Status and Roadmap of CernVM

G Ganis
For the CernVM team

16 April 2015
CHEP 2015, Okinawa, Japan
An ecosystem for virtualization in HEP

CernVM File system

General purpose Software Distribution System

CernVM Software Appliance

Contextualization tool

CernVM Online

CernVM Co-Pilot

Computing infrastructure toolkit
CernVM Virtual Appliance

Complete and portable environment for developing and running HEP applications

Use cases:

– Development environment
  • On desktops, laptops

– IaaS Clouds
  • Public, commercial, hybrid

– Volunteer Computing
  • See Tuesday’s talks and today I. Charalampidis’ talk

– Long-Term Analysis Preservation

– Outreach & Education
The bootloader technology

µCernVM bootloader (≈15 MB) + OS delivered by CernVM-FS (≈100 MB)

## Status of the building blocks

### Boot-loader (image)
- 3.10 Linux Kernel (LTS) + “guest additions”
- Root file system (/) from CernVM-FS + writable overlay
- Contextualization to choose OS, proxy, ...
- Stable (<5 security fix releases)

### OS on CernVM-FS
- Based on Scientific Linux
  - Configuration and tuning, Contextualization, Extra packages
- Strongly versioned: meta package determines all package versions
- Default repository: based on **SL 6.5** (CernVM 3.3)
  - Production grade, 2 major releases, 25 security hot fixes
- Also available:
  - **SL4**: prototype for LEP experiments
  - **SL5**: Stable, GUI and contextualization
  - **SL7**: prototype, under development
Hypervisor Support Status

<table>
<thead>
<tr>
<th>Hypervisor</th>
<th>Cloud Controller</th>
</tr>
</thead>
<tbody>
<tr>
<td>VirtualBox</td>
<td>Openstack</td>
</tr>
<tr>
<td>VMware</td>
<td>OpenNebula</td>
</tr>
<tr>
<td>KVM</td>
<td>Amazon EC2</td>
</tr>
<tr>
<td>Xen</td>
<td>Google Compute Engine</td>
</tr>
<tr>
<td>Microsoft Hyper-V</td>
<td>Docker</td>
</tr>
<tr>
<td>Parallels</td>
<td>Microsoft Azure</td>
</tr>
</tbody>
</table>

1. Unclear license of the guest additions
2. Only tested with ephemeral storage, not with EBS backed instances
3. Only amiconfig contextualization
4. Work in progress
CernVM Users 2015

≈ 27000 distinct IP addresses
CernVM running CMS on SL5

XFCE Graphical User Interface, CMS event display
Analysis Preservation / Outreach

ALEPH software in CernVM on SL 4

CernVM as enabling Technology for CERN

OpenData project

G.Ganis, Status of CernVM
## Future: virtual appliance

<table>
<thead>
<tr>
<th>CernVM 3.4</th>
<th>Q2/2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Based on SL 6.6</td>
<td></td>
</tr>
<tr>
<td>• Improved update/downgrade functionality</td>
<td></td>
</tr>
<tr>
<td>• Interface to downgrade after a faulty update</td>
<td></td>
</tr>
<tr>
<td>• Increased automation level for desktops VMs</td>
<td></td>
</tr>
<tr>
<td>• Full support for containers (➔ next slide)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CernVM 4</th>
<th>Q3/2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Based on SL 7.1</td>
<td></td>
</tr>
<tr>
<td>• Requires support for <strong>capabilities</strong> <em>(extended attributes)</em> in CernVM-FS ✔</td>
<td></td>
</tr>
<tr>
<td>• Requires CernVM-FS integration in <strong>systemd</strong> as <strong>low level storage daemon</strong> ✔</td>
<td></td>
</tr>
<tr>
<td>• Contextualization streamlining</td>
<td></td>
</tr>
<tr>
<td>• Deprecation of amiconfig</td>
<td></td>
</tr>
<tr>
<td>• Functionality provided via cloud-init for compatibility</td>
<td></td>
</tr>
<tr>
<td>• Prototype available</td>
<td></td>
</tr>
<tr>
<td>• Requires consolidation, tuning, in particular of <strong>systemd</strong> integration</td>
<td></td>
</tr>
</tbody>
</table>
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 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 talk)
2: Using Parrot CernVM-FS
   - Pure user-space (ptrace)
   - User-mode CernVM demonstrated (next slide)
   - Ongoing work on performance and stability

