



# Update on Software Environment Preservation with CernVM

J Blomer, G Ganis

CERN PH-SFT

8 June 2015  
DPHEP workshop, CERN

# An ecosystem for virtualization in HEP



CernVM  
File system

General purpose  
[Software Distribution System](#)



CernVM  
Software Appliance



CernVM  
Co-Pilot  
[Computing infrastructure toolkit](#)



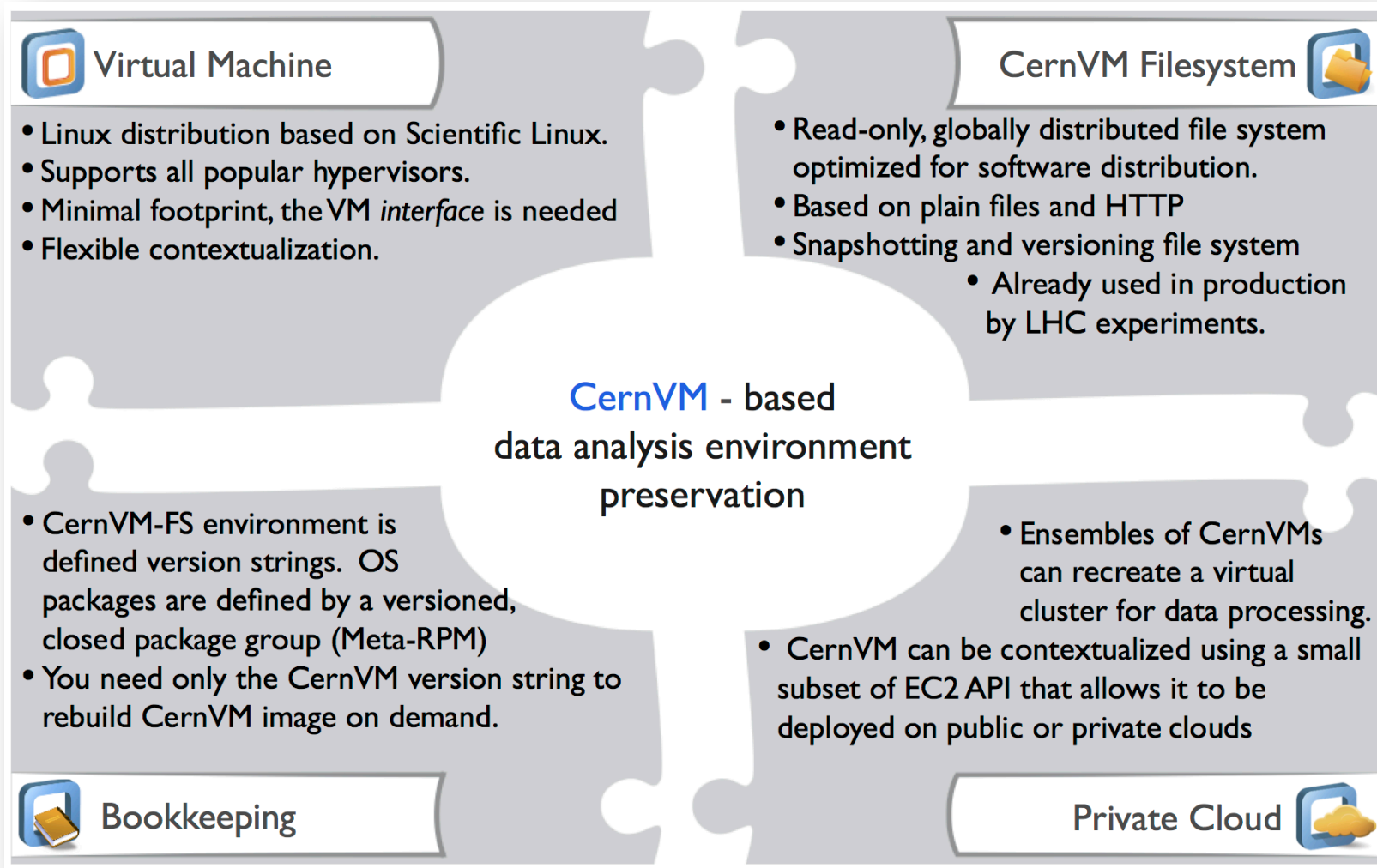
CernVM  
Online  
[Contextualization tool](#)

