Survey of HPC Container Tools

Presenters:
Bryan Whitehurst
Rachel Yamamoto

Mentors:
Eric Green
Martin Baltezore
David Fox

August 11, 2021
About

- Bryan
  - University of Alabama
  - Junior Computer Science Student
  - Club Tennis
About

- Rachel
  - Case Western Reserve University
  - Major: Computer Science BS
  - Expected Graduation 2024
What are Containers?

- Standard unit of code that packages up software and all its dependencies so that the application can be run quickly and reliably on multiple systems

- Podman, Singularity, Charliecloud, Sarus, Shifter
Why do Containers Matter to HPC?

- **Portable**
  - Ease of transporting software and its dependencies to different systems

- **Lightweight**
  - Containers are very lightweight compared to VMs
    - Containers use a fraction of the memory required to boot an OS

- **Scalable**
  - Can be distributed to many nodes easily
  - HPC workloads can face a spike in data processing requirements

- **Reproducible**
Why Not Docker?

- The Docker Runtime doesn't work well in HPC because....
  - Multitenancy
  - Networking

- Docker/OCI Compatible Containers can be run in HPC, just not with the Docker runtime
Our Goals

- Install and properly configure container runtimes optimized for HPC
- Run rootless containers using Singularity, Charliecloud, Sarus, Podman, and Shifter
- Configure MPI to work with containers
Roadmap: What We Compared

- MVAPICH2 Library performance vs OpenMPI Library performance
- Container performance vs Host System performance
- Singularity vs Charliecloud performance
MVAPICH2 vs. OpenMPI Runtimes

- MVAPICH2 performed **significantly** faster than OpenMPI for small and large message sizes
Container vs. Host MPIBench Runtimes

- Containers installed with MVAPICH2 were slower than the host system with MVAPICH2
Container vs. Host MPIBench Runtimes

• OpenMPI showed consistent results inside and outside Charliecloud/Singularity containers
Singularity vs. Charliecloud

- No significant difference between the performance of MPIBench inside Charliecloud and Singularity containers
Challenges

- **Sarus**
  - Not using the interconnect properly led to high runtimes

- **Running Podman Containers stored on NFS**
  - Setting up rootless podman to work with NFS
    - Stores images in an NFS based home directory
    - Podman containers cannot run on NFS so you must copy container storage over to each compute node manually

- **Establishing MPI and Slurm Compatibility**
  - Configuring Slurm, OpenMPI, and MVAPICH2 to work with PMI support
  - We had to install OpenMPI and MVAPICH2 from Source – not from the package manager

- **Installing Shifter**
  - Shifter uses Python 2.7 so it could not be installed on CentOS 8
Future Work

- Shifter
  - Testing runtimes

- Podman
  - Slurm and MPI compatibility

- Sarus high-speed infiniband interface
  - rather than ethernet

- E4s-cl Project
  - Extreme Scale Scientific Software Stack container launcher (e4s-cl)
  - a tool used to run MPI applications in containers
  - Use it to run MPI benchmarks inside the container
References

- https://www.redhat.com/sysadmin/rootless-podman-nfs
- https://podman.io/
- https://www.docker.com/resources/what-container
- https://chrisshort.net/docker-inc-is-dead/


- https://hpc.github.io/charliecloud/
- https://containerjournal.com/topics/container-management/containers-hpc-mutually-beneficial/
- https://www.netapp.com/devops-solutions/what-are-containers/
- https://cloud.google.com/containers
Questions