



**Hewlett Packard**  
Enterprise

# **Containerization Workflow for Weather and Climate Applications**

---

Paulo de Souza and Usama Anber

May 11, 2023

# Objective

---

Containers are no longer a second-class citizen for complex weather and climate workflows on HPC systems:

- Portability/Agility
- Reproducibility
- Performance



# Container base image

---

- We choose the container base image to match the OS running on our Cray supercomputer
- Since the host and the container OS are compatible, we can bind files from the host, like the complete Cray MPI and fabric libraries
  - Host OS: SUSE Linux Enterprise Server 15 SP4
  - Image OS: openSUSE Leap 15.4
- Here's the container SW stack:
  - Intel® MPI 2021.6.0
  - Intel® oneAPI Compilers/MKL 2021.4.0
  - hdf5-1\_13\_2
  - szip-2.1.1
  - netcdf-c-4.9.0
  - pnetcdf-1.12.3
  - netcdf-fortran-4.5.4
- Build with Docker and run with Apptainer

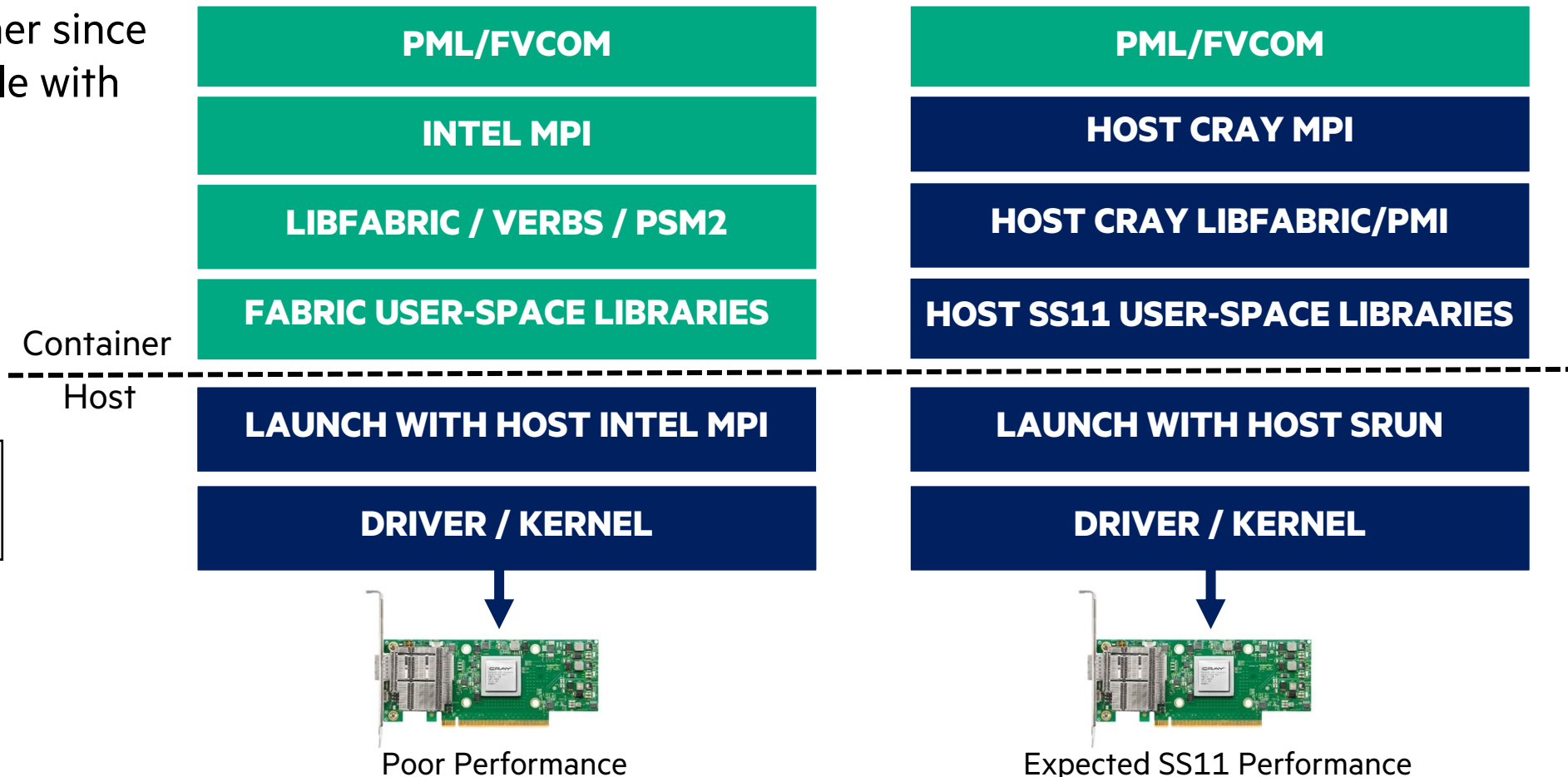


# Running Intel MPI Container on Slingshot 11

In this case we can bind host binaries into the container since SLES 15.4 is compatible with OpenSUSE 15.4

