UiO : Department of Physics
University of Oslo

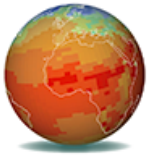
David Cameron
Riccardo Bianchi
Claire Adam Bourdarios
Andrej Filipcic
Eric Lançon
Efrat Tal Hod
Wenjing Wu
on behalf of the ATLAS Collaboration

CHEP 15, Okinawa

ATLAS@Home: Harnessing Volunteer Computing for HEP

What is volunteer computing?

Ordinary people voluntarily running scientific tasks on their PCs

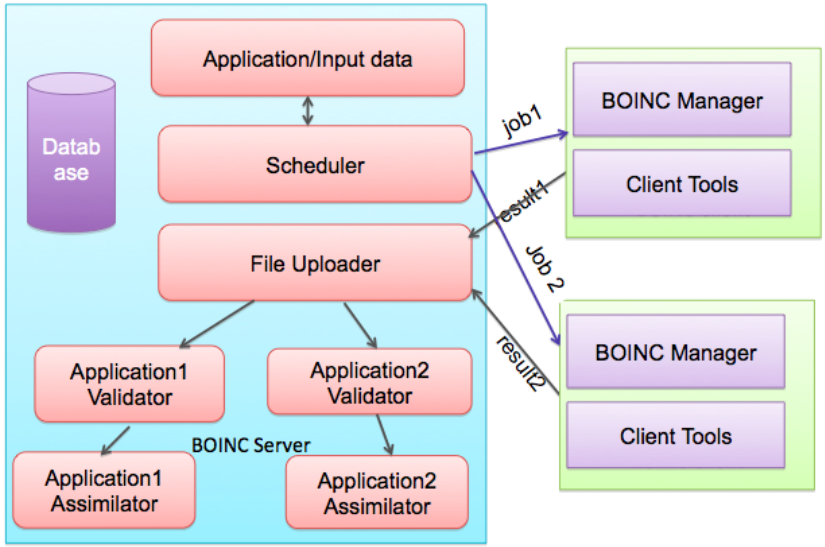


climateprediction.net
the world's largest climate modelling experiment for the 21st century

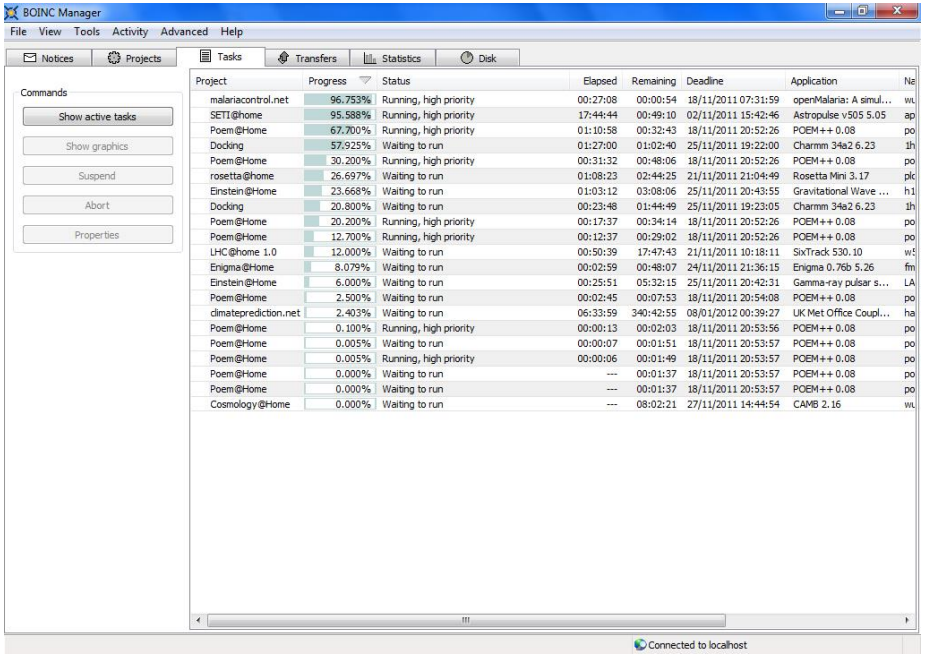
LHC@home
SixTrack

Berkeley Open Infrastructure for Network Computing (BOINC)

