Applied Containerization for Machine Learning in HPC Workshop

A Hands-on Guide to Running ML Workloads in HPC Environments

Parmanand Sinha, Computational Scientist (GIS+HPC)
University of Chicago Research Computing Center

Workshop Overview

- **Duration**: 2 hours (1 hour presentation, 1 hour hands-on)
- **Level**: Intermediate
- **Prerequisites**:
  - Basic programming knowledge
  - Linux CLI familiarity
  - Active RCC/cluster account
- **Repository**: GitHub

1. Introduction to Containerization in HPC

Why Containers for ML in HPC?

- **Dependency Management**:
  - Package complex ML software stacks with all dependencies.
  - Resolve version conflicts easily.
- **Portability**:
  - Run the same environment across different HPC systems, local machines, or cloud platforms.
  - "Build once, run anywhere."

Why Containers for ML in HPC? (Cont.)

- **Reproducibility**:
  - Ensure consistent results by capturing the exact software environment.
  - Crucial for research and collaboration.
- **Performance**:
  - Native-like performance with minimal overhead.
  - Direct hardware access (e.g., GPUs, high-speed interconnects).

Container Basics: What is a Container?

- An isolated, encapsulated user-space instance.
- Runs on a shared OS kernel but has its own:
  - Filesystem
  - Processes
  - Network interfaces (can be configured)
- Lightweight compared to Virtual Machines (VMs) as they don't require a full guest OS.

Container vs VM

Image source: apptainer.org

Container Basics: Key Components

- **Container Images**:
  - Read-only templates used to create containers.
  - Contain application code, libraries, dependencies, and metadata.
  - Examples: Docker Hub images, SIF files (Apptainer).
- **Runtime Environments**:
  - Software that runs containers (e.g., Apptainer, Docker, Charliecloud).
  - Manages container lifecycle, isolation, and resource allocation.

Containerization Workflow in HPC

```
flowchart TD
    A[Pull Container Image] --> B[Store SIF in $SCRATCH/$USER/]
    B --> C[Submit SLURM Job]
    C --> D[SLURM Allocates Compute Node]
    D --> E[Bind Data/Script Directories]
    E --> F[Run Container with Apptainer]
    F --> G[ML Training/Inference]
```

Container Basics: Key Components (Cont.)

- **Mount Points for Data (Bind Mounts)**:
    - Mechanism to make host directories/files accessible inside the container.
    - Essential for accessing datasets, scripts, and output directories.
- **Resource Allocation**:
    - Containers share host resources (CPU, memory, GPUs).
    - HPC schedulers (like SLURM) manage resource allocation for containerized jobs.

class: lead

2. Apptainer (formerly Singularity) Introduction

Apptainer & Singularity: A Quick Note

- **Apptainer** is the direct successor to **Singularity**.
- You might see `singularity` used in older documentation or as the module name on some HPC systems.
- Commands are largely interchangeable (e.g., `singularity pull` vs. `apptainer pull`).
- This workshop uses modern `apptainer` commands.

Hands-on with Apptainer (Introduction)

(This section introduces Apptainer. The actual hands-on will follow in the next hour.)

Apptainer: Getting Started

```
# Load Apptainer module (on HPC systems)
module load apptainer

# Pull a PyTorch container image from Docker Hub
apptainer pull docker://pytorch/pytorch:latest
# This creates a .sif file (e.g., pytorch_latest.sif)
```

- **SIF (Singularity Image Format)**: Apptainer's default, optimized image format.
- **Storage Tip**: For large images, consider pulling/storing them in your scratch directory.
  - On Midway3: `$SCRATCH/$USER/sif_files/` (where `$SCRATCH` is `/scratch/midway3`).

Apptainer: Basic Commands

- **Interactive Shell**:

```
# Start an interactive shell inside the container
# --nv enables NVIDIA GPU access
apptainer shell --nv pytorch_latest.sif
```

- **Run a Python Script**:

```
# Execute a command (e.g., a Python script) inside the container
apptainer exec --nv pytorch_latest.sif python my_script.py
```

Apptainer: Basic Commands (Cont.)

- **Bind Data Directories**:

```
# Run a container and mount /data on host to /data in container
apptainer run --nv --bind /path/on/host:/path/in/container pytorch_latest.sif

# Example: Mount current working directory ($PWD) to /mnt inside container
apptainer run --nv --bind $PWD:/mnt pytorch_latest.sif
```

  - `apptainer run` executes the default runscript defined in the image (if any).
  - Binding `$PWD` (current working directory) is very useful for accessing your scripts and local data within the container.

