



CernVM Status Update

Jakob Blomer

PH-SFT Group Meeting
14th December 2015



① Status of CernVM

② Status of CernVM-FS



- Much of the work this year was defined by the workshop in March
 - ▶ [Indico](#)
 - ▶ [Wrap-up Slides](#)
- Feedback from *experiments* and *sites*
- Invited guest speakers from Google, Amazon, Citrix, Basho, and University of Notre Dame
 - Sharpened our view on containers (e. g. Docker)
 - Strengthened the relationship to the *Parrot* development team at Notre Dame

→ Next workshop: 6th June 2016 at RAL



- CernVM 3.4, June '15
[Scientific Linux 6.6](#), [Docker](#) and [lxc](#), [image signatures](#)
- CernVM 3.5, August '15
[Scientific Linux 6.7](#), [Microsoft Azure](#), [CloudStack](#), [glideinwms-vm](#)
- CernVM-FS 2.1.20, March '15 (~50 JIRA tickets)
[S3](#), [Stratum 1](#) [Geo-API](#), [garbage-collected repositories](#)
- CernVM-FS 2.2.0 Server Only Pre-Release, September '15
(~15 JIRA tickets) [bugfix](#) and [consolidation release](#)
- Independent [libcvmfs-stable](#) summer '15

Being finalised:

- CernVM 4 [RHEL 7 compatible](#)
- CernVM-FS 2.2.0 (~100 JIRA tickets)
[semantic versioning](#), [extended attributes](#), [data federations](#), [HPC](#)



Journal paper together with D. Thain:

*The Evolution of Global Scale Filesystems for
Scientific Software Distribution,*

IEEE Computing in Science and Engineering **17**(6), Nov/Dec 2015

[▶ Link](#)



① Status of CernVM

② Status of CernVM-FS

197.4k visits



Boot of New CernVMs

- ~28 000 new CernVMs per day
- Covering ~75 countries
- ~40% from U.K.



**CernVM: complete and portable environment
for **developing** and **running** HEP data processing tasks**



**CernVM: complete and portable environment
for **developing** and **running** HEP data processing tasks**

Use Cases

- 1 IaaS Clouds
- 2 Development Environment
- 3 Volunteer Computing
- 4 Long-Term Analysis Preservation
- 5 Outreach & Education

Infrastructure-as-a-Service Cloud

Various clouds:

- ATLAS HLT farm
- Cloud resources seamlessly integrated with experiment task queues (e. g. ATLAS CloudScheduler, LHCb VAC)
- ALICE software release testing on CERN OpenStack
- Commercial providers ("Helix Nebula")
- ...



CernVM: complete and portable environment
for **developing** and **running** HEP data processing tasks

Use Cases

- 1 IaaS Clouds
- 2 Development Environment
- 3 Volunteer Computing
- 4 Long-Term Analysis Preservation
- 5 Outreach & Education

Interactive Users through CernVM Online

The screenshot shows the CernVM Online dashboard interface. The browser address bar displays `https://cernvm-online.cern.ch/dashboard`. The user is logged in as 'ichanala'. The dashboard features a sidebar with navigation options: 'About', 'Dashboard', 'Commands', 'Dashboard', 'Pair an instance', and 'Create Context'. The main content area is titled 'Dashboard' and 'Your context definitions', displaying a table of context definitions.

Name	ID	Operations
PrivateCloud-CatalogServer	22b13425ef7244c4b7de60dbbca64728	Remove Use as template
PrivateCloud-Worker	a632e7ad3ca64774ac9318fc9e086640	Remove Use as template
PrivateCloud-MakeflowPool	f7e9ba92a55146119ac3cd6141f6d957	Remove Use as template
Private-Desktop	79383bc71d7a4760a8397cc2e8d2a2ed	Remove Use as template
PrivateCloud4ALL	37bd937803b407e8fbd469d65fac74c	Remove Use as template

~300 users in 2015, out of which ~65 external users



**CernVM: complete and portable environment
for **developing** and **running** HEP data processing tasks**

Use Cases

- 1 IaaS Clouds
- 2 Development Environment
- 3 **Volunteer Computing**
- 4 Long-Term Analysis Preservation
- 5 Outreach & Education

Test4Theory and Computing Challenge

You are now creating virtual collisions Log-in and keep track of your progress

Event Rate	Progress	Jobs Completed	Ranking
How many events (particle collisions) your virtual machine is simulating per minute. 1,000 1,000 500	What fraction of the current job is completed. (Typically, 1 job = 100,000 events) 1 / 70 16 % 45 %	How many simulation jobs your virtual machine has processed. 0	Your ranking among the volunteers in this challenge, by number of jobs completed. 0

Learn more:

- Introduction to high energy physics simulations
- See the simulations produced by your computer
- Learn about the software that does the simulations
- View current activity on your virtual machine

Collisions analyzed with: CMS_2011_58968497

Generator: pythia8
Process: jets
Beam: pp
Producing: 100,000 events
Energy: 7000.00 GeV

While waiting... why don't you click your way to a Nobel prize?