# CernVM and Long-Term Analysis Preservation

- Solution to process data w/ any given software version and regenerated VM developed
  - [CHEP 2012](#), [DPHEP 2013](#), [CHEP 2013](#) (NA61 test)
- Main ingredients
  - CernVM-FS embedded versioning (à la time machine)
  - Well defined and versioned *recipes* to regenerate VMs



# Components of the CernVM blueprint





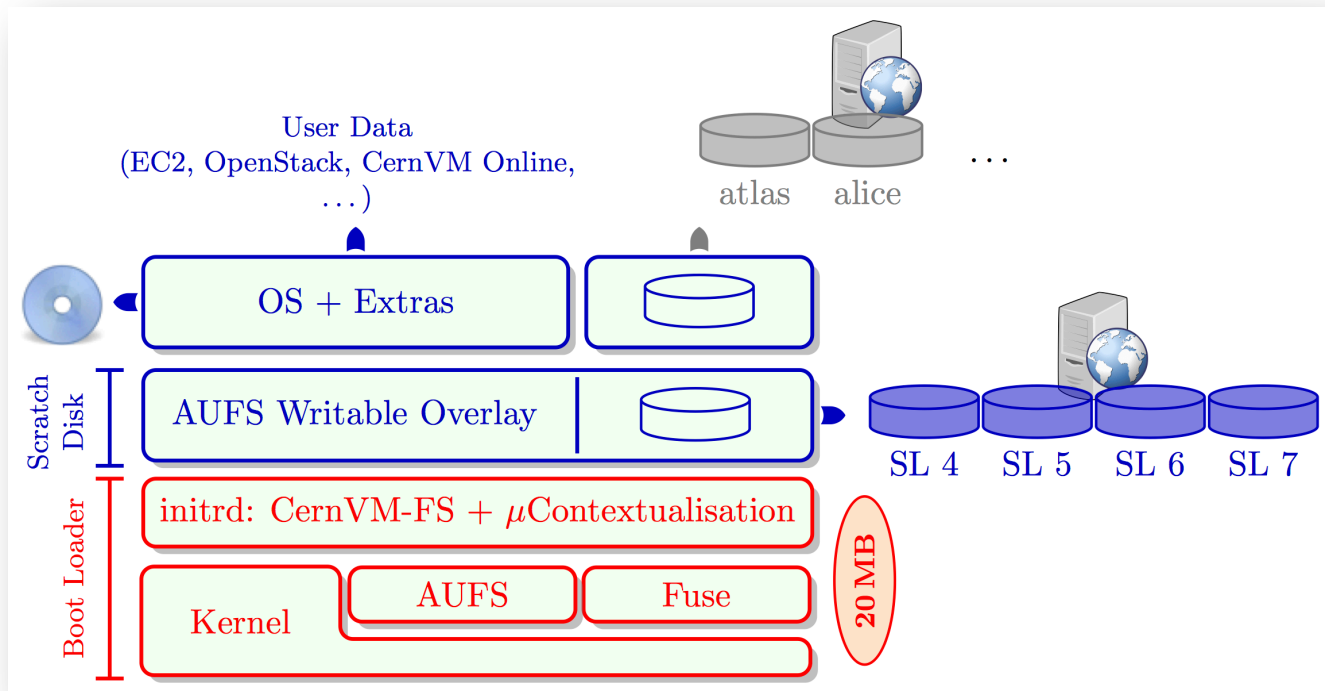
# CernVM and Long-Term Analysis Preservation

- Solution to process data w/ any given software version and regenerated VM developed
  - [CHEP 2012](#), [DPHEP 2013](#), [CHEP 2013](#) (NA61 test)
- Main ingredients
  - CernVM-FS embedded versioning (à la time machine)
  - Well defined and versioned *recipes* to regenerate VMs
- Novel *bootloader technology* (CernVM 3) opened new ways
  - Operating System template on CernVM-FS
    - Versioned and preserved as experiment software
  - Potential to recreate any older Linux environment
  - Support / Integration with container / Docker technology

This talk

# The *bootloader* technology

$\mu$ CernVM bootloader ( $\approx 20$  MB) + OS delivered by CernVM-FS ( $\approx 100$  MB)



