

# CVMFS Container Extractor



## The narrative

CVMFS solved the problem of software distribution.

Containers solved the problem of having a reproducible environment.

Serve containers over CVMFS bring the best of the two worlds.

How to automatically extract containers into CVMFS?

# Highlights

- CVMFS Ingest command
- CVMFS Container Extractor

## CVMFS ingest command

- Allow to ingest a generic tarball inside CVMFS
- Input:
  - The tarball to ingest
  - Where to ingest the tarball
- Allow to rapidly publish (Docker) containers images that are composed from layer tarballs.

## CVMFS Container Extractor

- Builds on top of the CVMFS ingest command
- Allows to automatically track and ingest into a CVMFS repository a set of docker images
- Those images can then be executed by Docker or by Singularity
- Input provided as a wish list in YAML format

# CVMFS Container Extractor File System Layout

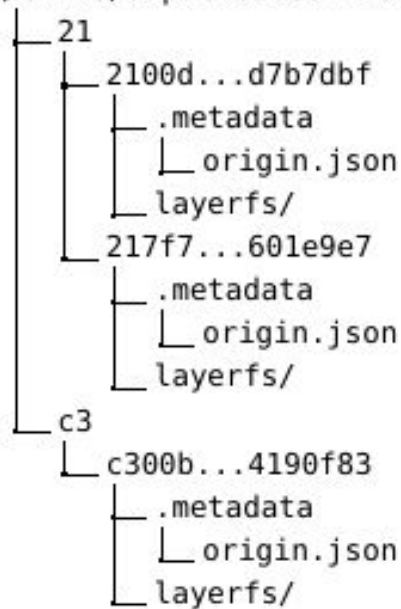
- The extractor keeps a directory structure to exploit CVMFS sub-catalogs
- It also keeps internal bookkeeping information to recover from possible errors

# CVMFS Container Extractor Singularity Layout

```
/cvmfs/unpacked.cern.ch
├── registry.hub.docker.com
│   ├── library
│   │   ├── centos:centos7 -> /cvmfs/unpacked.cern.ch/.flat/75/75835...0c4ab6d
│   │   ├── centos:latest -> /cvmfs/unpacked.cern.ch/.flat/75/75835...0c4ab6d
│   │   ├── debian:stable -> /cvmfs/unpacked.cern.ch/.flat/a4/a4274...ba594cb
│   │   ├── gcc:latest -> /cvmfs/unpacked.cern.ch/.flat/ce/ceccd...a75dd28
│   │   ├── openjdk:9 -> /cvmfs/unpacked.cern.ch/.flat/5a/5adaf...5344d70
│   │   ├── python:2.7 -> /cvmfs/unpacked.cern.ch/.flat/3c/3c43a...0e7c9d4
│   │   └── python:3.4 -> /cvmfs/unpacked.cern.ch/.flat/43/43953...ed73435
│   └── efajardo
│       └── docker-cms:tensorflow ->
│           /cvmfs/unpacked.cern.ch/.flat/2d/2d5b4...97d44fc
```

# CVMFS Container Extractor Docker Layout

```
/cvmfs/unpacked.cern.ch/.layers
```



- **Layers are shared between different images**
- **Layers are hidden from the user, only the docker daemon accesses them**
- **Layers are partitioned for name to enhance sub-catalog effectiveness**
- **Each layer contains metadata information**



## In operation at [unpacked.cern.ch](https://unpacked.cern.ch)

- For several years, there has been production use of CernVM-FS hosted container images through [singularity.opensciencegrid.org](https://singularity.opensciencegrid.org)
- This functionality is now part of upstream code
- In addition, the new code provides support both Singularity and Docker
- Runs on CVMFS CERN IT infrastructure
- We accept [pull requests](#) if a particular images is needed

## Future works

- Support for multiple architecture
- Enhance support for containerd

**Thanks!**