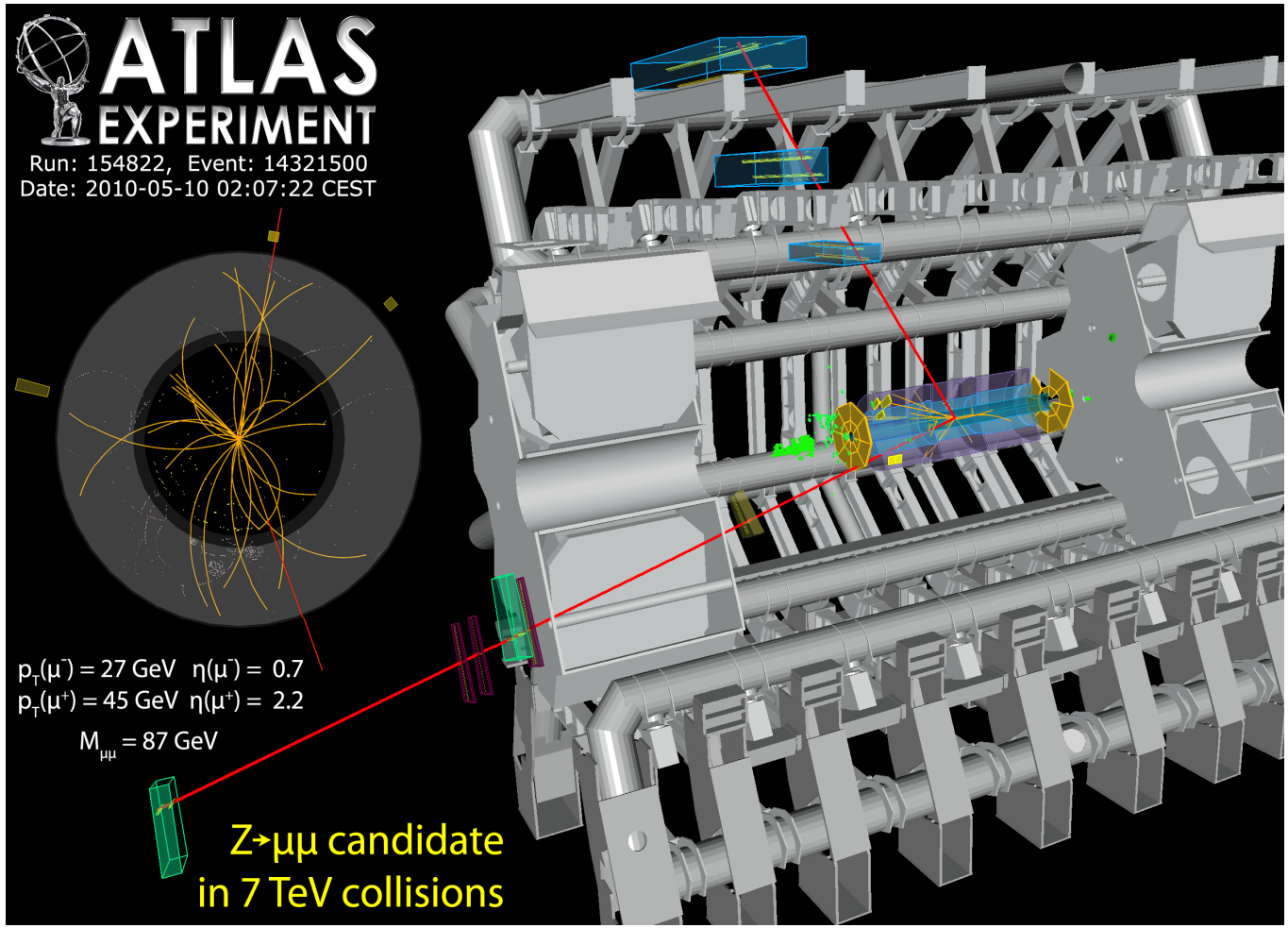
Atlas experience with ARC

Andrej Filipčič
Jozef Stefan Institute
Ljubljana, Slovenia

Atlas


ATLAS
EXPERIMENT
Run: 154822, Event: 14321500
Date: 2010-05-10 02:07:22 CEST
$p_T(\mu^-) = 27$ GeV $\eta(\mu^-) = 0.7$
$p_T(\mu^+) = 45$ GeV $\eta(\mu^+) = 2.2$
$M_{\mu\mu} = 87$ GeV
Z→μμ candidate
in 7 TeV collisions

