CernVM

SoFTware Appliance

CernVM R&D Project

Portable Analysis Environments using Virtualization technology

Predrag Buncic (CERN/PH-SFT)

Overview

- Introduction
  - Software packaging and distribution problem
  - Job execution environment problem
  - Long term software preservation and accountability
- Can virtualization technology help us solve all this problems at the same time?
- CernVM Project
  - Status and perspective
- Conclusions

CernVM
SoFTware Appliance

Problem #1

- Software @ LHC
  - Millions of lines of code
    - C++, C, Fortran, python, perl…
  - Different packaging and software distribution models
    - Complicated software installation/update/configuration procedure
  - Needs an expert knowledge and lots of patience by the end user (physicist) to keep up
    - And many just give up and use shared resources like lxplus
- Recent technology trends
  - Moore’s law will still hold but instead of faster CPU we will get multi and many cores in the same package

- **How can profit from (multi and many core) CPU power of user workstations and use them at least as development platform?**

Problem #2

- Distributed nature of data processing (Grid)
  - meant to be a solution and not part of the problem
- Heterogeneous set of resources
  - In terms of OS, QoS, configuration
  - Independent administration domains and evolution cycles
  - Potential source of many runtime errors
- Very complex and difficult to maintain middleware software stack
  - Trying to create illusion of Grid as homogeneous resource
  - Long and slow validation and certification process focused on middleware and not on applications, difficult to do OS upgrades

- **How to decouple application and infrastructure lifecycles and assure a homogeneous job execution environment compatible with one in which application was developed?**

Problem #3

- LHC experiments will run for many years
  - Hardware and software environment will inevitably (possibly dramatically) change in the lifespan of the experiments
  - Physicists like to constantly change their software (algorithms) but do not like when external changes in infrastructure require changes in the their applications
- In order to assure capability of the experiments some time in the future to (re)analyze their entire data sample we need to
  1. Conserve and transform all data on the media that can be affectively used in that point in time
  2. Be able to use exactly the same software as it was used when original files were created and be able to modify algorithms and apply new algorithms them on the same old data
- **How to preserve experiment software and keep it usable and accountable over many years?**

An example: NA49/NA61

- NA49 experience
  - Software written in 1994
  - data taken in 1996-2002 are still being used for physics analysis and publications,
  - most of the raw data lost: obsolete medium (Sony tapes) and tape drives not supported by CERN/ IT → new reconstruction is not possible
  - reconstruction/simulation software written in C, C++ and Fortran
  - very difficult (manpower/time) to maintain reconstruction/simulation software in a view of frequent changes of operating systems and compilers
- The above significantly limits the possibility to use the NA49 data for future analysis

An example: NA49/NA61

- NA61 (successor of NA49 in terms of hardware and software) started data taking in 2009, and recorded data (100 TB) for 9 different reactions
  - the data taking should continue over the next 5 years with up 40 reactions (1000 TB) to be recorded,
  - based on the NA49 experience we expect that physics interest in the NA61 data analysis
  - may continue for many years
- To avoid difficulties that NA49 had
  - raw data should be preserved over the whole analysis period
  - software should be made independent of underlying OS and physical hardware
  - the reconstruction/simulation results obtained in 2009 should be reproducible in 20 years from now
- In such case term “data preservation” means a lot more more than conserving 4-vectors but represents a practical problem seen by the running experiment

Can virtualization technology help us to solve all these problems at the same time?

