



CernVM-FS

CernVM-FS tutorial

6th EasyBuild User Meeting (EUM'21)

Jan 26-29 2021

<https://cvmfs-contrib.github.io/cvmfs-tutorial-2021>

<https://easybuild.io/eum/#cvmfs-tutorial>

Sessions

- [Mon Jan 25 2021, 17:00-17:30 UTC] **Introductory presentation** at EUM'21 (by Jakob Blomer, CERN)
See <https://www.youtube.com/watch?v=lxZLS3O9wo4>
- [Tue Jan 26 2021, 09:00-10:00 UTC] **CernVM-FS tutorial (part 1/4)**
 - 0. Azure resources
 - 1. Quick CernVM-FS introduction (concepts and terminology)
 - 2. Stratum-0 and client access
 - *hands-on exercise*
- [Wed Jan 27 2021, 09:00-10:00 UTC] **CernVM-FS tutorial (part 2/4)**
 - 3. Stratum-1 + proxies
 - *hands-on exercise*
- [Thu Jan 28 2021, 09:00-10:00 UTC] **CernVM-FS tutorial (part 3/4)**
 - 4. Publishing
 - *hands-on exercise*
- [Fri Jan 29 2021, 09:00-10:00 UTC] **CernVM-FS tutorial (part 4/4)**
 - 5. Advanced topics



CernVM-FS

Questions or problems?



CernVM-FS

- Support available throughout the day, via:
 - #eum21-cvmfs-tutorial in EasyBuild Slack
 - Email: eum@lists.ugent.be
- If necessary, an additional Zoom call can be set up
- Q&A at end of each live session

0. Azure resources

https://cvmfs-contrib.github.io/cvmfs-tutorial-2021/00_azure_cloud_resources



CernVM-FS

0. **Log in to CycleCloud** via <https://cvmfstutorial.westeurope.cloudapp.azure.com>
 - Use personal credentials (see email) + pick new password (and remember it!)
1. Add your **SSH public key** (must be RSA!)
2. **Start cluster** (and keep it running until you're done with all exercises)
 - IP addresses are NOT fixed, and will change on reboot

If you experience any problems, let us know (via eum@lists.ugent.be or Slack)!

