

CernVM and Volunteer Computing

Ivan D Reid

Brunel University London

Laurence Field

CERN

Volunteer Computing

- A type of distributed computing
- Origins in mid 1990s
- Computer owners donate computing capacity
 - To a cause or project
- Not necessarily only spare cycles on Desktops
 - Idle machines in data centers
 - Home clusters
- SETI@home and Folding@home
 - Launched 1999

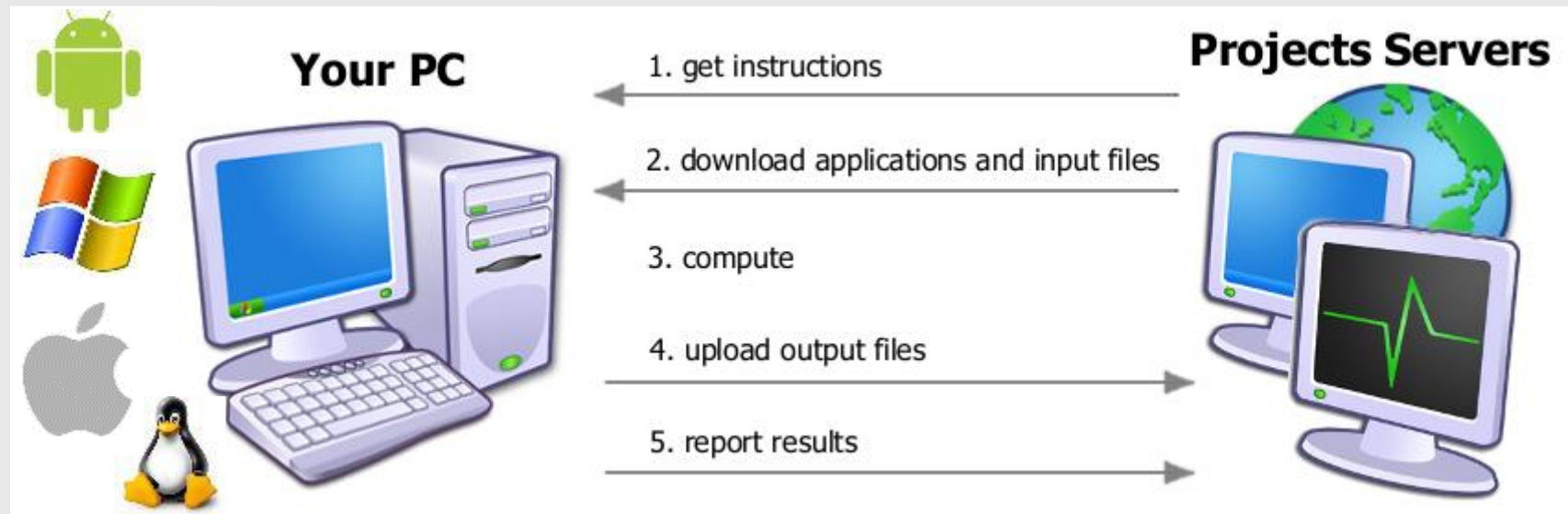
BOINC

- Berkeley Open Infrastructure for Network Computing
 - Started in 2002
 - Funded by the National Science Foundation (NSF)
 - Developed by a team based at the Space Sciences Laboratory
 - University of California, Berkeley
 - Led by David Anderson
- Provides the middleware for volunteer computing
 - Client (Mac, Windows, Linux, Android) with CLI
 - GUI
 - Application runtime system
 - Server software
 - Project Web site



Volunteer Perspective

- Download and run the BOINC client
- Choose a project
- Enter an email address and password
 - Or silent connection with a key
- Run the application and earn credit



