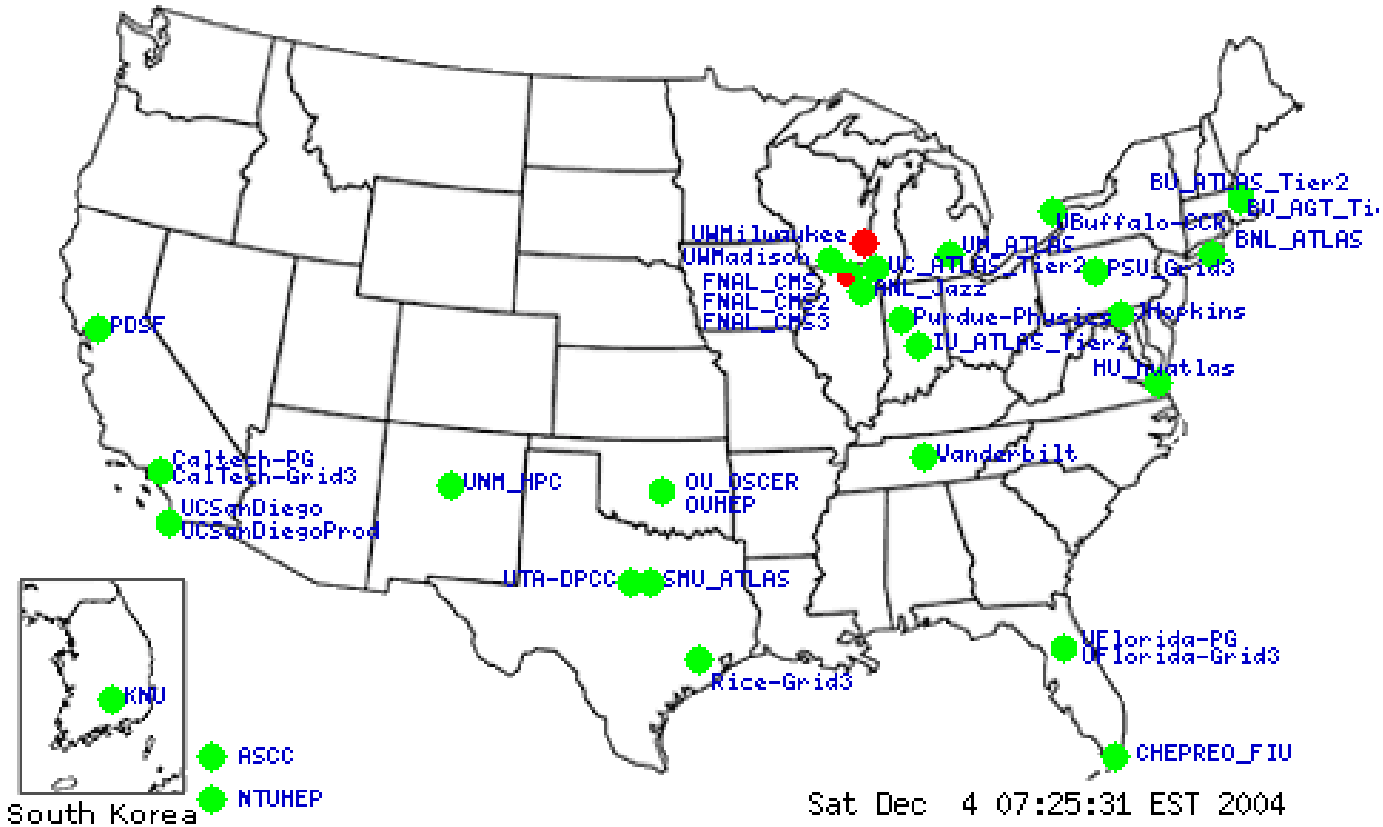
ATLAS Production system




Jabber communication pathway
supervisors
executors
Windmill
numJobsWanted
executeJobs
getExecutorData
getStatus
fixJob
killJob
1. lexor
2. nordugrid
3. capone
4. legacy
Don Quixote (file catalog)
Prod DB (jobs database)
execution sites (grid)
execution sites (grid)
Windmill & Capone – developed by Grid3 (see talk by M. Mambelli #35)

Grid3 Organization

- 35 sites; ~3500 CPUs (~1/4 dedicated to ATLAS)
- Used for U.S. DC2 production and Rome physics data generation



DC2 Phase I (June-Dec, 2004)