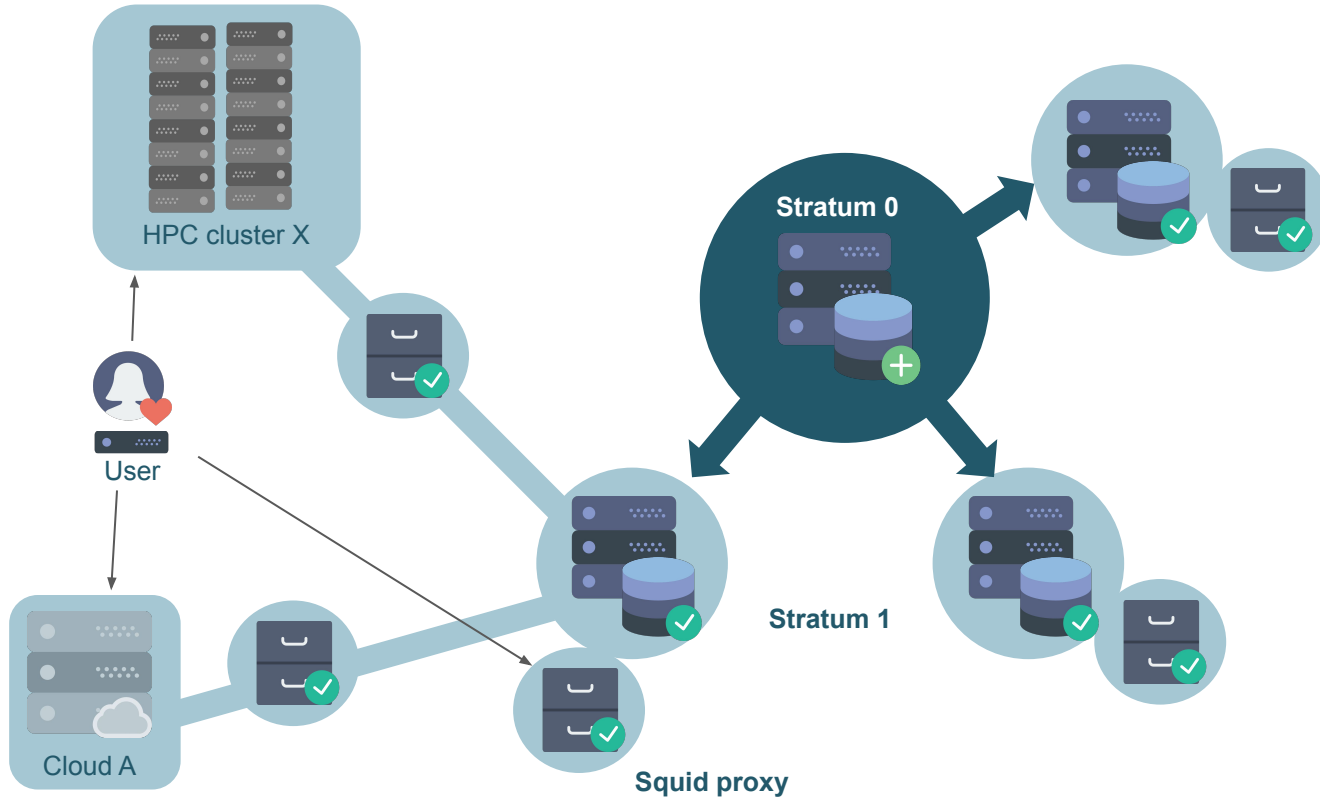
Sponsored by  Microsoft Azure

1. CernVM-FS introduction

https://cvmfs-contrib.github.io/cvmfs-tutorial-2021/01_introduction



CernVM-FS

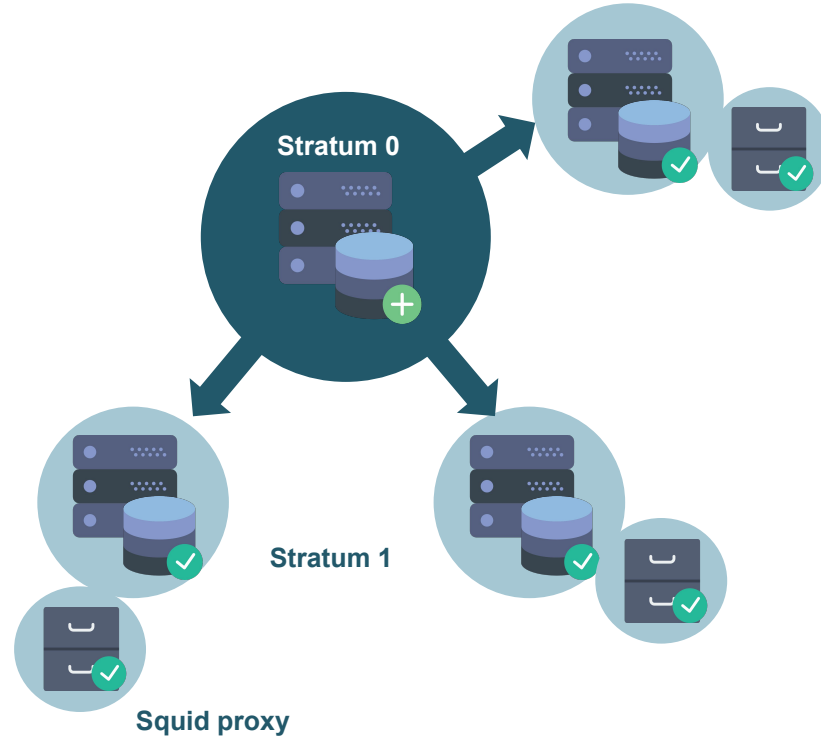


1. CernVM-FS introduction