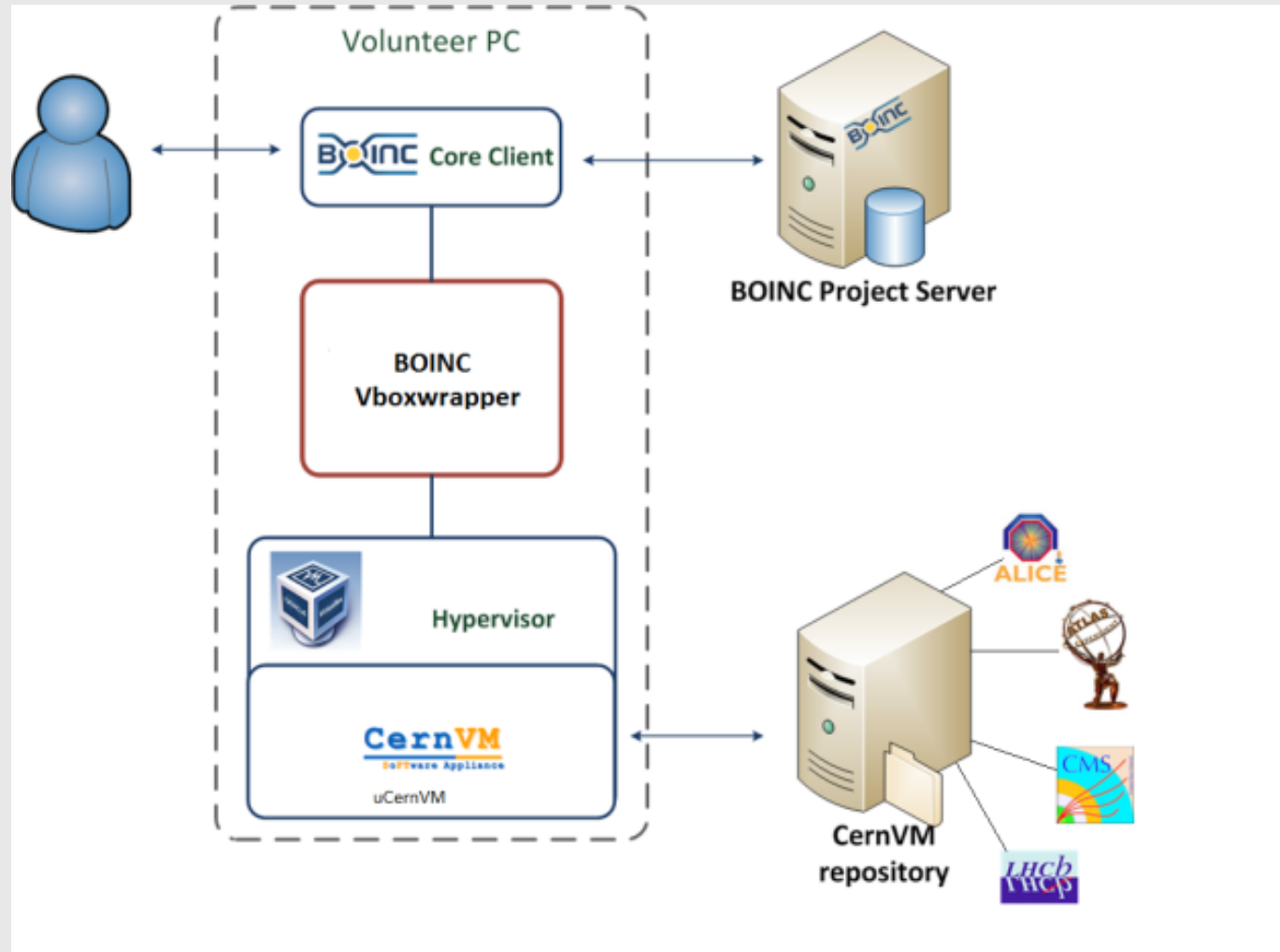
Motivation

- Free* resources
 - 100K hosts achievable for large projects
 - Actual job slot count (number of cores) is higher
 - Community engagement
 - Outreach channel
 - Explaining the purpose and value of the science
 - Participation
 - Offering people a chance to contribute
 - Engagement forms a strong bond
 - Community support
- * There are costs associated with their use

Challenges

- The cost of using the free resources
 - Initial integration requires investment
- Operations and Maintenance
 - Public facing support on all levels
 - Lowered by community supports
- Attracting and retention of volunteers
 - Advertisement and engagement
 - Communications cost for capacity building
- Low Level of Assurance
 - Anyone can register as a volunteer
 - Not the same level of trust as with Grid
- Running HEP software on Windows
 - 85% of the resources





















