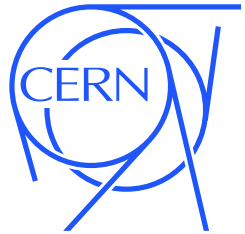


CernVM Status Update

Jakob Blomer

PH-SFT Group Meeting
14th December 2015

Agenda

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

Releases at a Glance

- 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

New Journal Publication

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

Agenda

1. Status of CernVM
2. Status of CernVM-FS

Boot of New CernVMs

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

Known Users

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

Known 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

Infrastructure-as-a-Service Cloud

Various clouds:

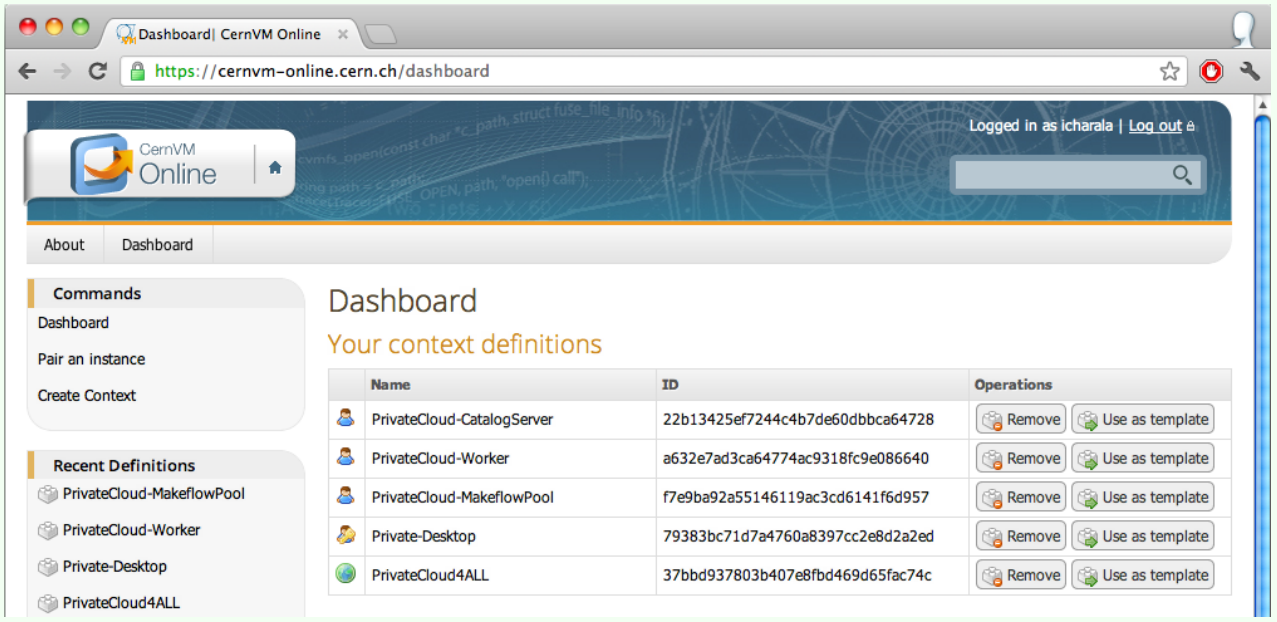
- ATLAS HLT farm
- Cloud resources seamlessly integrated with experiment task queues (e. g. ATLAS CloudScheduler, LHCb VAC)
- ALICE software release testing on CERN OpenStack
- Commercial providers ("Helix Nebula")
- ...

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



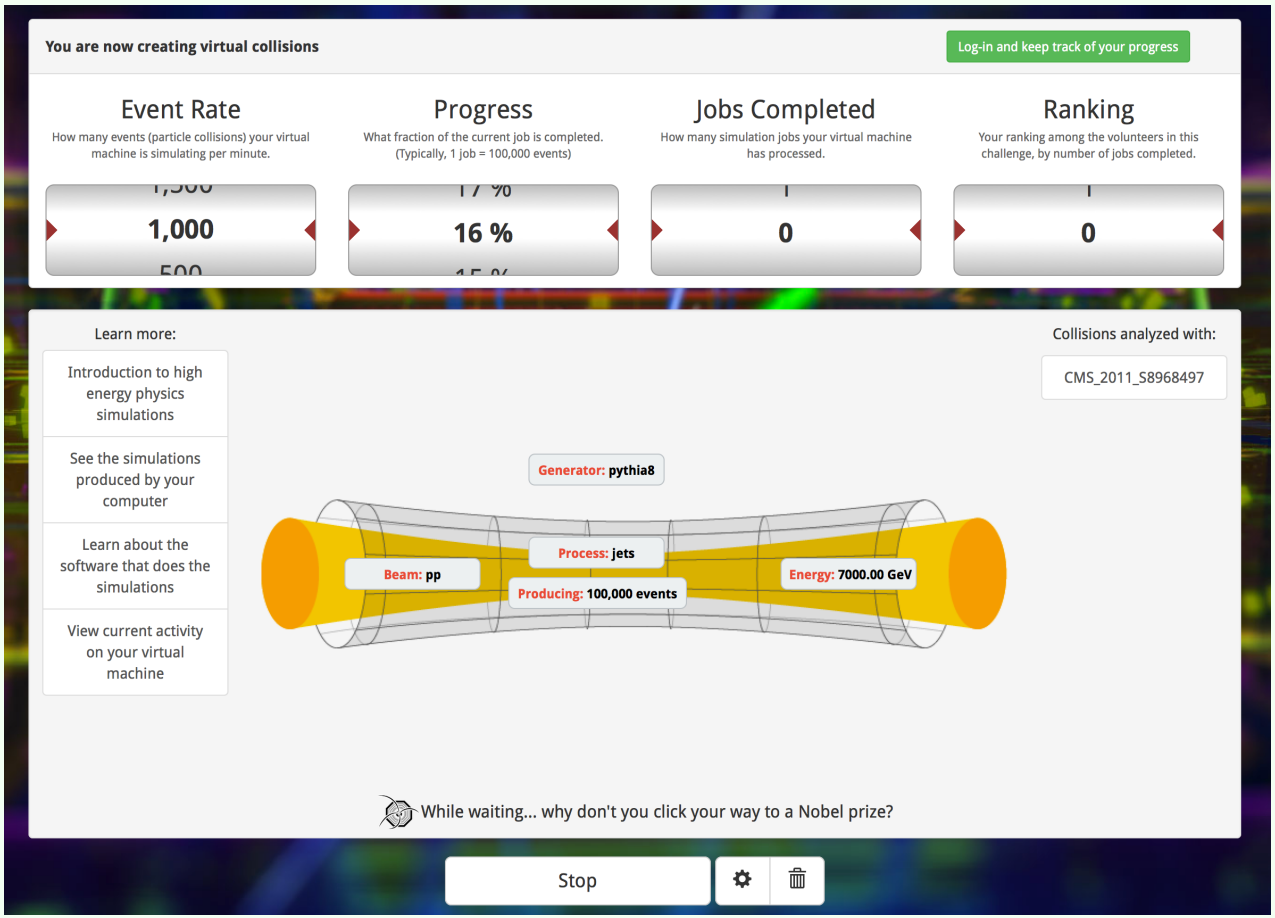
~300 users in 2015, out of which ~65 external users