https://cvmfs-contrib.github.io/cvmfs-tutorial-2021/01_introduction



CernVM-FS



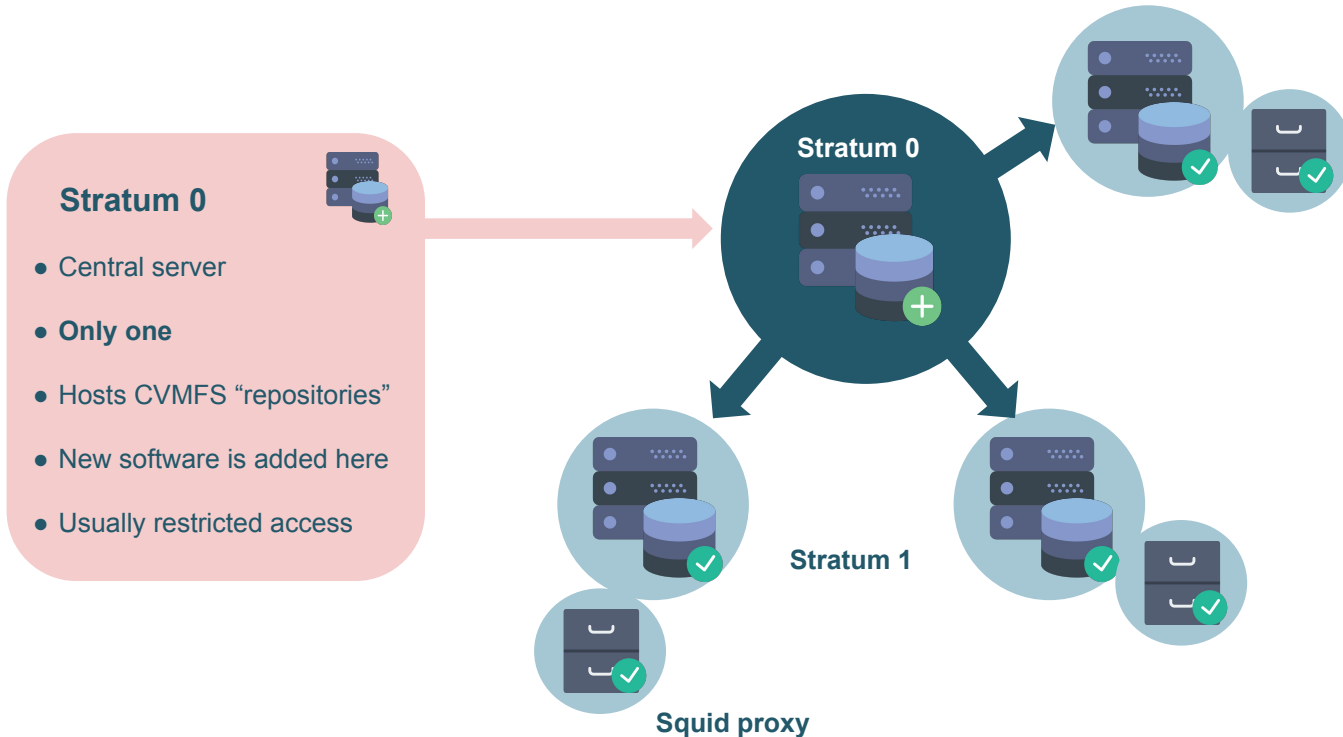
1. CernVM-FS introduction

https://cvmfs-contrib.github.io/cvmfs-tutorial-2021/01_introduction



CernVM-FS

(icons via <https://www.flaticon.com/authors/smashicons>)