CernVM

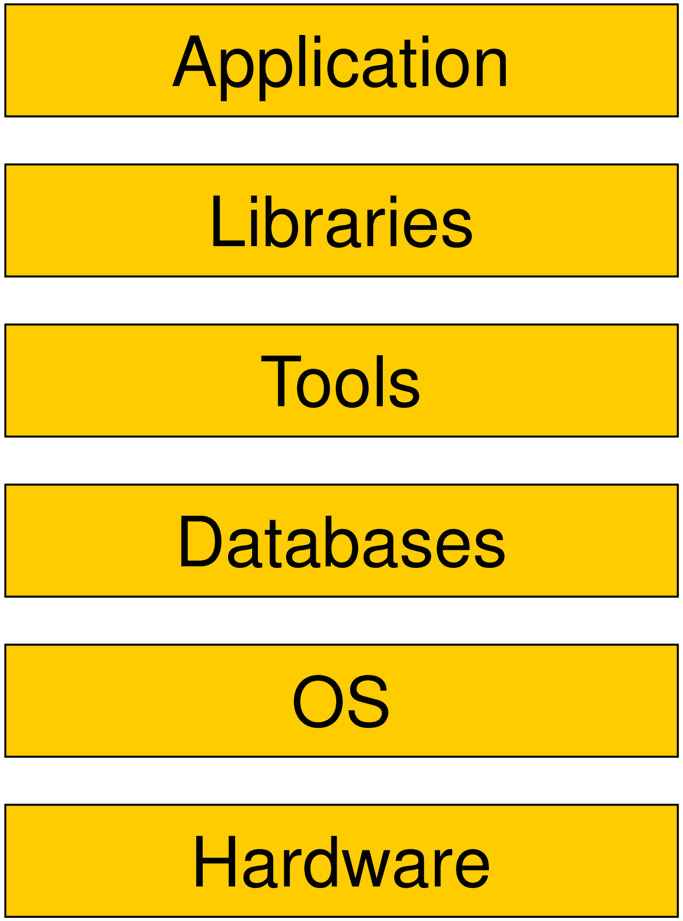
SoFTware Appliance

CernVM Project

- R&D Project in PH Department (WP9), started in 2007, 4 years
- Aims to provide a complete, portable and easy to configure **user environment** for developing and running LHC data analysis locally and on the Grid independent of physical software and hardware platform (Linux, Windows, MacOS)
  - Code check-out, edition, compilation, local small test, debugging, …
  - Grid submission, data access…
  - Event displays, interactive data analysis, …
  - Suspend, resume…

- Decouple application lifecycle from evolution of system infrastructure
- Reduce effort to install, maintain and keep up to date the experiment software
- Web site: http://cernvm.cern.ch

Application Deployment Today



Application

Libraries

Tools

Databases

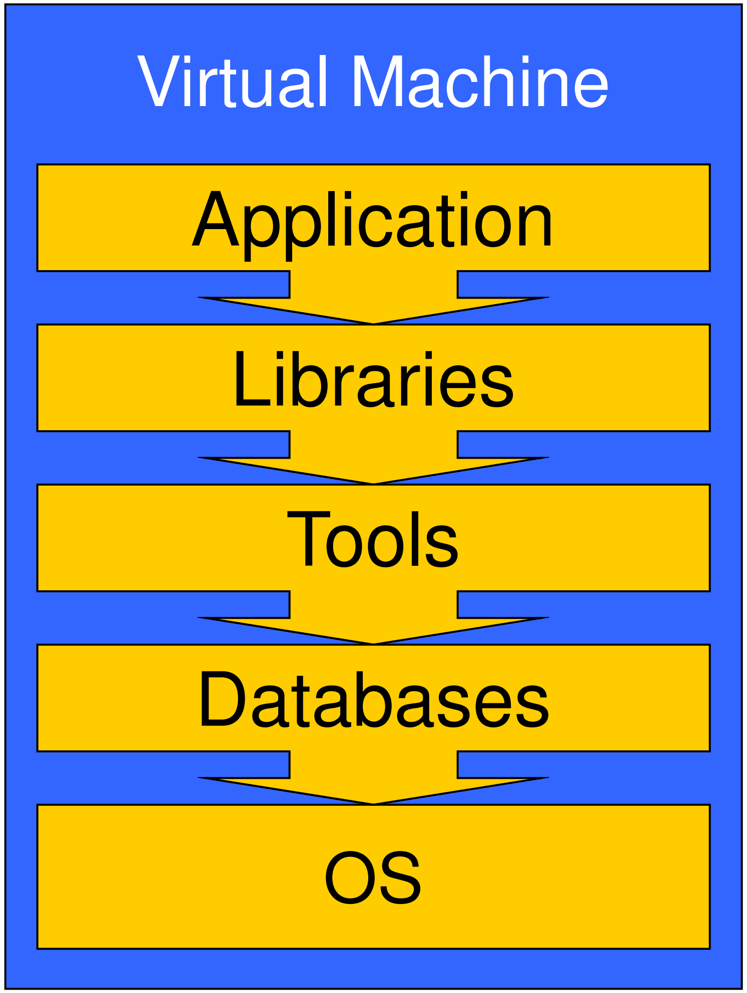
OS

Hardware

- Traditional model
  - Horizontal layers
  - Independently developed
  - Maintained by different groups
  - Different lifecycle
- Application is deployed on top of the stack
  - Breaks if any layer changes
  - Needs to be certified every time when something changes
  - Results in deployment nightmare
- Incompatible with “long term stability and accountability” requirement