Computing infrastructure

- CERN T0
- 10 clouds
  - T1 centers
  - Associated T2s, T3s
- 250 sites
- 70k cores
- 50PB of storage
- 3 grid flavors: gLite, OSG, ARC

Production system

- PanDA: central task definition, brokering, job control
  - Web interface
  - Client tools (pathena, ganga, …)
- Autopilot factory: (dummy) job submission to gLite/OSG/ARC clusters
- Pilot: running on WN, payload from PanDA, job control callbacks, middleware/site dependencies handled by 25k lines of python
  - Monte-Carlo production (1GB/core/day)
  - Data processing (20GB/core/day)
  - User analysis (200GB/core/day)
- DDM data management system (DQ2, SRM, LFC, FTS)
- Tasks brokered to clouds, jobs brokered to sites
- Typically per day:
  - 200k production jobs
  - 200-300k analysis jobs
  - 1-2PB data processed
  - 500TB data moved between sites
  - 10^7 catalog operations

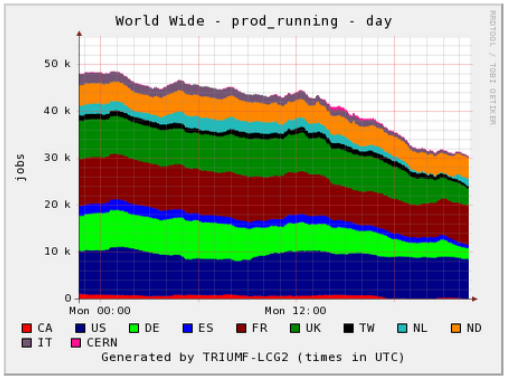
Production dashboard

Active tasks: CA:2 CERN:3 DE:16 ES:5 FR:12 IT:7 ND:13 NL:13 TW:6 UK:6 US:25
Bamboo task brokerage, job submissions, status over last 12 hours

Jobs updated >12 hrs ago: **activated:** 20140 **running:** none
Jobs updated >36 hrs ago: **transferring:** 1

Summary plots by cloud

Cloud efficiency history



Production job summary, last 12 hours (Details: errors, nodes)
Comments/requests on the new summary table to Torre (wenaus@gmail.com). Old version of page is here
Processing types: evgen(2050) merge(8632) pile(11496) prod_test(195) reco(60) reprocessing(11256) simul(223598) validation(961)
Users: Johanna.Fleckner(10) Michiru.Kaneda(60) borut.kersevan(246727) douglas(195) nick.barlow(6743) pavel.nevski(4410) strandbe(103)
Pilot counts are for the last 3 hours. Error rates above 5% are shown in red.

Cloud	Pilots	Latest	defined	assigned	waiting	activated	sent	running	holding	transferring	finished	failed	cancelled	%fail
ALL			0	879	6	46892	2	30628	1826	17922	107656	19490	32947	15%
CA	338	09-13 22:20	0	0	0	0	0	3	0	132	7967	890	1213	10%
CERN (brokeroff)	367	09-13 22:20	0	0	0	0	0	9	1	0	639	333	0	34%
DE	3719	09-13 22:20	0	0	0	796	0	2320	223	3482	20423	2513	5045	11%
ES	574	09-13 22:20	0	0	0	1258	0	760	77	1209	3071	1832	1383	37%
FR	2887	09-13 22:20	0	231	0	12756	2	8399	260	1452	7770	1796	17375	19%
IT	328	09-13 22:20	0	0	0	582	0	296	39	1490	7046	3576	633	34%
ND	504	09-13 22:20	0	0	0	3028	0	4531	175	214	12567	189	1433	1%
NL	1169	09-13 22:20	0	648	6	1432	0	1541	213	1230	7559	1461	687	16%
TW	567	09-13 22:20	0	0	0	0	0	550	37	10	3699	770	1369	17%
UK	2258	09-13 22:20	0	0	0	9740	0	3737	196	3517	11460	6100	2714	35%
US	3381	09-13 22:20	0	0	0	17300	0	8482	605	5186	25455	30	1095	0%

