

Podman für HPC-Systeme

NHR Container Workshop

13.12.2021

RWTHAACHEN
UNIVERSITY



TECHNISCHE
UNIVERSITÄT
DARMSTADT

NHR4
CES

NHR for
Computational
Engineering
Science

Ausgangslage

Motivation und erste Erfolge

Motivation

Spezifischer Bedarf eines Nutzers, der ausdrücklich nicht mit Singularity abgedeckt werden konnte

Passt zum Betriebssystem RedHat / CentOS

- Podman Features:
- Daemon less
 - Root less
 - user namespaces
 - Konfiguration im Nutzer-HOME

Erste Ergebnisse

- Einfache Installation aus dem OS Repository
- Root konnte erstmal ohne Probleme Container benutzen

- aber !

Umsetzung

Schwierigkeiten und Lösungen

GPUs (Nvidia) im Container

Nur mit Paketen aus einem Nvidia Repo möglich (nvidia-container-toolkit, <https://github.com/NVIDIA/nvidia-docker>)

Anpassungen in „/etc/nvidia-container-runtime/config.toml“ (einfach im Knotenimage)

→ Funktioniert recht problemlos

User Namespaces

Die Namespaces müssen in `/etc/subuid` und `/etc/subgid` für jeden Nutzer definiert werden.

Beim Anlegen lokaler Accounts könnte ein Eintrag generiert werden, allerdings gibt es eine solche Lösung nicht für eine Nutzerverwaltung mittels LDAP oder 389 DS.

Daher Podman Nutzung nur auf Nachfrage und Einpflegen des Namespaces.

Filesysteme

Nach umständlicher Suche gefunden: „Rootless Podman out of the box will not work well on NFS home directories because the protocol does not understand user namespaces.“ (gilt auch für GPFS und Lustre)

Es ist möglich Podman so zu konfigurieren (graphroot), dass „lokaler“ Storage genutzt wird, z.B. /tmp oder /dev/shm (unsere Systeme haben keine lokalen Platten mehr).

→ Kein permanenter Speicherplatz für Container

→ Kein einfacher Weg einen Container über mehrere Knoten zu nutzen


Nebenproblem „runroot“ für Slurm und andere Batchsysteme: „Getting a system assigned via slurm means that the login usually does not go through the normal processes and /run/user/ID does not exist.“

→ Einen anderen “runroot” konfigurieren, oder das Verzeichnis mit entsprechenden Rechten generieren (problematisch). Umgebungsvariablen wie XDG_RUNTIME_DIR werden bevorzugt ausgewertet.

Fehler in alten Podman Versionen (älter als 3.1)

Container funktionieren nur im "detached" (-d) Modus wenn die UserID länger als 7 Zeichen ist (alle unsere UserIDs).

Im detached Modus gibt es keinen direkten Output. Allerdings kann der Output des Containers über "podman logs" ausgelesen werden. → podman logs ...



TU Darmstadt
Hochschulrechenzentrum
Basisdienste, HPC-Gruppe

Alexanderstraße 2
64283 Darmstadt

hhr@hrz.tu-darmstadt.de

www.nhr4ces.de