ATLAS DC2 Production Status

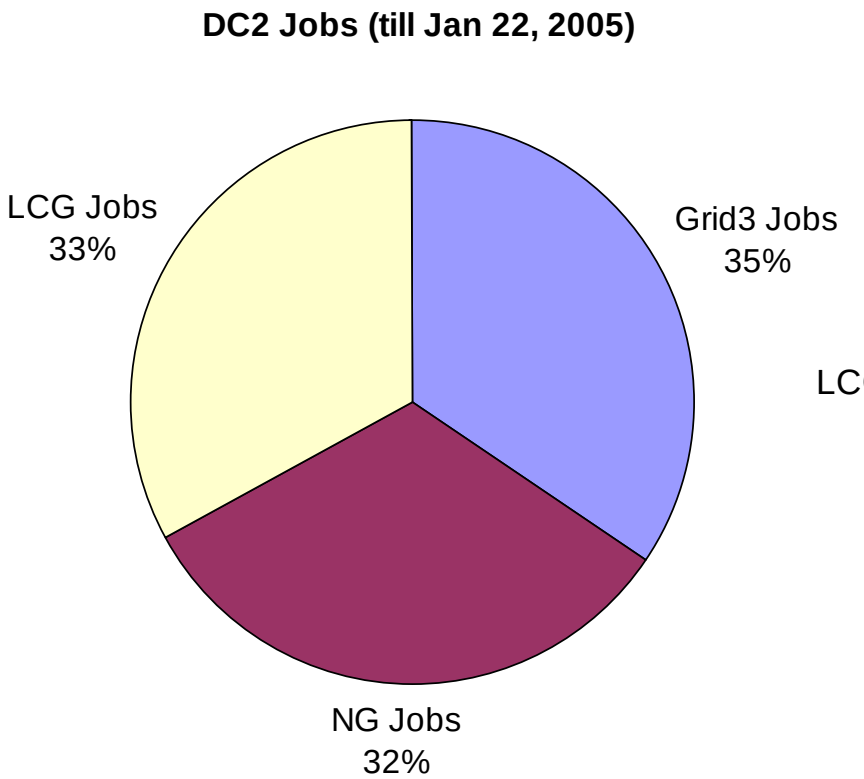
Overview of Grids

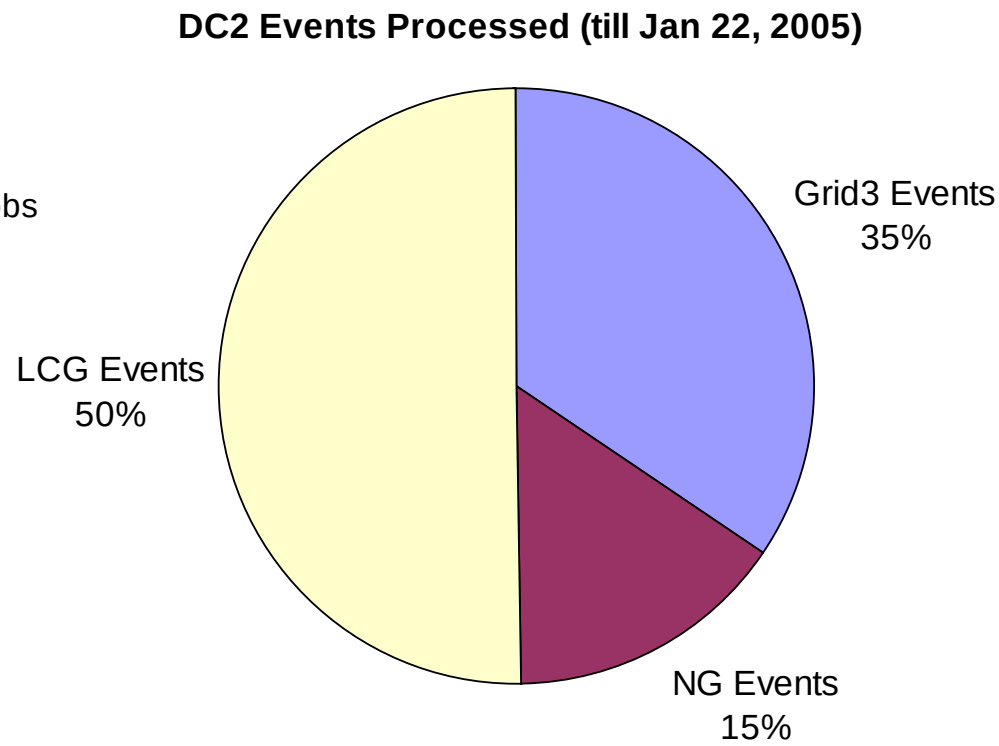
as of 2005-01-28 10:33:00

Grid	submitted	pending	running	finished	failed	efficiency
Grid3	27	10	215	122531	41385	75 %
NorduGrid	16			111108	66311	63 %
LCG	32	380	432	122688	215105	36 %
TOTAL	75	390	647	356327	322801	52 %

Grid3 – lowest failure rate among ATLAS grids
Above statistics does not include Rome Production

