Opportunistic resources: cloud providers

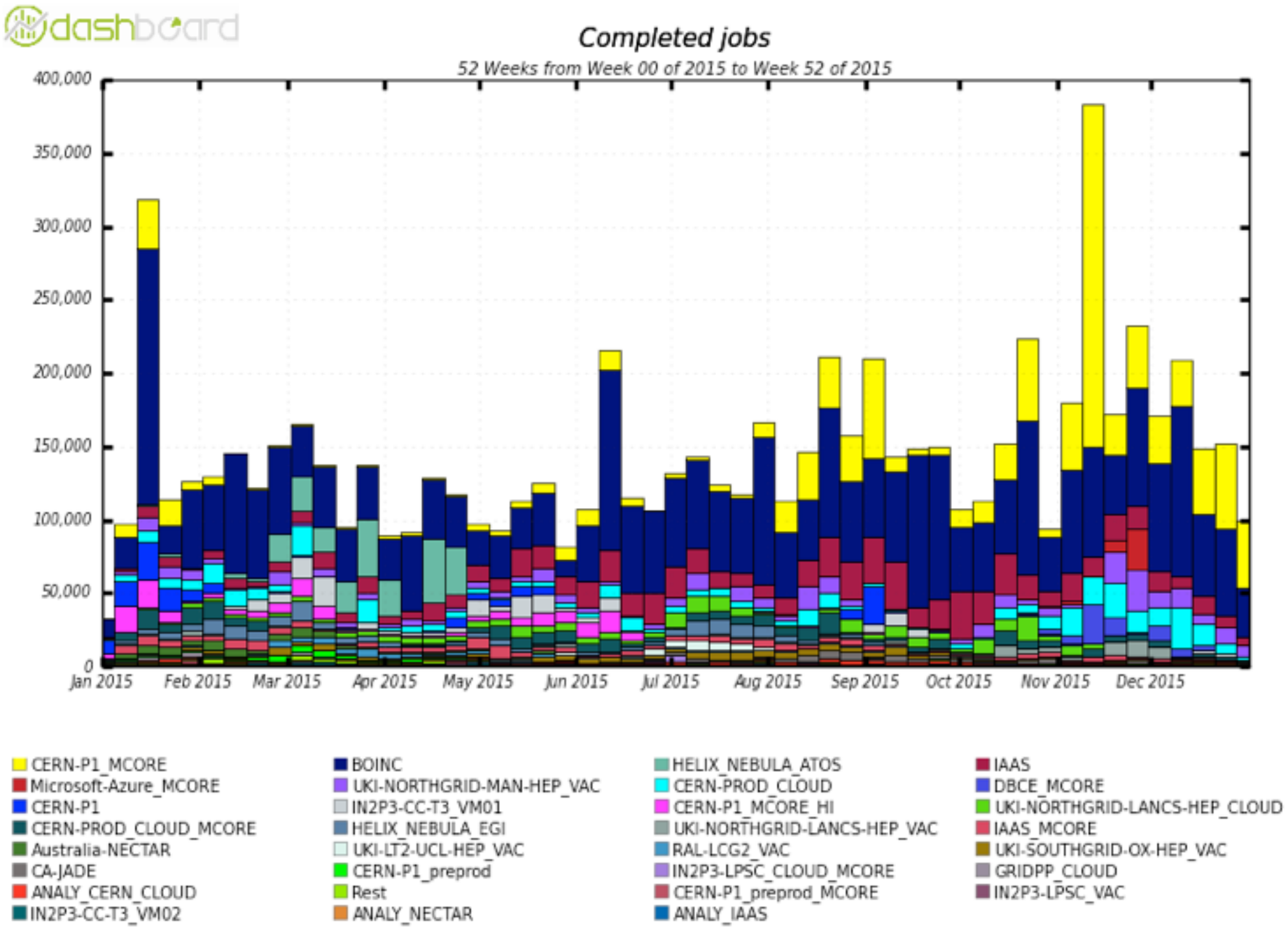
Peter Love
Lancaster University
28 Jan 2015 - CERN