Stop [Settings] [Trash]



**CernVM: complete and portable environment
for **developing** and **running** HEP data processing tasks**

Use Cases

- 1 IaaS Clouds
- 2 Development Environment
- 3 Volunteer Computing
- 4 Long-Term Analysis Preservation
- 5 Outreach & Education

ALEPH software in CernVM

```
↑ jakob — aleph@cernvm-aleph01:~/test/ALPHA — ssh — 66x18
pb-d-128-141-134-74:~ jakob$ ssh -X aleph@cernvm-aleph01
aleph@cernvm-aleph01's password:
[aleph@cernvm-aleph01 ~]$ source setaleph.sh
[aleph@cernvm-aleph01 ~]$ cd test/ALPHA/
[aleph@cernvm-aleph01 ALPHA]$ sh alpha.sh
*****
*****          ALPHA RUN          **** 11.6 ****
*****
*****

Wed Mar 19 16:10:27 CET 2014

*****
***   Compilation and creation of the makefile 6lep.mk
*****
gmake -f /home/aleph/test/ALPHA/6lep.mk
gmake: `6lep' is up to date.
```

Demonstrates that VMs can bridge 10+ years

► [DPHEP Status Report](#)



**CernVM: complete and portable environment
for **developing** and **running** HEP data processing tasks**

Use Cases

- 1 IaaS Clouds
- 2 Development Environment
- 3 Volunteer Computing
- 4 Long-Term Analysis Preservation
- 5 Outreach & Education

CERN OpenData Portal, CERN@School

The screenshot shows the CERN OpenData Portal interface. At the top, the logo 'opendata' is on the left, and navigation links 'ABOUT', 'SEARCH', 'EDUCATION', and 'RESEARCH' are on the right. Below the navigation is a search bar with a magnifying glass icon and the text 'Search'. The main content area is titled 'Virtual Machines' and contains a paragraph: 'CERN Virtual Machines allow you to run Scientific Linux on any operating system and access the CERN working environments and software tools.' Below this text are three large rectangular buttons, each with a logo and text: 'CMS Virtual Machines' (with the CMS logo), 'ALICE Virtual Machines' (with the ALICE logo), and 'LHCb Virtual Machines' (with the LHCb logo).



CernVM Hypervisor Support Status

The success of CernVM is largely based on the fact that it runs in practically **all** cloud environments.

Hypervisor / Cloud Controller	Status
VirtualBox	✓
VMware	✓
KVM	✓
Xen	✓
Microsoft Hyper-V	✓
Parallels	✗
Vagrant	✓
OpenStack	✓
OpenNebula	✓
CloudStack	✓
Amazon EC2	✓
Google Compute Engine	✓
Microsoft Azure	✓
Docker	✓ ¹

[support added this year](#)
[+ support for glideinWMS-VM](#)

¹ Release Candidate



- Facilitates the instantiation of reproducible and portable development environments using virtual machines.
- Wrapper around hypervisors and cloud controllers, e. g. around VirtualBox.
- For Linux, Windows, OS X

Example

```
> vagrant box add -name CernVM <cernvm image>.box  
> vagrant init CernVM  
> vagrant up  
> vagrant ssh
```

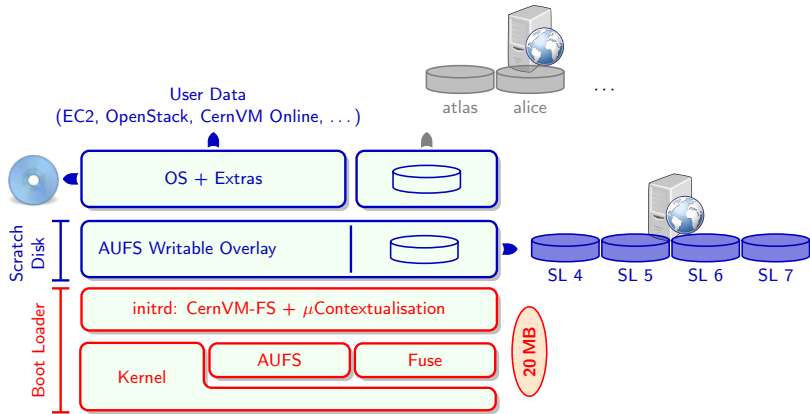
<https://www.vagrantup.com>

▶ [CernVM documentation](#)