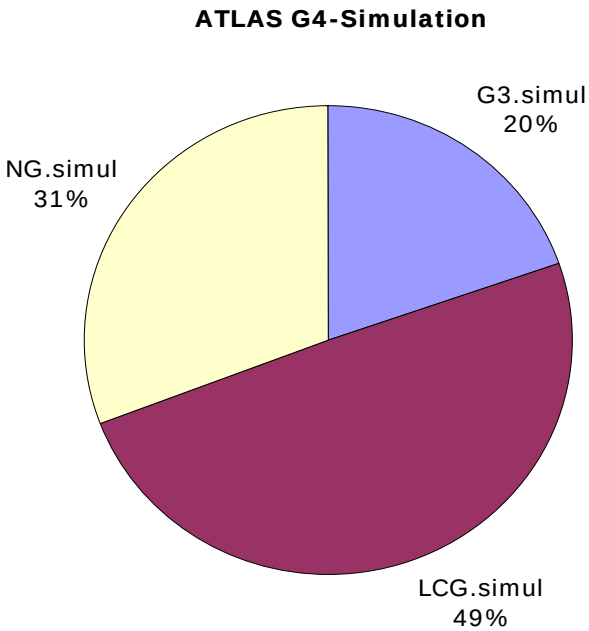
DC2 Total Jobs Per Grid

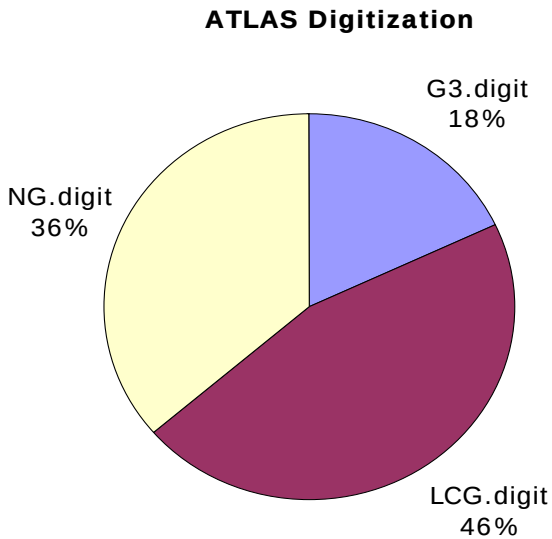


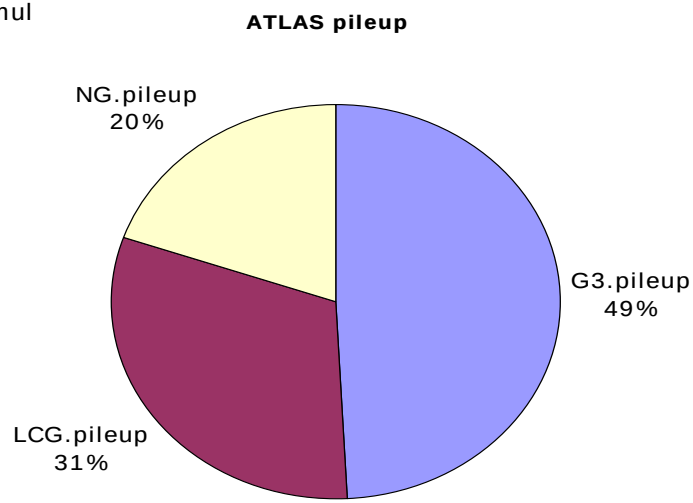


Different Job Types Per Grid

