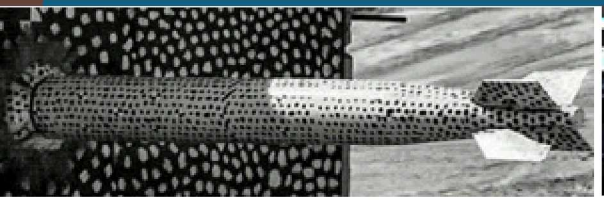
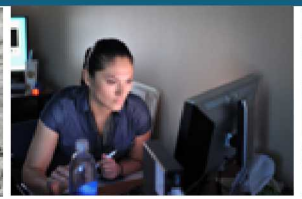
SAND2019-14091C

SUPERCONTAINERS

E4S and Supercontainers





PRESENTED BY Andrew J. Younge

Sandia National Laboratories

ajyoung@sandia.gov

E4S BOF @ SC19

UNCLASSIFED//UNLIMITED RELEASE

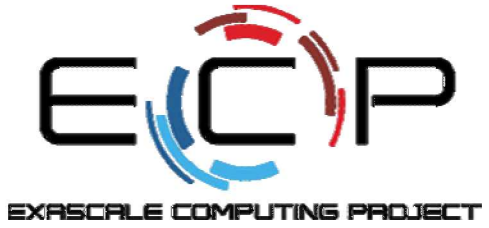
Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Initial HPC Container Vision

- Support HPC software development and testing on laptops/workstations
  - Create working container builds that can run on supercomputers
  - Minimize dev time on supercomputers
- Developers specify how to build the environment AND the application
  - Users just import a container and run on target platform
  - Have many containers, but with different manifests for arch, compilers, etc.
  - Not bound to vendor and sysadmin software release cycles
- Performance matters
  - Use mini-apps to “shake out” container implementations on HPC
  - Envision features to support future workflows (ML/DL/in-situ analytics)

ECP Supercontainers

- Joint DOE effort - Sandia, LANL, LBNL, LLNL, U. of Oregon
- Ensure container runtimes will be scalable, interoperable, and well integrated across DOE
  - Enable container deployments from laptops to Exascale
  - Assist ECP applications and facilities leverage containers most efficiently
- Three-fold approach
  - Scalable R&D activities
  - Collaboration with related ST and AD projects
  - Training, Education, and Support
- Activities conducted in the context of interoperability
  - Portable solutions
    - Optimized E4S container images for each machine type
    - Containerized ECP that runs on Astra, A21, El-Capitan, …
  - Work for multiple container implementations
    - Not picking a “winning” container runtime
  - Multiple DOE facilities at multiple scales



E4S Container Images

SUPERCONTAINERS

- Today: E4S Container images available
  - Include full gambit of software found in E4S
  - Leverages Spack package management
  - Can be used for system software, library, and app development within ECP
  - Images available on several container runtimes
    - Singularity, Shifter, Charliecloud, …
- Next: Optimized container images targeting Exascale platforms
  - Base container images for target platforms
  - Will leverage advanced Spack features
    - Environments, Binary installs, multi-stage build optimizations…
  - Supported by centralized container repository
  - Developers leverage base images for custom deployments

E4S

Spack environments help with building containers

- We recently started providing base images with **Spack** preinstalled.
- **Very** easy to build a container with some Spack packages in it:

spack-docker-demo/
- Dockerfile
- spack.yaml

```
FROM spack/centos:7

WORKDIR /build
COPY spack.yaml .
RUN spack install
```

Base image with Spack in PATH

Copy in spack.yaml
Then run spack install

```
spack:
  specs:
    - hdf5 @1.8.16
    - openmpi fabrics=libfabric
    - nalu
```

List of packages to install, with constraints

Build with **docker build .**

Run with Singularity
(or some other tool)

Emerging workloads on HPC with Containers

- Support emerging AI/ML/DL frameworks on HPC
  - Containers useful to adapt ML software to HPC
  - Already supported and heavily utilized in industry
- Extreme-scale Scientific Software Stack (**E4S**)
  - Includes TensorFlow & Pytorch in container image
  - Find Sameer Shende for more details! – e4s.io
- Working with DOE app teams to deploy custom ML tools in containers
- Investigating scalability challenges and opportunities

TensorFlow

PYTORCH

Thanks!

ajyoung@sandia.gov