Reminder: Building Blocks of CernVM

Twofold system: μ CernVM boot loader + OS delivered by CernVM-FS





Boot Loader (Image)

- Linux Kernel + “guest additions”
- Extra features:
zRAM, x32 ABI, AUFS, ...
- This year: 3.10 → 3.18 → 4.1
- Cryptographic signature added to images

OS on CernVM-FS

- Based on Scientific Linux
 - + Configuration and tuning
 - + Contextualization
 - + Extra packages
- Available repositories:
 - SL4 Prototype, LEP experiments
 - SL5 Stable,
GUI and contextualization
 - SL6 Production, in 2015
2 major releases,
16 security hotfixes
 - SL7 Release candidate
- Production repository:
 - 6 G, 210 k files
 - 800 curated packages
 - 1 600 total packages
 - 30 custom packages



Docker

Docker and 1xc (Linux containers) available in CernVM

```
[root@localhost ~]# service docker start
Starting Docker:                                     [ OK ]
[root@localhost ~]# docker run -i ubuntu /bin/bash
Unable to find image 'ubuntu:latest' locally
ubuntu:latest: The image you are pulling has been verified
511136ea3c5a: Pull complete
fa4fd76b09ce: Pull complete
1c8294cc5160: Pull complete
117ee323aaa9: Pull complete
2d24f826cb16: Pull complete
Status: Downloaded newer image for ubuntu:latest
cat /etc/issue
Ubuntu 14.04.2 LTS \n \l
[root@localhost ~]#
```

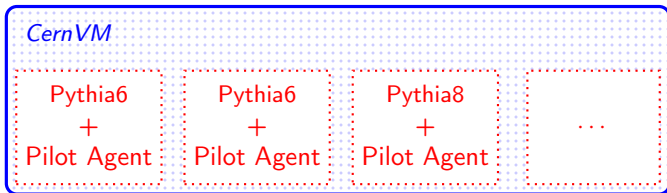


Resource Containers and Isolated Environments

The `cernvm-fork` utility allows for simple creation of *resource containers* with *isolated environments*

```
$ cernvm-fork fork03 --new --fast --cvdfs=sft.cern.ch \  
  --run=/cvdfs/sft.cern.ch/my_app
```

- Facilitates **sharing of multi-core** virtual machines
- CernVM-FS cache shared across containers





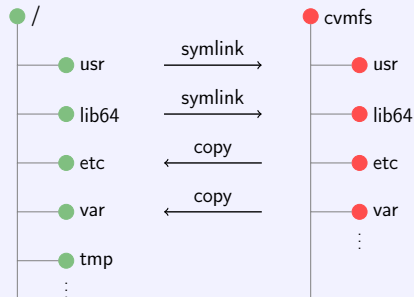
CernVM (Docker) container

- 1 Make CernVM-FS repositories available inside a container
- 2 Work around missing union file system for writable overlay

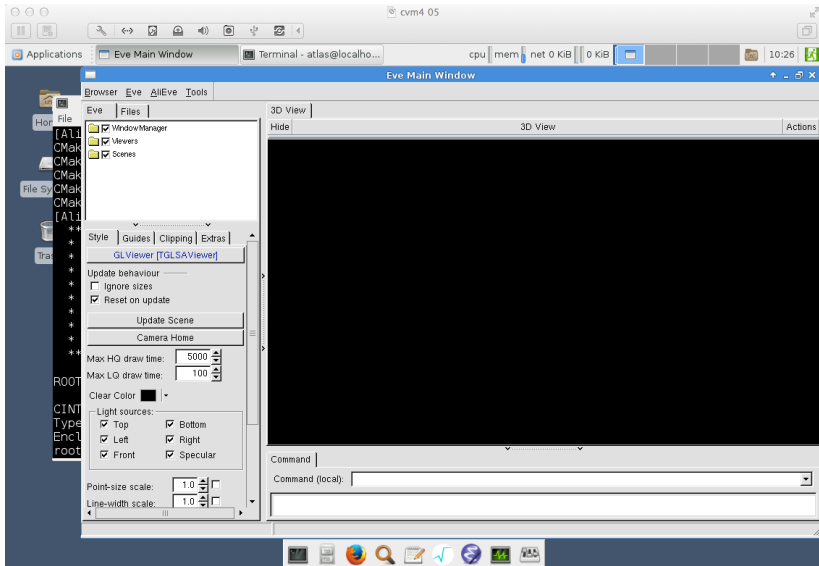
1 Options for CernVM-FS

- 1 Fuse, mapped from host
 - Shared cache
 - Collaboration from host
- 2 Using Parrot-Cvmfs
 - Pure user-space (ptrace)
 - Less performant, some limitations (e.g. suid binaries)

