Container in HPC - 2022 Edition

FOSDEM'22 talk about what to look out for in 2022

Christian Kniep, 05.Feb.2022

I did talk about the runtime aspect in 2021.

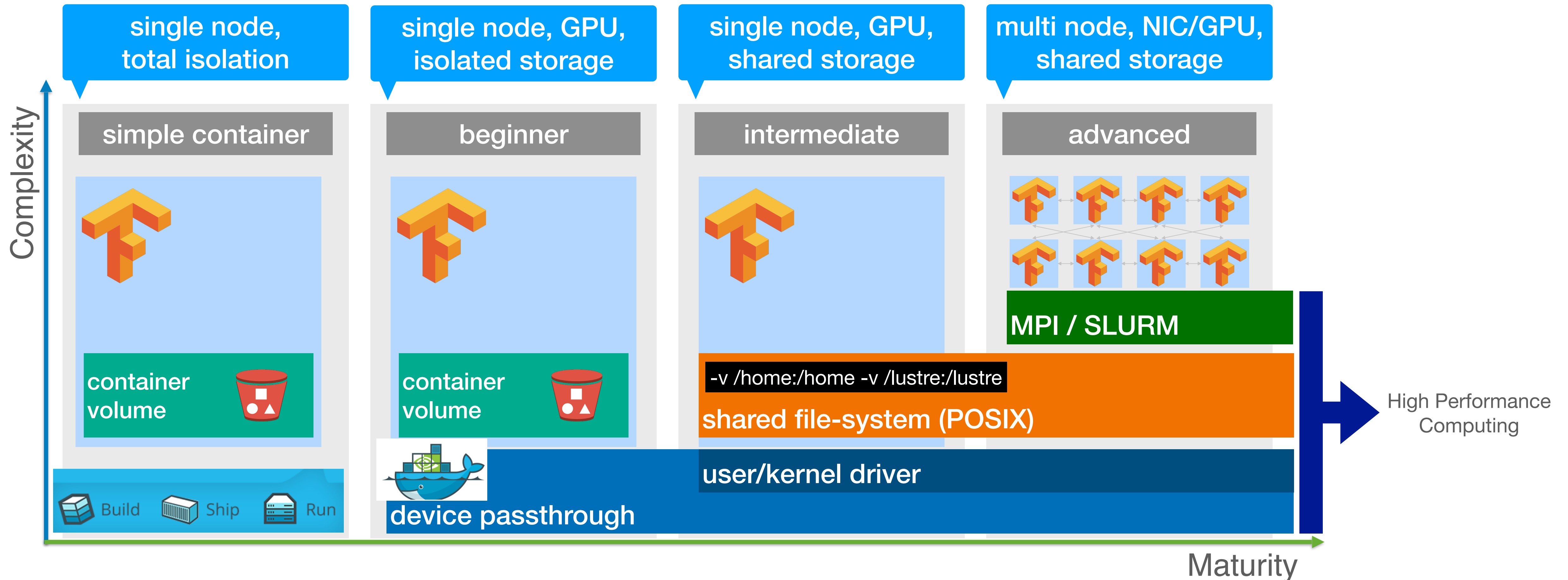**ReCap: Do not fight over which runtime to use; they all do the same.**

Insist on OCI runtime and image spec!

Overview

Challenges for High Performance Containers

… through the lens of (distributed) AI/ML

Complexity →

single node, total isolation	single node, GPU, isolated storage	single node, GPU, shared storage	multi node, NIC/GPU, shared storage
simple container	beginner	intermediate	advanced

container volume

container volume

-v /home:/home -v /lustre:/lustre

shared file-system (POSIX)

MPI / SLURM

user/kernel driver

device passthrough

Build Ship Run

High Performance Computing

Maturity →

Container Ecosystem Segments

From the bottom up…

Who	What	Examples
Runtime	Create a (containerized) process which had (at least) an isolated file-system view.	runc,crun,youki
Engine	Lifecycle of container images and containers on a single node. Create snapshot of image, setup network and generate config to be picked up by the runtime.	Containerd, Sarus, Podman, Singularity
Scheduler	Orchestrates container placment (usually across a cluster of multiple indtances). Provides API to interact.	Docker (Swarm), Kubernetes, Slurm
Build Tools	Creates container images to encapsulate software stacks. Either based on deriving a Dockerfile to be build or create a snapshot of a filesystem.	Spack/EasyBuild, HPCM, Dockerfile, OpenHPC
Distribution	Tooling, APIs ro distribute OCI container images.	DockerHub, Github Container Registry, ECR

Build/Distribute/Schedule

Things to Do/Not Do in 2022

Do **NOT** focus on Runtimes!

It's a pre-COVID discussion!

Some Runtimes

Runtimes

- youki (rust)

- crun (C)

- runc (GOLANG)

Engines

- Sarus

- containerd

- Podman

All-in-One

- Singularity

- NROOT(?)

- Shifter

Do **NOT** start your container journey on Kubernetes!

Unless you are ‚born' in K8s. Otherwise use the environment you know and love…

If you know SLURM, do get started there.

Running an HPC runtime of your (admins) choosing. It's easier than you think!

Do build your image with AUTOMATION!

Or at least do it reproducibly

Do build your image with AUTOMATION!

Or at least do it reproducibly

Do think about how to annotate your images and compute resources!

So that your runtime, scheduler or registry gets some clues about preferences, constraints.