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



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 — 66×18

```
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

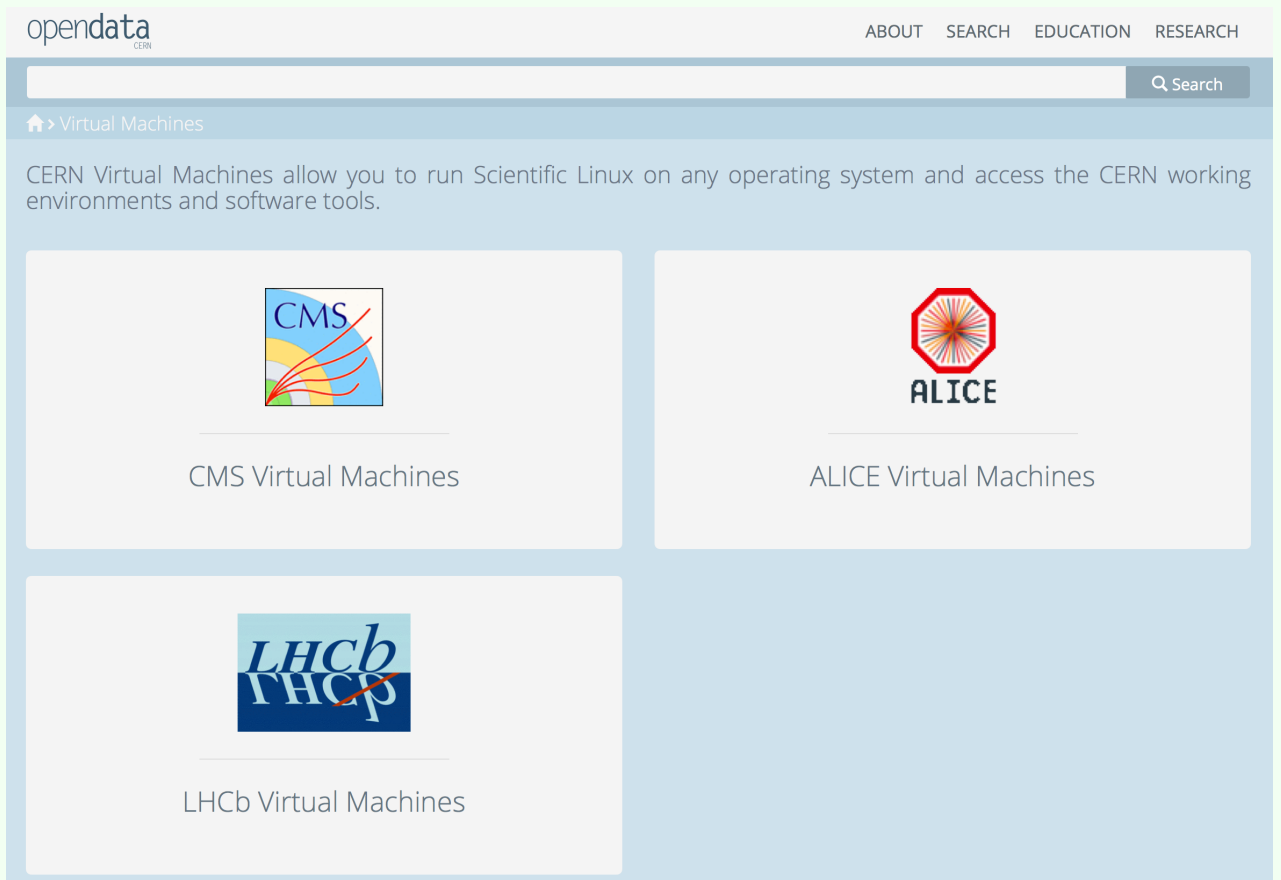
▸ DPHEP Status Report

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



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

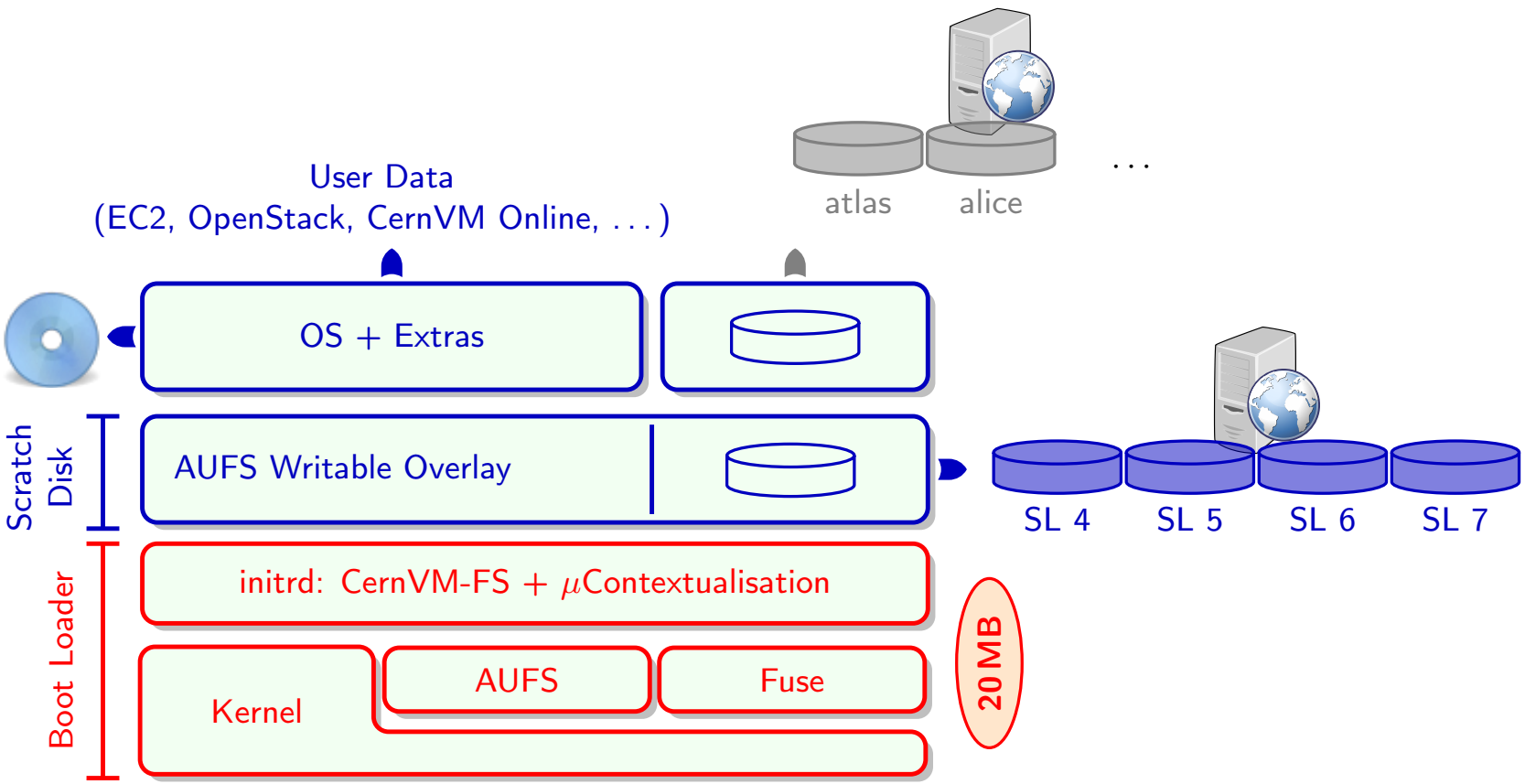
```
> vagrant box add -name CernVM <cernvm image>.box
> vagrant init CernVM
> vagrant up
> vagrant ssh
```

https://www.vagrantup.com

▸ CernVM documentation

Reminder: Building Blocks of CernVM

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



Status of the Building Blocks

Boot Loader (Image)

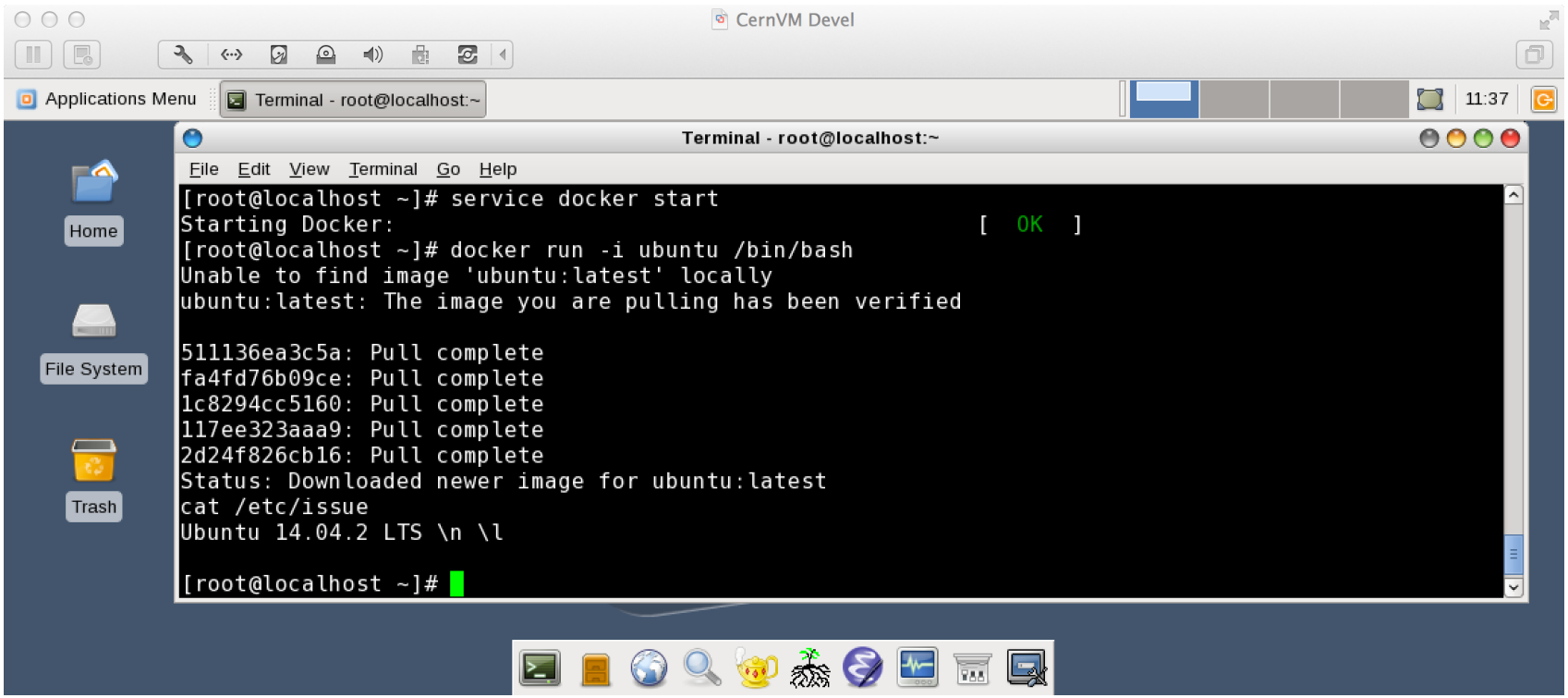