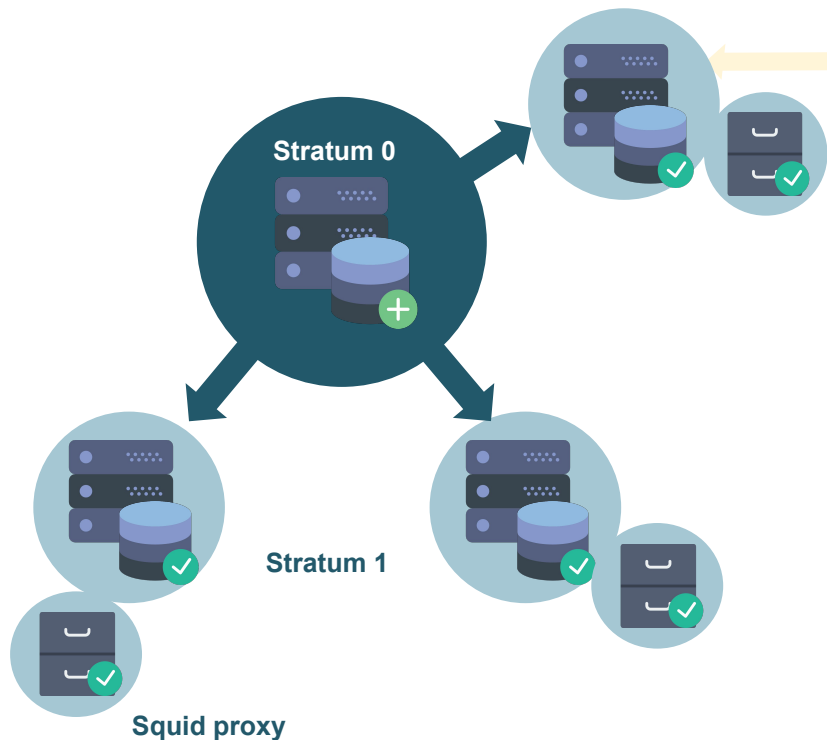


1. CernVM-FS introduction

https://cvmfs-contrib.github.io/cvmfs-tutorial-2021/01_introduction



CernVM-FS



Stratum 1



- CernVM-FS “replica”
- Full copy of repositories (mirror)
- **Read-only**
- **Multiple servers**
- Geographically distributed
- Standard web server (HTTP)
- Reduce load on Stratum 0
- Improve reliability
- Clients *may* connect here

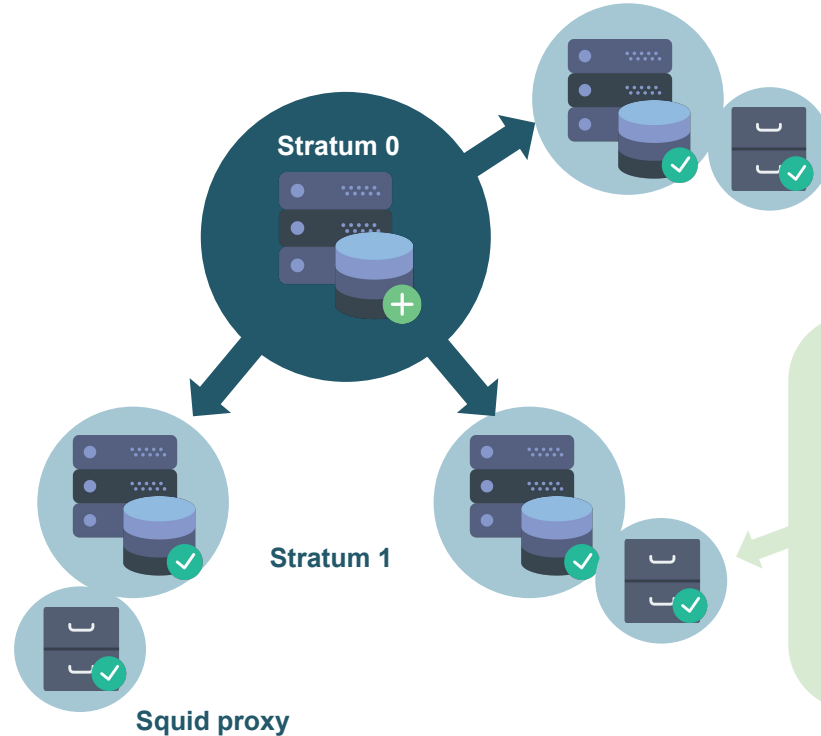
1. CernVM-FS introduction

https://cvmfs-contrib.github.io/cvmfs-tutorial-2021/01_introduction



CernVM-FS

(icons via <https://www.flaticon.com/authors/smashicons>)



Squid proxy

- Reverse proxy for Stratum 1 servers
- Caching to improve client I/O performance
- Reduce load on Stratum 1 servers
- Can be used for load balancing
- Clients usually connect to one of these

