Commissioning the CERN IT Agile Infrastructure with experiment workloads

Ramón Medrano Llamas
IT-SDC-OL

14.10.2013

Agenda

- The Agile Infrastructure
- Workload Management Systems
- Dynamic provisioning
- Conclusions

The Agile Infrastructure

- Private IaaS cloud
- OpenStack based
- *Federates* Meyrin and Wigner
- 15,000 hypervisors by 2015
- 300,000 VMs by 2015
- Configuration management tools
  - Puppet, Foreman

Workload Management Systems

- Pilot based systems
- ATLAS: PanDA
- CMS: glidein WMS
- Both have an HTCondor backend
- Using Nova and EC2 APIs

PanDA integration

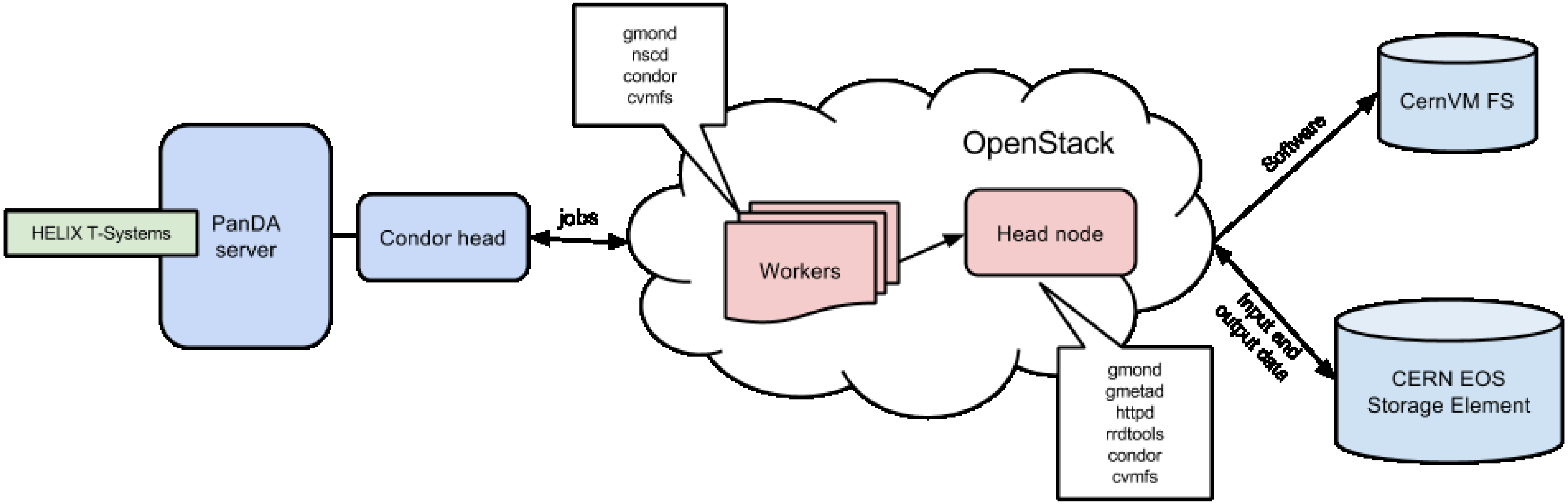
- Manual HTCondor cluster deployment
- Long lived worker nodes
- Condor for pilot submission
- CVMFS + EOS

- Same setup on HLT, Helix Nebula, Rackspace…

glidein integration

- Dynamic cluster deployment (EC2)
- Worker node automatically managed
- Condor for batch orchestration
- CVMFS + EOS

Deployment


gmond
nscd
condor
cvmfs
OpenStack
HELIX T-Systems
PanDA server
Condor head
jobs
Workers
Head node
Software
CernVM FS
Input and output data
CERN EOS Storage Element
gmond
gmetad
httpd
rrdtools
condor
cvmfs

Support from the AI Team

- Got 1,600 cores from the OpenStack team
  - Testing Essex, Folsom, Grizzly
- Complete freedom to access resources ☺
- Consultancy at any time
- Rapid bug report-solution cycle

Testing strategy

- Standard HammerCloud benchmark
- Compared with other clouds, bare metal
- 690,000 ATLAS jobs
- 337,000 CMS jobs

