Creating a Container Runtime from Scratch

Who is Ajitem?

- **Lead Consultant** at **Technogise**, working for nearly 15 years, 8+ yrs Docker/K8s, deep in Go + runtimes.

- **GitHub**: https://github.com/asahasrabuddhe

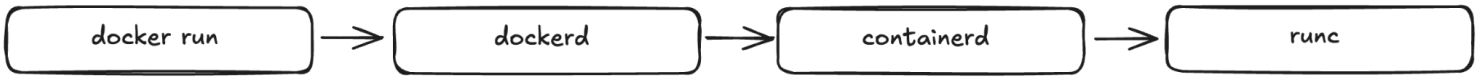
- **LinkedIn**: https://linkedin.com/in/ajitem

What is a Container Runtime?

- A container runtime is the foundational software that allows containers to operate within a host system.
- It is responsible for pulling the container images from a container registry, managing their life cycle, and running the containers on your system.
- Docker is not a container runtime. It is a container engine.
- Docker includes a container runtime (*containerd*).
- Docker also offers a suite of tools for building, shipping, and running containerized applications.

Docker is not a Runtime

- Docker uses **containerd** and **runc**.



What is a Container, really?

- A container is a process
- which runs in an isolated environment (**cgroups** + **namespaces**)
- with a rootfs.

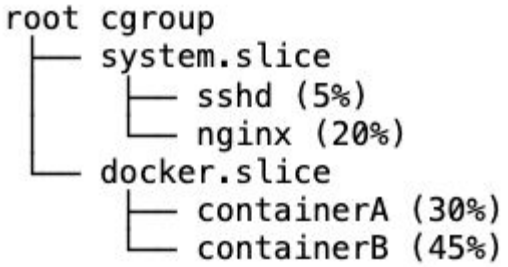
The container essentially *contains* the **filesystem** which *signifies* **what to run. Namespaces (isolation)** and **Cgroups (sharing)** *signify* **how exactly to run it.**

What are CGroups (Control Groups)?

- Kernel feature for resource accounting & limiting.
- Think of it as “*budgets*” for processes.
- You can control:
  - **CPU** (shares, quotas)
  - **Memory** (hard/soft limits)
  - **I/O** (disk/network bandwidth)

How CGroups Work?

- Processes are grouped into hierarchies.
- Limits are applied per group, not per process.
- Kernel enforces the rules.



What are Namespaces?

- Kernel feature for isolation of system resources.
- Each namespace gives processes their own view of the world.
- Together, they make containers feel like mini-systems.

Namespaces

- **pid** → own process tree (separate init, PIDs start at 1).
- **mnt** → isolated filesystem mounts.
- **user** → UID/GID mapping, privilege isolation.
- **uts** → separate hostname & domain name.
- **net** → isolated interfaces, routes, ports.
- **ipc** → isolated shared memory, semaphores.

Demo

technognise

Thank you!

Code

Contact Details

Ajitem Sahasrabuddhe
Lead Technology Consultant, Technogise
ajitem@technogise.com
https://github.com/asahasrabuddhe
https://www.linkedin.com/in/ajitem
https://medium.com/@ajitem