- Linux Kernel + "guest additions"
- Extra features:
  zRAM, x32 ABI, AUFS, ...
- This year: 3.10 → 3.18 → 4.1
- Cryptographic signature
  added to images

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 `lxc` (Linux containers) available in CernVM

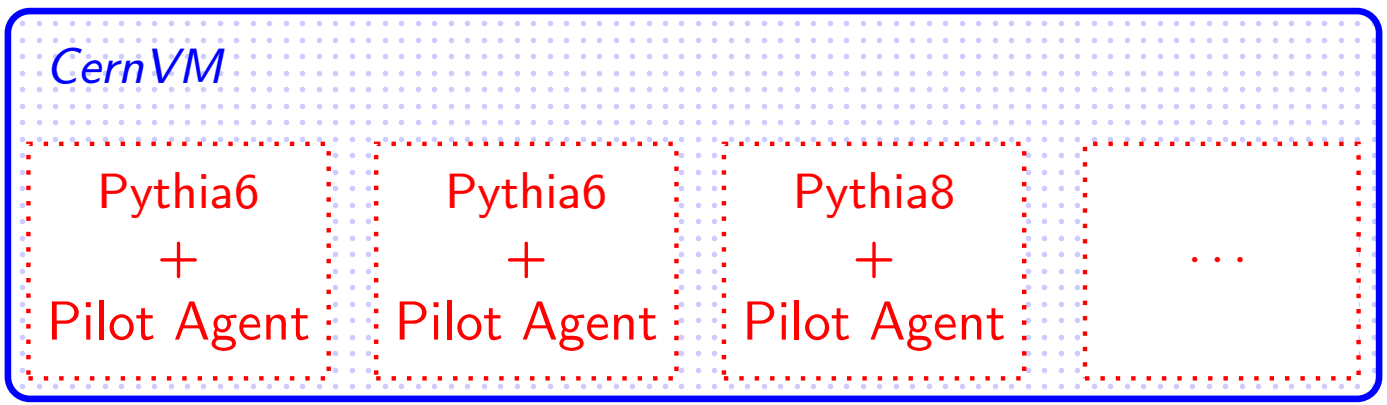


Resource Containers and Isolated Environments

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

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

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



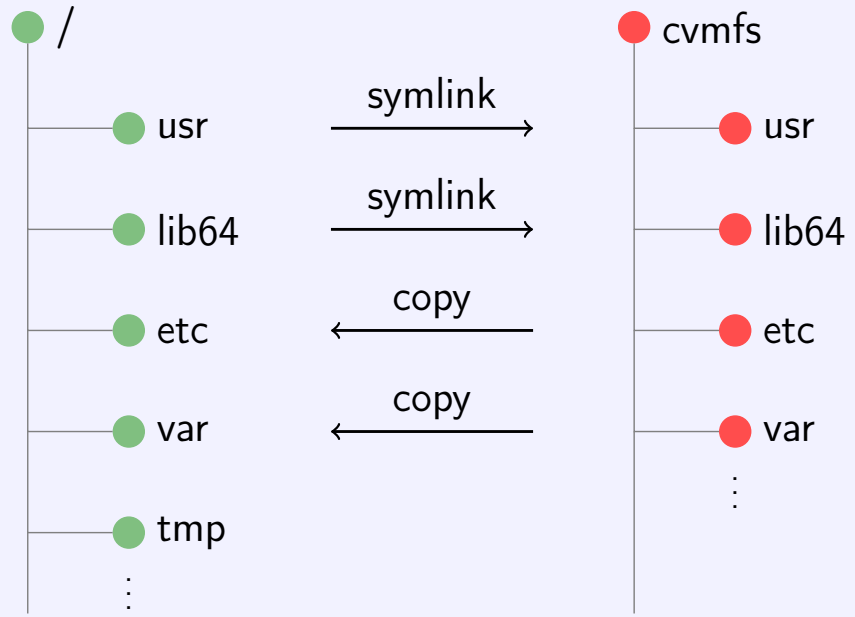
CernVM (Docker) container

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

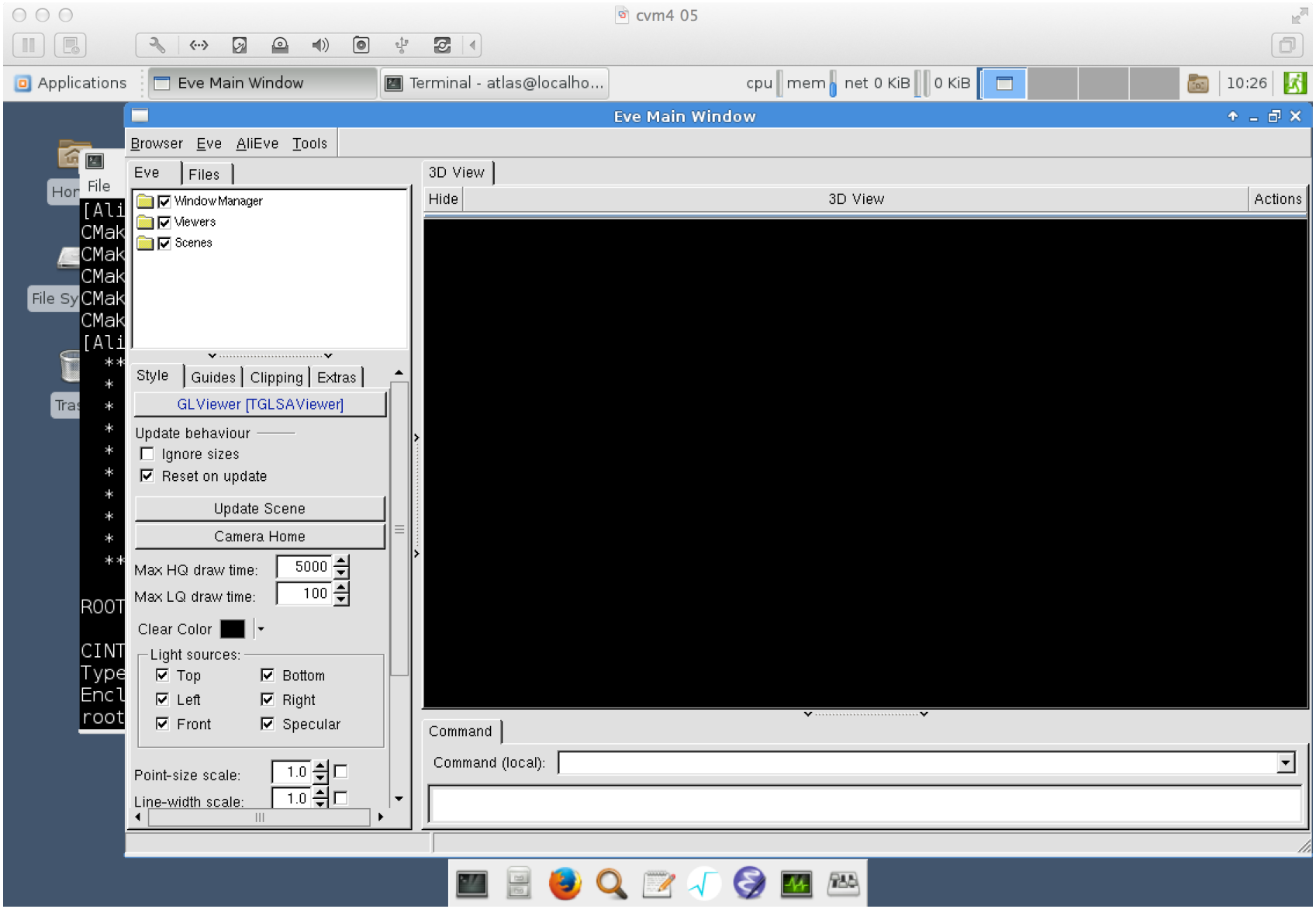
❶ Options for CernVM-FS

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)