BOINC Architecture



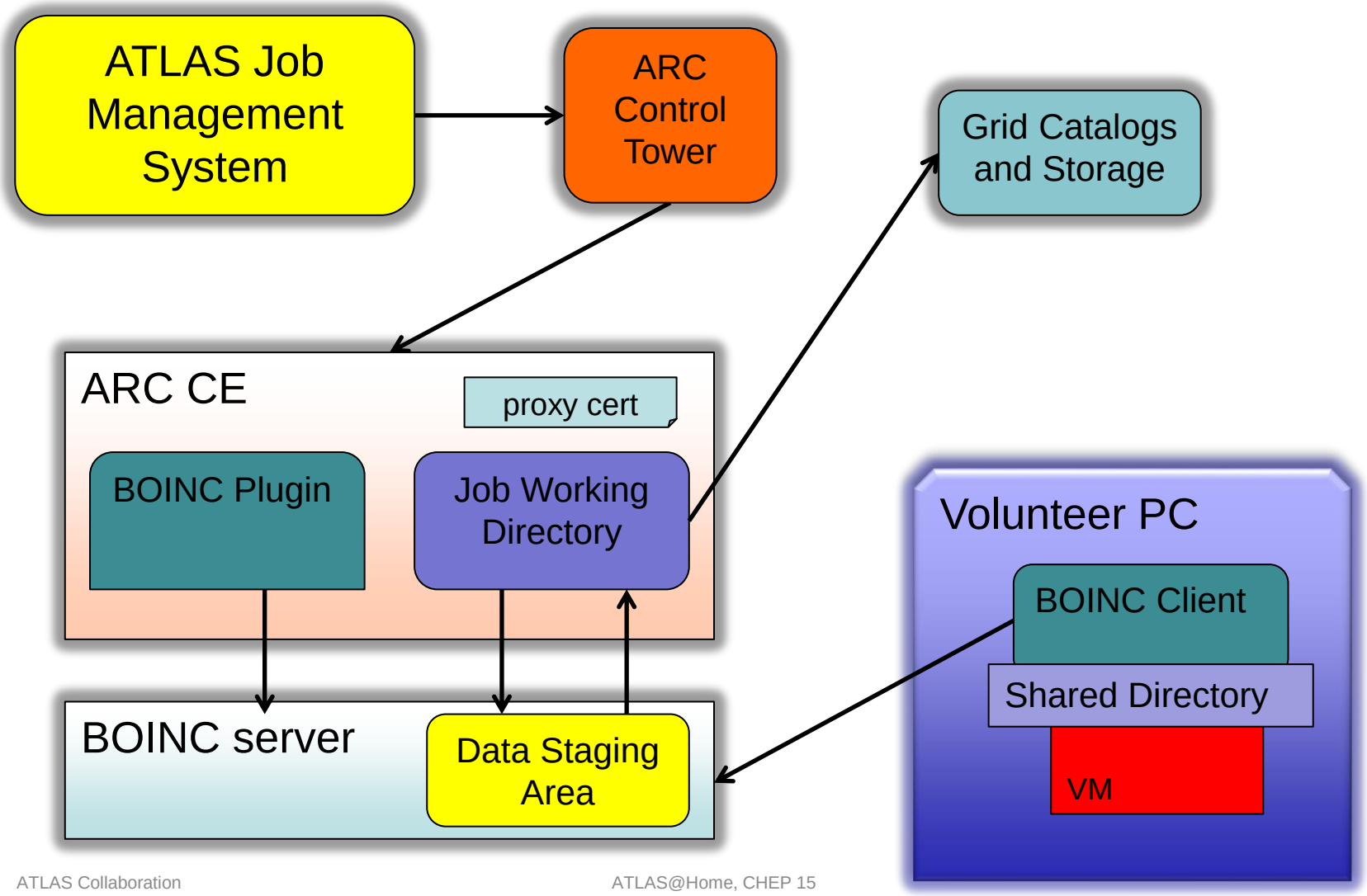
BOINC Client



ATLAS@Home

- Why use volunteer computing for ATLAS?
  - It's free! (almost)
  - Public outreach
- Considerations
  - Low priority jobs with high CPU-I/O ratio
    - Non-urgent Monte Carlo simulation or specific tasks
  - Need virtualisation for ATLAS sw environment
    - CERNVM image and CVMFS
  - No grid credentials or access on volunteer hosts
    - ARC middleware for data staging
  - The resources should look like a regular ATLAS computing resource
    - ARC Control Tower