CernVM
SoFTware Appliance

CernVM Approach

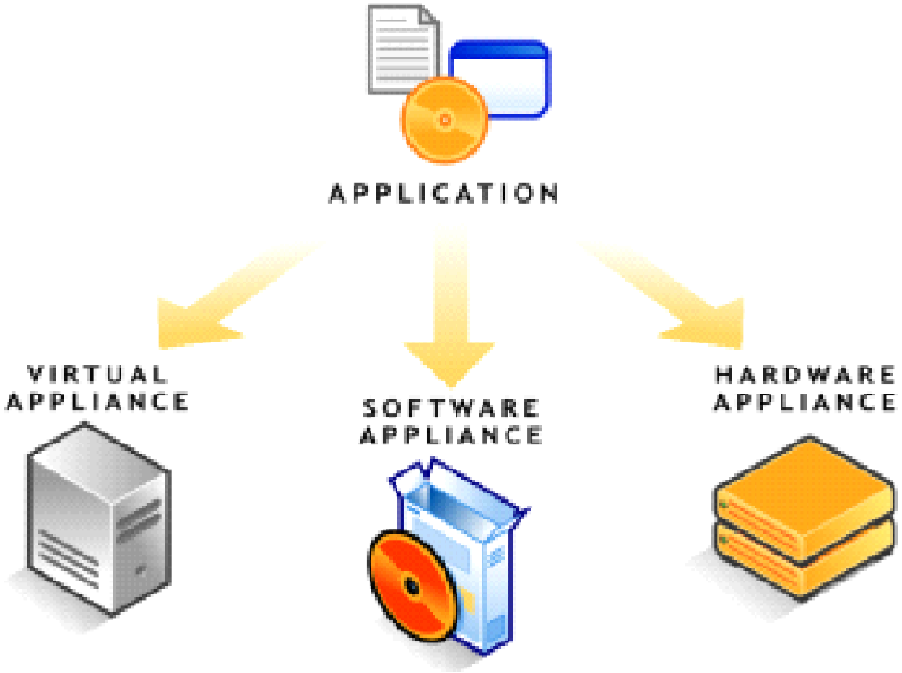


- Vertical integration
  - Starting from the application
  - Finding its requirements and dependencies
  - Adding required tools and libraries
  - Building minimal OS
  - Bundling all this into Virtual Machine image
- Application is in control
  - Virtual Machines can be versioned like applications
- Virtualization technology
  - Enables decoupling of OS running in VM from the physical hardware

CernVM
SoFTware Appliance

Virtualizing Applications

Starting from experiment software…



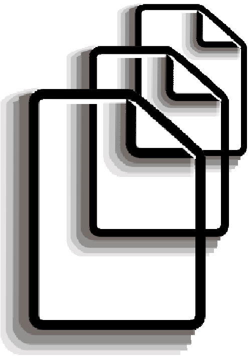
…ending with a custom Linux specialised for a given task