2 Root file system (/) layout









Tasks for Scientific Linux 7 Support

- ✓ CernVM-FS support for file capabilities (extended attributes)
- ✓ CernVM-FS systemd integration as a “low level storage daemon”
- ✓ System configuration and tuning with systemd
 - ~25 minor issues open

Boot time optimization:

Scenario	CernVM 3	CernVM 4
ADSL	120 s	55 s
CERN	60 s	40 s
Warm Cache	37 s	17 s

Speedy boot is particularly important for the cloud use case

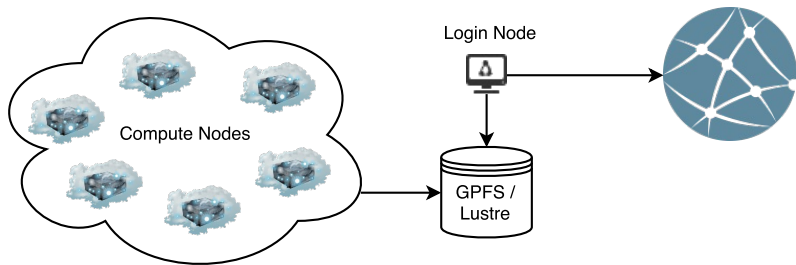


① Status of CernVM

② Status of CernVM-FS



- 1 Access to CernVM-FS on HPC resources
- 2 Exploring data distribution
- 3 Monitoring
- 4 Code housekeeping

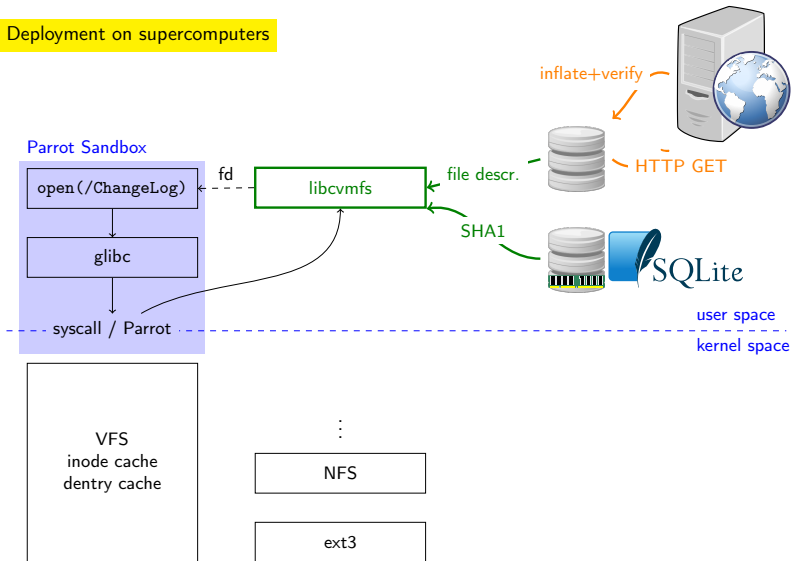


Issues

- 1 No Fuse on compute nodes → Parrot connector
- 2 No Internet access → preload CernVM-FS cache on the cluster file system