Testing summary

ATLAS	
Scheduler	PanDA
Cluster management	Static
Cluster size	770 cores
Jobs submitted	694,698
Failure rate	9.95%
Job type	Simulation
Typical job duration	31 min.
Duration variance	17.8 min.
Most common error	Failed to read LFC

CMS	
Scheduler	glideinWMS
Cluster management	Dynamic
Cluster size	200 cores
Jobs submitted	337,080
Failure rate	0.31%
Job type	Simulation
Typical job duration	9 min.
Duration variance	4.8 min.
Most common error	App. Error 8020

Performance testing: ATLAS, Ibex

Site	Wallclock (s)	CPU efficiency(%)	Failure rate (%)
OPENSTACK_CLOUD	3,114	78.8	2.1
BNL_CLOUD	1,505	80.2	-
IAAS	1,539	61.5	-
CERN-PROD	1,540	78.5	-
BNL_CVMFS_1	1,660	67.5	-

- Tested Late 2012
- Over commission of resources
  - CPU efficiency not reliable in this context

Performance testing: ATLAS, Grizzly

Site	Wallclock (s)	CPU efficiency(%)	Failure rate (%)
OPENSTACK_CLOUD	1,827	82.3	13.7
BNL_CLOUD	1,960	69.9	-
IAAS	1,417	67.5	-
CERN-PROD	1,499	82.3	-
BNL_CVMFS_1	1,611	72.6	-

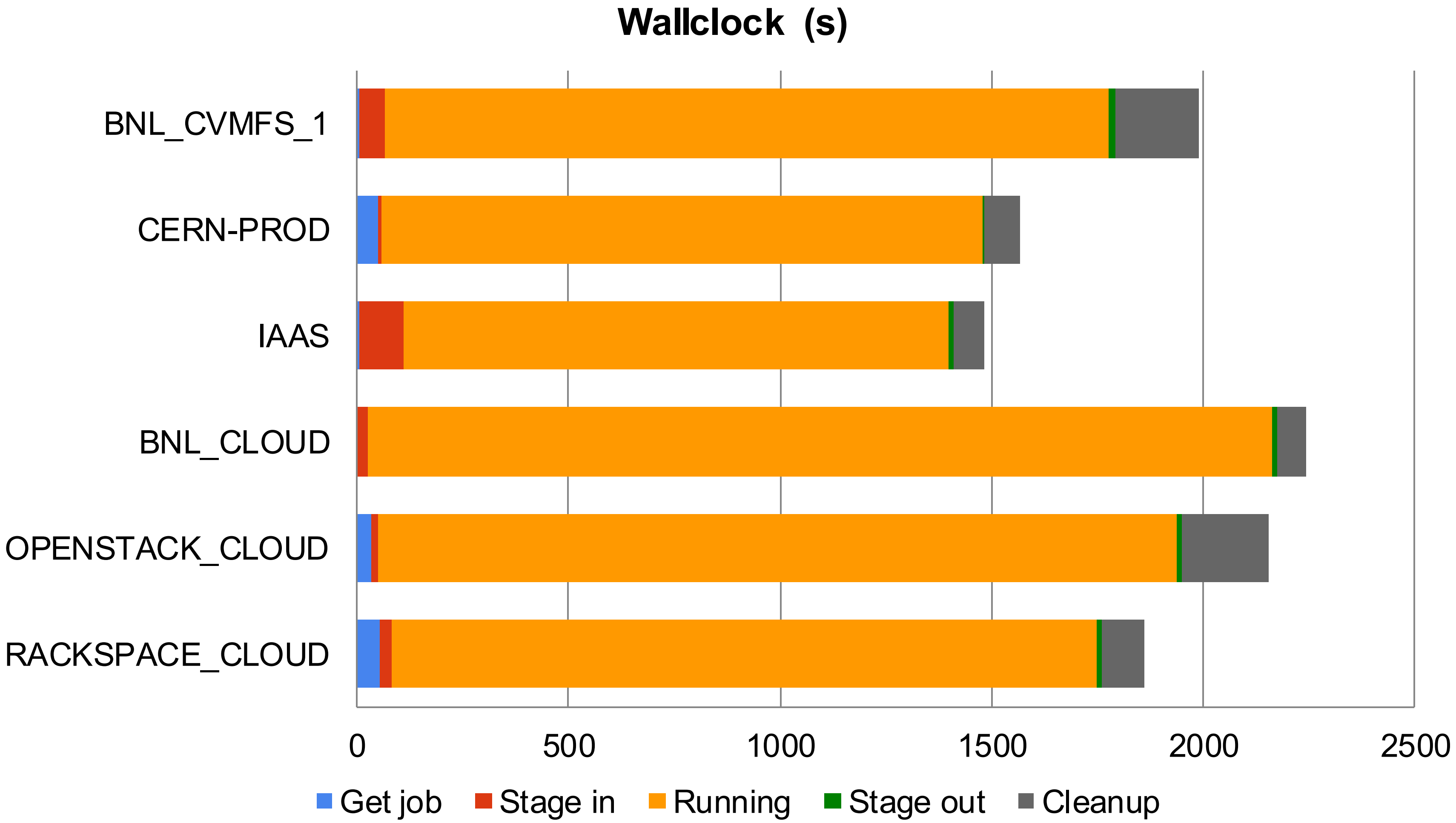
- Tested Late 2013
- Good improvement in performance
  - And predictability