rBuilder™

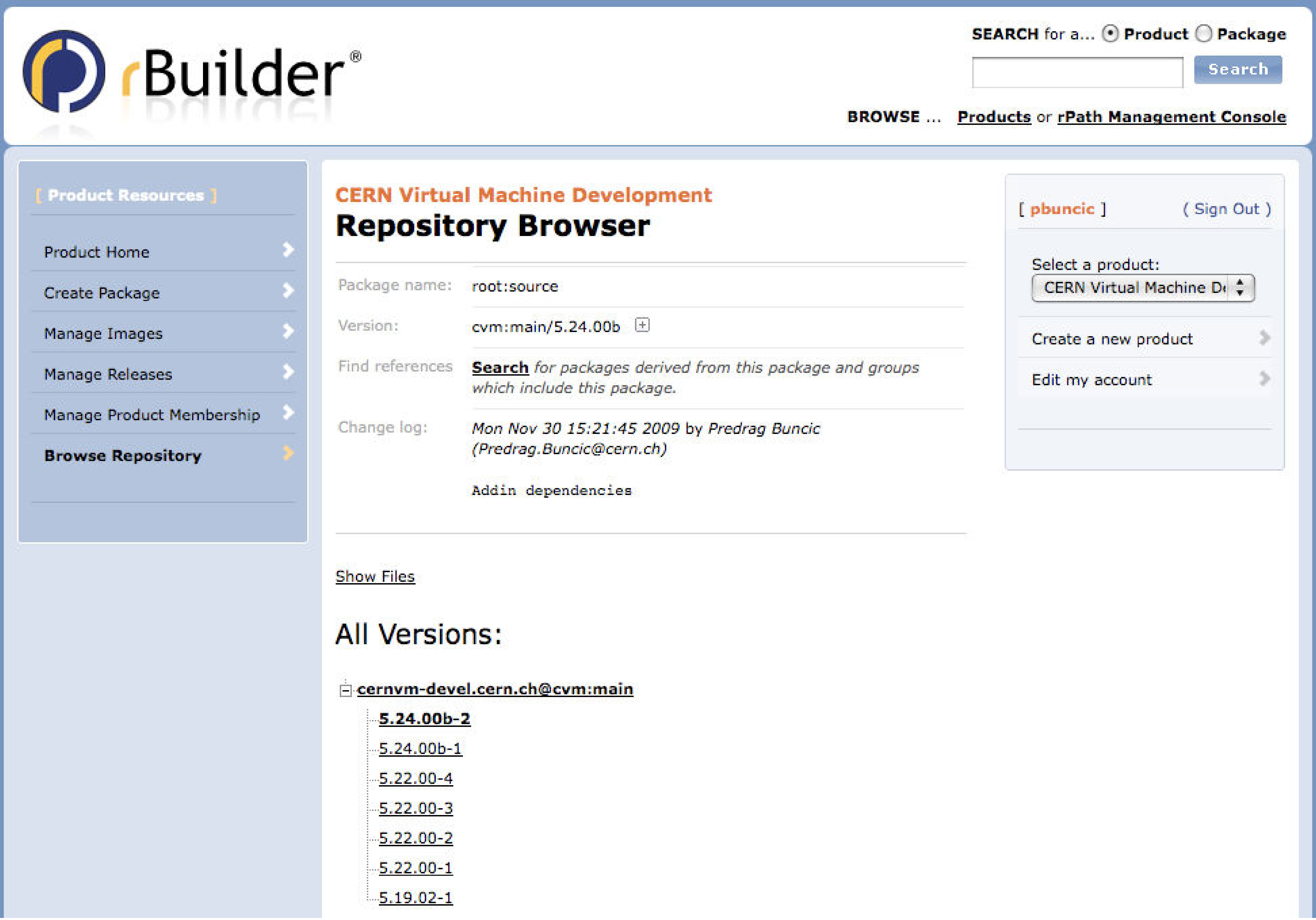
- Installable CD/DVD
- Stub Image
- Raw Filesystem Image
- Netboot Image
- Compressed Tar File
- Demo CD/DVD (Live CD/DVD)
- Raw Hard Disk Image
- Vmware ® Virtual Appliance
- Vmware ® ESX Server Virtual Appliance
- Microsoft ® VHD Virtual Apliance
- Xen Enterprise Virtual Appliance
- Virtual Iron Virtual Appliance
- Parallels Virtual Appliance
- Amazon Machine Image
- Update CD/DVD
- Appliance Installable ISO

Conary Package Manager

my linux
1 part = love & devotion
2 parts - hamsterdance
1 part = final fantasy X
1/3 cup - methyl ethyl ketone
1 tsp. - expresso beans (whole)
1. Find what you need
2. Build your recipe
3. Cook it!