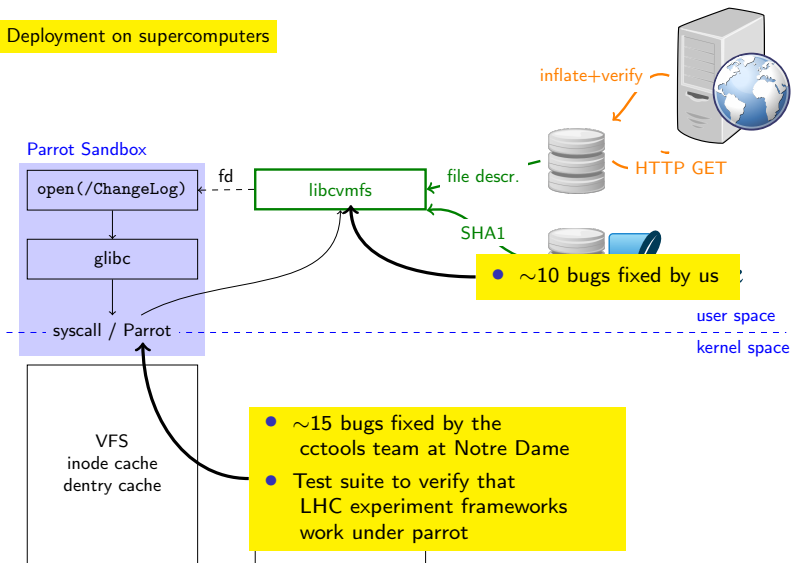


Deployment on supercomputers





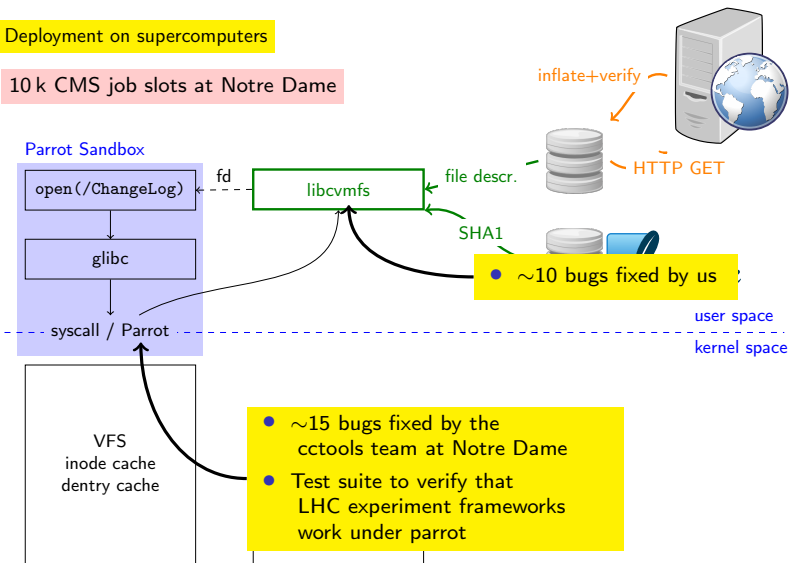
Deployment on supercomputers





Deployment on supercomputers

10 k CMS job slots at Notre Dame





cvmfs_preload

- Self-extracting binary, can be copied to the login node
- Same code than stratum 0 → stratum 1 replication
- Very efficient in transferring change sets (seconds to minutes)
- Allows for partial preloading of the namespace

Example Invocation

```
> cvmfs_preload -u http://hcc-cvmfs.unl.edu:8000/cvmfs/alice.cern.ch \  
-r /shared/cache -d <dirtab>
```

▶ [CernVM documentation](#)



cvmfs_preload

- Self-extracting binary, can be copied to the login node
- Same code than stratum 0 → stratum 1 replication
- Very efficient in transferring change sets (seconds to minutes)
- Allows for partial preloading of the namespace

Example Invocation

```
> cvmfs_preload -u http://hcc-cvmfs.unl.edu:8000/cvmfs/alice.cern.ch \  
-r /shared/cache -d <dirtab>
```

▶ [CernVM documentation](#)

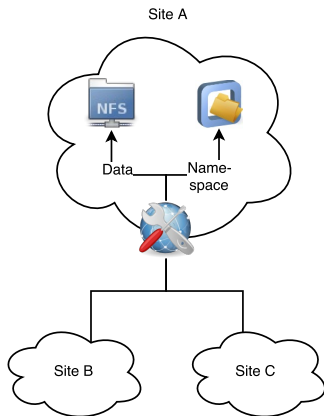
Prototyped with ALICE on NERSC, next step: scale test



CernVM-FS for Data Federations

Contribution from Brian Bockelman & Derek Weitzel / OSG

Use CernVM-FS as a POSIX compliant, consistent, cryptographically secured name space for data files.