Performance testing: CMS, Ibex

Site	Wallclock (s)	CPU efficiency(%)	Failure rate (%)
T2_CH_CERN_AI	616	91.1	0.0
T2_CH_CERN	914	82.8	-
T1_US_FNAL	742	91.8	-
T1_DE_KIT	783	91.6	

- Tested Mid 2013
- Reliability is incredible good

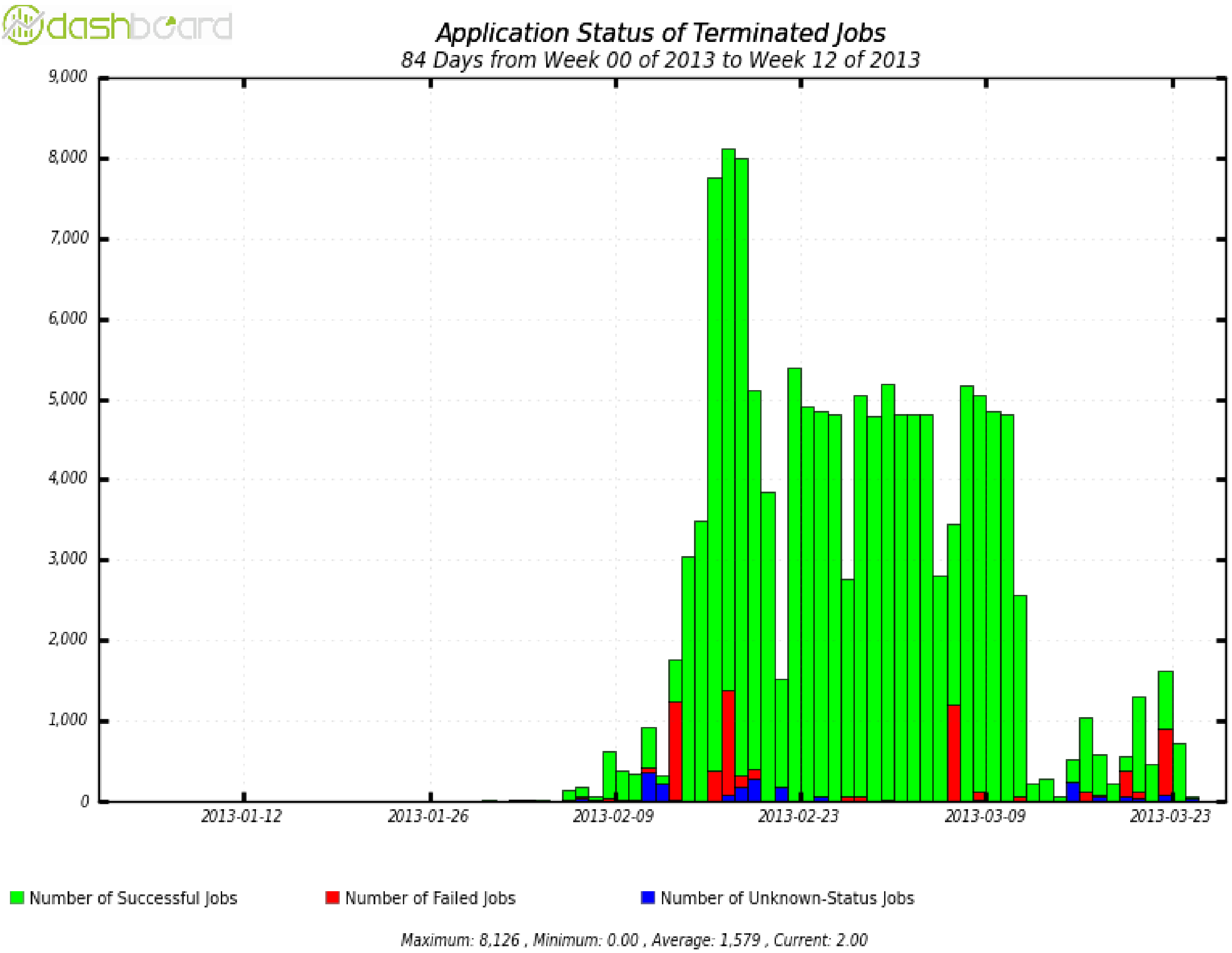