ND cloud

- NDGF T1
  - 2.5PB distributed dCache storage (disk, tape)
- T1, T2, T3
  - 15 clusters: NO, SE, DK, SI, CH, UK
  - 4-6k cores available to Atlas (of 16k)
- 10Gb/s OPN for storage + mostly 10Gb/s Geant for clusters

ARC queues

ND Tasks Jobs: activated:12038 running:2647 holding:49 transferring:715							
ARC	ANALY_ARC	ANALY_ARC	4			online	manual
	ARC-arc-ce.smokerings.nsc.liu.se	ARC	4			online	manual
	ARC-arc-ce01.pdc.kth.se	ARC	4			online	manual
	ARC-arc.bccs.uib.no	ARC	4			online	manual
	ARC-arc01.lcg.cscs.ch	ARC	4			online	manual
	ARC-ce.lhep.unibe.ch	ARC	4			online	manual
	ARC-ce01.titan.uio.no	ARC	4			online	manual
	ARC-gateway01.dcsc.ku.dk	ARC	4			online	manual
	ARC-grad.uppmax.uu.se	ARC	4			online	manual
	ARC-grid.uio.no	ARC	4			online	manual
	ARC-grid03.unige.ch	ARC	4			online	manual
	ARC-jeannedarc.hpc2n.umu.se	ARC	4			online	manual
	ARC-nordugrid.unibe.ch	ARC *NoComment*	4			online	auto
	ARC-pikolit.ijs.si	ARC	4			online	manual
	ARC-siri.lunarc.lu.se	ARC	4			online	manual
ARC-T2	ARC-T2-arc-ce.smokerings.nsc.liu.se	ARC-T2	4			online	manual
	ARC-T2-arc-ce01.pdc.kth.se	ARC-T2	4			online	manual
	ARC-T2-arc.bccs.uib.no	ARC-T2	4			online	manual
	ARC-T2-arc01.lcg.cscs.ch	ARC-T2	4			online	manual
	ARC-T2-arc02.lcg.cscs.ch	ARC-T2	4			online	manual
	ARC-T2-ce.lhep.unibe.ch	ARC-T2	4			online	manual
	ARC-T2-ce01.titan.uio.no	ARC-T2	4			online	manual
	ARC-T2-gateway01.dcsc.ku.dk	ARC-T2	4			online	manual
	ARC-T2-grad.uppmax.uu.se	ARC-T2	4			online	manual
	ARC-T2-grid.uio.no	ARC-T2	4			online	manual
	ARC-T2-grid03.unige.ch	ARC-T2	4			online	manual
	ARC-T2-jeannedarc.hpc2n.umu.se	ARC-T2	4			online	manual
	ARC-T2-nordugrid.unibe.ch	ARC-T2 *Massive failure*	4			online	manual
	ARC-T2-pikolit.ijs.si	ARC-T2	4			online	manual
	ARC-T2-siri.lunarc.lu.se	ARC-T2	4			online	manual