❷ Root file system (/) layout




cvm4 05
Applications
cpu
mem
net 0 KiB
0 KiB
10:10
Home
File System
CernVM
Software Appliance
ROOT
http://root.cern.ch


cvm4 05
Applications
Eve Main Window
Terminal - atlas@localho...
cpu
mem
net 0 KiB
0 KiB
10:26
Eve Main Window
Browser
Eve
AliEve
Tools
Eve
Files
Window Manager
Viewers
Scenes
Style
Guides
Clipping
Extras
GLViewer [TGLSAViewer]
Update behaviour
Ignore sizes
Reset on update
Update Scene
Camera Home
Max HQ draw time:
5000
Max LQ draw time:
100
Clear Color
Light sources:
Top
Bottom
Left
Right
Front
Specular
Point-size scale:
1.0
Line-width scale:
1.0
3D View
Hide
3D View
Actions
Command
Command (local):

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:

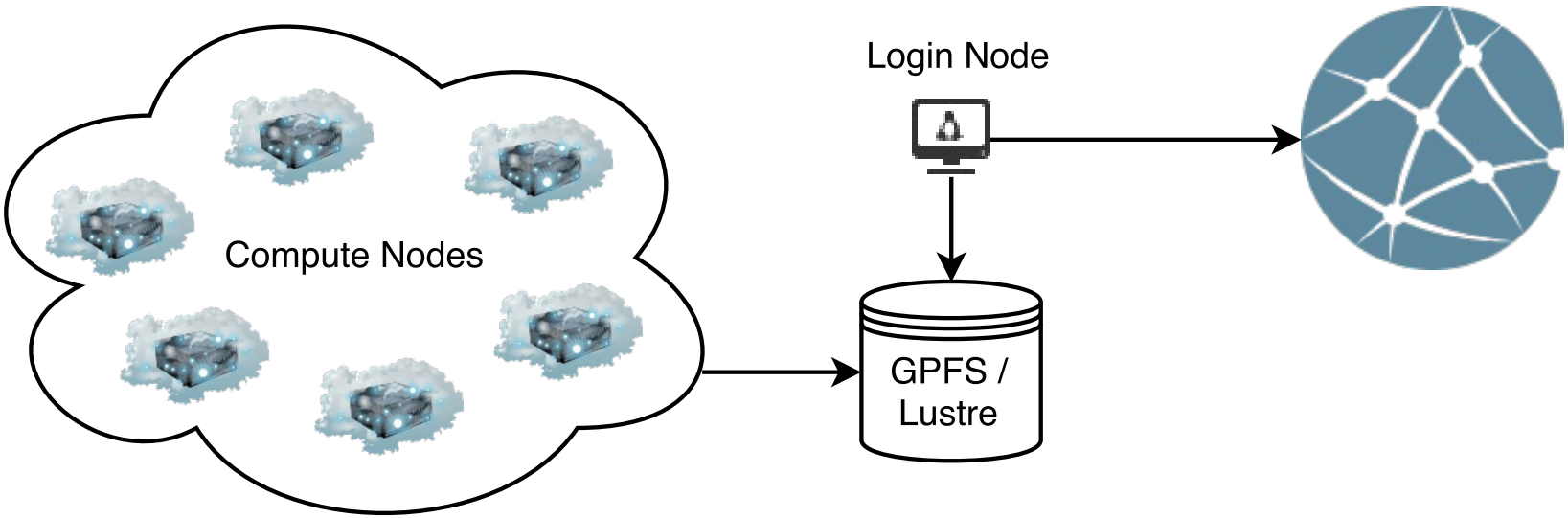