J Blomer et al., "Micro-CernVM: slashing the cost of building and deploying Virtual Machines", [2014 J Phys Conf Ser 513 032009](https://arxiv.org/abs/1307.1101) (CHEP 2013)

# The *bootloader* technology (2)

- Stable, **Long-Term Supported** Linux kernel (3.10, 3.18)
- **Operating System templates** on CernVM-FS
  - Installed with OS package manager
- After boot, basically a **'chroot'** to the chosen template
  - Writable area provided by a overlay FS (AUFS)
  - Like in *live CD* trial installations
- Possibility to **choose a flavor and its version**
  - Exploiting CernVM-FS versioning features
- Default template **SL 6**
  - Available: **SL 4, SL 5, SL 7** (prototype)
  - Not limited to RedHat-based systems

# *Demonstrators for* Software Environment Preservation

- **ALEPH**
  - Software was last validated on Scientific Linux 4
- **CMS Open Data Pilot**
  - Released data were taken in 2011
  - Requires Scientific Linux 5 from same period
- CernVM 3 is an enabling technology for the CERN Open Data Project
  - <http://opendata.cern.ch>



# ALEPH: regenerating SLC4

```

/cvmfs/cernvm-slc4.cern.ch      (OS template)
/cvmfs/aleph.cern.ch           (ALEPH software)
/cvmfs/sft.cern.ch/lcg/external/cernlib (CERNlib)

```

Instances

| Instance Name                           | Image Name   | IP Address     | Size                                      | Keypair | Status | Task | Power State | Uptime            | Actions  |
|---|--------------|----------------|---|---------|--------|------|-------------|-------------------|--|
| <input type="checkbox"/> cernvm-aleph01 | ucernvm-slc4 | 188.184.134.26 | m1.small   2GB RAM   1 VCPU   20.0GB Disk | -       | Active | None | Running     | 3 months, 2 weeks | <input type="button" value="Create Snapshot"/> <input type="button" value="More"/> |

**cernvm-aleph01**  
VM machine on  
CERN openstack

```

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.

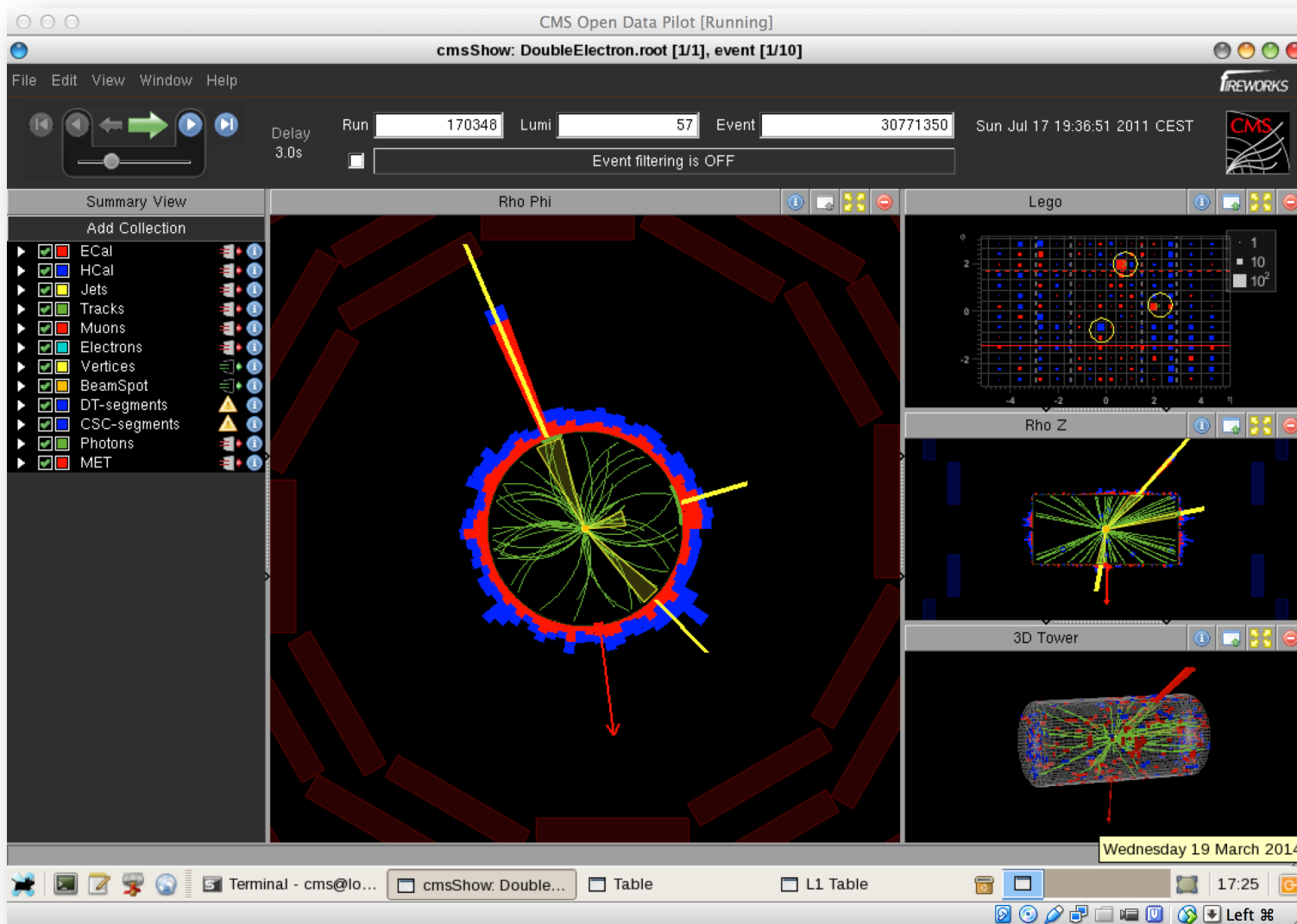
```