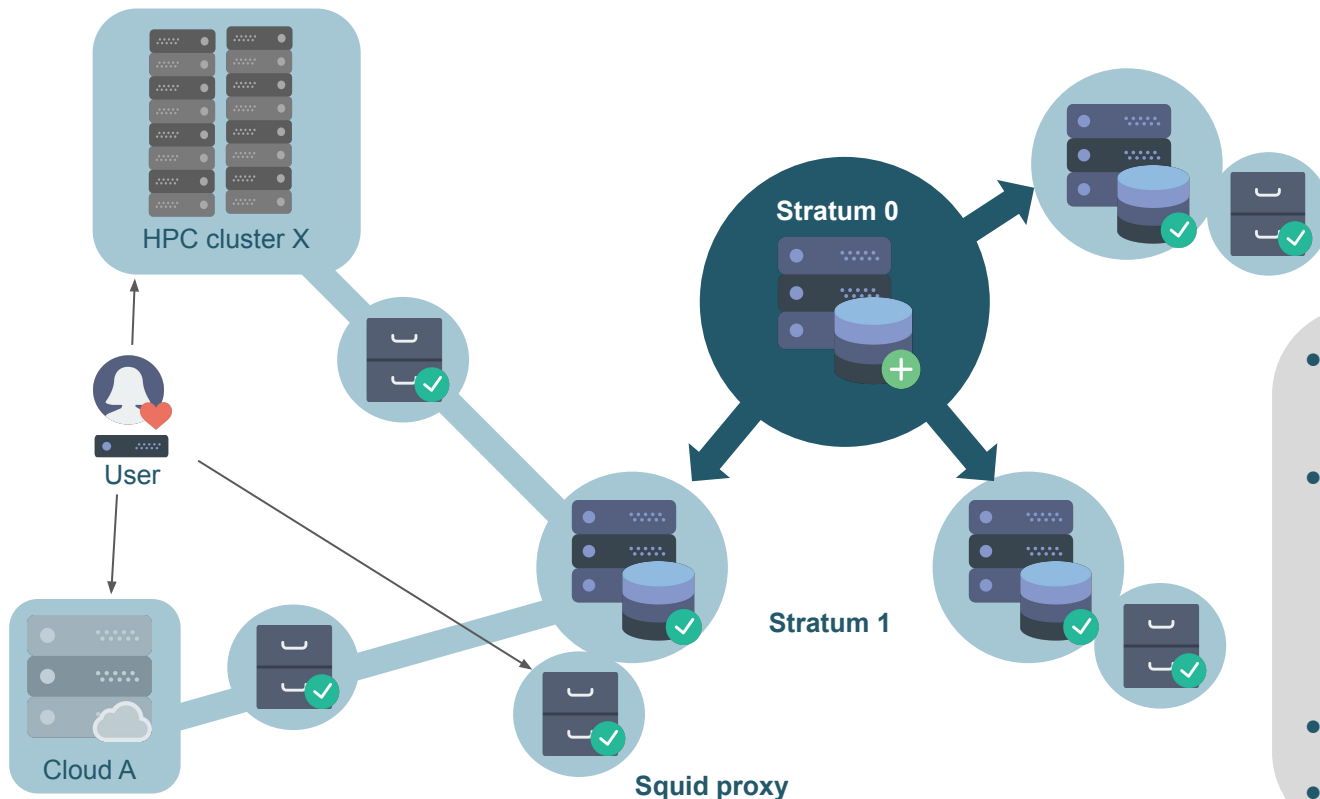
1. CernVM-FS introduction

https://cvmfs-contrib.github.io/cvmfs-tutorial-2021/01_introduction



CernVM-FS

(icons via <https://www.flaticon.com/authors/smashicons>)



- Clients access software stack via (squid proxy of) a Stratum 1
- Clients include:
 - Laptops
 - Personal workstations
 - HPC clusters
 - Cloud instances
- Clients also use local filesystem cache
- **Same software stack everywhere!** 10

2. Getting started: Stratum 0 and client

https://cvmfs-contrib.github.io/cvmfs-tutorial-2021/02_stratum0_client

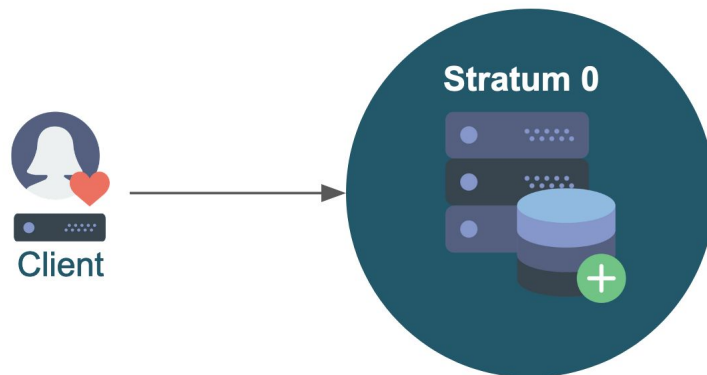


CernVM-FS

- Deploy a fully operational Stratum 0 server
- Make a repository
- Add a file
- Set up a client that connects to the Stratum 0

(WARNING: you should not do this in production!)