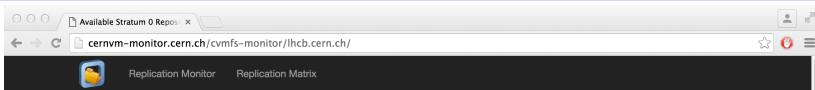


Contributions

- 1 Grafting files [\[merged\]](#)
Describe name space without processing of files
- 2 Uncompressed files [\[review\]](#)
- 3 External data [\[review\]](#)
Separate downloads of file catalogs and data files
- 4 HTTPS/VOMS support [\[review\]](#)
Login to HTTPS data server with grid certificate



CernVM-FS Monitor & CernVM-FS Browser



Main : [Stratum0](#) : [lhcb.cern.ch](#)



LHCb ([lhcb.cern.ch](#))

Project Information

Browse: [lhcb.cern.ch](#)

Stratum 0

Stratum0 Revision:	33736
Oldest Stratum1 Revision:	33736
Last Modified:	12/11/2015 3:20:34 PM
Whitelist Expiry Date:	12/24/2015 1:11:35 PM (12 days left)
Root Catalog Hash:	8d42d76c820f0356697ad95f1f9f618a9b6658be
Number of known Stratum 1 Replicas:	6

Stratum 1 Replicas



CERN Replica

Revision: 33736 (Last Replication: 12/11/2015 12:10:57 PM)
URL: <http://cvmfs-stratum-one.cern.ch/opt/lhcb>



RAL Replica

Revision: 33736 (Last Replication: 12/11/2015 3:44:19 PM)
URL: <http://cernvmfs.gridpp.rl.ac.uk/opt/lhcb>



CernVM-FS Monitor & CernVM-FS Browser

Cloud Browser - /lib x

cernvm-monitor.cern.ch/cvmfs-monitor/cb/browser/lhcb.cern.ch/latest//lib

(Home) = lib

Date: 11/12/2015 Number: latest

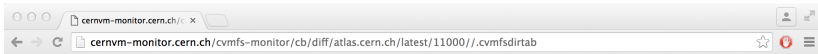
	Name	Content Type	Content Hash	Size	Last Modified	
	.install	--	cbd2051078d1cb362d4ade1e96c398f9a4a59e40 Sha1	292 bytes	June 19, 2008, 12:05 p.m.	generic-2015-12-11T14:19:44Z trunk
	GangaConfig	--	--	3 bytes	Sept. 5, 2010, 11:03 p.m.	generic-2015-12-11T14:15:02Z
	LHCb_config.py	text/x-python	4d35df97a8f2eed6fa371cf06e82efaa6a93962f Sha1	4.4 KB	May 25, 2009, 10:50 a.m.	trunk-previous
	LHCb_config.pyc	application/x-python-code	d0065edd9b5116a45ed061f2976fbf35df3e4af8 Sha1	3.1 KB	May 25, 2009, 10:50 a.m.	generic-2015-12-11T13:19:44Z
	LbLogin.bat	text/plain	--	51 bytes	Sept. 17, 2015, 5:51 p.m.	generic-2015-12-11T12:19:43Z
	LbLogin.csh	application/x-csh	--	51 bytes	Sept. 17, 2015, 5:51 p.m.	generic-2015-12-11T12:15:02Z
	LbLogin.sh	application/x-sh	--	50 bytes	Sept. 17, 2015, 5:51 p.m.	generic-2015-12-11T10:15:02Z
	RootConfig	--	--	4.0 KB	July 22, 2015, 9:29 a.m.	generic-2015-12-11T09:19:49Z
	conf	--	--	4.0 KB	March 6, 2012, 10:10 a.m.	generic-2015-12-11T09:19:49Z
	contrib	--	--	4.0 KB	Nov. 1, 2010, 10:32 a.m.	generic-2015-12-11T08:19:49Z
	etc	--	--	14 bytes	Sept. 17, 2015, 5:51 p.m.	generic-2015-12-11T08:15:02Z
	etc_extra_settings.csh	application/x-csh	7ff366d10b4f7e3efe5c4e8ca4449cf90ea7b556 Sha1	526 bytes	March 16, 2015, 4:18 p.m.	generic-2015-12-11T08:15:02Z
	group_extra_settings.sh	application/x-sh	36aa655c296d02357a8f0f865244cde66bb1b7 Sha1	540 bytes	March 16, 2015, 4:17 p.m.	generic-2015-12-11T07:59:49Z
	group_login.csh	application/x-csh	1916834e0f1e978acdfe5c9e313653ca57c3942e Sha1	390 bytes	June 14, 2011, 1:28 p.m.	generic-2015-12-11T06:59:48Z
	group_login.sh	application/x-sh	6e31f9515600577043d35567b74356d28892a87f Sha1	377 bytes	June 14, 2011, 1:29 p.m.	generic-2015-12-11T06:15:03Z
	group_shell.csh	application/x-csh	bf0fbcc057e93bc4e5b87ae0249fb9457e816828 Sha1	498 bytes	June 14, 2011, 1:29 p.m.	generic-2015-12-11T05:59:48Z
	group_shell.sh	application/x-sh	e03812e90650beb528d392929b5769afa802fa6b Sha1	482 bytes	June 14, 2011, 1:30 p.m.	generic-2015-12-11T04:59:48Z
	html	--	--	3 bytes	Dec. 7, 2015, 2:32 p.m.	generic-2015-12-11T04:15:02Z
	install_project.py	text/x-python	7f07939f85d53ac88276a695e04a65a536f64d Sha1	107.6 KB	Sept. 17, 2015, 5:50 p.m.	generic-2015-12-11T04:15:02Z
	install_project.py.old	--	fc98cb268ac8b144c453dbf42e9838fb89047e4 Sha1	107.6 KB	Sept. 17, 2015, 5:50 p.m.	generic-2015-12-11T03:59:50Z
	latest_install_project.py	text/x-python	7f07939f85d53ac88276a695e04a65a536f64d Sha1	107.6 KB	Dec. 11, 2015, 3:19 p.m.	generic-2015-12-11T02:59:48Z
	lcb	--	--	3 bytes	Aug. 27, 2015, 10:38 a.m.	generic-2015-12-11T02:59:48Z
	lhcb	--	--	3 bytes	March 6, 2015, 9:01 a.m.	generic-2015-12-11T02:15:02Z
	log	--	--	5 bytes	Dec. 7, 2015, 2:32 p.m.	generic-2015-12-11T02:15:02Z
	slc4_amd64_gcc34	--	--	3 bytes	April 30, 2009, 10:30 a.m.	generic-2015-12-11T02:15:02Z
	slc4_ia32_gcc34	--	--	5 bytes	March 19, 2009, 10:22 a.m.	<< >>
	targz	--	--	3 bytes	Dec. 7, 2015, 2:32 p.m.	Regular files 16100272
	tmp	--	--	3 bytes	Dec. 11, 2015, 3:19 p.m.	Directories 2357862
Revision 33736 - generic-2015-12-11T14:19:44Z						Symlinks 116583