# CMS Open Data Pilot context

Provide an easy-to-use virtual machine with CMS computing environment for CMS Open Data

- Data
  - Frozen set, accessed remotely (XRootD, DPHEP portal)
- Software
  - Frozen CMS software framework
    - CMSSW.4.2.8.patch7
  - Complete analysis environment required (compile + run)
  - Requires Scientific Linux 5 compatible VM
- Virtual machine, user interface
  - Graphical environment
  - Easy-to-install, easy-to-use

# CMS Open Data Pilot: event display



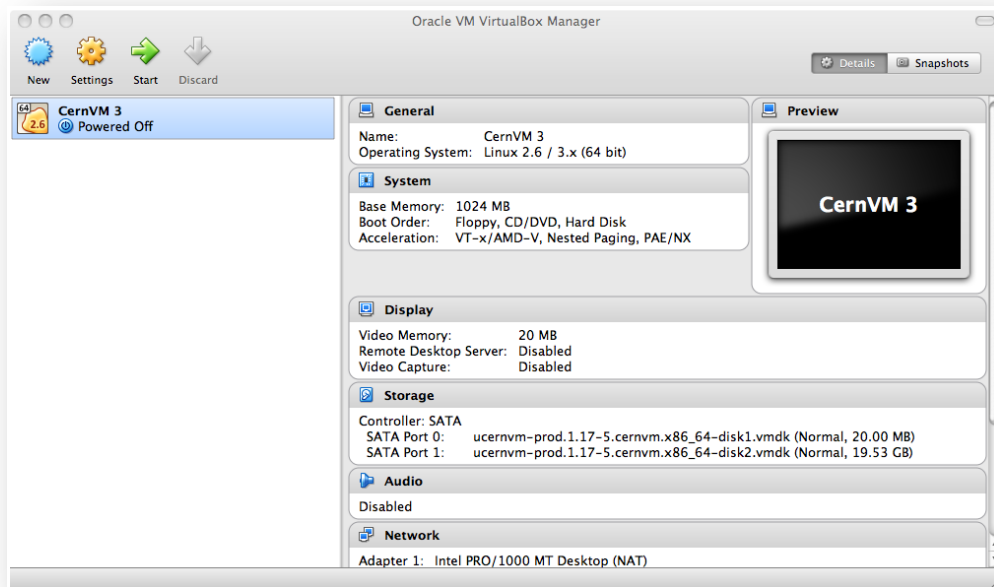
# CMS Open Data Pilot: deployment

|   |                |
|---|----------------|
| <code>/cvmfs/cernvm-slc5.cern.ch</code> | (OS template)  |
| <code>/cvmfs/cms.cern.ch</code>         | (CMS software) |