Basic ATLAS@Home Architecture



ATLAS@Home jobs

- Real simulation tasks
  - Full athena jobs
  - 50 events/job
- Runs in CERNVM with pre-cached software
- But some data still needs to be downloaded at runtime
- Image is 1.1GB (500MB compressed) and downloaded only once
- Input files (data file + small scripts) is 1-100MB
- Output is ~100MB
- VM memory is now 2GB (was 1GB initially, but now more complex jobs)
- Jobs take from few hours up to a few days on fast (single) core

Validation of results

- Several steps of validation
  - Per work unit, that correct output is produced (just that file exists)
  - The contents of the output is verified in a later merging process
  - Physics validation comparing results to regular Grid task
- Full physics validation of the BOINC Monte-Carlo simulation was successfully verified in March 2015
- In addition to validation, 20 million top quark pairs were simulated to extend the statistics to study the top quark properties in detail.

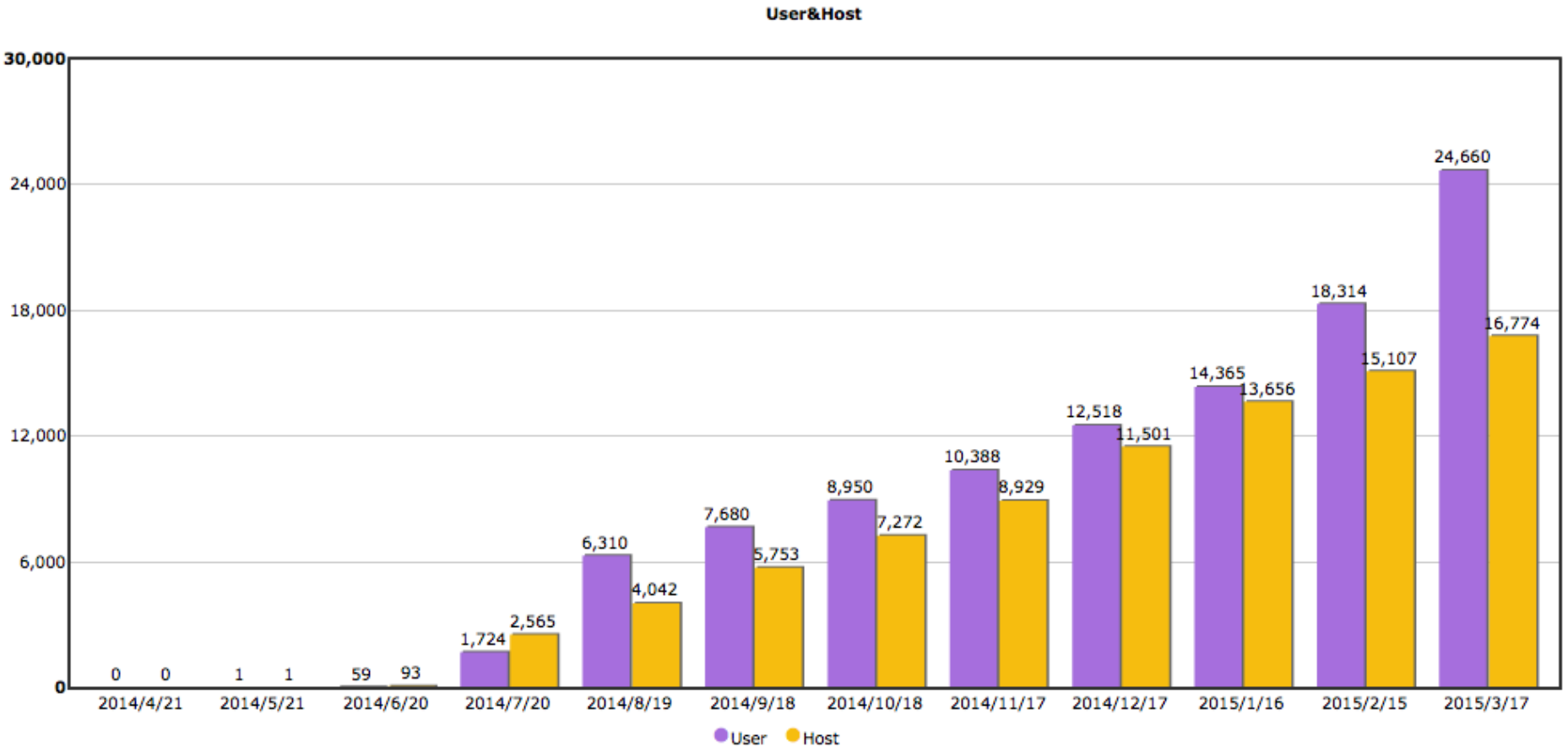
How does it work for volunteers?