Every build and every file installed on the system
is automatically versioned and accounted for in database



CERN Virtual Machine Development

Repository Browser

Package name: root:source

Version: cvm:main/5.24.00b ⊞

Find references: **Search** *for packages derived from this package and groups which include this package.*

Change log: *Mon Nov 30 15:21:45 2009* by *Predrag Buncic (Predrag.Buncic@cern.ch)*

```
Addin dependencies
```

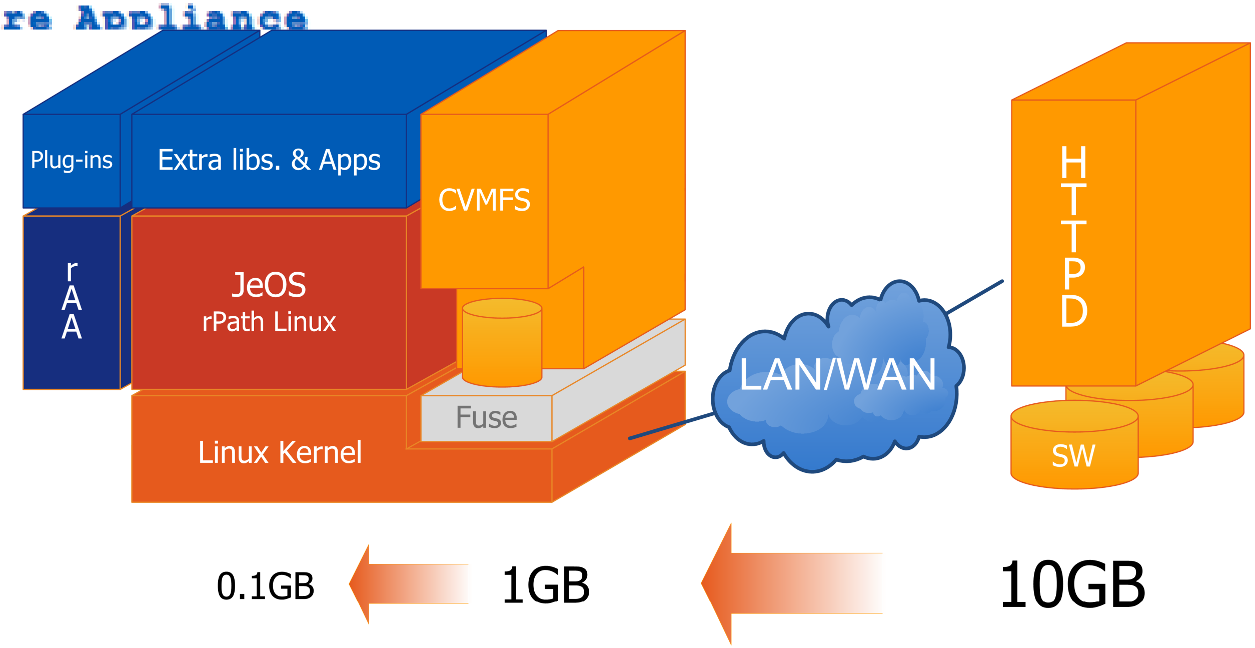
Show Files

All Versions:

⊟ **cernvm-devel.cern.ch@cvm:main**
- **5.24.00b-2**
- 5.24.00b-1
- 5.22.00-4
- 5.22.00-3
- 5.22.00-2
- 5.22.00-1
- 5.19.02-1