Intro

- Review of cloud resources in 2015
- Challenges
- Outlook

In this context “cloud” covers VM-based resources

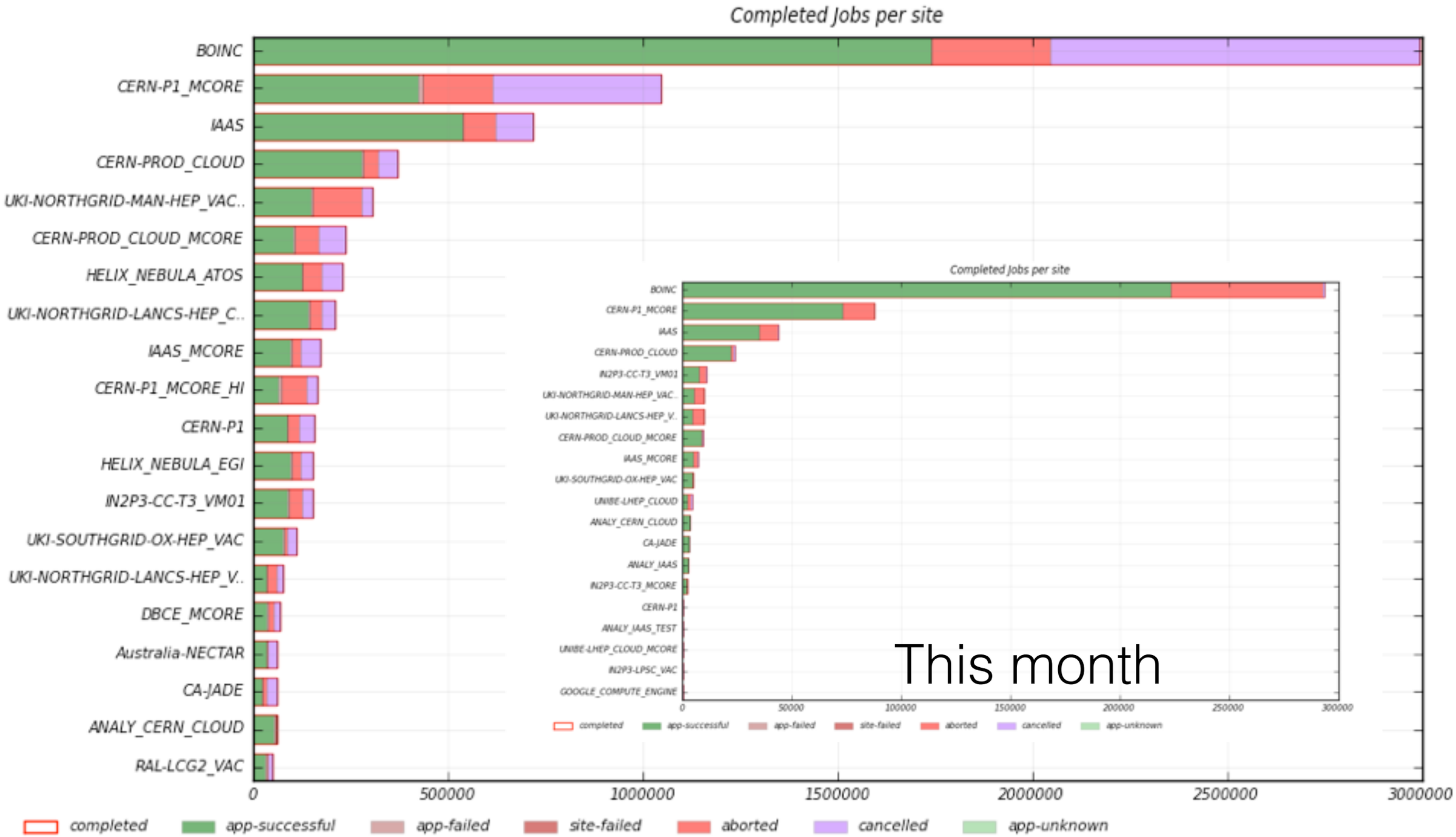
A rainbow of activity



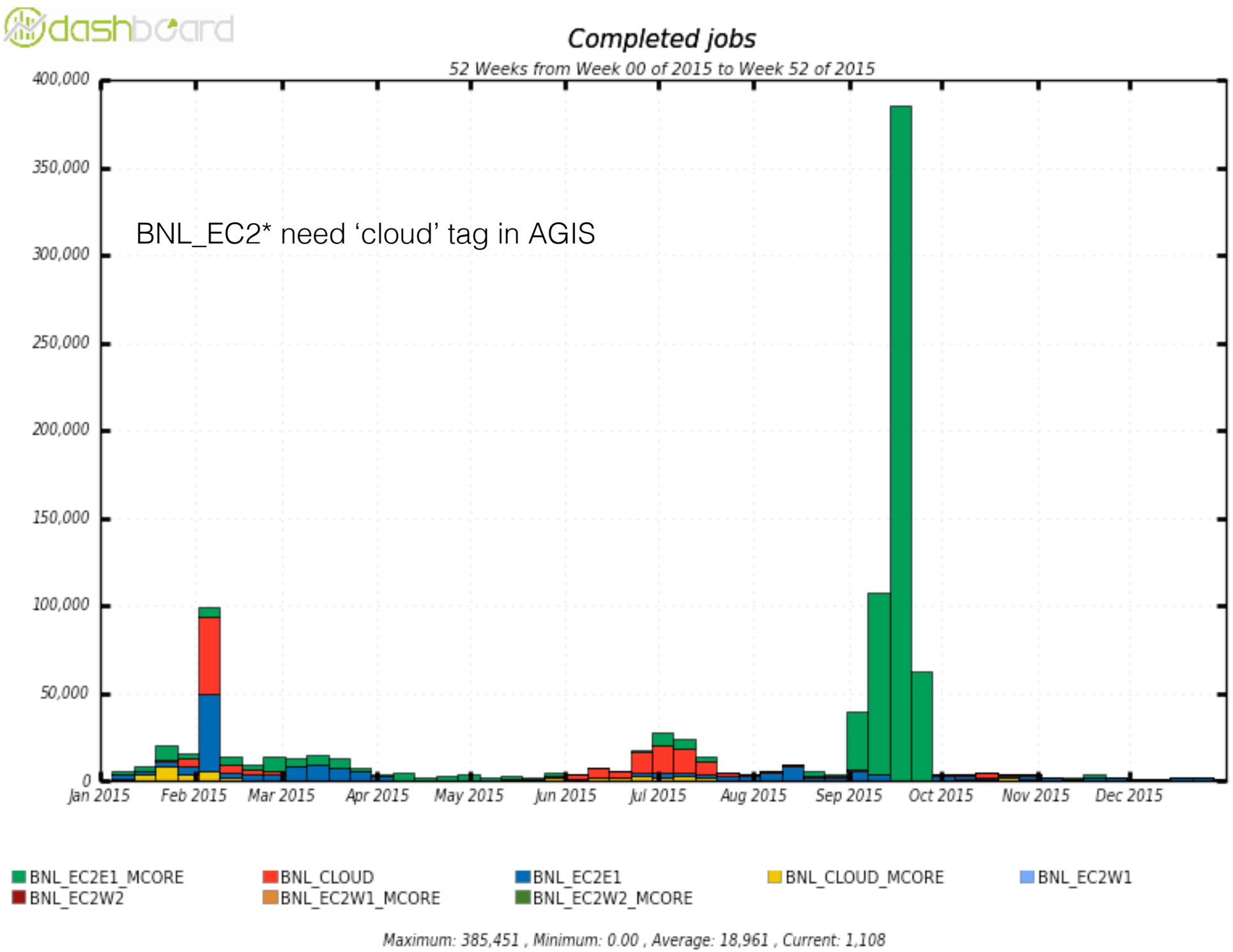
Mix of stable, ad-hoc, and experimental

- BOINC
- HLT
- Commercial
- Agile
- VAC
- Openstack

Cloud jobs in 2015



+BNL on EC2

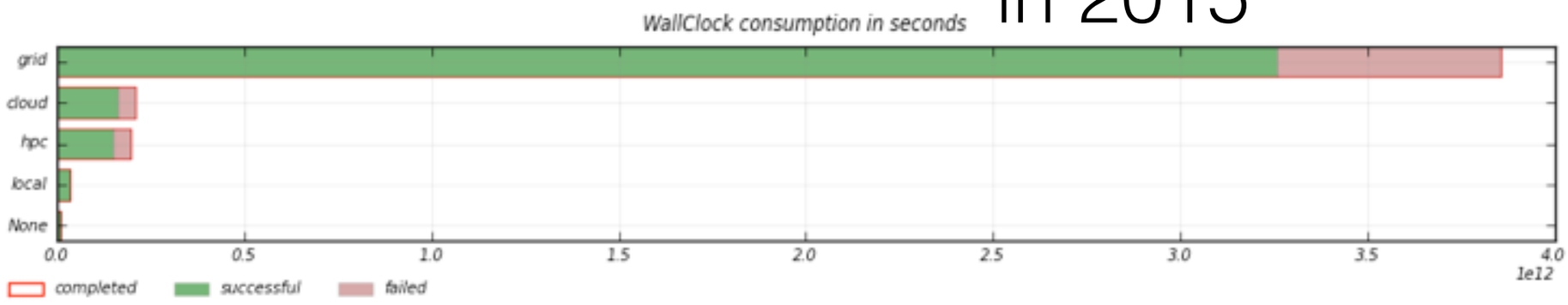


BNL_EC2* need ‘cloud’ tag in AGIS

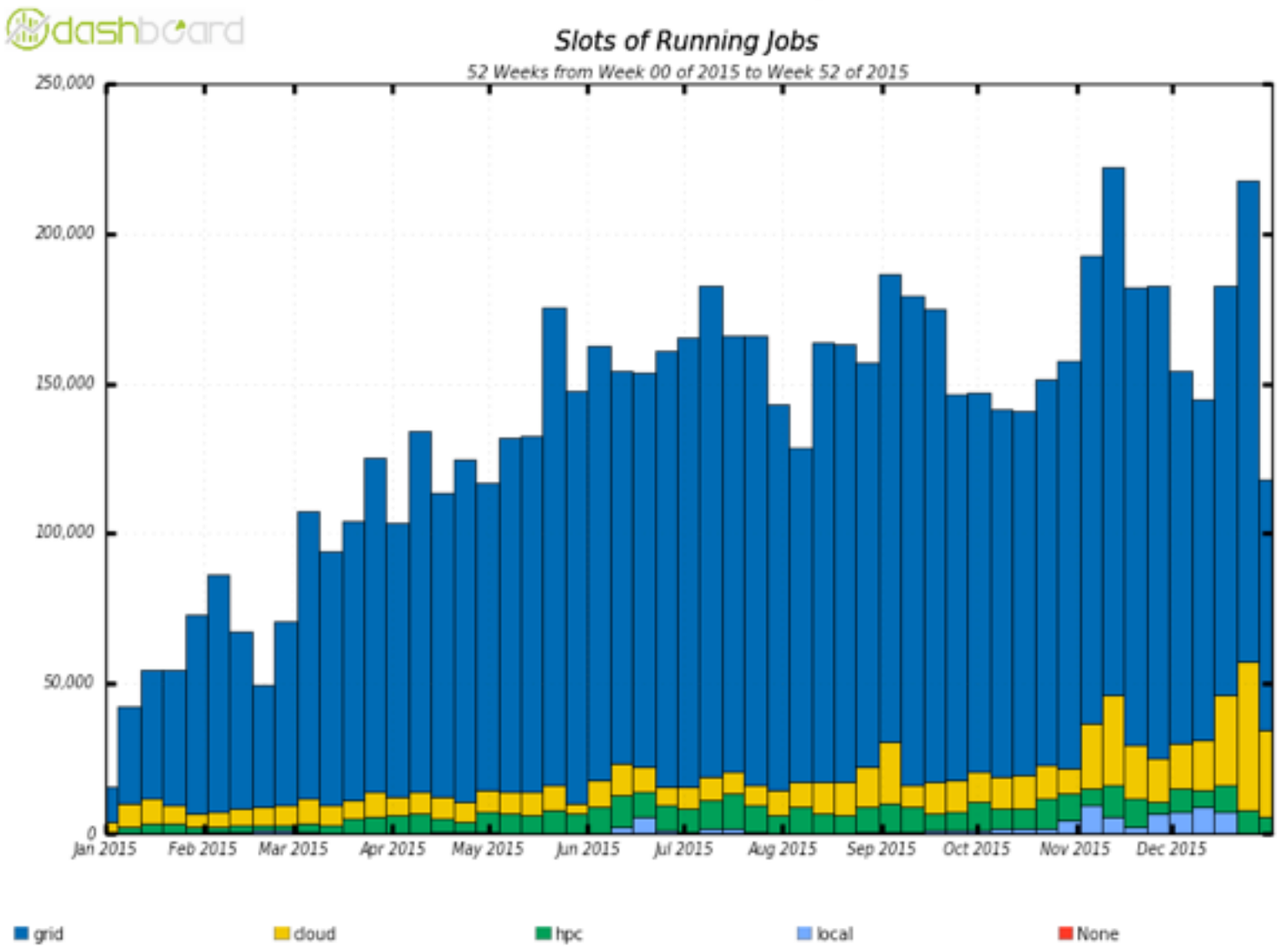
Capacities

- Commercial - $$/££/€€
- AWS - 40,000 (Feb expected100k)
- BOINC - 10,000 (growing)
- HLT - 10,000 (stable)
- IAAS - 1500 (growing +6500 this year)