- Install BOINC client and VirtualBox
  - Linux, Mac and Windows supported
  - Currently 80% of hosts have Windows
- In BOINC client choose ATLAS@Home and create an account
- That's it!
- BOINC client can be configured to run whenever is convenient, e.g.:
  - After computer is idle for 5 mins
  - Only between 5pm and 8am
- More info in backup slides

Volunteer growth

Currently 25k volunteers, 1200 active

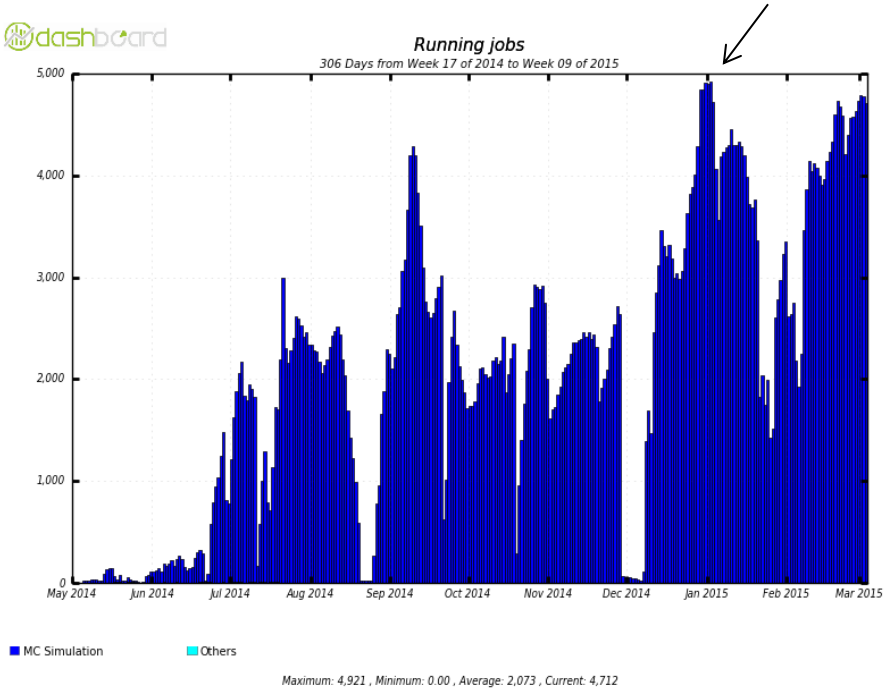
Einstein@Home: 300k volunteers, 47k active
Seti@Home: 5 million volunteers, 150k active