- Exercise



2. Getting started: Stratum 0 and client (recap)

Recap of yesterday's exercise



CernVM-FS

- A client can mount multiple repos, even from different servers
(i.e. from different domains/projects/organizations)
- The repository name can basically be anything, but make sure to pick a suitable one
 - The domain name part doesn't really have to exist (but it usually does exist)
 - You're free to pick a domain (or even repository) name that doesn't belong to you
- `cvmfs_config setup` **should be run before** `cvmfs_config chksetup` (*fixed in tutorial*)
- `CVMFS_DEBUGLOG` (see Debugging section on the Advanced topics page)
should point to a location that is *writable* for the `cvmfs` user
- Make sure that you put the public key at the right location!
 - Note the difference between its original location on the Stratum 0 (`/etc/cvmfs/keys`) and the recommended location on the clients (and Stratum 1)

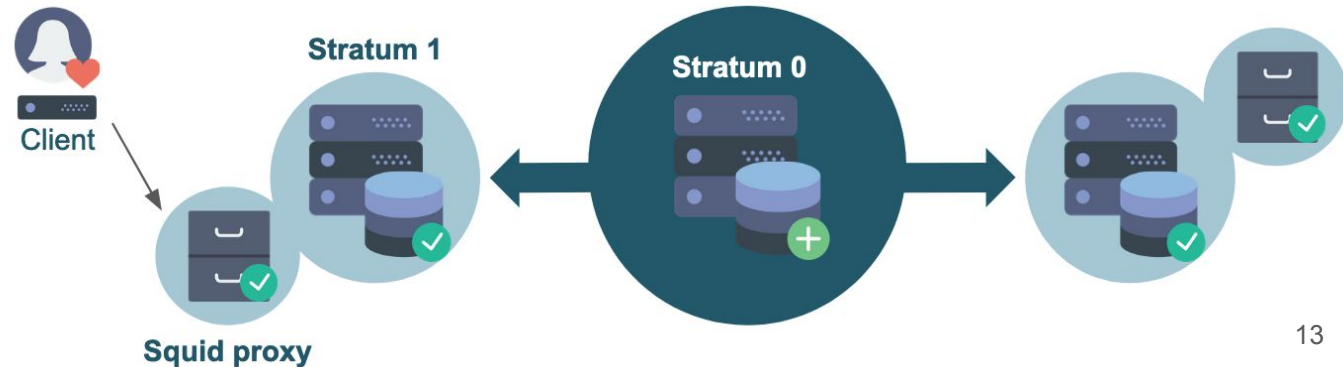
3. Stratum 1 and proxies

https://cvmfs-contrib.github.io/cvmfs-tutorial-2021/03_stratum1_proxies



CernVM-FS

- Deploy a Stratum 1 server
- Set up a replica of the repository
- Synchronize the replica
- Deploy a Squid caching proxy server
- Reconfigure your client, so that it uses the proxy and Stratum 1
- Exercise



3. Stratum 1 and proxies

Recap



CernVM-FS

- Make sure that all the keys in `/etc/cvmfs/keys/yourdomain.tld/` are correct!
 - `cvmfs_server` checks (and may use) them all
 - If one of them is incorrect, you may get weird issues, even for other repos
- You can always enable the use of the Geo API later on
 - The next snapshot will then update the Geo database
- Request to add a TLDR to each page/section
 - Quick summary of which commands have to be run and which files have to be created/modified
- Request to add a section about running CVMFS in user space
 - Can be done with Singularity and [cvmfsexec](#) for the client
 - Not possible for the Stratum 0/1 servers
- `cvmfs_config stat -v <repo>` only works when the repo is mounted

4. Publishing

https://cvmfs-contrib.github.io/cvmfs-tutorial-2021/04_publishing



CernVM-FS

- Transactions
- Ingesting tarballs
- Tags and rollbacks
- (Nested) catalogs
- Exercise