Scenario	CernVM 3	CernVM 4
ADSL	120 s	55 s
CERN	60 s	40 s
Warm Cache	37 s	17 s

Speedy boot is particularly important for the cloud use case

Agenda

Topics for CernVM-FS 2.2.0

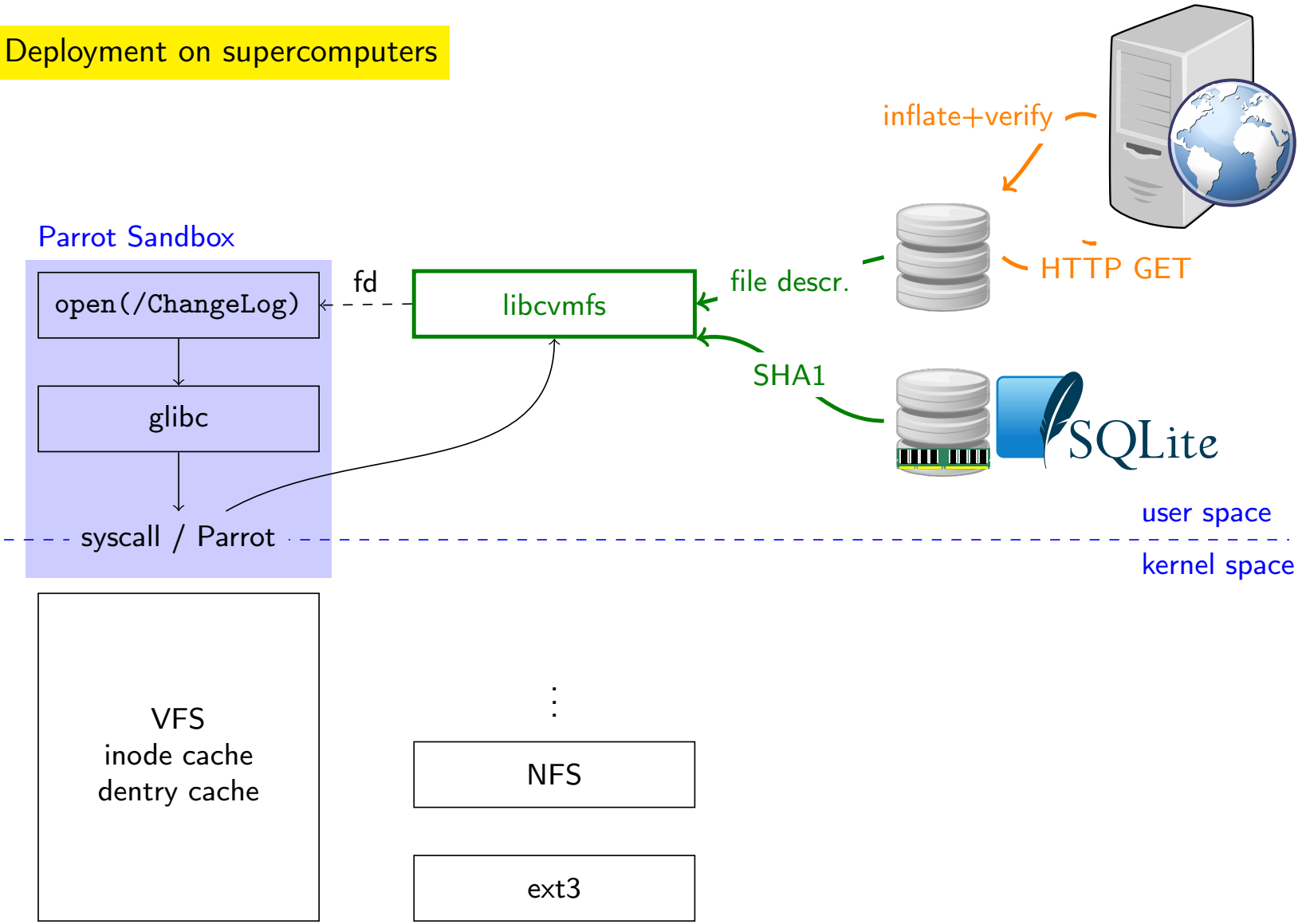
1. Access to CernVM-FS on HPC resources
2. Exploring data distribution
3. Monitoring
4. Code housekeeping



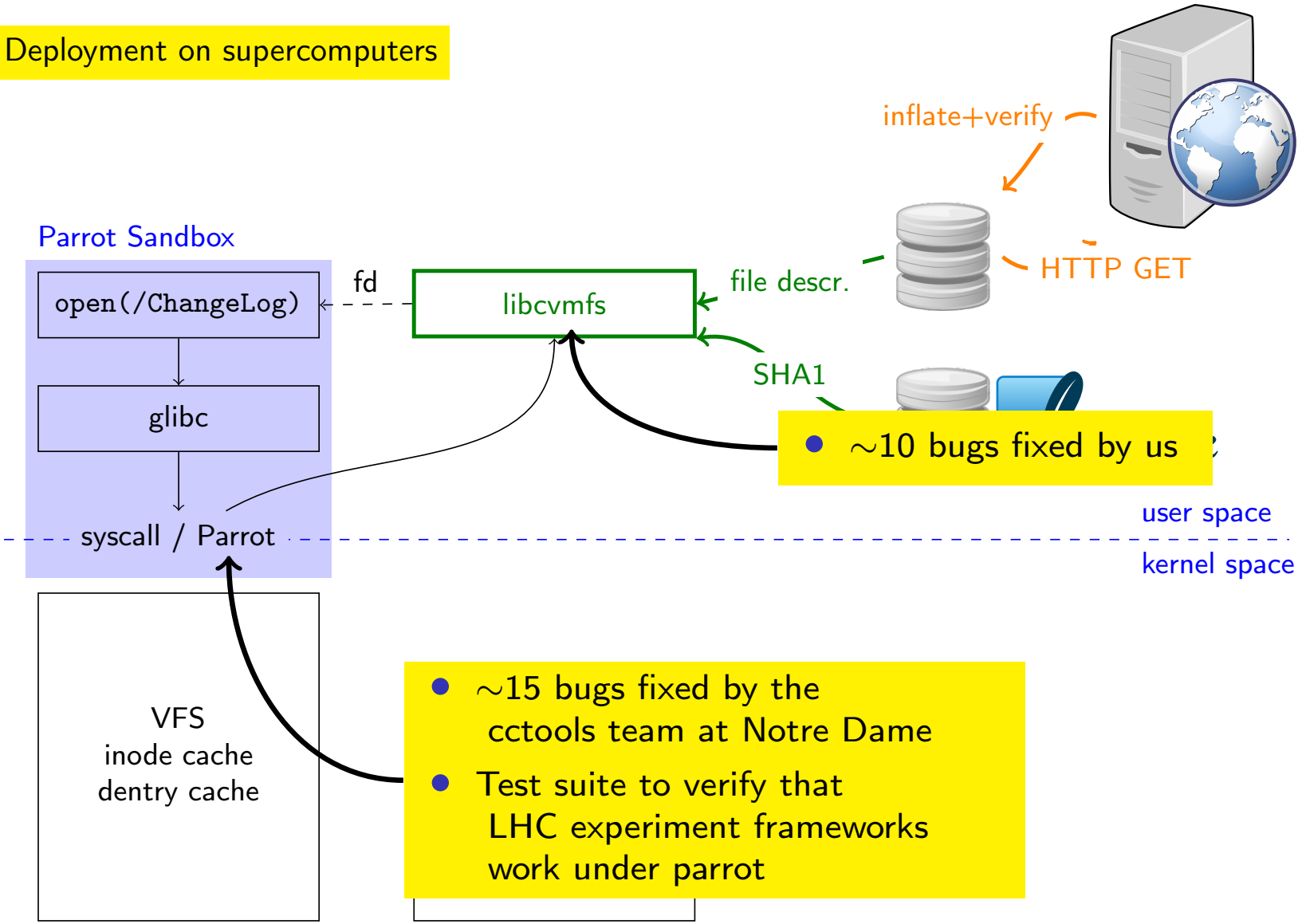
Issues

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

Deployment on supercomputers



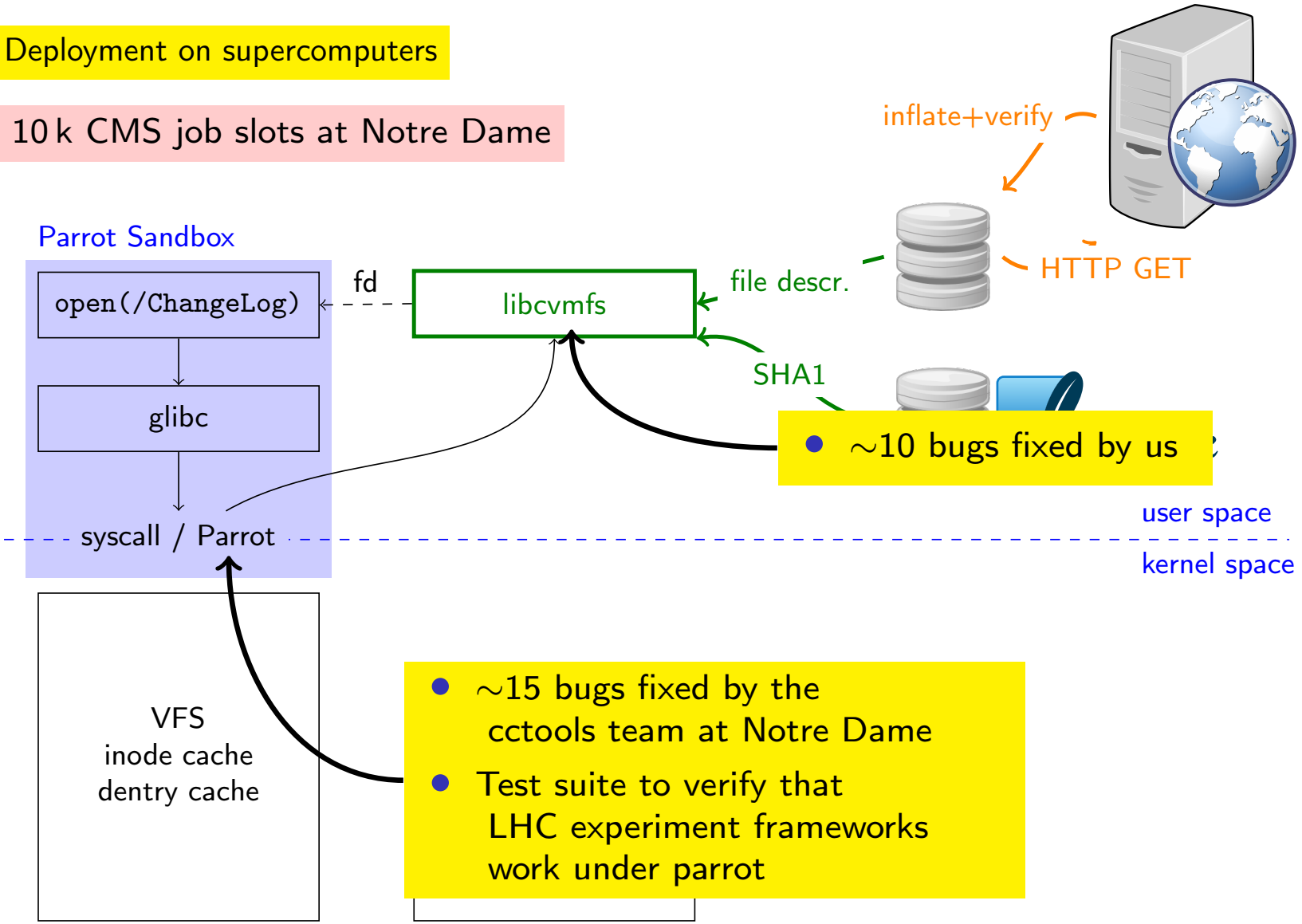
Deployment on supercomputers



Parrot

Deployment on supercomputers

10 k CMS job slots at Notre Dame



Parrot

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

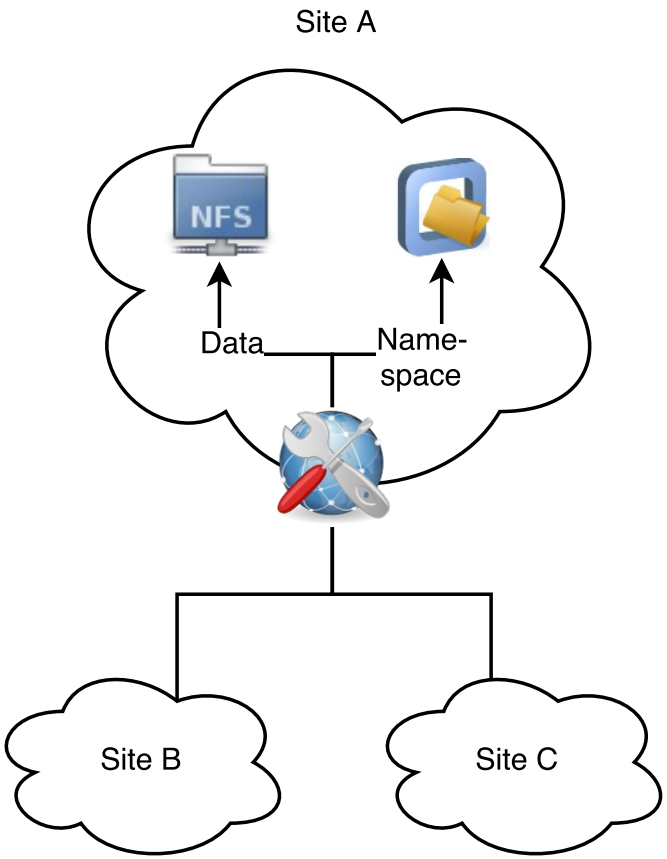
Example Invocation