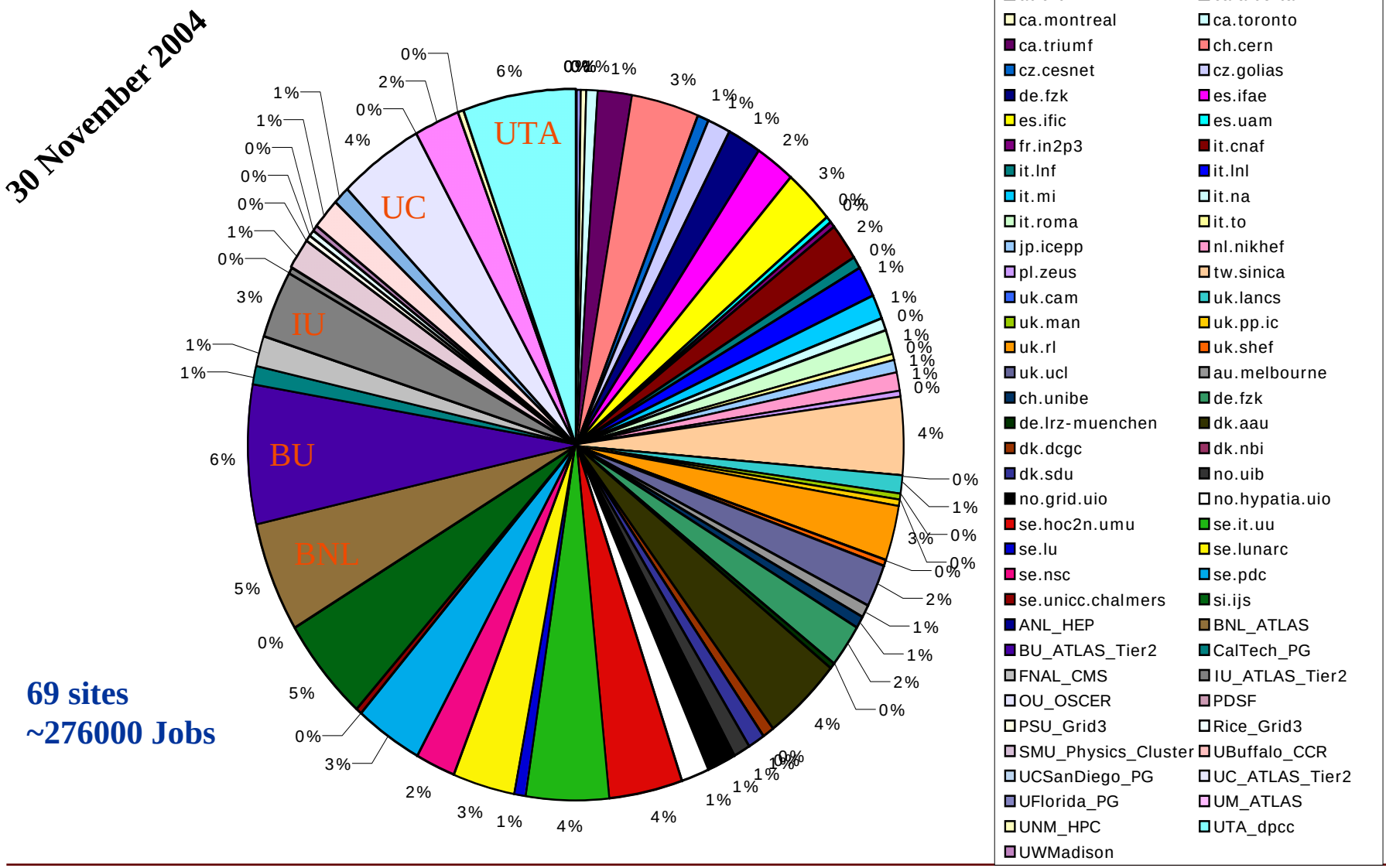
Physics channels only
30 November 2004



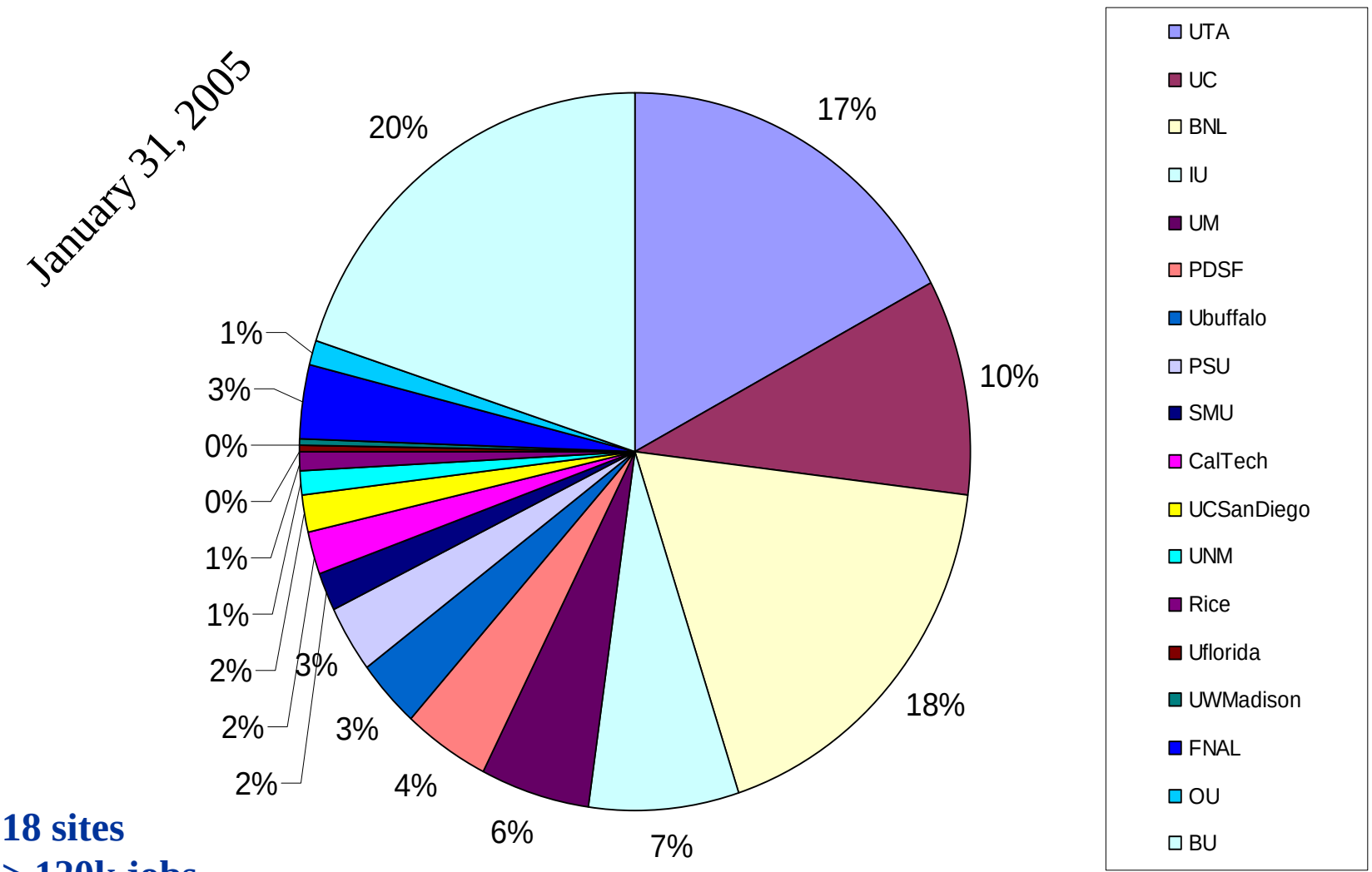




DC2 Jobs Per Site



U.S. DC2 Production by Site



18 sites
> 120k jobs

