New CernVM-FS use cases at CMS

Andrea Valenzuela Ramírez
andrea.valenzuela.ramirez@cern.ch

CERN - CMS Core Software
CernVM Workshop 2024
Geneva, Switzerland

September 17th, 2024

Outline

CernVM-FS Use Cases at CMS

CMS Offline & Computing deploys to CernVM-FS under different use cases:

- Distribution of experiment **production software** (CMSSW).
- Distribution of **Integration Builds** (IBs).
- **Continuous Integration** (CI) purposes.

Repository Name	Size	Garbage Collection	Parallel Setup	Publishing (ops/day)	Year
`/cvmfs/cms.cern.ch`	23 TB	No	No	$\sim$ 5-30	2009
`/cvmfs/cms-ib.cern.ch`	3.77 TB	Yes (weekly)	Yes	$\sim$ 40	2016
`/cvmfs/cms-ci.cern.ch`	883 GB	Yes (weekly)	No	$\sim$ 1-40	2020

Table: CMS main repositories and their characteristics in terms of size, garbage collection frequency, publication setup, number of commits and year of creation.

- Distribution of **CMSSW environment images** in `unpacked.cern.ch`.

CernVM-FS Workshop 2022

CMS Parallel Deployment Workflow

- We moved to a parallel publishing setup on late February 2023 to speed-up IB deployments (`cms-ib.cern.ch`).
- Multi-release manager setup with **3 publishers** for `amd64` and **one (native)** for `aarch64`.
- Parallelization based on architecture.
  - Git mirrors also have independent paths.
- Job orchestration with Jenkins.
  - Feature to avoid running two deployment jobs with the same architecture parameter.
  - Wrapper to start a transaction that keeps the job pending until there is no lease.
  - Independent job that locks the top level directory and triggers GC.

Results

- Deployment waiting times have been reduced a $\sim$ 70% in high-demand periods (Sundays).
- Using a native `aarch64` publisher has reduced publication time by a $\sim$ 60%.

▸ CernVM-FS Parallel Publishing at CMS

New CernVM-FS Use Cases at CMS

- Distribution of Gridpacks.
- Usage of HPC Resources.
- Deployment of RISC-V Integration Builds.

Distribution of Gridpacks

- CMS high precision analyses require very precise Monte Carlo generators. For example, to guarantee Next to Leading Order (NLO) calculations.
- MadGraph generates the outcomes of particle interactions, which can be latter used to speed up computations.
- Concretely, MadGraph produces the so-called Gridpacks.
- Gridpacks are "pre-computed diagrams" used speed-up Monte Carlo generation.
- Distributed in tarballs, they are uncompressed for every generator job on local disk. **Many sites do not support such operation**.
- The proposed solution was serving already-untarred Gridpacks via CernVM-FS.

It is a new use-case of distribution of lookup files at CMS.

- At the moment, content is synchronized using `rsync` from `/eos` to `/cvmfs`, but it seems a nice use-case for `cvmfs_server ingest` utility.

New CernVM-FS Use Cases at CMS

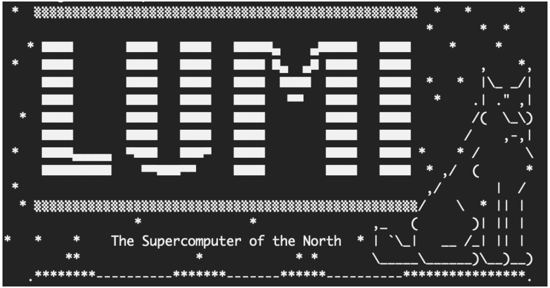
- Distribution of Gridpacks.
- Usage of HPC Resources.
- Deployment of RISC-V Integration Builds.

Usage of HPC Resources

- Access to AMD GPUs at LUMI (Finland) through the project *Exploring the Use of AMD GPUs for High-Performance Computing in the CMS Reconstruction*.

- Access to cvmfs using `singcvmfs exec`.

- CMS container images deployed to the LUMI user node from Dockerhub.



- Use `SINGCVMFS_REPOSITORIES` to indicate which repositories to load.

```
export SINGCVMFS_REPOSITORIES=cms.cern.ch, cms-ib.cern.ch, cms-ci.cern.ch,
grid.cern.ch, unpacked.cern.ch, patatrack.cern.ch
```

Usage of HPC Resources

- Selection of different cache directory for each invocation by setting SINGCVMFS_CACHEDIR.

```
if !  [ -f $SINGCVMFS_CACHEIMAGE ]; then
   mkdir -p $(dirname $SINGCVMFS_CACHEIMAGE)
   /usr/sbin/mkfs.ext3 -m 0 -E root_owner $SINGCVMFS_CACHEIMAGE 50G
fi
export SINGCVMFS_CACHEIMAGE=$SCRATCH/cvmfscache.ext3
```

- Submission of the resource allocation via SLURM.

```
srun -pty -time=08:00:00 -partition=small-g -hint=multithread -nodes=1
-ntasks=1 -cpus-per-task=14 -gpus=1 -mem=60G -
/project/$SLURM_ACCOUNT/cvmfsexec/singcvmfs exec -bind
/opt,/project/$SLURM_ACCOUNT,/scratch/$SLURM_ACCOUNT -bind
$SINGULARITY_SCRATCH:/workspace:image-src=/ -env PS1="$SINGULARITY_PROMPT"
$SINGULARITY_CACHEDIR/cmssw_el8.sif $SHELL
```

New CernVM-FS Use Cases at CMS

- Distribution of Gridpacks.
- Usage of HPC Resources.
- Deployment of RISC-V Integration Builds.

Deployment of RISC-V Integration Builds

- CMS is also pushing towards exploring new hardware architectures.
- Risc-V (Milk-V) machine at Bologna (64 cores, 128 GB RAM).
- Building the CMS Offline Software stack (CMSSW) using Fedora 39 as base container.
  ▸ CMSSW on Risc-V
- CernVM-FS client built locally on the nodes (with root privileges).
- Deployment of the IBs to `cms-ib.cern.ch` using emulation on the publisher nodes.
  - The `qemu` emulator is needed since we run `rpm` commands that are architecture-dependent.
- Deployment time < 1h.

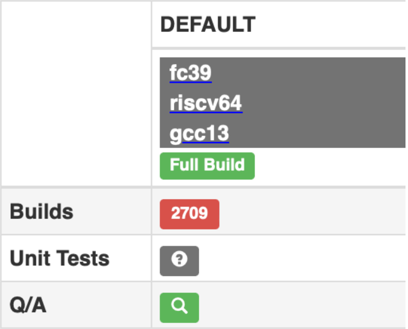


Figure: First Risc-V Integration Builds for CMSSW.

Conclusion

- CernVM-FS is crucial for CMSSW.
  - It helps in development, distribution and preservation of the software.
- The parallel publishing setup for IB deployment continues succeeding in speeding-up the IB delivery.
  - The current setup allows horizontal scaling.
- As new use-cases appear, we find utilities on the CernVM-FS side to support them.

Finally, we would like to thank the CernVM-FS team for the support provided.

Thank you!