```
> 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

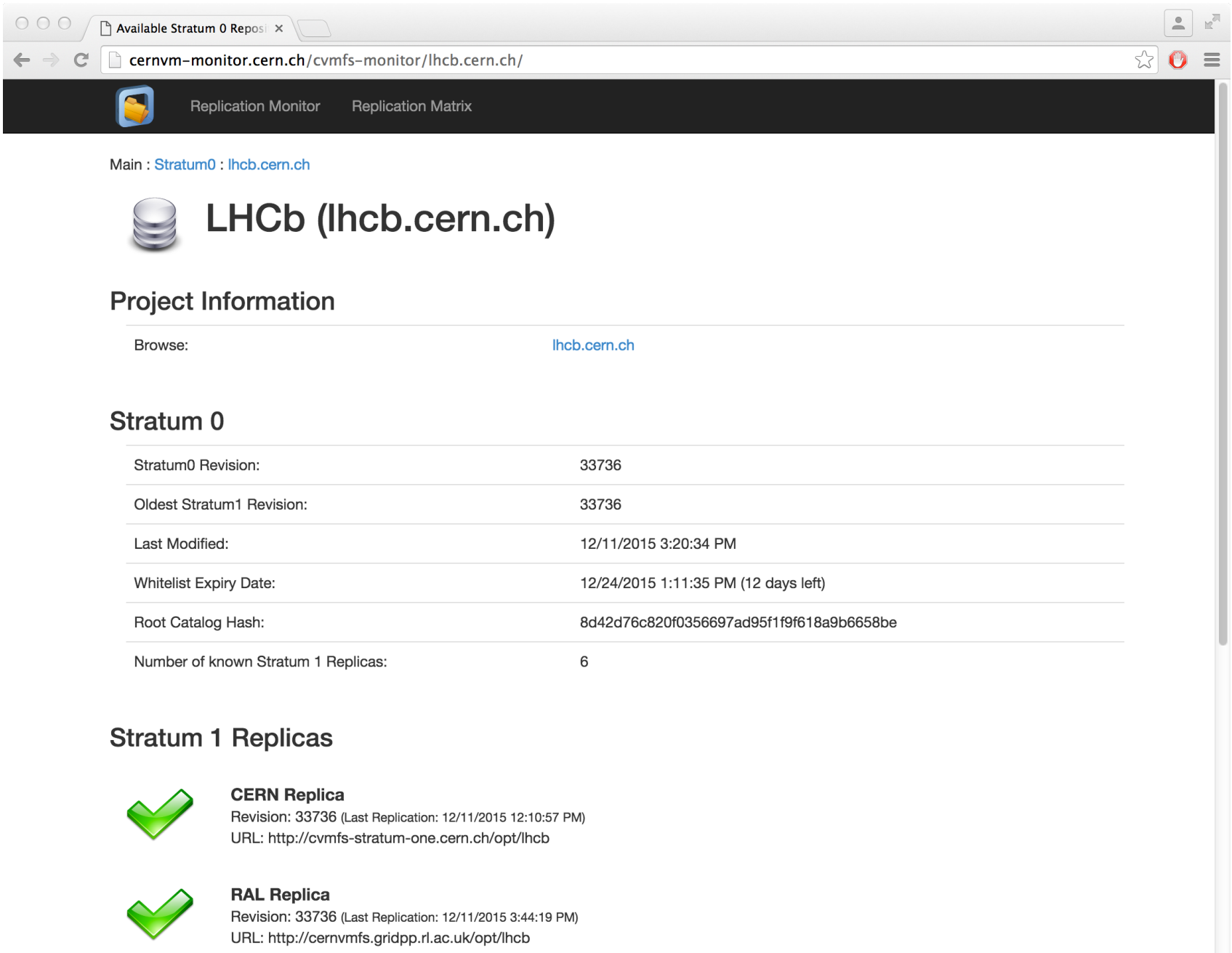
Use CernVM-FS as a POSIX compliant, consistent, cryptographically secured name space for data files.