- Deployed as OVF/OVA bundle<sup>1</sup>
  - Open specification for bundling VMs, stable since 2009
  - OVA: tarball with a hard disk image and an XML specification

Easy auto-installation  
in [VirtualBox](#)

Same solution for  
LHCb @ OpenData  
<http://opendata.cern.ch>



<sup>1</sup> Open Virtualization Format / Open Virtual Appliance: <http://www.dmtf.org/standards/ovf>



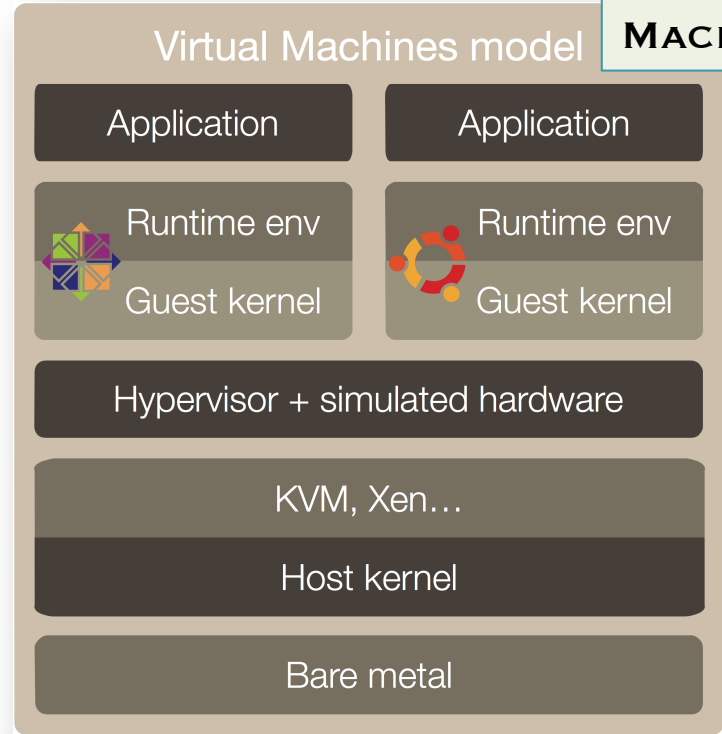
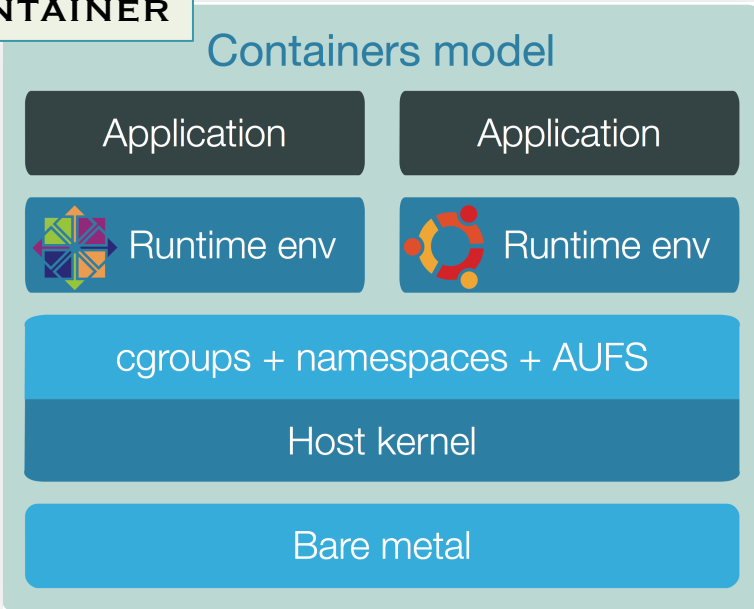


# Containers

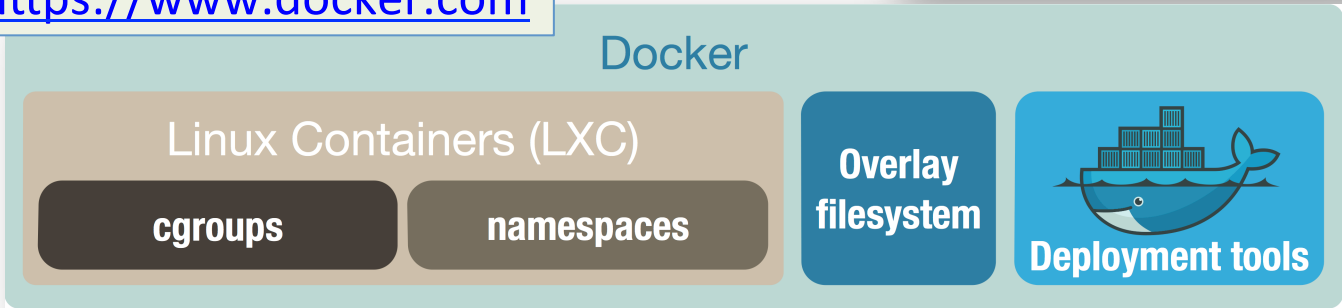
# Containers and Docker

CONTAINER

VIRTUAL MACHINE



<https://www.docker.com>



DOCKER

# Lot of interest ...

- ... as way to optimize use of resources
- Kubernetes, Mesos, Openstack, ...
- All LHC experiments gave a try
  - Basic approach: pack everything in the container
    - Huge containers
  - ALICE investigating the concept of pilot container
  - CernVM-FS efficient integration/support essential

# CernVM as container

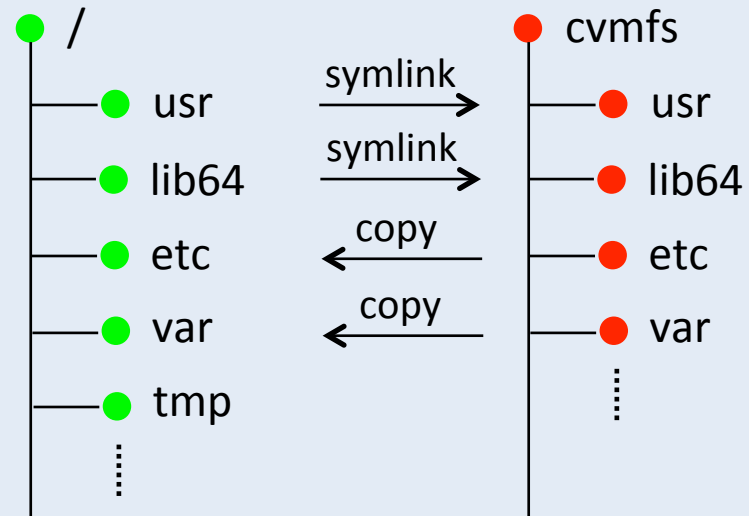
## Roadmap

- 1: Make CernVM-FS repositories available inside a container
- 2: Writable overlay via union file system

## Options for CernVM-FS (1)

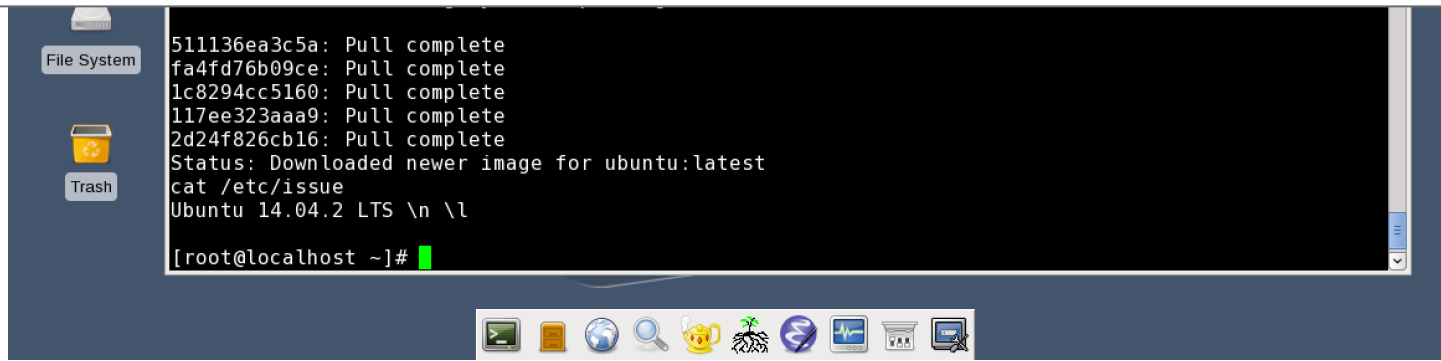
- 1: Fuse, mapped from host
  - Shared cache
  - Promising prototype in progress
    - → [D Berzano's CHEP 2015](#)
- 2: Using Parrot CernVM-FS
  - Pure user-space (ptrace)
  - User-mode CernVM demonstrated
    - → [G Ganis's CHEP 2015](#)
  - Ongoing work on performance and stability

## Root file system (/) layout



# CernVM 3.4: full support for `lxc` and Docker

Allows to run other linux flavours in CernVM, e.g. Ubuntu



Facilitates sharing of multi-core virtual machines

[cernvm-fork](#): create containers with isolated environments

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

CernVM-FS cache shared across containers



# CernVM-docker exercise: ALEPH/SLC4

1. CernVM 3.4 with **generic context**
  1. E.g. CernVM-SLC4 from the marketplace
2. Start the **docker daemon** inside the VM
3. Run the **ucvm-slc4-docker container**, bind mounting the OS template and software dirs
4. Compile and run the ALEPH software in the container

Details at <http://cernvm.cern.ch/portal/cvm3-slc4-docker>

# CernVM-docker SLC4: step 1,2

The screenshot shows the CernVM Marketplace interface. At the top, there is a navigation bar with links for About, Dashboard, Marketplace, Documentation, Downloads, and Publications. Below the navigation bar, the main heading is "CernVM Marketplace" with a sub-heading "Pick one of the public contextualization information and pair you CernVM instance." The main content area displays several contextualization options. One option is highlighted, showing a Docker icon and the text "CernVM 3 for getting to SLC4 via Docker." To the right, the details for the "CernVM-SLC4" package are shown, including the author "ganis", tags "slc4 +- sl4 +- dphep +- ", and access level "Open". There are also buttons for "Pair", "Clone", and other actions.

```
Terminal - userslc4@localhost:~
File Edit View Terminal Go Help
[userslc4@localhost ~]$ ls /cvmfs/cernvm-slc4.cern.ch
cvm3 new_repository update-packs
[userslc4@localhost ~]$ ls /cvmfs/aleph.cern.ch
Linux etc i386_redhat42 reference shared
[userslc4@localhost ~]$ sudo service docker start
[sudo] password for userslc4:
Starting Docker: [ OK ]
[userslc4@localhost ~]$
```

# CernVM-docker SLC4: step 3

```
docker run -i -t \  
  -v /cvmfs/cernvm-slc4.cern.ch/cvm3:/cernvm \  
  -v /cvmfs/aleph.cern.ch:/cvmfs/aleph.cern.ch \  
  -v /cvmfs/sft.cern.ch:/cvmfs/sft.cern.ch \  
  -v $HOME/local/aleph/MIT:/MIT \  
  ucvm-slc4-docker /init
```

- Bind mount relevant volumes from /cvmfs
- Also working area, for convenience
- The init script contains the instruction how to setup the system



# CernVM-docker SLC4: step 3,4

```
$ ./local/etc/start-slc4.sh
*****
*** Welcome to SLC4 in CernVM+Docker ***
*****

Scientific Linux SL release 4.9 (Beryllium)

root@ucvm-slc4:/MIT $ ls
Compile.sh  MIT.cards  MIT.mk      MIT_MC.cards  aleph-env.sh  zm4000_43.edir
MIT.F       MIT.job    MIT.o       ZM4000.43.AL  get_file.sh
```

```
root@ucvm-slc4:/MIT $ source aleph-env.sh
root@ucvm-slc4:/MIT $ ./Compile.sh
root@ucvm-slc4:/MIT $ ./MIT

-----
                A L P H A      126.22      14.52.47  07/06/15
                    running on LINUX
                    ALEPHLIB Version  316.20
-----

0_QUIBOS_ Init BOS with 5000000 words working space
  *** ALFMT*** BANKAL.FMT on unit 77 contains 1331 bank names
OBREADC----- NAME.....NR
+-----+
```

# Summary

- CernVM ecosystem: natural potential for software environment preservation
  - CernVM-FS embedded versioning of SW and OS
- The *bootloader* technology has simplified the way we can regenerate the target VM
  - Demonstrated with ALEPH and CMS Open Data Pilot
- Investigating the use of container technology
  - Potential to add smoothness to the whole process
  - Promising tests with ALEPH/SLC4
- On the way for a full provenance system