BOINC with Virtualization



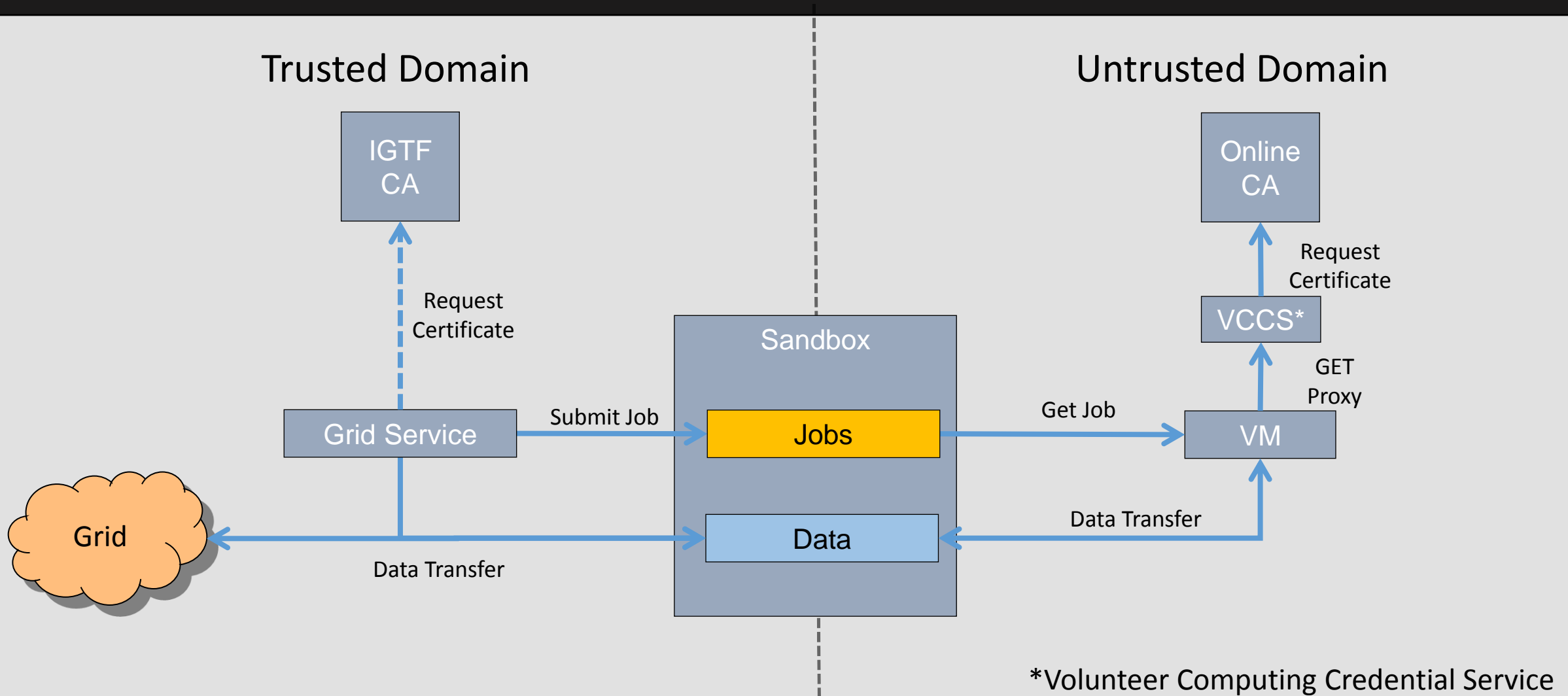
Squid Proxies

- Currently have one service
 - lhchomeproxy.cern.ch
 - Two instances
- But using CVMFS_PAC_URLS
 - <http://lhchomeproxy.cern.ch/wpad.dat>

Visitors domains/countries (Top 10) - Full list

Domains/Countries			Pages	Hits	Bandwidth	
	Network	net	16521519	16521519	1388.31 GB	
	Unknown	ip	14755964	14755964	954.11 GB	
	Switzerland	ch	13336665	13336665	764.86 GB	
	Germany	de	8491345	8491345	448.70 GB	
	United Kingdom	uk	6001174	6001174	141.45 GB	
	Commercial	com	4145887	4145887	230.47 GB	
	Canada	ca	1896722	1896722	170.49 GB	
	Italy	it	1769052	1769052	92.04 GB	
	Netherlands	nl	900438	900438	55.83 GB	
	Austria	at	759248	759248	32.87 GB	
	Others		5299347	5299347	398.90 GB	