- CERN AI - 1000
- IN2P3-CC - 700
- VAC - 500 (growing)

in 2015



non-Grid growth



FTE for 2015

4.5 Cloud computing Operations					1.27		
x	R. Taylor	Jan-Dec	12	35		0.35	
M	L. Field		0	0		0.00	
x	M. Paterson	Jan-Dec	12	10		0.10	
x	A. DaSilva	Jan-Dec	12	2		0.02	
x	P. Love	Jan-Dec	12	15		0.15	
x	R. Sobie	Jan-Dec	12	10		0.10	
x	C. Cordeiro	Jan-Dec	12	5		0.05	
x	E. Tal Hod	Jan-Dec	12	50		0.50	Boinc

Techno mix

- **BOINC** stable delivery and growing
- **Cloudscheduler** used for non-US activity, CERN, CA, UK
- **AutoPyFactory** used for US cloud provision
- **Vac** CPU-only sites where manpower limited
- **Vcycle** provison on commercial openstack

These are all mature technologies. Sharing VM and contextualization details. The challenges are not here.

Shared monitor http://agm.cern.ch

Challenges

- Technical problems mostly solved
- How can we get more Cloud resources?
  - **beg?** Commercial/private commissioning
  - **steal?** Spot pricing / ES <1hr
  - **borrow?** BOINC
  - pay? CERN/BNL assessing this
- Encourage sites to do all these, ATLAS will keep the resource busy.

Cloud cloud summary

- Referring to yesterday's (regional) cloud reports
    - UK: RAL objectstore developments, Datacentred commercial collaboration, Vac CPU-only for T2 evolution, preemtive ES queue
    - FR: Openstack at CC inc. devel work, using Cloudscheduler, benchmarking
    - US: AWS grants, EC2 scaling tests, Spot pricing, APF provisioning, Ceph deployment BNL/MWT2
    - DE: what benefit of running Openstack
- Also, see WLCG perspective at next week's WLCG meeting
- Role of (private) cloud facility still unclear in most regions

Commercial providers

- UK collaboration with Datacentred during 2015. Successful activity (5-10% site) but now need to pay.
- CERN procurement programme continues
- US AWS grants, what plans after all the credits have been spent?
- There must be more out there…approach a provider near you!