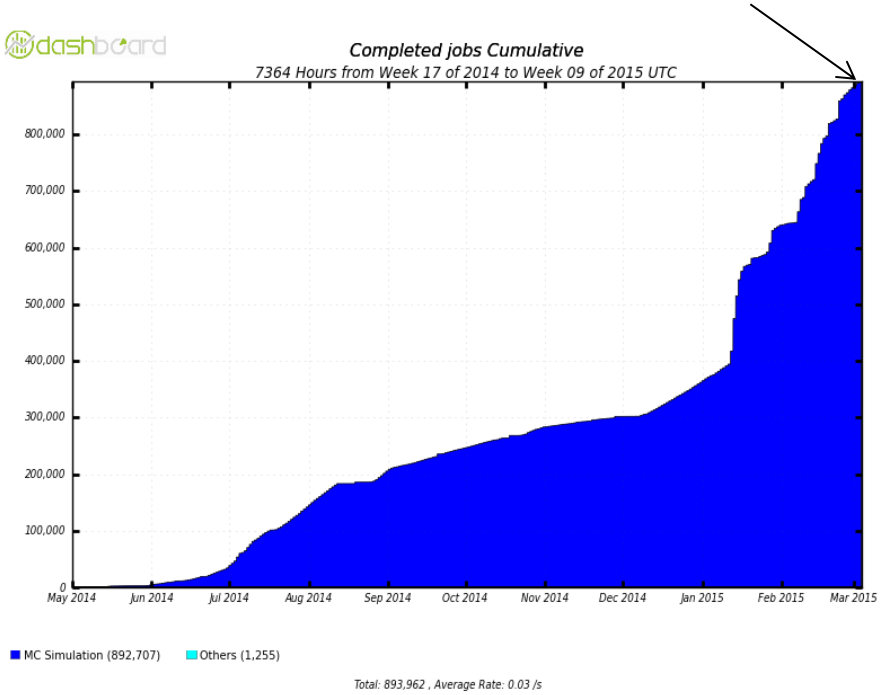


Job statistics since May 2014

5000 running jobs

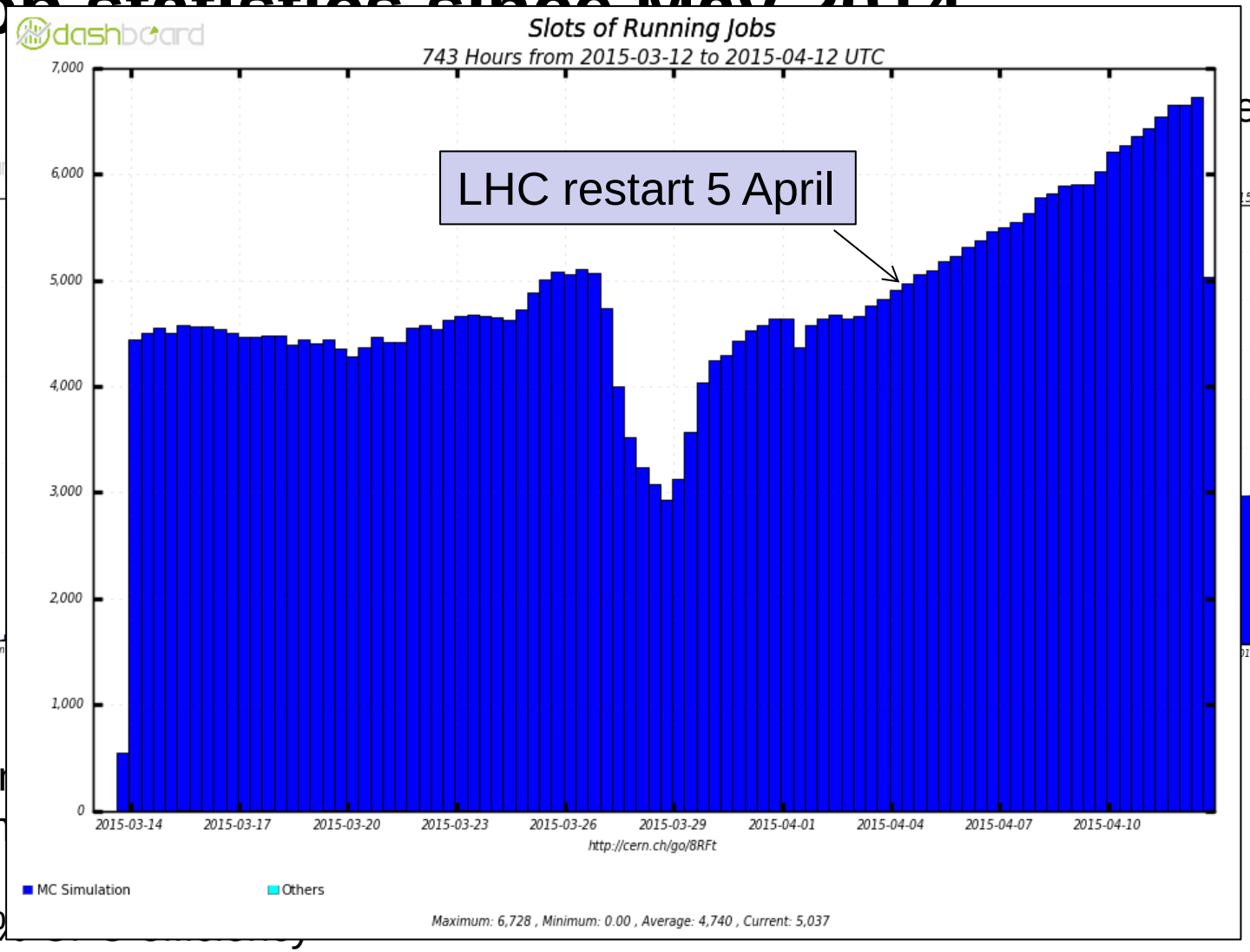
900k completed jobs

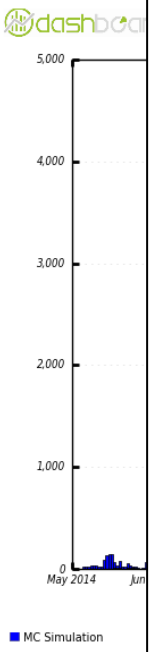




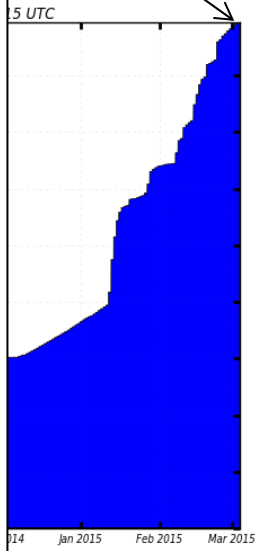
- Continuous 4000-5000 running jobs
- almost 900k completed jobs
- 1M CPU hours, 20M events