Apptainer: Key Features

- **GPU Support**: `--nv` flag for seamless NVIDIA GPU access.
- **Data Binding**: Mount host directories inside the container (`--bind` or `-B`).
    - Crucial for accessing datasets and saving results.
- **Environment Variables**: Pass configuration through the container boundary.
    - Apptainer typically inherits host environment; can be controlled.
- **MPI Support**: Run distributed workloads across nodes.
    - Often requires MPI compatibility between host and container.

class: lead

3. Charliecloud Overview

Charliecloud vs. Apptainer

Feature	Apptainer (formerly Singularity)	Charliecloud
Security Model	set-UID (optional), rootless execution	Fully unprivileged (user namespaces)
Image Format	SIF (optimized, single file)	Directory trees, SquashFS
Build Process	`.def` files, build from Docker Hub	Direct Dockerfile support
Best Use Case	Complex ML workflows, ease of use	Security-critical environments, Docker

Charliecloud: Key Benefits

- **Minimal Attack Surface**:
    - Fully unprivileged operation using user namespaces.
    - Reduces security risks on shared HPC systems.
- **Docker Compatibility**:
    - Direct use of Dockerfiles for building images (`ch-convert`).
    - Easier transition for users familiar with Docker.
- **Lightweight & Simple**:
    - Simpler architecture and deployment compared to some other runtimes.
    - Focuses on core containerization features.

class: lead

4. Practical ML Container Deployment (Examples)

TensorFlow Example with Apptainer

```
# 1. Pull TensorFlow container (GPU version)
apptainer pull docker://tensorflow/tensorflow:latest-gpu

# 2. Run training script
# Assumes train.py and data are accessible via bind mounts
apptainer exec --nv tensorflow_latest-gpu.sif \
  python /path/to/your/train.py --data /path/to/data
```

- `--nv`: Enables NVIDIA GPU access.
- Bind mount your script directory and dataset directory.

PyTorch Example with Apptainer (Multi-GPU)

```
# Pull PyTorch container (if not already done)
# apptainer pull docker://pytorch/pytorch:latest

# Multi-GPU training using torchrun (formerly torch.distributed.launch)
apptainer exec --nv pytorch_latest.sif \
  torchrun --nproc_per_node=4 /path/to/your/distributed_train.py
```

- Assumes `distributed_train.py` is set up for PyTorch DDP.
- `--nproc_per_node` should match available GPUs.

Distributed Training Topology (Multi-Node/Multi-GPU)

```
flowchart TB
    subgraph Node1 [Compute Node 1]
        GPU1[GPU 0]
        GPU2[GPU 1]
        GPU3[GPU 2]
        GPU4[GPU 3]
    end
    subgraph Node2 [Compute Node 2]
        GPU5[GPU 0]
        GPU6[GPU 1]
        GPU7[GPU 2]
        GPU8[GPU 3]
    end
    UserScript([distributed_train.py])
    UserScript -- Launches via torchrun/srun --> GPU1
    UserScript -- Launches via torchrun/srun --> GPU5
    GPU1 <--> GPU2 <--> GPU3 <--> GPU4
    GPU5 <--> GPU6 <--> GPU7 <--> GPU8
    GPU1 <--> GPU5
    GPU2 <--> GPU6
    GPU3 <--> GPU7
    GPU4 <--> GPU8
```

class: lead

5. SLURM Integration (Brief Overview)

SLURM Job Lifecycle

```
flowchart TD
    A[User Submits sbatch/srun] --> B[SLURM Scheduler]
    B --> C[Resources Allocated (Nodes/GPUs)]
    C --> D[Job Starts on Compute Node(s)]
    D --> E[Container Launched (Apptainer)]
    E --> F[ML Script Executes]
    F --> G[Job Output Written]
    G --> H[Job Completes]
```

SLURM: Single-Node Job Example

```
#!/bin/bash
#SBATCH --job-name=ml-training
#SBATCH --gres=gpu:1          # Request 1 GPU
#SBATCH --partition=gpu       # Specify GPU partition
#SBATCH --cpus-per-task=4     # Request 4 CPUs
#SBATCH --mem=16G             # Request 16GB RAM

# Load Apptainer module
module load apptainer

# Define paths (replace with your actual paths)
CONTAINER_IMAGE=/path/to/pytorch_latest.sif
SCRIPT_DIR=/path/to/your/scripts
DATA_DIR=/path/to/your/data
OUTPUT_DIR=/path/to/your/output

# Ensure output directory exists
mkdir -p $OUTPUT_DIR

# Run the containerized job
apptainer run --nv \
  --bind $SCRIPT_DIR:/scripts \
  --bind $DATA_DIR:/data \
  --bind $OUTPUT_DIR:/output \
  $CONTAINER_IMAGE python /scripts/train.py --data_dir /data --output_dir /output
```