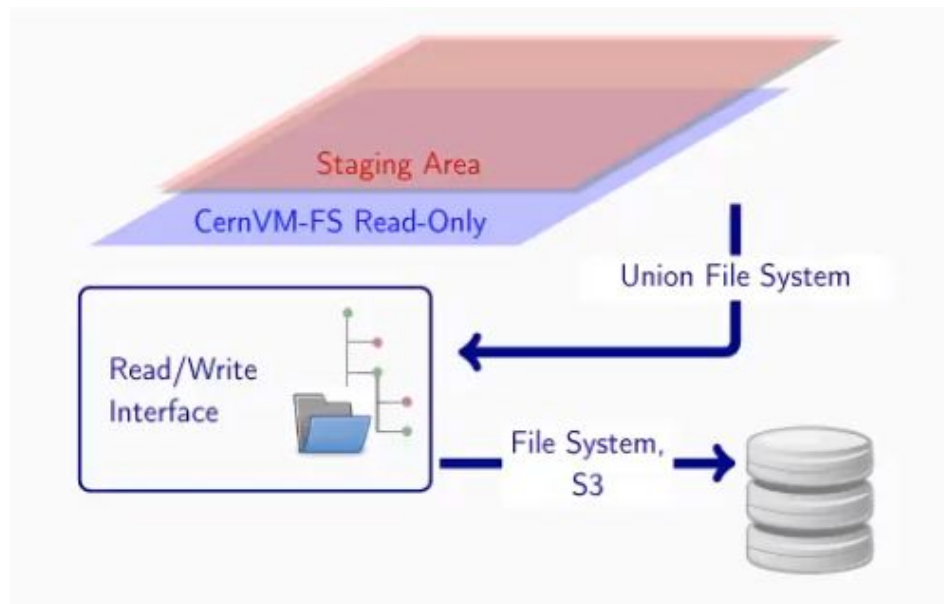


Image source:

CernVM-FS: Overview and Roadmap (EUM'21) by Jakob Blomer

<https://www.youtube.com/watch?v=IxZLS3O9wo4>

4. Publishing

Recap



CernVM-FS

- `.cvmfscatalog` are empty files, and will stay empty
- The solution of the exercise also makes some small (8 or 24 entries) nested catalogs
 - This does not comply to the recommendations
 - Partly because of the dummy installations, in practice there will be more files
 - Still it would make sense to have, for instance, at least one per architecture
- You may have to remove `.cvmfscatalog` files after changing the `.cvmfsdirtab` file
- Large publish operations require a lot scratch space (`/var/spool/cvmfs`)
 - Tarballs are directly ingested to the CVMFS storage area (`/srv/cvmfs`)

5. Advanced and other topics

https://cvmfs-contrib.github.io/cvmfs-tutorial-2021/05_advanced



CernVM-FS

- Automating the deployment of CernVM-FS servers and clients
- Debugging issues
- Garbage collection
- Gateway and publisher nodes
- Running the client as normal user
- Configuration repositories

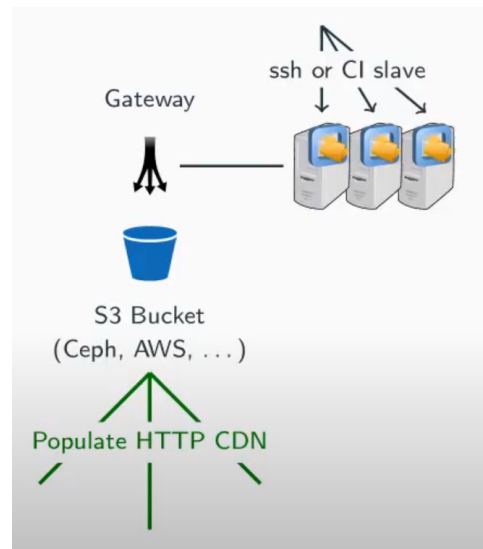


Image source:

CernVM-FS: Overview and Roadmap (EUM'21) by Jakob Blomer

<https://www.youtube.com/watch?v=IxZLS3O9wo4>

Wrap-up



- VMs will be removed today (Fri Jan 29th) at 8pm CET
- Please terminate your cluster yourself as soon as you're done using it!

Connecting to the CernVM-FS community

- JIRA (bugs, feature requests): <https://sft.its.cern.ch/jira/projects/CVM/issues>
- Mattermost (chat, similar to Slack): <https://mattermost.web.cern.ch> (CernVM)
 - Requires CERN lightweight account: <https://account.cern.ch/account>
- Discord forum (brand new!): <https://cernvm-forum.cern.ch>
- **CernVM-FS workshop (Feb 1-2 2021):** <https://indico.cern.ch/e/cvm21>
 - Incl. talks on CernVM-FS roadmap, cvmfsexec, EESSI, containers and CernVM-FS, etc.