



ESD HPC Containers

Eric Müller on behalf of the joint ESD working groups ("normal" and "HPC")



2025-03-13



Virtualization?



• VMs? Hardware virtualization, Bootup/Kernel, CPU support, ...



"Containerization"!



"Containerization"!

- VMs? Hardware virtualization, Bootup/Kernel, CPU support, ...
- 20+ years of "namespaces" (in Linux), approx. 15 years of containers (LXC in 2008)
 - mount, pid, net, ipc, uts, user, cgroup, time, (syslog?)...
- Docker (since 2013)/OCI Containers aim to look like VMs
 - \rightarrow service encapsulation (migration, scaling, \ldots)
- Do we need all that for software deployment?
 - No, this is about software deployment only!

ESD HPC

"Containerization"!

- VMs? Hardware virtualization, Bootup/Kernel, CPU support, ...
- 20+ years of "namespaces" (in Linux), approx. 15 years of containers (LXC in 2008)
 - mount, pid, net, ipc, uts, user, cgroup, time, (syslog?)...
- Docker (since 2013)/OCI Containers aim to look like VMs
 - \rightarrow service encapsulation (migration, scaling, . . .)
- Do we need all that for software deployment?
 - No, this is about software deployment only!
- On HPC: site-specific software components (MPI, CUDA, optimized libraries) → site-specific configuration (expert knowledge)

Co-funded by the European

Apptainer

ESD HPC Containers

Co-funded by the European

- \bullet Used to be called singularity...
- Available on EBRAINS 2.0 HPC sites and (almost?) all EuroHPC sites
- "Transparency" by default... (mnt namespace!)
- E.g., bind mounts for HOME, PWD, or all host filesystems; /dev, /proc, /sys, etc.
- (Does support many of the other namespaces too, ...)
- suid-based spawning vs. user namespaces
- (btw. fallback for image building: proot (faking namespaces in userspace via ptrace Or seccomp)

Reminder: (Full) Build Process for the Lab



EBRAINS

Co-funded by the European Unio



Reminder: (Full) Build Process for Container Images



Reminder: (Full) Build Process for Virtual Machines



Reminder: (Full) Build Process for HPC Optimized Images



Current State: WIP!

"feat(HPC): build image on JUSUF"

https://gitlab.ebrains.eu/ri/tech-hub/platform/esd/ebrains-spack-builds/-/ merge_requests/630

- introducing new CI step \rightarrow (personal) gitlab runner on HPC
- next steps:
 - validating build by performance benchmark comparison against a trusted (flat) installation
 - finalize milestone M6.4 on "HPC containers"
 - "generically optimized for HPC" containers (MPI trampoline functionality, multi-arch, multi-CUDA builds?)
 - (test-drive in multi-site workflows)

ESD HPC Containers