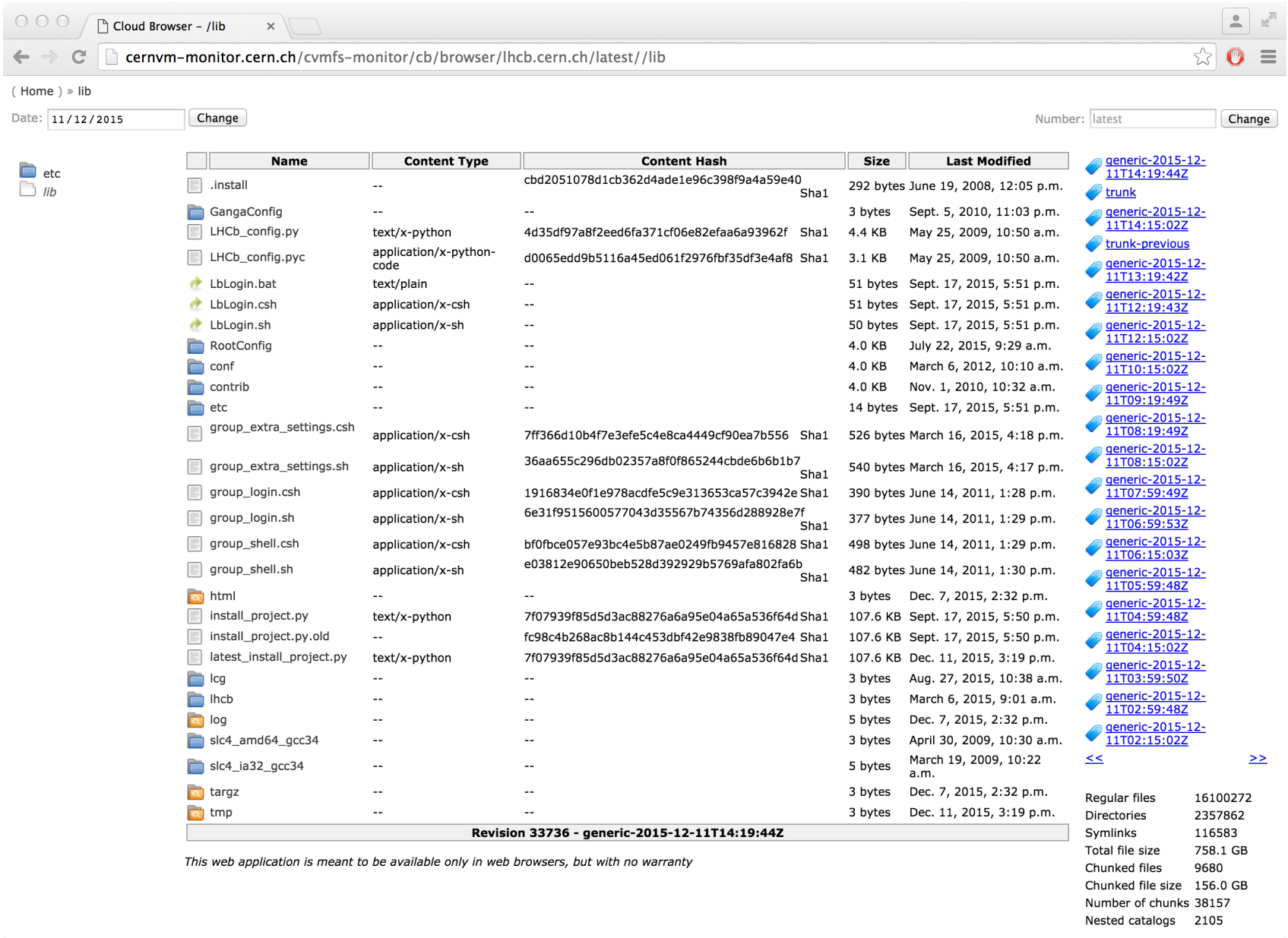
Contributions

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



CernVM-FS Monitor & CernVM-FS Browser



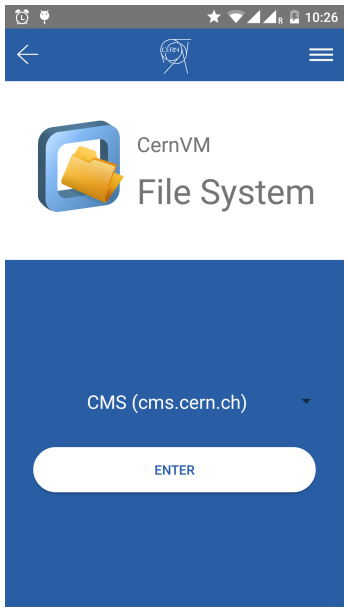
CernVM-FS Monitor & CernVM-FS Browser

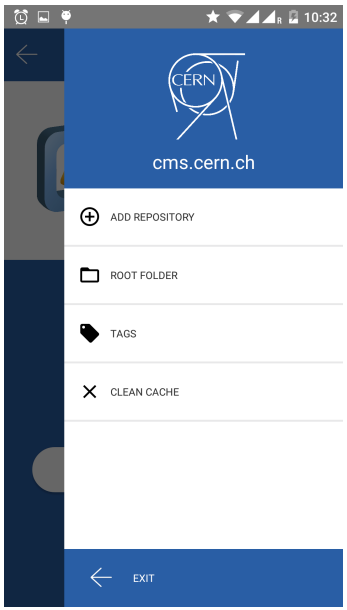
cernvm-monitor.cern.ch/c

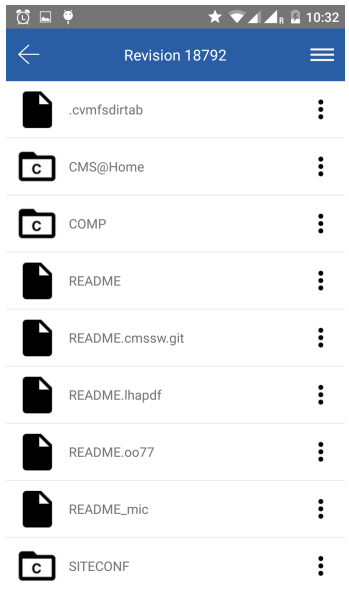
cernvm-monitor.cern.ch/cvmfs-monitor/cb/diff/atlas.cern.ch/latest/11000//.cvmfsdirtab