Sandboxing and Authentication

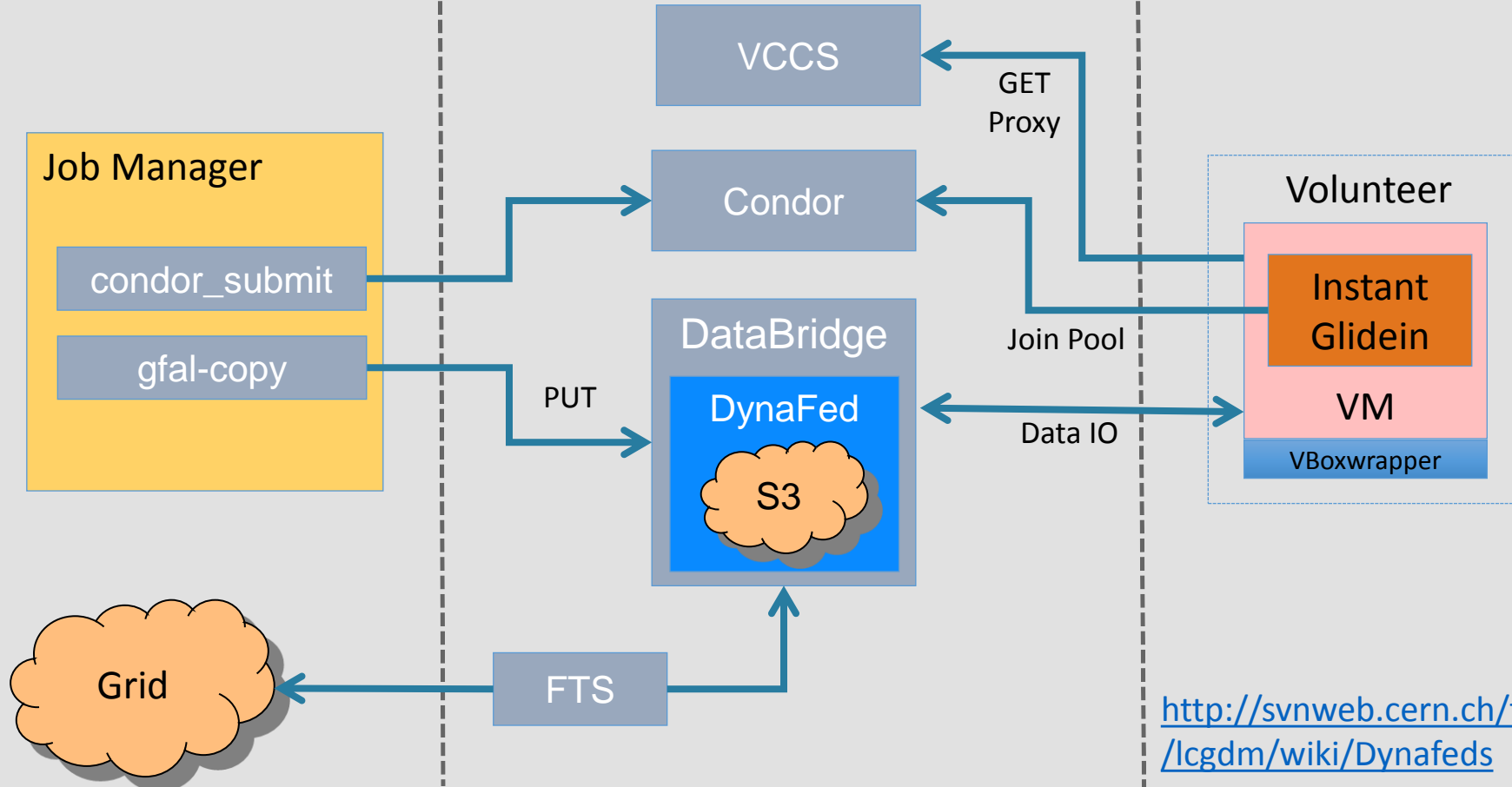


Implementation

Application Server



Common Infrastructure




<http://svnweb.cern.ch/trac/lcgdm/wiki/Dynafeds>

Common Platform

- Coordinated outreach efforts
 - Maximize the potential resource pool
 - Fair share the resources
 - Volunteers typically configure multiple projects
- Development, Maintenance and Operations
 - Share the costs
- Build upon a common approach
 - Reuse components and services
 - Provided centrally as an infrastructure
- Towards a common platform
 - BOINC
 - Web presence
 - Outreach
 - Data Bridge
 - Condor

CERN Accelerating science Sign in Directory




LHC@home

Volunteer computing for the LHC

Search

[HOME](#) [ABOUT](#) [PROJECTS](#) [JOIN US!](#) [HELP & FAQ](#) [CONTACT](#)