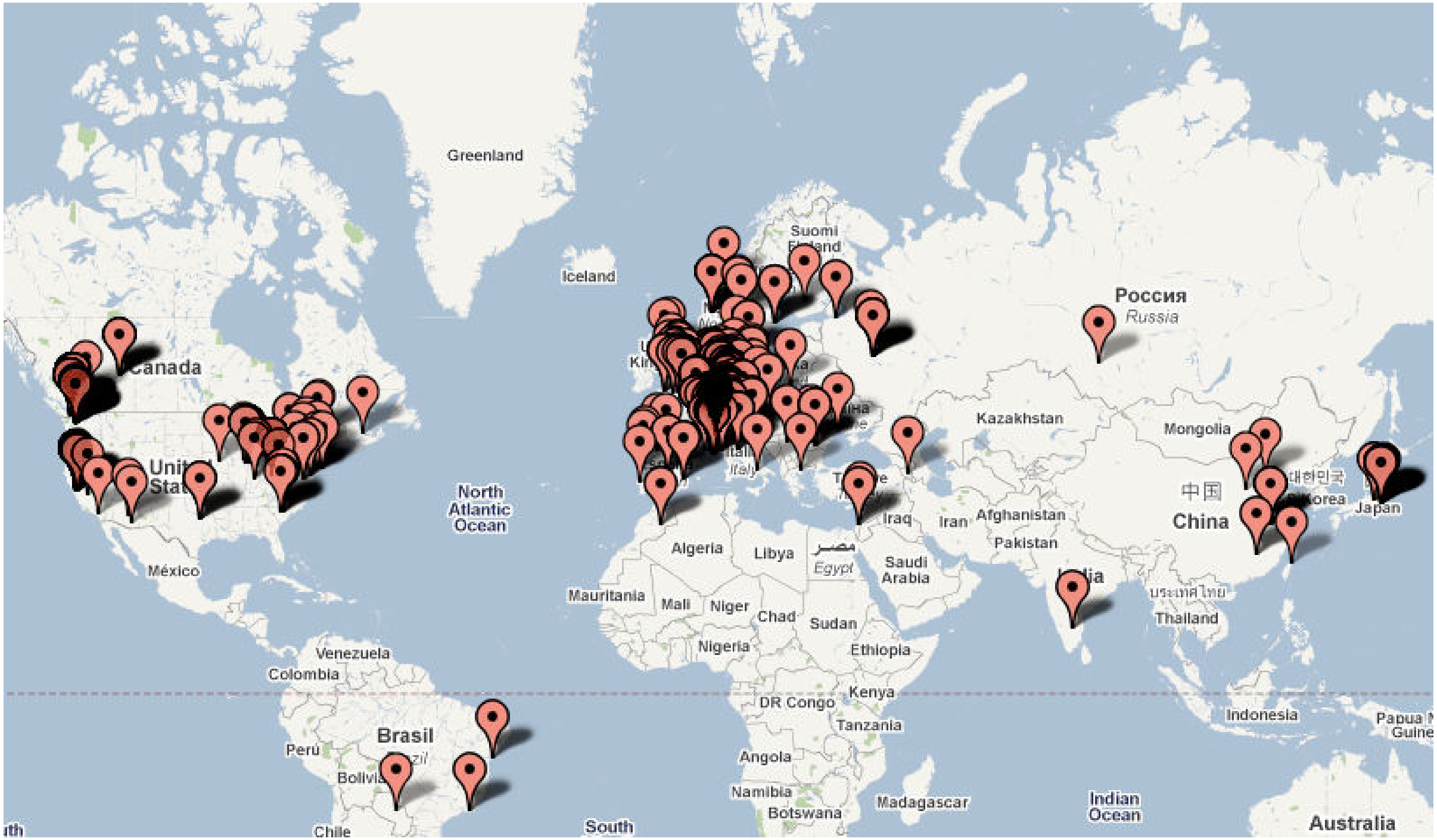
CernVM
SoFTware Appliance

CernVM architecture



- The experiments are packaging gigabytes of code per release
  - but really use only fraction of it at runtime
- CernVM downloads only what is really needed and puts it in the cache
  - Does not require persistent network connection (offline mode)
  - Minimal impact on the network
- Defines common **platform** that can be used by all experiments/projects