- 50% CPU efficiency
- Gaps are due to technical issues, not lack of volunteers

Job statistics since May 2014



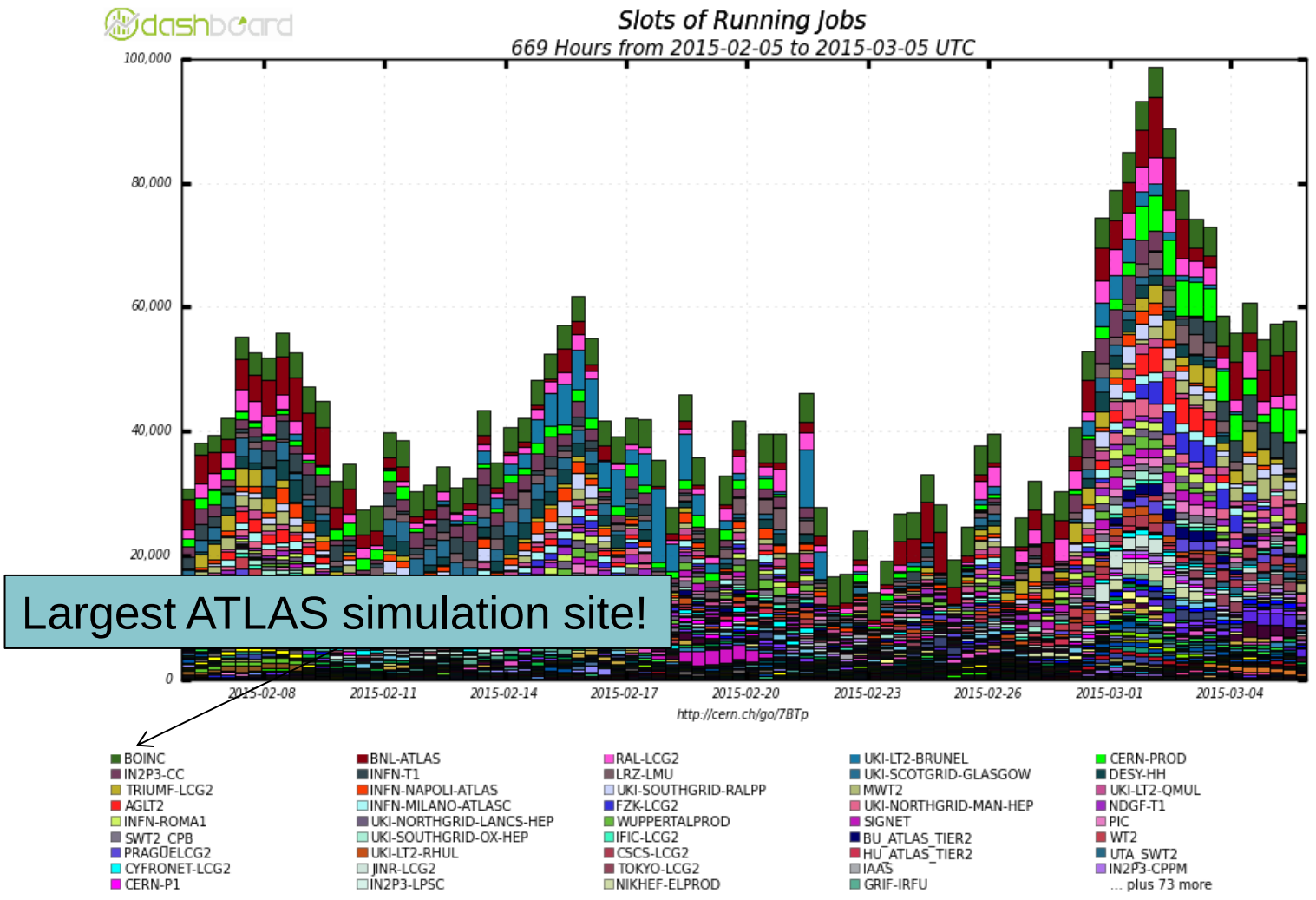


eted jobs



- Co
- alm
- 1M
- 50%
- Gaps are due to technical issues, not lack of volunteers

Scale of ATLAS@Home



Largest ATLAS simulation site!