Performance testing: ATLAS wallclock



Reliability testing

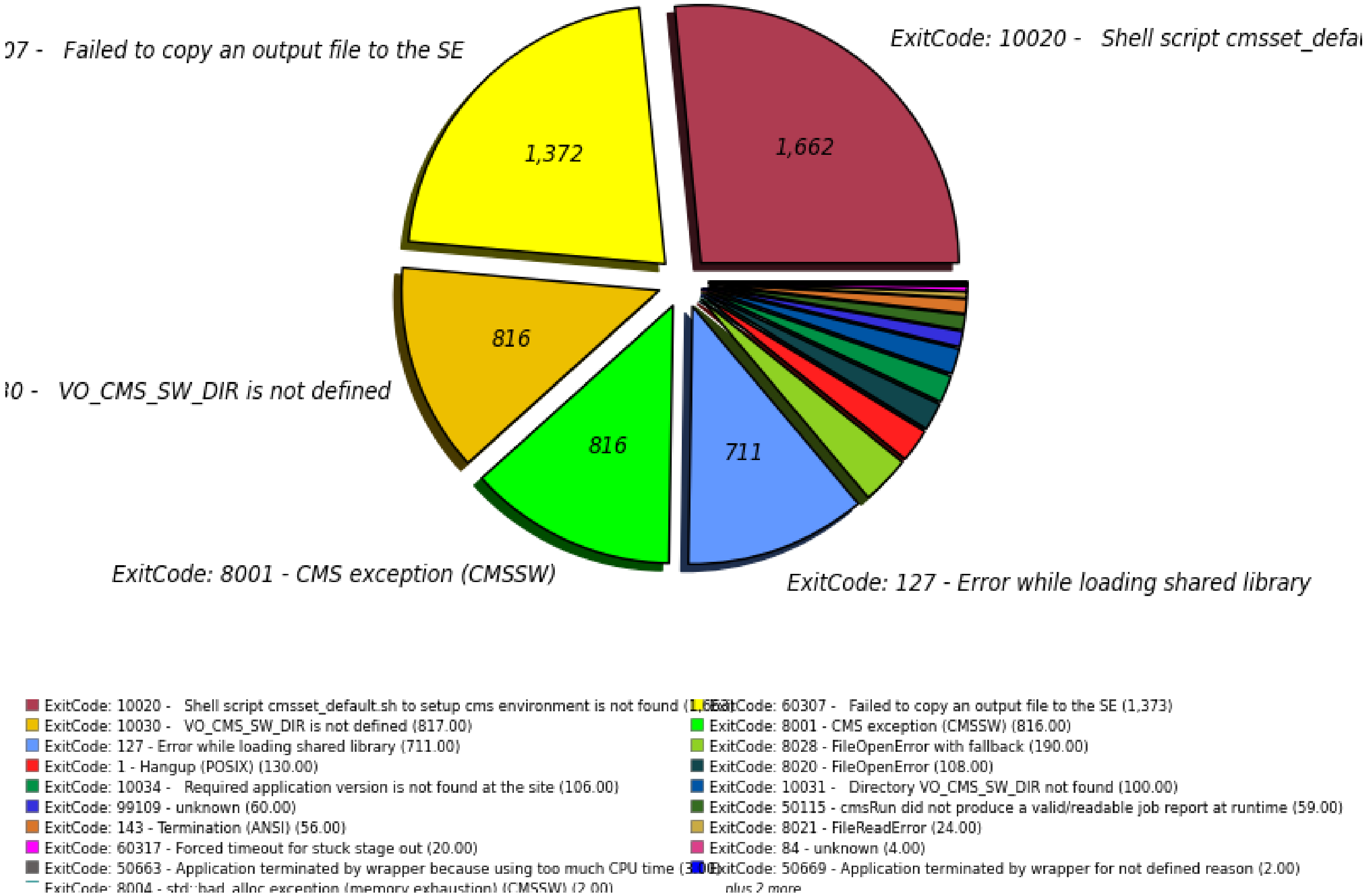
- Few failures
- Infrastructure vs. WMS failure
  - Need new monitoring techniques
  - Difficult to measure with state of the art tools