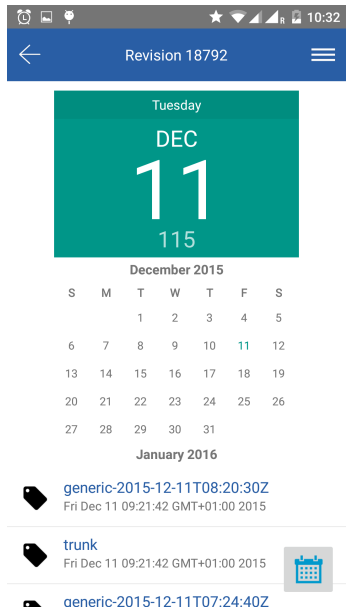
		latest		11000
f	1	/repo/dev/atlas-gcc/*	f	/repo/dev/atlas-gcc/*
	2	/repo/sw/agis/*/*		/repo/sw/agis/*/*
	3	/repo/sw/atlas-gcc/*		/repo/sw/atlas-gcc/*
	4	/repo/sw/cctools/*		/repo/sw/cctools/*
	5	/repo/sw/ddm/*		/repo/sw/ddm/*
	6	/repo/sw/ddm/rucio-clients/*		/repo/sw/ddm/rucio-clients/*
	7	/repo/sw/database/DBRelease/*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

Add-On: Android App









https://drive.google.com/file/d/0B8QL3WDIY49cV3FGOEhrcnJGUEU

CernVM-FS Code Base

- ∼50 kSLOC production code
- ∼40 kSLOC unit and integration tests
- Code base ages

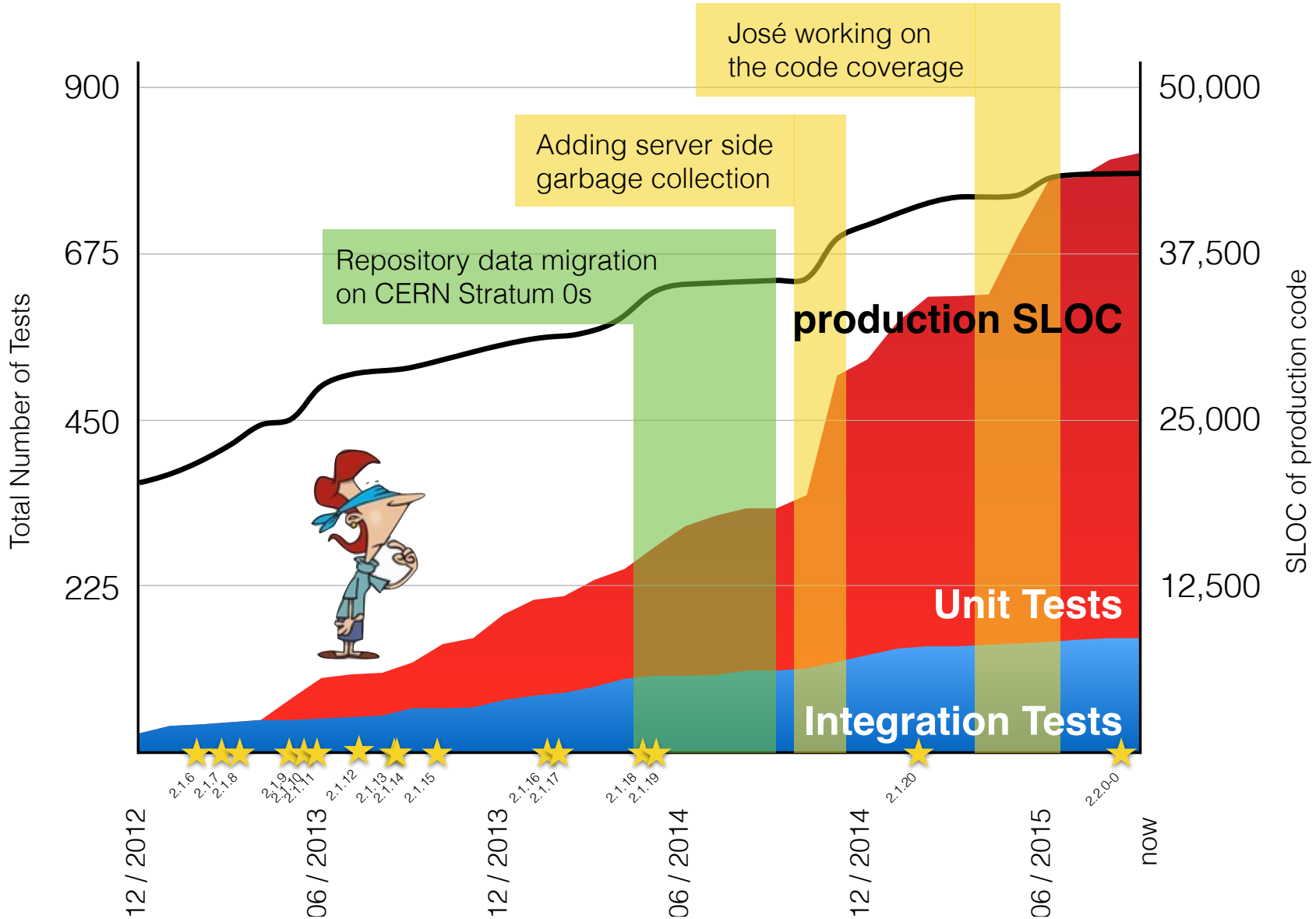
→ Testing and continuous code cleanup

Order by

CVM-689
Code Housekeeping

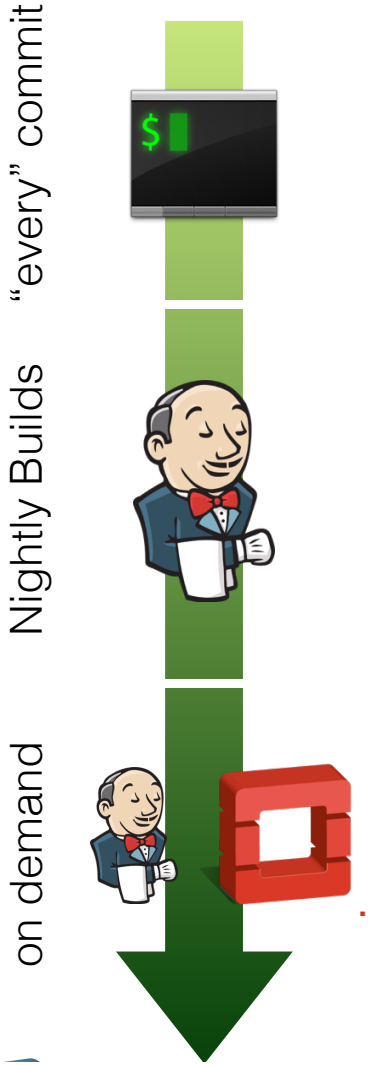
Sub-Tasks

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



- Development unit test set
  - ~ 600 tests (< **2 minutes**)
  - **39%** code coverage
- Full multi-platform unit test set
  - ~ 660 tests (~ **15 minutes**)
  - runs on **all build platforms**
  - **43%** code coverage
- Multi-platform integration and unit test test
  - ~ 820 tests in **various configurations** (~ 16 hours)
  - **80%** code coverage (73% integration + unit tests)

Source: Meusel

Build Targets

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

Summary

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