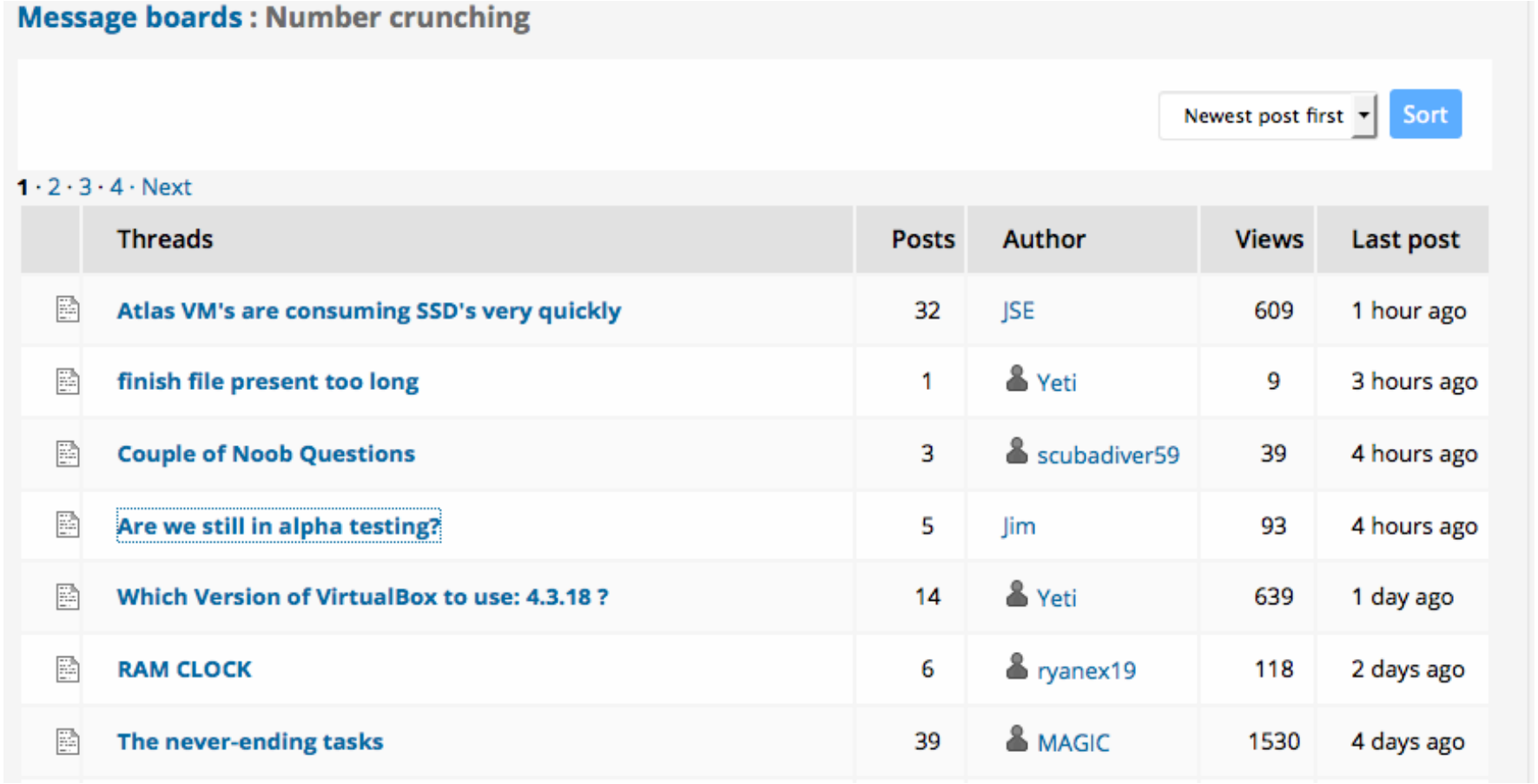
Top participants

Very roughly 3 credits/event

Rank	Name	Recent average credit	Total credit	Country	Participant since
1	David	42,771	7,877,473	Switzerland	20 Jun 2014, 11:19:25 UTC
2	JSE	16,466	354,253	Netherlands	14 Feb 2015, 13:47:00 UTC
3	Yeti	14,126	2,225,057	Germany	20 Jul 2014, 10:50:08 UTC
4	Tom*	14,064	1,607,819	United States	28 Jun 2014, 15:54:05 UTC
5	Snow Crash	11,928	719,379	United States	26 Jul 2014, 12:57:53 UTC
6	Andy_Taximan	10,075	236,941	United Kingdom	2 Feb 2015, 21:37:16 UTC
7	rbpeake	9,721	1,451,265	United States	27 Jun 2014, 23:07:23 UTC
8	Georg Stoifl	5,398	246,614	Austria	28 Oct 2014, 23:36:49 UTC
9	Mogens Dam	4,412	651,674	Denmark	1 Jul 2014, 10:02:08 UTC
10	Mumps [MM]	4,410	247,882	United States	4 Jan 2015, 2:53:29 UTC
11	Kosice@CERN	4,093	823,727	Switzerland	23 Jul 2014, 11:06:00 UTC
12	Filla	3,837	402,885	Germany	25 Jun 2014, 11:38:53 UTC
13	Toby Broom	3,803	1,112,591	United States	
14	Gary	3,487	158,507	United States	
15	MattB	3,485	293,699	United Kingdom	
16	ROG	3,448	383,429	United States	
17	Spatzthecat	2,843	392,212	United Kingdom	19 Jul 2014, 16:18:12 UTC

Wide variety and distribution of volunteers

Very active message boards



Message boards : Number crunching

Newest post first | Sort

1 · 2 · 3 · 4 · Next

	Threads	Posts	Author	Views	Last post
	Atlas VM's are consuming SSD's very quickly	32	JSE	609	1 hour ago
	finish file present too long	1	Yeti	9	3 hours ago
	Couple of Noob Questions	3	scubadiver59	39	4 hours ago
	Are we still in alpha testing?	5	Jim	93	4 hours ago
	Which Version of VirtualBox to use: 4.3.18 ?	14	Yeti	639	1 day ago
	RAM CLOCK	6	ryanex19	118	2 days ago
	The never-ending tasks	39	MAGIC	1530	4 days ago

ATLAS@Home potential

- It is not possible to run all ATLAS jobs on ATLAS@home
  - See earlier considerations about I/O, unreliability etc
- But ~50% of jobs could feasibly run on this platform
  - Event generation, MC simulation, other non-data intensive tasks