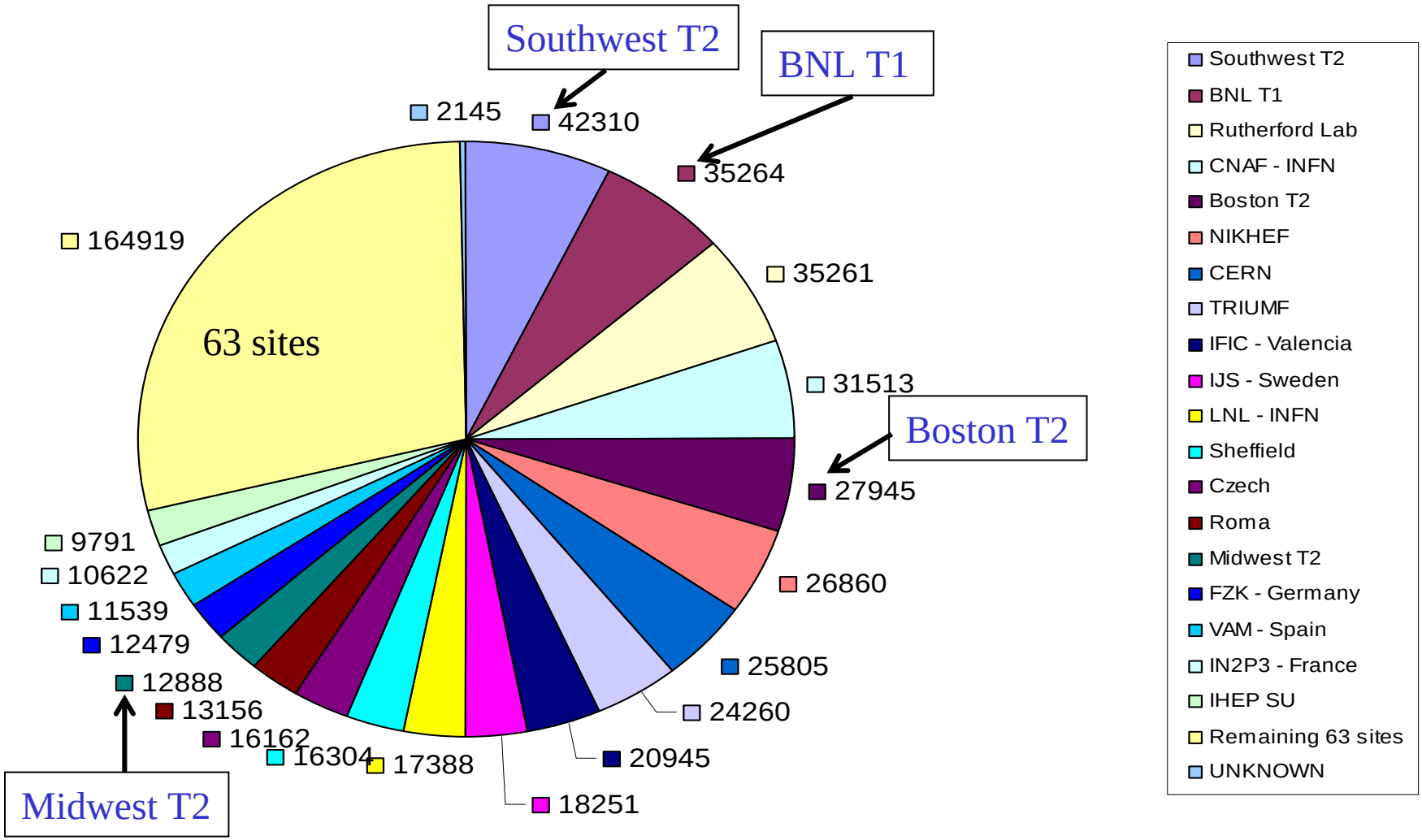
Rome Production Summary - Jobs

Overview of Grids

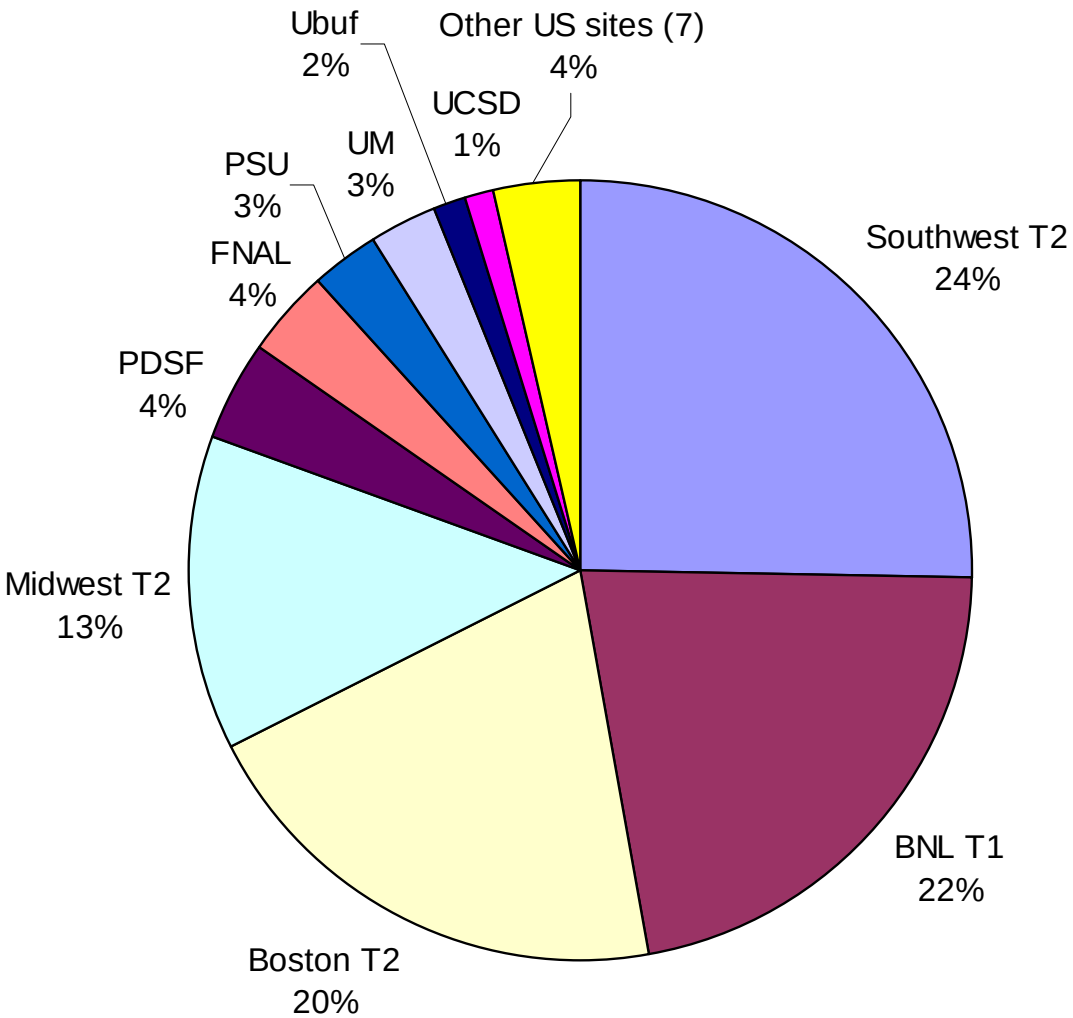
as of 2005-05-23 19:08:24

Grid	submitted	pending	running	finished	failed	efficiency
Grid3	14		345	253362	92777	73 %
condorg	1	2563	1006	135034	165111	45 %
NorduGrid	21	242	117	165744	113630	59 %
LCG	251	633	2145	305178	443420	41 %
zorro	2			1		100 %
TOTAL	289	3438	3613	859319	814938	51 %