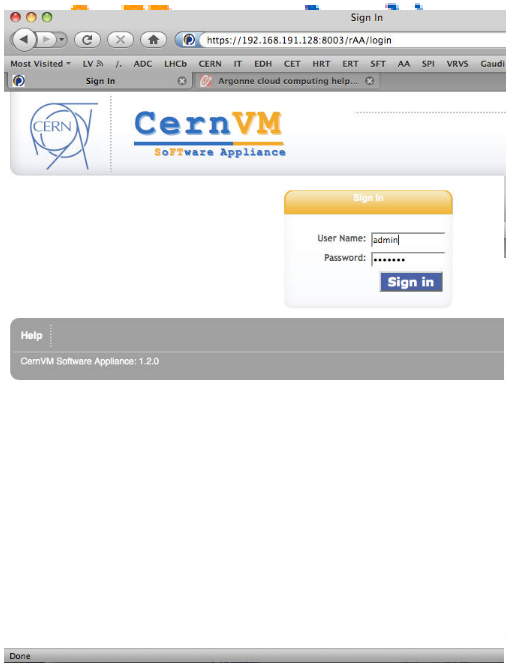
CernVM users



~1500 different IP addresses

CernVM

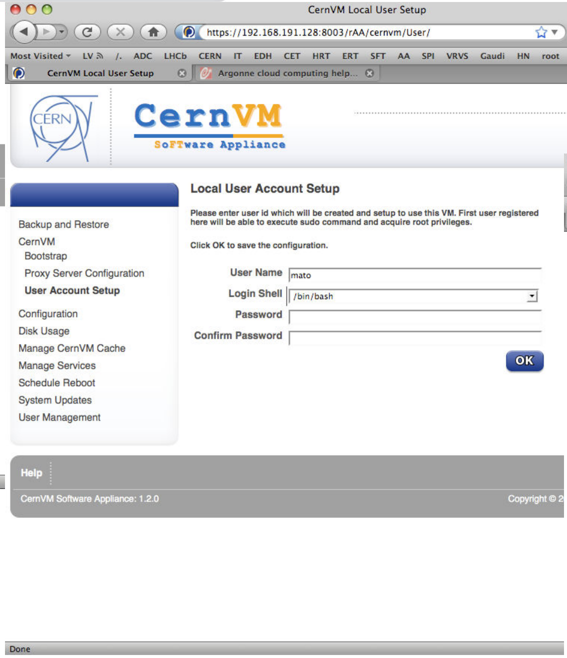
As easy as 1,2,3



User Name: admin
Password: •••••••

CernVM Software Appliance: 1.2.0

1. Login to Web interface



Local User Account Setup

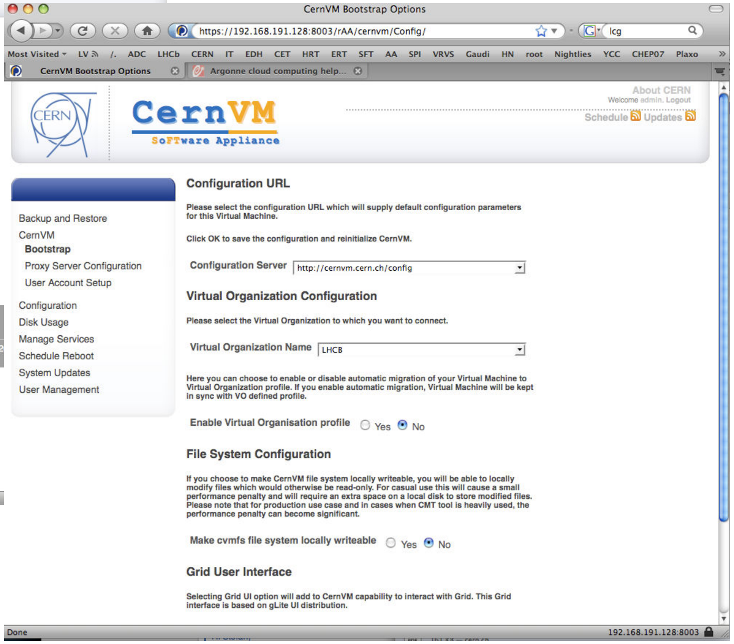
Please enter user id which will be created and setup to use this VM. First user registered here will be able to execute sudo command and acquire root privileges.

Click OK to save the configuration.

User Name: mato
Login Shell: /bin/bash
Password:
Confirm Password:

CernVM Software Appliance: 1.2.0 Copyright © 2

2. Create user account



Configuration URL

Please select the configuration URL which will supply default configuration parameters for this Virtual Machine.

Click OK to save the configuration and reinitialize CernVM.

Configuration Server: http://cernvm.cern.ch/config

Virtual Organization Configuration

Please select the Virtual Organization to which you want to connect.

Virtual Organization Name: LHCB

Here you can choose to enable or disable automatic migration of your Virtual Machine to Virtual Organization profile. If you enable automatic migration, Virtual Machine will be kept in sync with VO defined profile.