Total file size 758.1 GB
 Chunked files 9680
 Chunked file size 156.0 GB
 Number of chunks 38157
 Nested catalogs 2105

This web application is meant to be available only in web browsers, but with no warranty



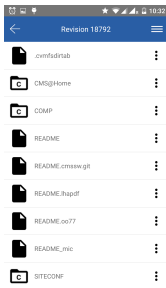
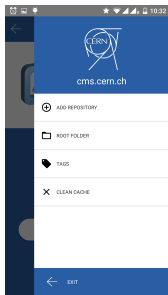
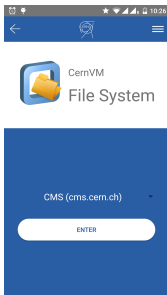
CernVM-FS Monitor & CernVM-FS Browser



	latest		11000
1	/repo/dev/atlas-gcc/*	1	/repo/dev/atlas-gcc/*
2	/repo/sw/sgis/*/*	2	/repo/sw/sgis/*/*
3	/repo/sw/atlas-gcc/*	3	/repo/sw/atlas-gcc/*
4	/repo/sw/octools/*	4	/repo/sw/octools/*
5	/repo/sw/ddm/*	5	/repo/sw/ddm/*
6	/repo/sw/ddm/rucio-clients/*	6	/repo/sw/ddm/rucio-clients/*
7	/repo/sw/database/DBRelease/*	7	/repo/sw/database/DBRelease/*
8	/repo/sw/external/*	8	/repo/sw/external/*
9	/repo/sw/JEM-WN/*	9	/repo/sw/JEM-WN/*
10	/repo/sw/muon/LCDS/*	10	/repo/sw/muon/LCDS/*
11	/repo/sw/python/*	11	/repo/sw/python/*
12	/repo/sw/software/*-opt/*	12	/repo/sw/software/*-opt/*
13	/repo/sw/software/Atbt/*-opt/*	13	/repo/ATLASLocalRootBase/x86_64/root/*
14	/repo/ATLASLocalRootBase/x86_64/root/*	14	/repo/ATLASLocalRootBase/x86_64/python/*
15	/repo/ATLASLocalRootBase/x86_64/python/*	15	/repo/ATLASLocalRootBase/x86_64/boost/*
16	/repo/ATLASLocalRootBase/x86_64/boost/*	16	
17		17	