Rome Grid Production
Successful Job Count at 83 ATLAS sites



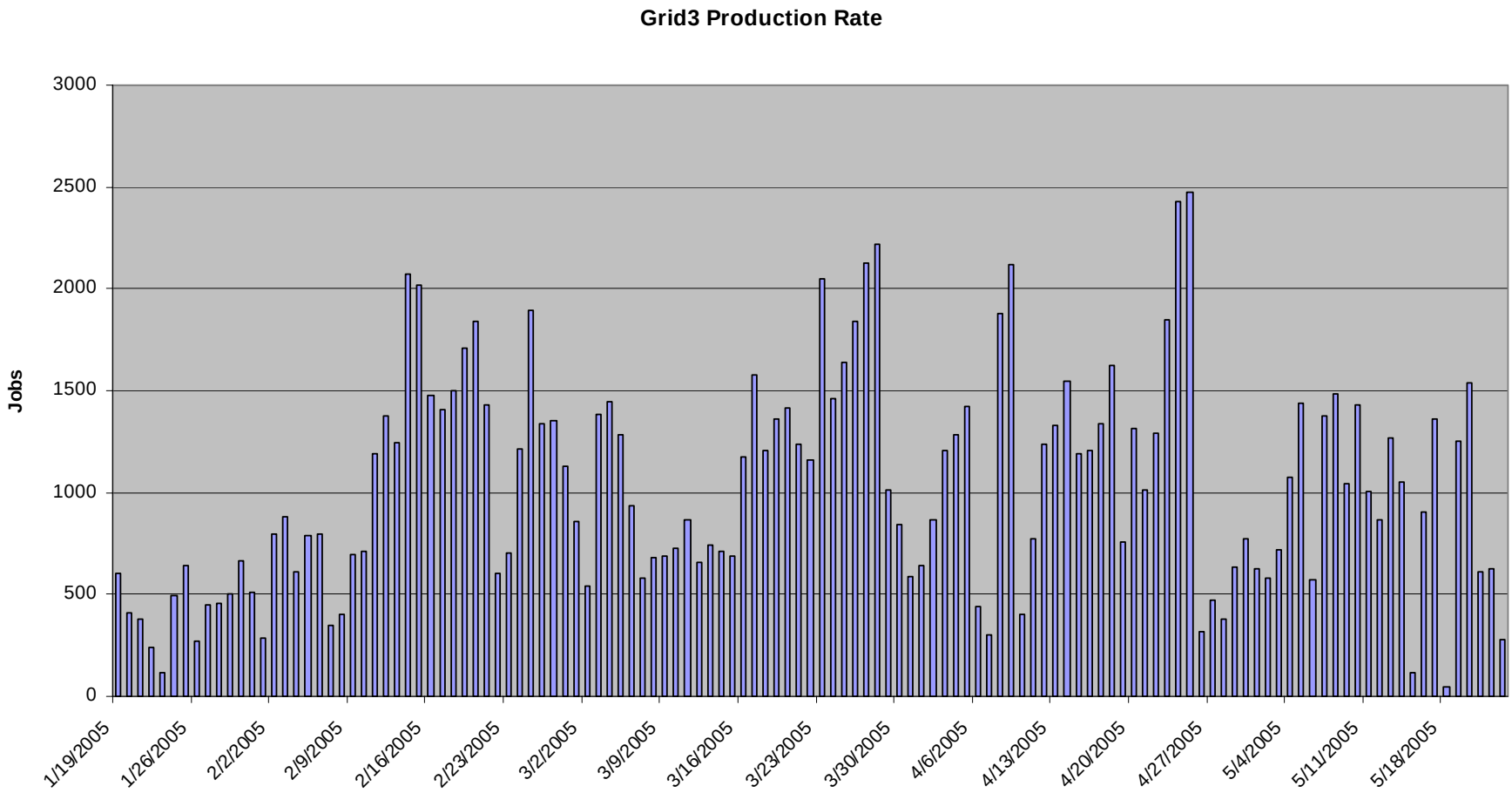
U.S. Grid Production (Rome/DC2 combined)