Atlas Monitor

Grid Monitor

2010-09-14 CEST 05:38:24

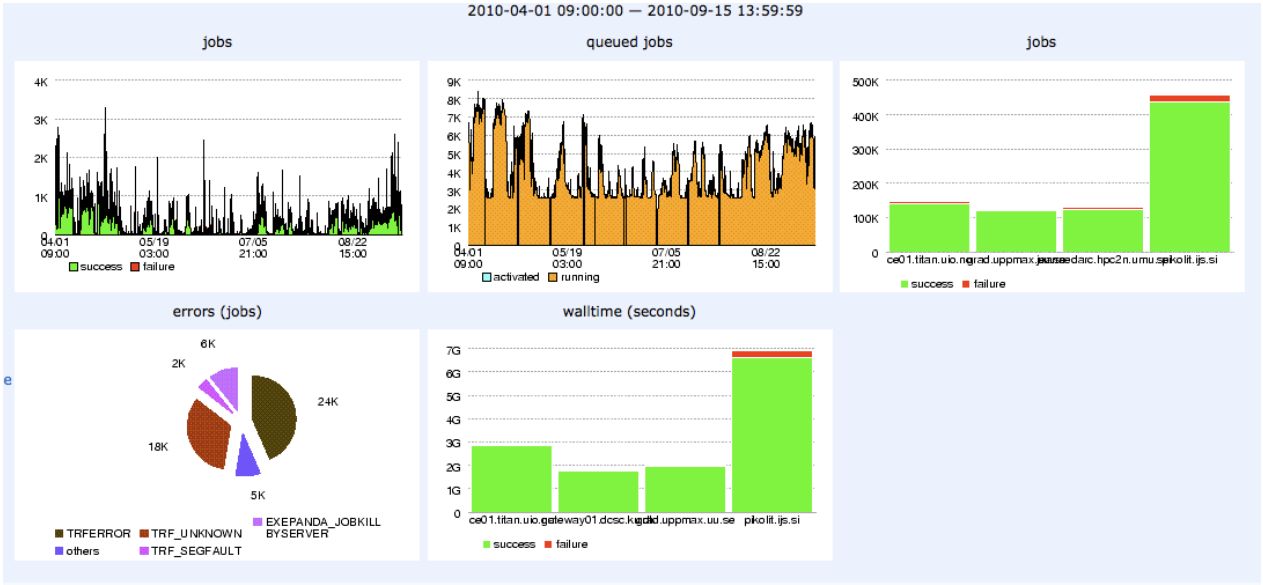
Processes: Grid Local

Country	Site	CPUs	Load (processes: Grid+local)	Queueing
Denmark	Steno (DCSC/KU)	2744	430+1747	1880+6
Norway	EPF (UiO/FI)	20	0+6	1+0
	Titan A (UiO/USIT)	4624	550+2432	116+6474
Slovenia	Arnes	264	261+3	158+0
	SiGNET	1152	1064+86	1383+1
Sweden	Grad (SweGrid, Uppmax)	512	504+0	1963+28
	Ritsem (SweGrid, HPC2>)	512	485+2	3045+0
	Ruth (SweGrid, PDC)	286	277+0	1366+-2
	Siri (SweGrid, Lunarc)	496	251+232	2770+51
	Smokerings (NSC)	512	183+305	2711+0
Switzerland	Bern ATLAS T3	212	144+0	113+0
	Bern UBELIX T3 Cluster	1072	264+499	184+327
	Geneva ATLAS T3	222	120+50	63+0
	Manno PHOENIX T2	1536	195+473	89+3
	Manno PHOENIX T2	1536	300+704	66+4
UK	UKI-SCOTGRID-GLASGOW	1980	0+240	0+0
TOTAL	***16 sites***	***17680***	***5028 + 6779***	***15908 + 6892***

arcControlTower

- Unique ARC features, not possible with (true) pilots
  - Walltime, memory, installed software (RTE) brokering
  - Dynamic Atlas GIIS, automatic cluster brokering/submission
  - Data transfer performed on ARC CE frontend
  - Cache + ACIX index
- aCT: Intermediate submission system
  - Autopilot factory
  - Pilot functionality: payload, panda callbacks
  - Submission of fully defined jobs
  - ARC job control
  - Data transfer + registration
- ARC pseudo pilot jobs on WN: no middleware required on WNs
  - All files available locally
  - Pilot in local mode: pure job execution

Production last 6 months



cluster	defined	assigned	waiting	activated	running	holding	transferring	success	failure	efficiency
pikolit.ijs.si	0	0	0	0	2149	17	164	435864	20508	95.5%
ce01.titan.uio.no	0	0	0	0	615	0	13	138628	6986	95.2%
jeannedarc.hpc2n.umu.se	0	0	0	0	403	7	12	122447	8993	93.2%
grad.uppmax.uu.se	0	0	0	0	101	0	0	118228	4536	96.3%
arc-ce01.pdc.kth.se	0	0	0	0	23	9	1	81658	3675	95.7%
arc-ce.smokerings.nsc.liu.se	0	0	0	0	161	0	5	83895	1354	98.4%
siri.lunarc.lu.se	0	0	0	0	503	7	3	74505	3179	95.9%
gateway01.dcsc.ku.dk	0	0	0	0	707	52	73	63307	432	99.3%
nordugrid.unibe.ch	0	0	0	0	552	22	17	46122	1947	95.9%
arc01.lcg.cscs.ch	0	0	0	0	32	0	0	45287	1982	95.8%
grid03.unige.ch	0	0	0	0	577	0	43	30704	2273	93.1%
lheppc50.unibe.ch	0	0	0	0	4	2	0	0	0	-
condor.titan.uio.no	0	0	0	0	6	0	0	0	0	-
total	*0*	*0*	*0*	*0*	*5833*	*116*	*331*	*1240645*	*55865*	*95.7%*