Add-On: Android App



<https://drive.google.com/file/d/0B8QL3WDIY49cV3FG0EhrcnJGUEU>



CernVM-FS Code Base

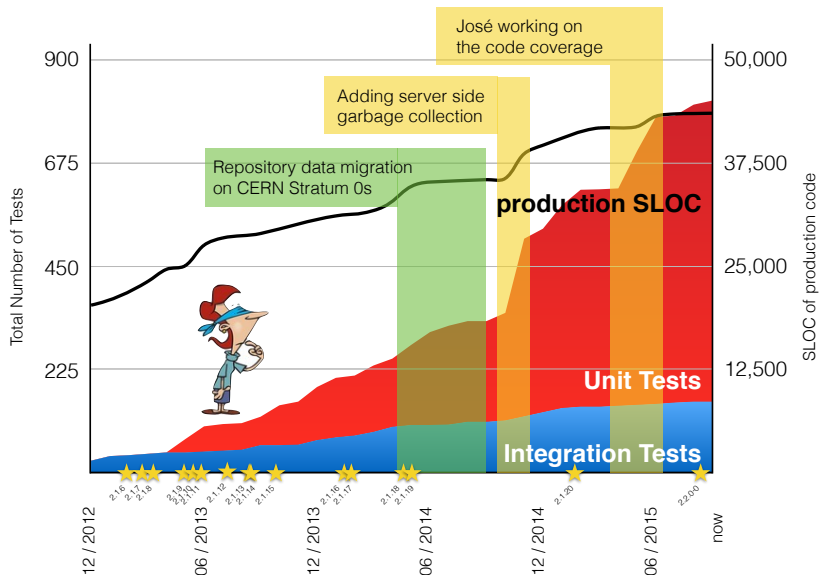
- ~50 kSLOC production code
 - ~40 kSLOC unit and integration tests
 - Code base ages
- Testing and continuous code cleanup

Order by | ▾

CVM-689
Code Housekeeping

Sub-Tasks

- | | | | | |
|-----|--|--|----------|----------------------------------|
| 1. | ✓ Verify and fix code style | | RESOLVED | Jakob Blomer |
| 2. | ✓ Add make test | | RESOLVED | Jose Molina Colmenero ↑ ...
↓ |
| 3. | ✓ Add make doc | | RESOLVED | Jose Molina Colmenero |
| 4. | ✓ Add pre-commit hooks framework | | CLOSED | Jakob Blomer |
| 5. | Consistently use lock guards | | OPEN | Jakob Blomer |
| 6. | Overhaul cvmfs_swissknife | | OPEN | Rene Meusel |
| 7. | ✓ Use Database::SetProperty<> where appropriate | | RESOLVED | Rene Meusel |
| 8. | Use CVMFS_* config variables directly in cvmfs_server | | OPEN | Rene Meusel |
| 9. | Eliminate unclear Catalog::path() and Catalog::root_prefix() | | REOPENED | Jakob Blomer |
| 10. | ✓ Consistently Sort Include Directives | | RESOLVED | Jakob Blomer |



Source: Meusel



Testing Strategy in Practise

“every” commit



- Development unit test set
 - ~ 600 tests (< **2 minutes**)
 - **39%** code coverage

Nightly Builds



- Full multi-platform unit test set
 - ~ 660 tests (~ **15 minutes**)
 - runs on **all build platforms**
 - **43%** code coverage

on demand



- Multi-platform integration and unit test test
 - ~ 820 tests in **various configurations** (~ 16 hours)
 - **80%** code coverage (73% integration + unit tests)

Source: Meusel



Configuration Matrix	docker-i386	docker-x86_64	bare-armv7hl	osx10-x86_64	arm64	cvmfs-power8
slc4						
slc5						
slc6						
cc7						
fedora21-power8						
fedora22						
sles11						
opensuse13						
ubuntu1404						
mac						

Recently added: Power8, ARM64, OS X El Capitan



CernVM

- Close to 30 000 new VMs / day
- Gives access to practically all open-source and commercial cloud environments
- Release candidates:
 - CernVM 4
 - CernVM container

CernVM-FS

- Tight collaboration with OSG and University of Notre Dame
- Screened by CERN KT group for commercial applications, interest from Mesosphere
- Parrot and the new preloader allow for access to CernVM-FS on Supercomputers
- Exploring CernVM-FS as a name space for data repositories
- Ongoing work on monitoring