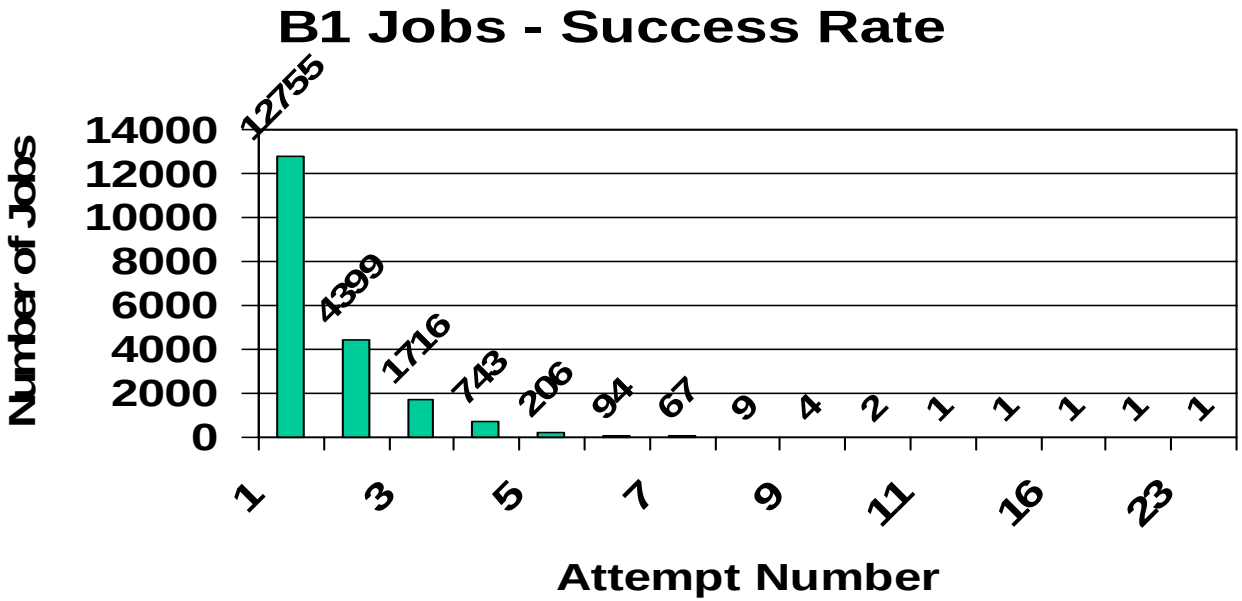
Includes stats from >250k jobs

20 different sites used in the U.S.

U.S. Production Rate



Discussion of Efficiency



- B1 example (20,000 jobs – one million events – done in U.S.)
- Windmill automatically resubmits failed jobs
  - For this sample, up to 3 retries was automatic (defined in production DB)
  - Production manager manually increased this to 9 attempts in database
  - Jobs that failed greater than 9 attempts checked by hand
  - Most failed attempts are infrastructure problems or middleware limitations

Job Success Rate on GRID3

	Passed	Failed	Success Rate
July	8799	6676	57%
August	17083	9448	64%
September	17283	7717	69%
October	26600	5186	84%
November	21869	5038	81%

- Key factors in improved success rate:
  - Experienced team using common submit hosts, shorter jobs
  - Quicker response to large scale site/network/hardware failures
- Can we improve more?
  - Some shifts >95% success, others <50%
  - Automatic throttle for failures at problem sites? But still lose all running jobs
  - Do we care?

Error Analysis on Grid3

Failures	Cumulative till end of Nov. 2004	Sep. 2004	Oct.-Nov. 2004
Submission	894	472	422
Exe check	428	428	0
Run-End	10131	1147	8984
StageOut	10833	8037	2796
RLS reg	1065	989	76
Capone host interruption	3975	2725	1250
WM verification	564	57	507
Other	5225	5139	86
TOTAL	33165	19303	13862

Lessons Learned

- Scalability
  - New problems discovered every few weeks as we increased scale
  - Needed constant interaction with software developers
  - Dependencies - solution to some problems introduced new ones
- Testing of new ATLAS software releases
  - Need to plan for 2 week validation of new releases on grid
  - Required 2-8 new transformations after deployment on grid
- Continuous validation and physicist involvement
  - Some problems found after dataset completion
  - Many datasets aborted or restarted
  - Need involvement of physicists for Quality Control

People Involved in Grid3 Produciton

- SHANK, James (Boston University)
  DE, Kaushik (University of Texas at Arlington)
- SMIRNOV, Yuri (Brookhaven National Laboratory)
  SOSEBEE, Mark (UNIVERSITY OF TEXAS AT ARLINGTON)
  VASSILAKOPOULOS, Vassilios (Hampton University)
  ALEXANDRE, Vaniachine (Argonne National Laboratory)
  WLODEK, Tomasz (Brookhaven National Laboratory)
  ZHAO, Xin (Brookhaven National Laboratory)
  OZTURK, Nurcan (UNIVERSITY OF TEXAS AT ARLINGTON)
  DENG, Wensheng (Brookhaven National Laboratory)
  GARDNER, Robert (University of Chicago)
  GERIALTOWSKI, Jerry (Argonne National Laboratory)
  JOFFE, David (Southern Methodist University)
  MAMBELLI, Marco (University of Chicago)
  HINCHLIFFE, Ian (Lawrence Berkeley Laboratory)
  COSTANZO, Davide (Brookhaven National Laboratory)
  SHIN, Taeksu (Hampton University)

Conclusion

- Impressive success in spite of many problems – thanks to hard working Grid3 production team
- Grid3 has now evolved into OSG (many talks at CHEP06)
- ATLAS has started Computer System Commissioning (CSC)
- New production and distributed analysis system developed in U.S. based on DC2 and Rome production experience called Production and Distributed Analysis (PanDA) See CHEP 06 talk by K. De #347
  - Tightly coupled to the ATLAS Distributed Data Management system
    - See CHEP 06 Talk D. Cameron #75
  - Relies on proven (at scale) grid middleware
  - Distributed Data Analysis is a fundamental component