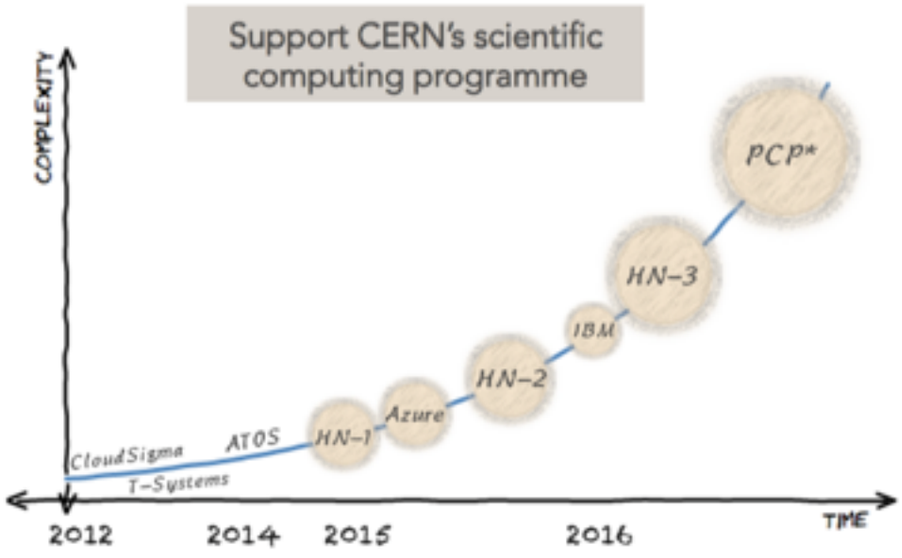
CERN activities

Slides: Cristovao

Roadmap

- ☑ *First Procurement*, March '15
- ☑ *Second Procurement*, November '15
- ☑ Microsoft Azure IaaS evaluation phase I, April '15
- ☐ IBM *Bare metal server* Evaluation (until May '16)
- ☐ *Third Procurement*
  - Production activity to start during Spring '16, Target multi VOs, full chain processing
  - Procurement phase is ongoing
  - Crucial components: Cloud Storage and Network

Past, ongoing & future commercial activities@CERN



HELIX NEBULA THE SCIENCE CLOUD

HN - Helix Nebula

- Partnership between research organization and European commercial cloud providers

* EC co-funded joint Pre-Commercial Procurement (PCP) project: https://indico.cern.ch/event/319753
** Other work has been conducted outside CERN, such as the Amazon Pilot project at BNL for ATLAS

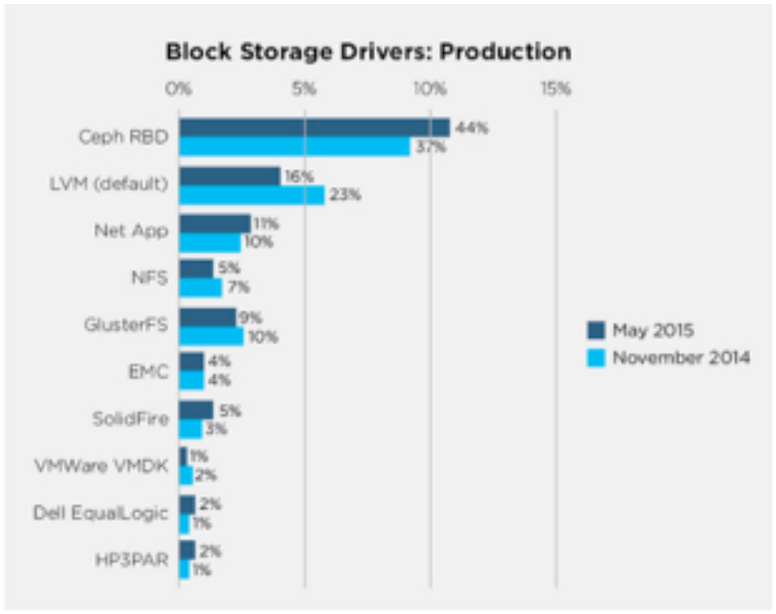
DBCE (Helix Nebula)

Summary (Slide: Jarka, Ale)

- Successfully ran MC Simul, Reconstruction and Reprocessing jobs
  - Cvmfs and network connectivity essential
- A separate site configured in AGIS
- Output to distant storage (EOS@CERN) OK
  - Output to close storage requires no dev on ATLAS side, configuration to be done in AGIS
- Job submission: using standard ATLAS workflow
  - Direct HTCondor jobs: ATLAS jobs ran OK
  - HTCondor-CE: job submission worked, *VM not configured for ATLAS so no jobs ran on ALICE VMs* ➔ need a VO-neutral VM
- Network: performance essential
  - Able to run with VMs behind NAT via HTCondor CCB

Object stores

- Expect to see more Objectstore deployment coming with Openstack deployments
- Openstack dominating private cloud deployment
- Ceph becoming preferred block storage
- http://superuser.openstack.org/articles/openstack-users-share-how-their-deployments-stack-up



Watch this space

- Things for 2016
  - EC2 100k core scale test, temporarily duplicate entire US ATLAS processing capacity
  - combined with HLT and HPC(?) to break record of concurrent running jobs (scheduled for end of February)
  - Compute Canada cloud system "GP1" at UVic coming online in 2016Q2, 6500+ cores with ATLAS getting initial access for stress/scale testing
  - Also Compute Canada plans on distributed objectstore
  - Vac ramp-up expected as UK consolidates effort
  - ES in production will prompt some questions for private clouds and preemtable VMs. Will anything change here?


Lancaster
University