SLURM: Multi-Node Distributed Training Example

```bash
#!/bin/bash
#SBATCH --job-name=dist-ml-train
#SBATCH --nodes=2                  # Request 2 nodes
#SBATCH --ntasks-per-node=4        # 4 tasks (processes) per node
#SBATCH --gres=gpu:4               # 4 GPUs per node (total 8 GPUs)
#SBATCH --cpus-per-task=2          # 2 CPUs per task
#SBATCH --partition=gpu-multi      # Example partition for multi-node GPU jobs

module load apptainer

CONTAINER_IMAGE=/path/to/your/ml_container.sif
# Script should handle distributed setup (e.g., using torchrun environment variables)

# srun will launch 'ntasks-per-node' copies of this command on each node
srun apptainer run --nv \
  --bind /path/to/data:/data \
  $CONTAINER_IMAGE \
  torchrun --nnodes=$SLURM_NNODES \
           --nproc_per_node=$SLURM_NTASKS_PER_NODE \
           --rdzv_id=$SLURM_JOB_ID \
           --rdzv_backend=c10d \
           --rdzv_endpoint=$SLURM_STEP_NODELIST:29500 \
           /path/to/your/distributed_train.py
```

class: lead

6. Best Practices

Best Practices: Container Management

- **Version Control & Tagging**:
  - Tag containers with specific versions (e.g., `myimage:1.0.0`, `myimage:latest-cuda11.8`).
  - Store definition files (`.def`, Dockerfiles) in version control (Git).
- **Data Management**:
  - Use bind mounts for large datasets to avoid including them in images.
  - Keep images small and focused on software environment.
- **Resource Allocation**:
  - Match container resource needs to SLURM (or other scheduler)

Best Practices: Performance Optimization

- **GPU Access**: Always use `--nv` (Apptainer) or equivalent for GPU workloads.
- **I/O Optimization**:
  - Bind fast storage for temporary files. On Midway3, compute nodes often have a high-throughput SSD directory at `$TEMP` (e.g., `/scratch/local/$USER/`) ideal for this.
  - For frequently accessed small files or datasets, consider staging them to such temporary storage.
  - Be mindful of I/O patterns within the container.
- **MPI Configuration**:

Best Practices: Security Considerations

- **Unprivileged Execution**:
  - Run containers without root access whenever possible.
  - Apptainer runs as user by default. Charliecloud is designed for unprivileged execution.
- **Data Protection**:
  - Use appropriate bind mounts; be specific about what host paths are exposed.
  - Avoid overly broad mounts (e.g., binding `/`).
- **Resource Limits**:
  - Rely on the HPC scheduler (SLURM) to enforce resource limits

class: lead

7. Additional Resources

Resources: Documentation

- **Apptainer Official Documentation**:
  Apptainer Documentation

- **Charliecloud User Guide**:
Charliecloud User Guide
- **Workshop Repository**:
Workshop Repository
(Contains this presentation, examples, and hands-on materials)

Resources: Example Repositories & Further Learning

- **ML Framework Containers (Official Hubs)**:
    - Docker Hub:

      Docker Hub

(e.g., `pytorch/pytorch`, `tensorflow/tensorflow`) - NVIDIA NGC:
NVIDIA NGC
- Docker Hub: hub.docker.com (e.g., `pytorch/pytorch`, `tensorflow/tensorflow`) - NVIDIA NGC: ngc.nvidia.com - **SLURM Integration Scripts**: Check your HPC center's documentation for specific

Key Takeaway

> “ **Containerization enables reproducible, portable, and efficient ML workflows in HPC environments.** ”

- Choose between **Apptainer** and **Charliecloud** based on your specific needs for:
  - Security model
  - Ease of use
  - Image format preferences
  - Integration requirements with existing Docker workflows

Q&A and Thank You!

Next: Hands-on Session!

References & Further Reading

- Apptainer Documentation
- Charliecloud User Guide
- Singularity/Apptainer on RCC
- Docker Documentation
- NVIDIA NGC Containers
- PyTorch Containers
- TensorFlow Containers
- SLURM Documentation
- HPC Best Practices
- Research Computing Center, UChicago