Root file system (/) layout

```
/               cvmfs
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>usr</td>
<td>usr</td>
</tr>
<tr>
<td>symp</td>
<td>symp</td>
</tr>
<tr>
<td>lib64    -&gt;</td>
<td>lib64        -&gt;</td>
</tr>
<tr>
<td>etc</td>
<td>etc</td>
</tr>
<tr>
<td>var</td>
<td>var</td>
</tr>
<tr>
<td>tmp</td>
<td>...</td>
</tr>
</tbody>
</table>
```
Parrot + CernVM-FS: User-Mode CernVM

Using `parrot_run` with CernVM-FS support

```
[lxplus0026]$ ./cernvm-parrot.sh -c cms.cern.ch
CernVM-Lite: Downloading latest CVMU boot specifications
CernVM-Lite: Downloading required software
CernVM-Lite: Preparing root filesystem
CernVM-Lite: Starting CernVM in userland
CernVM-Lite: Welcome to CernVM v3.3.52.0
<cvmu>[I have no name!@lxplus0026 ~]$ cat /etc/cernvm/metarm

cernvm-system
<cvmu>[I have no name!@lxplus0026 ~]$ . /cvmfs/cms.cern.ch/cmsset_default.sh
<cvmu>[I have no name!@lxplus0026 ~]$ scram list

Listing installed projects available for platform >> slc6_amd64_gcc472 <<

<table>
<thead>
<tr>
<th>Project Name</th>
<th>Project Version</th>
<th>Project Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>CORAL</td>
<td>CORAL_2_3_21</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>---&gt; /afs/cern.ch/cms/slc6_amd64_gcc472/</td>
</tr>
<tr>
<td>cms/coral/CORAL_2_3_21</td>
<td>CORAL_2_3_21-cms</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>---&gt; /afs/cern.ch/cms/slc6_amd64_gcc472/</td>
</tr>
</tbody>
</table>
```
Summary

• Bootloader technology consolidated
• Demonstrated technology for wide range of cases
• SL 7-based soon available
• Support containers on its way

http://cernvm.cern.ch
Backup

- CernVM-FS
- CernVM-FS Parrot
- Security updates
- Contextualization
# CernVM-FS

Scalable Software Distribution System

- Central publication point, trusted release manager, read-only clients
- Minimal protocol requirements (HTTP), aggressive hierarchical cache
- POSIX mount point on clients (FUSE, local NFS share, Parrot)

### Main features

- **Content Delivery Network**
  - High avail by horizontal scaling
- **Content Addressable Storage**
  - Automatic integrity, deduplication
  - Directory structure w/ symlinks
- **User-assisted Merkel trees**
  - Optimized meta-data, scalability
  - Embedded versioning
- **File-chunking**

### Latest developments/plans

- **Consolidation of base use-cases**
  - Profiling, testing
  - Full SL 7 support (XFS, capabilities, overlay-FS)
  - Flexible config / key distribution
- **Address emerging use-cases**
  - History-less repositories
  - HPC, containers, ...
    - Consolidate access in user-space (Parrot driver)

See also R Meusel at [CernVM workshop](#)
Parrot & CernVM-FS: client in user-space

http://csl.cse.nd.edu/software/parrot
## Security Updates

### Options for updating CernVM

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1: <strong>stay</strong> (short-lived VM)</td>
<td></td>
</tr>
<tr>
<td>2: <strong>diverge</strong> (puppet managed)</td>
<td></td>
</tr>
<tr>
<td>3: <strong>rebase</strong> (Desktop)</td>
<td></td>
</tr>
</tbody>
</table>

### Rebase process

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. On first boot, CernVM selects and pins newest available versionShared cache</td>
<td></td>
</tr>
<tr>
<td>2. <strong>Automatic update notifications</strong></td>
<td></td>
</tr>
<tr>
<td>3. <strong>Running</strong> cernvm-update -a</td>
<td></td>
</tr>
<tr>
<td>4. Reboot virtual machine</td>
<td></td>
</tr>
</tbody>
</table>

### Planned improvements

- CernVM downgrade interface after a faulty security update
- More pervasive notification (e.g. MOTD, icon in XFCE status bar, …)
- Automate everything except reboot? Desktops only
Contextualization

User-Data sources

• Well-known web server (EC2, GCE, OpenStack)
• ISO (OpenNebula, HEPiX) or HDD (VirtualBox OVA format) image
• CernVM Online User-Data Unique ID
• Baked into the image

User-Data Formats

• cloud-init
• μCernVM bootloader format
• amiconfig
• Mixable in MIME multipart user-data

Deprecation of amiconfig

• amiconfig developed by rPath, maintained by us
• Available/required functionality can be provided by cloud-init \(^1\)
• Break existing contexts for CernVM 4?

\(^1\) CernVM-FS, HT-Condor, cctools, CernVM main user, CernVM GUI (desktop icons, autostart, ...), inject Grid certificate, Grid UI version