Average efficiency: 95.7%

Failures mostly due to Athena crashes (validation of new releases, reconstruction failures, …)

Cluster running only stable MC simulation: 99.3% efficiency

gLite/ARC in Atlas

- gLite:
  - True pilot jobs: scheduling controlled by PanDA -> fast, controlled job throughput
  - High failure rate (20%) of gLite dummy jobs, hidden to Atlas monitoring
  - Still using WMS, CREAM is in preparation
  - Submission to one site (cluster), one or several queues
  - Strict SE/CE mapping: jobs must have inputs on local SE
  - All remote transfers done by DDM
- ARC
  - Two virtual sites submitting to all the clusters (one queue per cluster)
  - Direct job submission: additional waiting time in the batch system
  - High reliability, better core efficiency
  - Suitable for custom tasks (large memory, long walltime, selected architecture)
  - Some difficulties to fit the current PanDA organization, monitoring not transparent to shifters
  - Remote transfers done by CE from/to NDGF-T1 dCache: fast transfers required
  - Clusters with slow transfer limited to low I/O jobs
  - Analysis throughput is fast when inputs are pre-cached on CE, slower for fresh data (80% of jobs have inputs pre-cached on average)
  - Automatic job recovery with downtimes

How does ARC CE fit Atlas

- Several compromises had to be done to fit current rigid job workflow and data model
- Custom submitter developed, still some missing functionality (software autoinstallation)
- Some jobs do not run yet (TAG analysis with on-demand inputs)
- Transfers done by ARC middleware, in some cases it is quicker to add new functionality in pilot code
- In general, very good performance thanks to ARC and NDGF dCache
  - cluster downtimes do not stop ND cloud job processing
  - dCache pool (site) downtimes do not stop jobs due to recent file auto-replication
  - Critical service downtimes are short
- Several non-standard OSes used: Ubuntu, Gentoo (64-bit from 2004)
- Many batch systems in use: torque, SGE, LL, SLURM, easy
- Critical middleware bugs solved promptly, requested features implemented quickly
- Testing new releases requires only one server

Difficulties

- Traditional model does not scale
  - Amount of data grows, storage space limited, data replication will be reduced: many smaller sites will be able to run only a limited subset of analysis tasks
  - Large high-priority tasks need several experts to speed-up the job distribution and data migration
  - Clusters are far from uniform (expected) performance: not enough to count cores to distribute jobs
- Still too much human intervention needed in case of:
  - High-failure rate
  - Data distribution, space management
  - High-priority jobs (MC validation, reprocessing)
  - Error recovery (data corruption)
  - Network failures (partially working transfers)

Plans for the future

- ARC CE for T3 (T2) analysis
  - Cache-aware brokering in PanDA: run on site with inputs in ARC cache
  - If not, pre-staging job to selected cluster: fill the ARC cache with inputs, downloads from storage within the same cloud (preferred sites)
  - Real job activated after pre-staging job is completed: inputs are pre-cached
  - First attempt to relax the strict data model

Missing parts

- gLite + ARC distributions not compatible on the same host: standalone ARC client on gLite installation
- Finalize the instructions to setup ARC CE
- Provide the documentation
  - to fully include ARC CE site within EGI infrastructure
  - for gLite (OSG) experts to quickly and better understand the ARC CE model

More missing parts

- Compatible software repositories
- Unified clients: tools, submission interface
- Job priorities, advanced reservations
- Transparent data management
  - Index, metadata
  - Data migration
  - Error recovery
  - Replication
- Efficient use of resources
  - Cores not used since data not available
  - Lack of free space due to over-replication
- Last but not least, clusters are expensive and making them useless by canceling support for older systems is a bad idea, especially for sites with limited funding. Instead, middleware should gain a broader platform support.

The final goal…

- Users do not care where jobs run as long as they run quickly and reliably, and where the files are stored as long as they can be retrieved fast
- Many Atlas users moved their analysis to grid since the (Atlas) interface is far simpler and much quicker than anything else
- The goal of EGI should be to provide the basis for the above to be true for everybody else

Conclusions

- ARC on Nordugrid/NDGF is a stable Atlas resource provider for more than 5 years
- A cluster with low resources can significantly contribute to Atlas production
- ARC CE can provide a significant role in overcoming the rigidity of the current production and data model