- The high entry barrier may limit general public participation for now
- Can it replace small Grid sites?
  - For example a CPU-only T3 site or small university cluster
  - Instead of setting up all the Grid infrastructure just install BOINC on the worker nodes
  - Standard Grid accounting in APEL is provided by ARC CE
- Idle administrative desktops
  - eg now available as part of NICE, to put on CERN administrative PCs

Lessons Learned and Conclusions

- Setting up ATLAS@Home has been an exciting and fruitful experience
- Hardware is free and manpower running costs are a fraction of an equivalent-sized computing centre
- But the volunteers don't come completely for free
  - Some volunteers are extremely competent and knowledgeable and help others
  - But some expect support/feedback/rewards
- The number of running jobs is rising slowly
  - ATLAS@Home is heavy to run compared to most projects
  - Still a beta project
- BOINC developers very enthusiastic to help us
  - They give us fixes/new features in days
- We have a few more things to fix before ATLAS@Home can move out of beta
  - Adding "screensaver" visualisation of events and physics info
- There is a lot of potential as long as we keep the public interested

Acknowledgements

- Our CERN IT colleagues in LHC@Home for providing the BOINC infrastructure and storage space
- BOINC developers for rapid response to our questions and problems
- .. and please join us!

  http://atlasathome.cern.ch

  http://atlasphysathome.web.cern.ch

Backup

How to join ATLAS@Home in a few clicks

- http://boinc.berkeley.edu/download.php
- Install the package and start it
- Tools -> Add project or account manager…
- Add project
- Select ATLAS@Home (2nd on the list)
- Enter email address and a password
- Click Finish
- Optionally add information to your profile

Configure BOINC Client

- BOINC by default will run one task per core
- ATLAS@Home is too heavy for this on most PCs
- Preferences -> Computing preferences -> % of the processors
  - eg to use 1 core of a 4 core PC: