



CernVM: Overview and Future Plans

D Berzano, J Blomer, P Buncic, I Charalampidis, G Ganis,
A Harutyunyan, G Lestaris, B Segal
jblomer@cern.ch

CERN PH-SFT

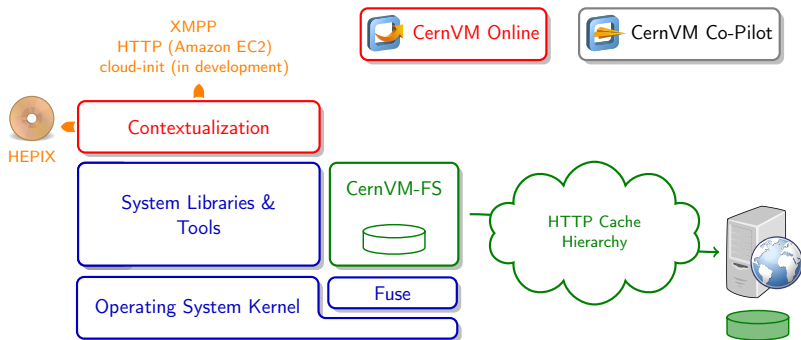
- ① Introduction to the CernVM Virtual Appliance
- ② CoPilot: Connecting CernVMs to the Grid
- ③ CernVM Use Cases
- ④ μ CernVM: Slashing the Cost of Building and Deploying VMs

① Introduction to the CernVM Virtual Appliance

② CoPilot: Connecting CernVMs to the Grid

③ CernVM Use Cases

④ μ CernVM: Slashing the Cost of Building and Deploying VMs



- 1 Generic Stable Platform
- 2 CernVM-FS for Software Delivery
- 3 Flexible Contextualization
- 4 Co-Pilot Task Queue Connection
- 5 CernVM Online for Context Bookkeeping

Application

Libraries

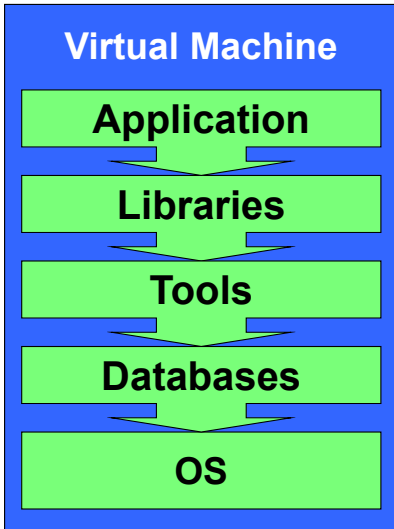
Tools

Databases

OS

Hardware

- Traditional worker node:
General purpose operating system and libraries
- Independently developed and deployed
- Different life cycles
- Applications break if any layer changes
- Difficult to support multiple applications



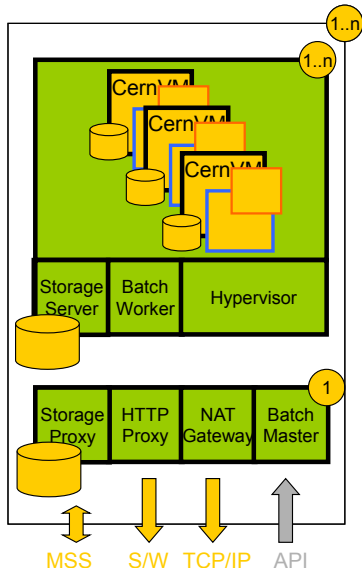
- Application dependencies are analyzed
- Virtual machine is defined by these dependencies
- Results in a minimal operating system
- Dependencies can be versioned and stored as recipes

Simple API

- Instantiate + Contextualize
- Terminate
- List instances, list images

Contextualization plug-ins


- Credentials (ssh, X.509)
- Condor head & batch services
- Squid server
- XrootD storage proxy
- CernVM-FS
- Monitoring & directory service agents
- Network configuration & tuning



Dashboard | CernVM Online

<https://cernvm-online.cern.ch/dashboard>

Logged in as icharala | [Log out](#)


CernVM Online

[About](#)
[Dashboard](#)

Commands
















- Dashboard
- Pair an instance
- Create Context

Recent Definitions

- PrivateCloud-MakeflowPool
- PrivateCloud-Worker
- Private-Desktop
- PrivateCloud4ALL
- PrivateCloud-CatalogServer



Dashboard

Your context definitions

Name	ID	Operations
 PrivateCloud-CatalogServer	22b13425ef7244c4b7de60dbbca64728	 Remove  Use as template
 PrivateCloud-Worker	a632e7ad3ca64774ac9318fc9e086640	 Remove  Use as template
 PrivateCloud-MakeflowPool	f7e9ba92a55146119ac3cd6141fd957	 Remove  Use as template
 Private-Desktop	79383bc71d7a4760a8397cc2e8d2a2ed	 Remove  Use as template
 PrivateCloud4ALL	37bbd937803b407e8fbd469d65fac74c	 Remove  Use as template

Create new context

Your virtual machines


Machine	CernVM	Context	Operations
 128.141.235.19 (b5609c63-a2bb-4ba5-901a-25066733df9c)	2.1.0	PrivateCloud-CatalogServer	 Unmanage

Pair an instance of CernVM

© Copyright CERN 2012 - PH Department - SFT - CernVM Software appliance

Context Definition | CernVM

← → ↻ https://cernvm-online.cern.ch/context/new



CernVM Online

Logged in as icharala | Log out

About

Dashboard

Commands

Dashboard

Pair an instance

Create Context

Recent Definitions

PrivateCloud-MakeflowPool

PrivateCloud-Worker

Private-Desktop

PrivateCloud4ALL

PrivateCloud-CatalogServer

Context template

Please fill the following parameters and click create in order to create a new virtual machine context definition

General

Context name:

PrivateCloud-Monitor

Description:

A monitoring entity for my private cloud.

☐ Make this context visible on the public lists

☒ Enable CernVM Agent infrastructure

Secret key:

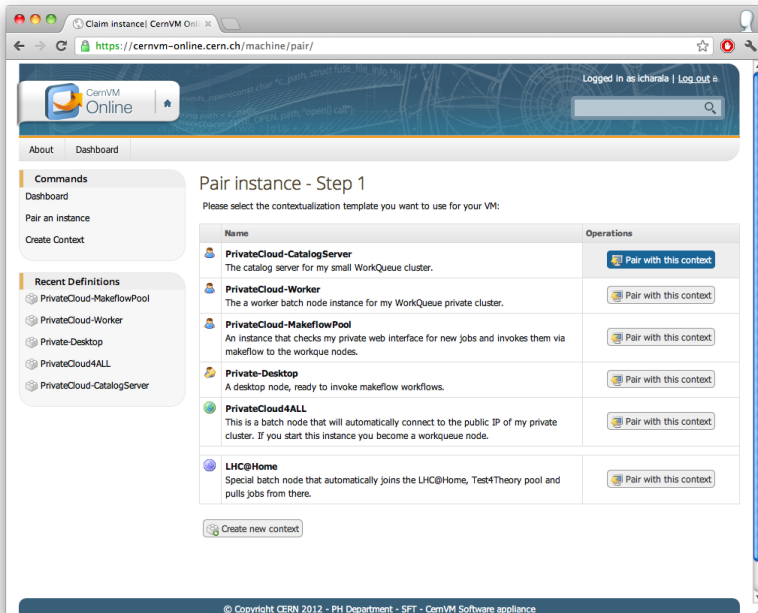
protected-file

☒ Protect this context with a secret key

Repository

Users

Contextualization



Claim instance| CernVM Online

https://cernvm-online.cern.ch/machine/pair/

Logged in as icharala | Log out

CernVM Online

About Dashboard

Commands

Dashboard

Pair an instance


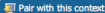

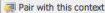

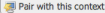

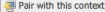

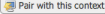

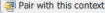
Create Context

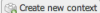
Recent Definitions

- PrivateCloud-MakeflowPool
- PrivateCloud-Worker
- Private-Desktop
- PrivateCloud4ALL
- PrivateCloud-CatalogServer

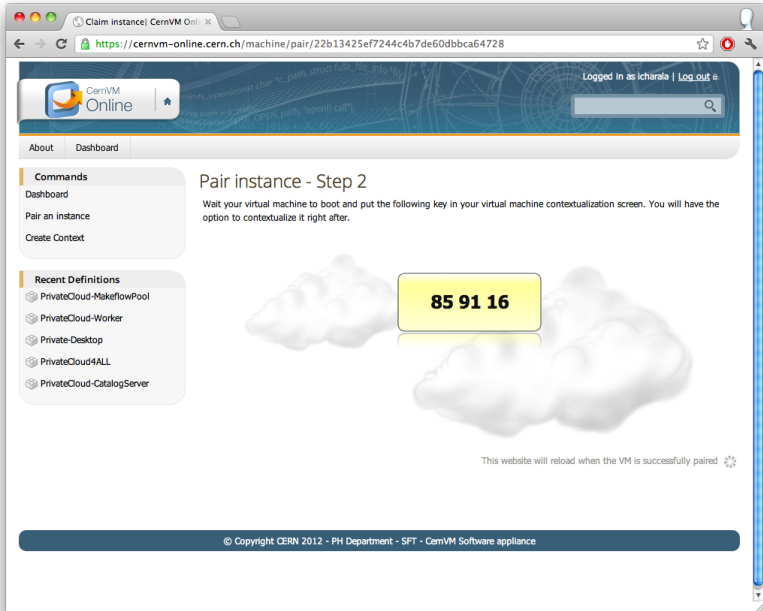
Pair instance - Step 1

Please select the contextualization template you want to use for your VM:

Name	Operations
 PrivateCloud-CatalogServer The catalog server for my small WorkQueue cluster.	
 PrivateCloud-Worker The a worker batch node instance for my WorkQueue private cluster.	
 PrivateCloud-MakeflowPool An instance that checks my private web interface for new jobs and invokes them via makeflow to the workqueue nodes.	
 Private-Desktop A desktop node, ready to invoke makeflow workflows.	
 PrivateCloud4ALL This is a batch node that will automatically connect to the public IP of my private cluster. If you start this instance you become a workqueue node.	
 LHC@Home Special batch node that automatically joins the LHC@Home, Test4Theory pool and pulls jobs from there.	



© Copyright CERN 2012 - PH Department - SFT - CernVM Software appliance



Claim instance | CernVM Online

https://cernvm-online.cern.ch/machine/pair/22b13425ef7244c4b7de60dbbca64728

Logged in as icharala | Log out

CernVM Online

About Dashboard

Commands

Dashboard

Pair an instance

Create Context

Recent Definitions

- PrivateCloud-MakeflowPool
- PrivateCloud-Worker
- Private-Desktop
- PrivateCloud4ALL
- PrivateCloud-CatalogServer

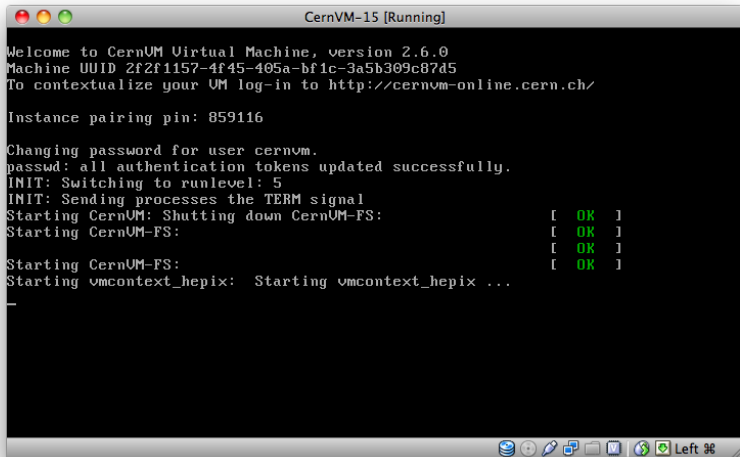
Pair instance - Step 2

Wait your virtual machine to boot and put the following key in your virtual machine contextualization screen. You will have the option to contextualize it right after.

85 91 16

This website will reload when the VM is successfully paired

© Copyright CERN 2012 - PH Department - SFT - CernVM Software appliance



```

CernVM-15 [Running]
Welcome to CernVM Virtual Machine, version 2.6.0
Machine UUID 2f2f1157-4f45-405a-bf1c-3a5b309c87d5
To contextualize your VM log-in to http://cernvm-online.cern.ch/

Instance pairing pin: 859116

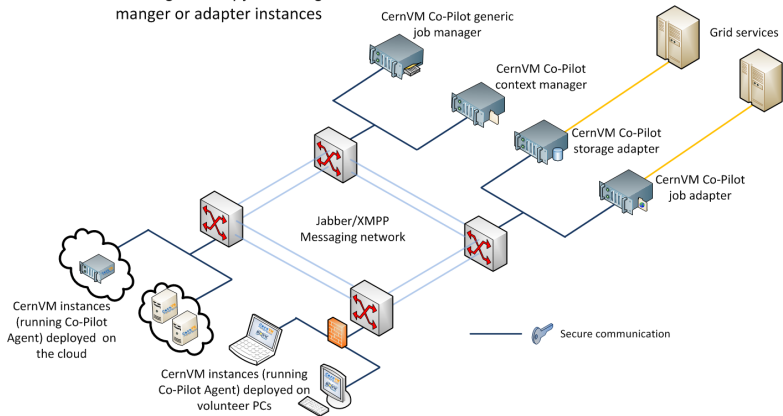
Changing password for user cernvm.
passwd: all authentication tokens updated successfully.
INIT: Switching to runlevel: 5
INIT: Sending processes the TERM signal
Starting CernVM: Shutting down CernVM-FS:           [ OK ]
Starting CernVM-FS:                                  [ OK ]
Starting CernVM-FS:                                  [ OK ]
Starting CernVM-FS:                                  [ OK ]
Starting vmcontext_hepix: Starting vmcontext_hepix ...

```

- ① Introduction to the CernVM Virtual Appliance
- ② CoPilot: Connecting CernVMs to the Grid
- ③ CernVM Use Cases
- ④ μ CernVM: Slashing the Cost of Building and Deploying VMs

Challenges: untrusted environment, transparency for grid users

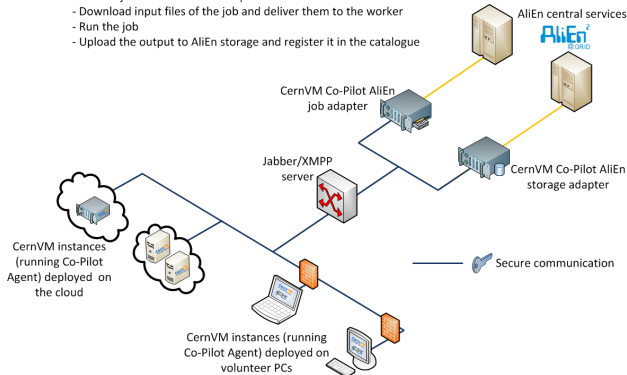
Message driven system can be scaled up
in case of high load by just adding new
manger or adapter instances



Source: Harutyunyan

Co-Pilot is used to:

- Get the job from AliEn central queue
- Download input files of the job and deliver them to the worker
- Run the job
- Upload the output to AliEn storage and register it in the catalogue



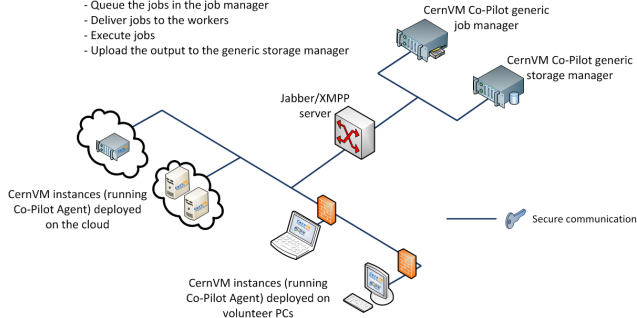
Source: Harutyunyan

A similar adapter has been developed for PanDA/ATLAS

Very simple, file based I/O queue

Co-Pilot is used to:

- Queue the jobs in the job manager
- Deliver jobs to the workers
- Execute jobs
- Upload the output to the generic storage manager



Source: Harutyunyan

- Runs volunteer computing for CERN theory group in “auto-pilot mode” for over a year
- ... more in the following slides

- ① Introduction to the CernVM Virtual Appliance
- ② CoPilot: Connecting CernVMs to the Grid
- ③ CernVM Use Cases
- ④ μ CernVM: Slashing the Cost of Building and Deploying VMs

- **Portable Development and Analysis Environment**
Identical environment for development and execution of HEP applications.
- **Virtual Analysis Facility**
Medium-size data sets, fast response time:
CernVM + PROOF (+...), see Dario's talk
- **Better Use of High-Level Trigger Farms**
HLT nodes are delicate resources.
VMs provide a non-intrusive means to do offline computing on HLT nodes.
Instant switch from online to offline computing and vice versa.
- **Volunteer Computing**
Part of experiment's outreach program.
Use of computing resources of volunteers without the need to change or port applications.
- **Preservation of Historic Data Processing Environment**
Historic data are useless without the ability to interpret them. The strong versioning in CernVM and CernVM-FS allows for respawning a historic *data processing environment*.

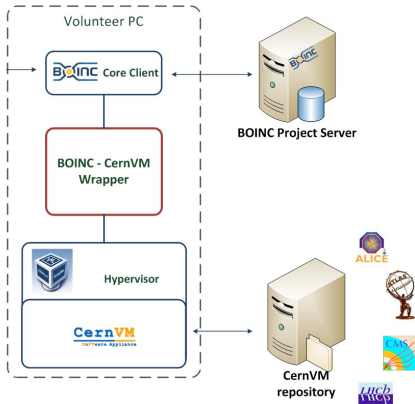
- Portable Development and Analysis Environment
Identical environment for development and execution of HEP applications.
- Virtual Analysis Facility
Medium-size data sets, fast response time:
CernVM + PROOF (+...), see Dario's talk
- Better Use of High-Level Trigger Farms
HLT nodes are delicate resources.
VMs provide a non-intrusive means to do offline computing on HLT nodes.
Instant switch from online to offline computing and vice versa.
- Volunteer Computing
Part of experiment's outreach program.
Use of computing resources of volunteers without the need to change or port applications.
- Preservation of Historic Data Processing Environment
Historic data are useless without the ability to interpret them. The strong versioning in CernVM and CernVM-FS allows for respawning a historic *data processing environment*.

- Portable Development and Analysis Environment
Identical environment for development and execution of HEP applications.
- Virtual Analysis Facility
Medium-size data sets, fast response time:
CernVM + PROOF (+...), see Dario's talk
- Better Use of High-Level Trigger Farms
HLT nodes are delicate resources.
VMs provide a non-intrusive means to do offline computing on HLT nodes.
Instant switch from online to offline computing and vice versa.
- Volunteer Computing
Part of experiment's outreach program.
Use of computing resources of volunteers without the need to change or port applications.
- Preservation of Historic Data Processing Environment
Historic data are useless without the ability to interpret them. The strong versioning in CernVM and CernVM-FS allows for respawning a historic *data processing environment*.

- Portable Development and Analysis Environment
Identical environment for development and execution of HEP applications.
- Virtual Analysis Facility
Medium-size data sets, fast response time:
CernVM + PROOF (+...), see Dario's talk
- Better Use of High-Level Trigger Farms
HLT nodes are delicate resources.
VMs provide a non-intrusive means to do offline computing on HLT nodes.
Instant switch from online to offline computing and vice versa.
- Volunteer Computing
Part of experiment's outreach program.
Use of computing resources of volunteers without the need to change or port applications.
- Preservation of Historic Data Processing Environment
Historic data are useless without the ability to interpret them. The strong versioning in CernVM and CernVM-FS allows for respawning a historic *data processing environment*.

Monte-Carlo simulations, parameter tuning

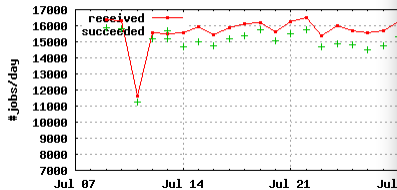
First BOINC project using virtual machines



Numbers

- At any point in time
600–700 VMs connected
- Overall: $9 \cdot 10^{11}$ events created

Jobs rates



Source: Harutyunyan

Geographic distribution, 28. May 2013, 2500 distinct IPs

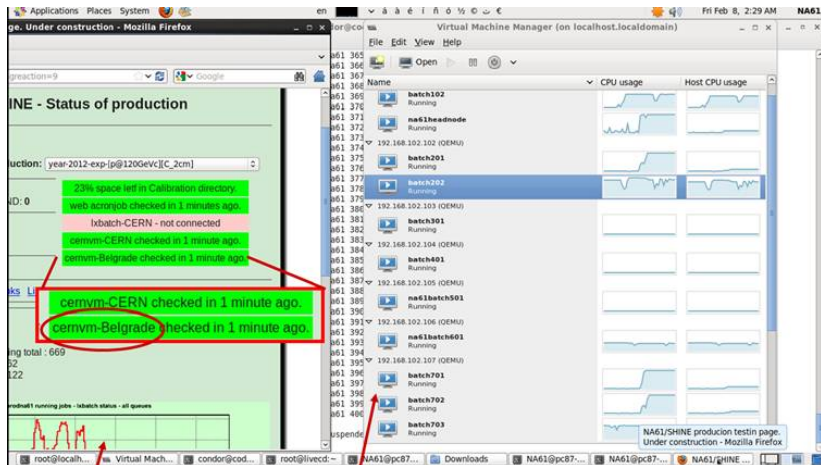


- **Portable Development and Analysis Environment**
Identical environment for development and execution of HEP applications.
- **Virtual Analysis Facility**
Medium-size data sets, fast response time:
CernVM + PROOF (+...), see Dario's talk
- **Better Use of High-Level Trigger Farms**
HLT nodes are delicate resources.
VMs provide a non-intrusive means to do offline computing on HLT nodes.
Instant switch from online to offline computing and vice versa.
- **Volunteer Computing**
Part of experiment's outreach program.
Use of computing resources of volunteers without the need to change or port applications.
- **Preservation of Historic Data Processing Environment**
Historic data are useless without the ability to interpret them. The strong versioning in CernVM and CernVM-FS allows for respawning a historic *data processing environment*.



NA61 Production Jobs in Belgrade

Integration of a CernVM cloud with a data provenance system

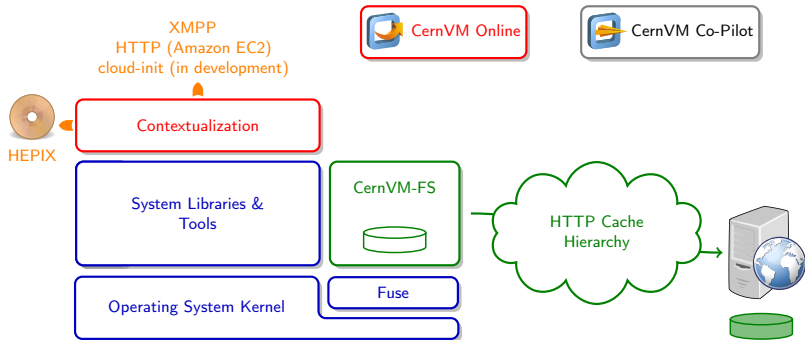


virtual machines manager on Belgrade's cluster.

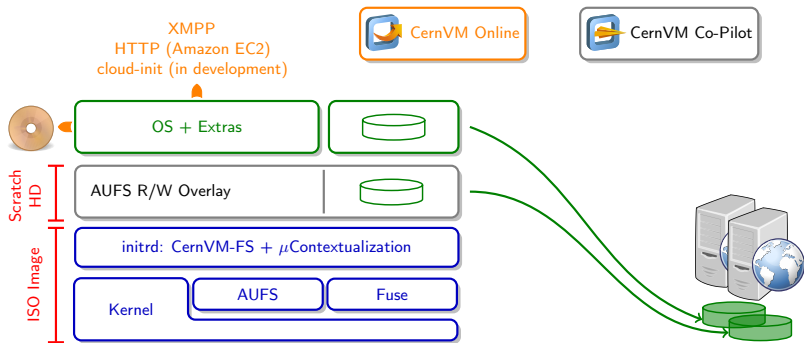
Belgrade's cluster registered and monitored on
NA61/SHINE production page (*under construction*)
<http://dmaletic.web.cern.ch/dmaletic/cgi-bin/na61prod>

- ① Introduction to the CernVM Virtual Appliance
- ② CoPilot: Connecting CernVMs to the Grid
- ③ CernVM Use Cases
- ④ μ CernVM: Slashing the Cost of Building and Deploying VMs

Classic CernVM



- Uniform and portable environment for physics data processing
- Minimal operating system derived from application dependencies
- Easy to maintain and to distribute

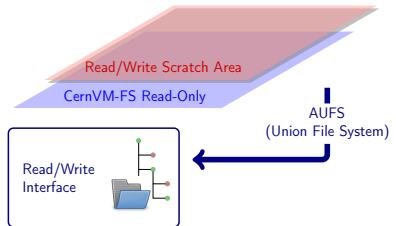
μ CernVM

Idea: Operating system on CernVM-FS

Instead of 400 MB hard disk image: 10 MB ISO image + 100 MB cache.

- *Not* a LiveCD, *not* a diskless node
- ⇒ Operating System on Demand

- AUFS well-maintained kernel module
- < 5 % performance loss (untar)
- Some use cases faster due to CernVM-FS meta-data handling



-
- Root file system created in early user space by init ramdisk script
 - **Difficulty:** shutdown and proper unfolding of the stack
Required a few twists to SL6 halt script and CernVM-FS

```

MicroCernVM

ISOLINUX 4.06 2012-10-23 ETCD Copyright (C) 1994-2012 H. Peter Anvin et al
early console in decompress_kernel

Decompressing Linux... Parsing ELF... done.
Booting the kernel.
[    1.277254] acpiphp_ibm: ibm_acpiphp_init: acpi_walk_namespace failed
[    1.459769] sd 2:0:0:0: [sda] Assuming drive cache: write through
[    1.460199] sd 2:0:0:0: [sda] Assuming drive cache: write through
[    1.461662] sd 2:0:0:0: [sda] Assuming drive cache: write through

* Welcome to micro-CernVM
* Beta release 1.2

[INF] Setting up environment... check
[INF] Loading predefined modules... check
[INF] Starting networking... check
[INF] Mounting root filesystem... check
[INF] Starting CernVM File System... check

mount: mount point /proc/bus/usb does not exist
      Welcome to Scientific Linux
Starting udev: _

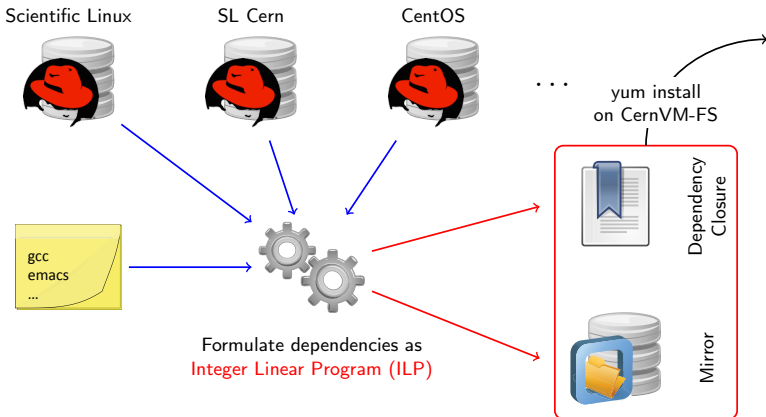
```

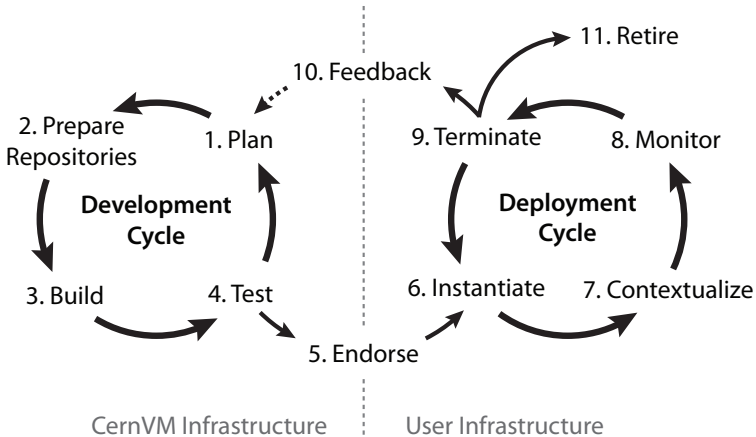
Build Process: Scientific Linux on CernVM-FS

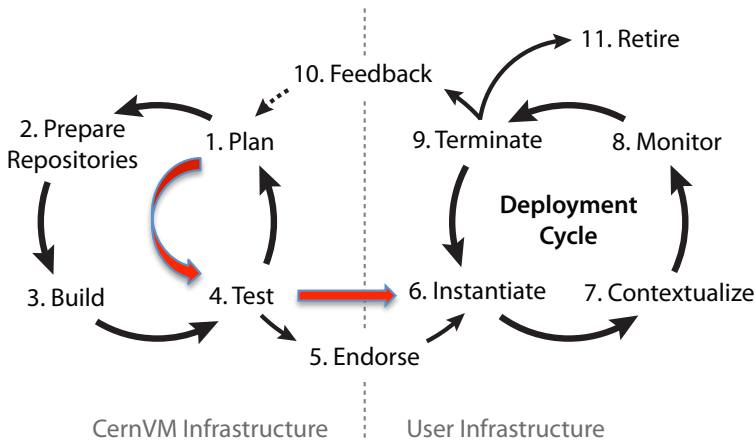
Maintenance of the repository **must not** become a Linux distributor's job

But: should be reproducible and well-documented

Idea: Automatically generate a **fully versioned, closed** package list
from an unversioned "shopping list" of packages
(*Standard package managers are not designed for preservation!*)







Avoids: Image Building

Solves: Image Distribution

Options for updating: **stay, diverge, rebase**

- For a virtualized infrastructure: development environment is production environment
- By encapsulating the runtime environment in light-weight virtual machines, applications can be sent to volunteers and “interested citizens”
- Strongly-versioned, VM encapsulated runtime environments facilitate long-term data preservation
- The CernVM appliance aims at avoiding image proliferation through
 - A small base image defined by application dependencies
 - CernVM-FS to distribute experiment software
 - Flexible contextualization means
- μ CernVM avoids the need to distribute hard disk images altogether

μ CernVM Technology Preview:

<http://cernvm.cern.ch/portal/ucernvm>