Reliability testing

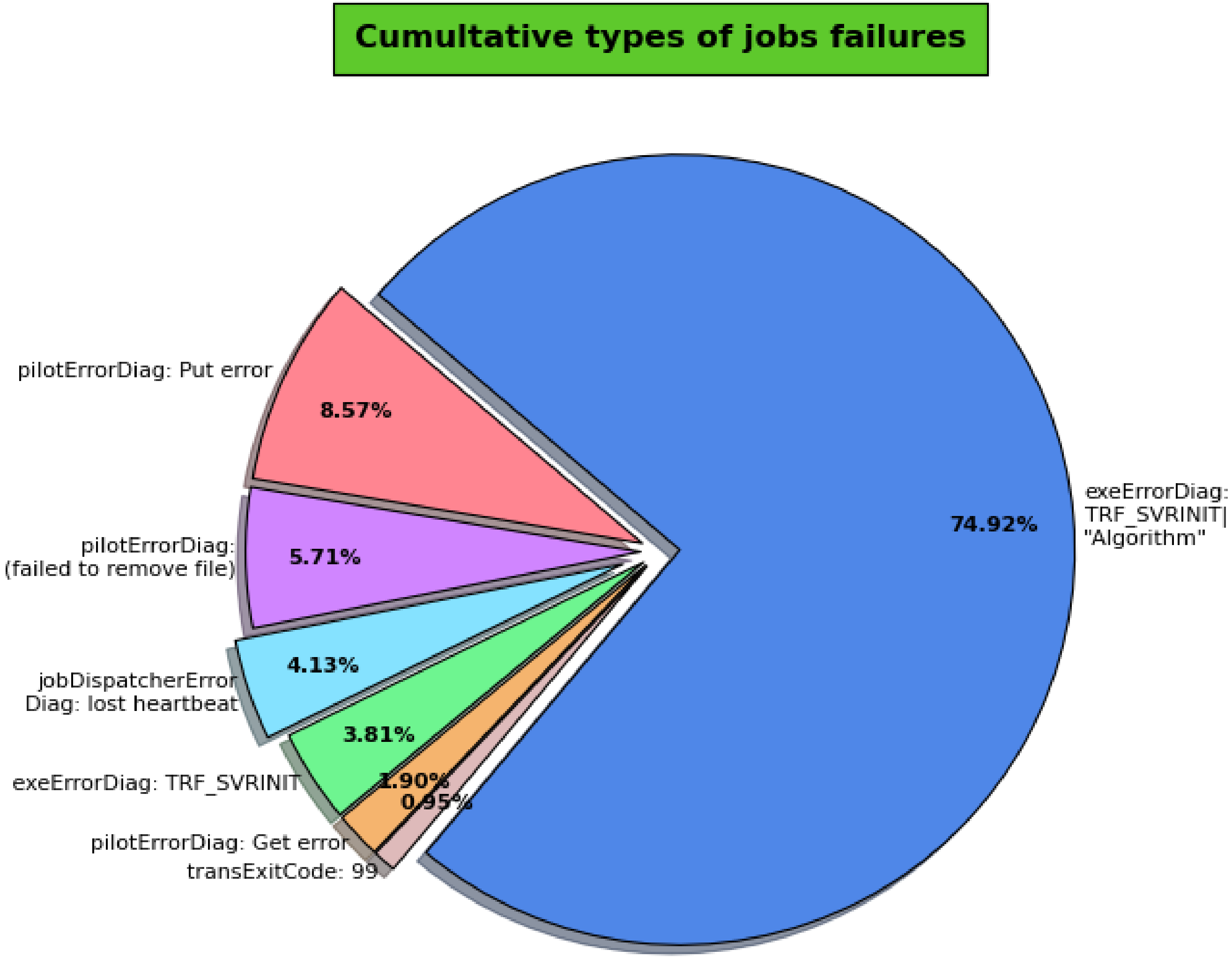


Reliability testing

Cumulative failed jobs exit codes



Reliability testing



Dynamic provisioning

- gLidein scales clusters automatically,
  - Slowness with non-batch requests

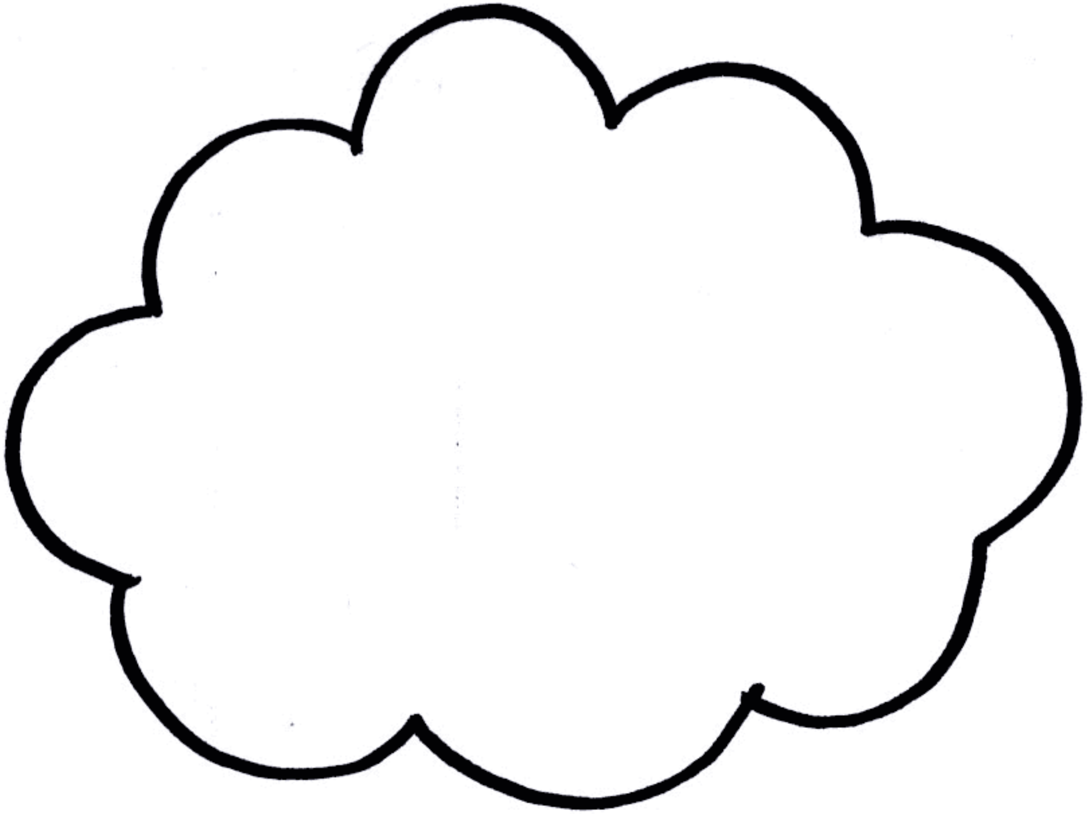
- PanDA was still not ready for it
  - Studying APF and Cloud Scheduler

Conclusions

- Being able to successfully use the infrastructure
- Scalability tests passed ☺

Future work

- Unification of image lifetime
- Federation of other OpenStack clouds
- Accounting tools for clouds
- Better understanding of failures



Questions?