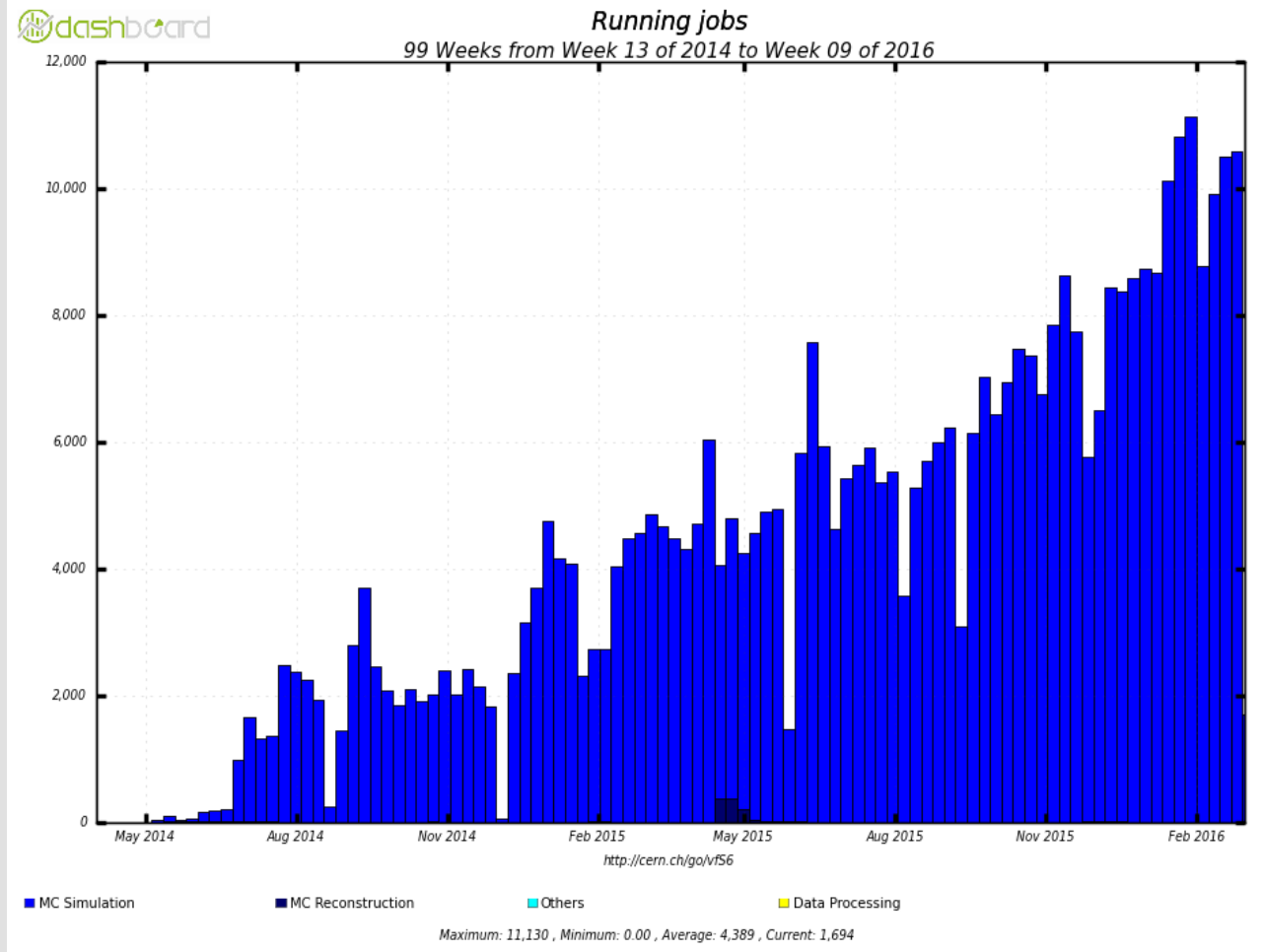
Antimatter
Exotic particles
Proton beam physics

Help CERN explore our Universe.

● ○ ○ ○ ○ ○

Applications

 <p>ATLAS @HOME</p>	<p>ATLAS@home</p> <p>"Known" physics and "new" phenomena - want to create alternative models of the universe?</p>
	<p>Beauty</p> <p>"b" is for "Beauty" - the gorgeous little particle in antimatter physics.</p>
 <p>CMS</p>	<p>CMS@Home</p> <p>CMS is on the lookout for completely new, unpredicted phenomena.</p>
	<p>SixTrack</p> <p>Help CERN accelerator engineers to run intensive simulations to check the stability of the twin proton beams circulating in the LHC machine.</p>
	<p>Test4Theory</p> <p>Simulate high-energy particle collisions and help tune the theory to the experimental results.</p>



Summary

- Volunteer Computing can and is providing
 - Significant additional computing resources
 - Potentially O(100K) machines
- Virtualization enables HEP applications
 - To run on multiple x86 platforms
 - Can therefore reach more volunteers
 - And hence resources
- Using CernVM
 - Baked VM to reduce downloading each time the VM is restarted
 - But can automatically update the image
 - Squid management and placement is an important operational concern.
- Condor is used for job management
 - The VM's join a pool and authenticate with the volunteer's x509 proxy
 - Instant glideins provide a short lived tenancy
- vLHC@home is a common platform
 - Supporting multiple applications
- Come and join the fun!
 - <http://lhcatome.web.cern.ch/join-us>