Enable Virtual Organisation profile ○ Yes ● No

File System Configuration

If you choose to make CernVM file system locally writeable, you will be able to locally modify files which would otherwise be read-only. For casual use this will cause a small performance penalty and will require an extra space on a local disk to store modified files. Please note that for production use case and in cases when CMT tool is heavily used, the performance penalty can become significant.

Make cvmfs file system locally writeable ○ Yes ● No

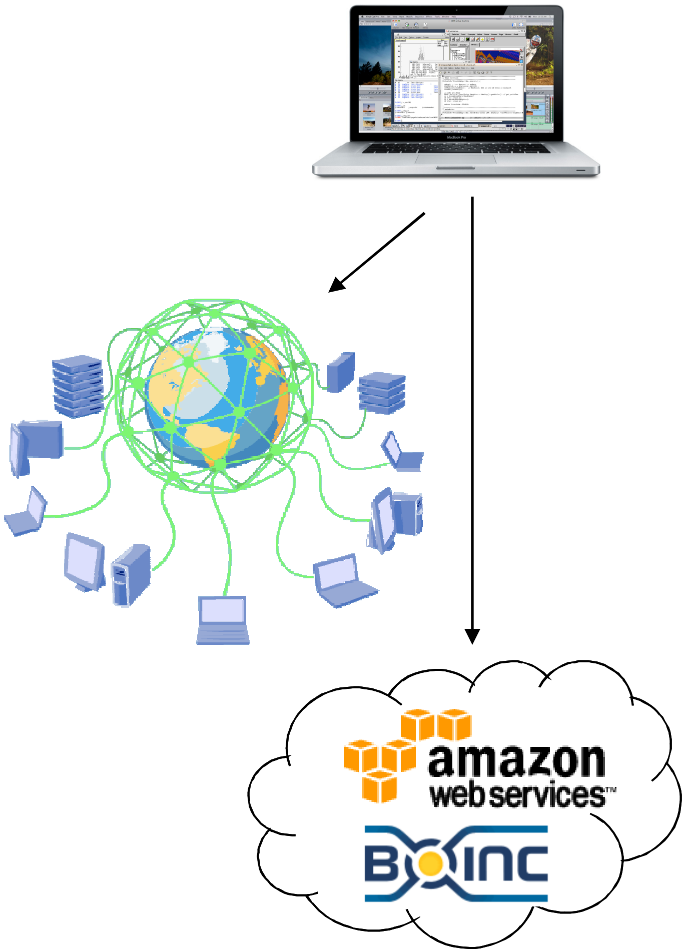
Grid User Interface

Selecting Grid UI option will add to CernVM capability to interact with Grid. This Grid interface is based on gLite UI distribution.

3. Select experiment, appliance flavor and preferences

CernVM
SoFTware Appliance

CernVM@Grid



- CernVM as job hosting environment on Cloud/Grid
  - Ideally, users would like to run their applications on the grid (or cloud) infrastructure in exactly the same conditions in which they were originally developed
- CernVM already provides development environment and can be deployed on cloud (EC2)
  - One image supports all four LHC experiments
  - Easily extensible to other communities
- In this model, it is essential to allow experiments run their own virtual images (=applications) on Grid/Cloud

CernVM

SoFTware Appliance

Conclusions

- Long term data preservation without equivalent mechanism for software preservation is of limited usefulness to the experiments
- Technology exists today to aide this process and is likely to stay with us
  - Virtualization enables Cloud computing and that seems to be where all industry is going
- Using virtualization to achieve long term software preservation requires to
  - Make end users (physicist) feel comfortable to use virtual machine on their desktop/laptop to carry out daily work on developing and using experiment software
  - Convince resource providers to accept this technology and run virtual machines in places of batch jobs that they run today
  - Give application central role and derive OS from it while keeping track of all system components in an external version control system
- CernVM Software Appliance
  - Used by ATLAS, LHCb and to lesser extent ALICE and CMS
  - Already deployable on the cloud (EC2, Nimbus)
  - Can be deployed on unmanaged infrastructure like BOINC
  - Based on innovative second generation package manager
  - Provides versioning of every build and build products installed on the system
  - It is sufficient to record only a full version string of top level package group to be able to reconstruct a full VM