**OpenSUSE** Container  
Without binding any files from Host

**Bind Cray MPI**, transport layer and fabric runtime libraries from host



Legend

- Host Binaries **SLES 15.4**
- Container Binaries **OpenSUSE 15.4**

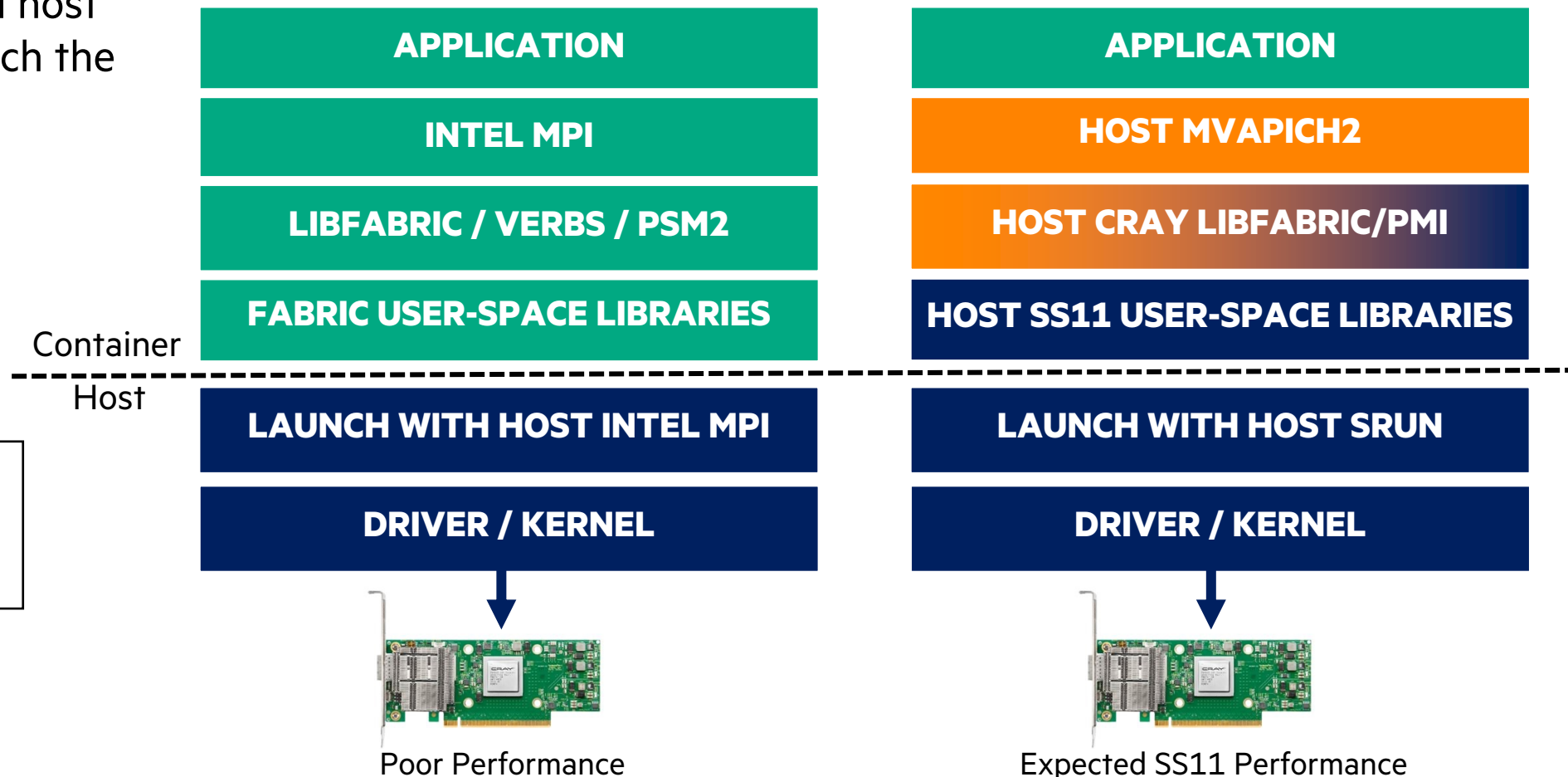


# Running Intel MPI Container on Slingshot 11

In this case we bind host binaries built to match the container OS

**Ubuntu** Container  
Without binding any files from Host

**Bind MVAPICH2**, transport layer and fabric runtime libraries from host



Legend

- Host Binaries SLES 15.4
- Host Binaries for Ubuntu **20.04**
- Container Binaries Ubuntu **20.04**



# Running Intel MPI Container on Slingshot 11

---

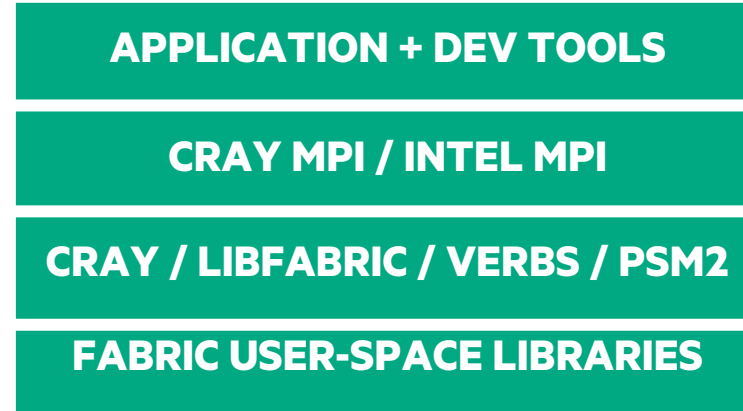
- If you don't use host pre-built binaries for the container OS, you might get into compatibility issues
- Got the error below running a CentOS 8.4.2105 container on a SUSE Linux Enterprise Server 15 SP4 host:

```
mpi_isend_recv: /lib64/libm.so.6: version `GLIBC_2.29' not found (required by /opt/mvapich2/lib/libmpi.so.12)
```



# Multistage build

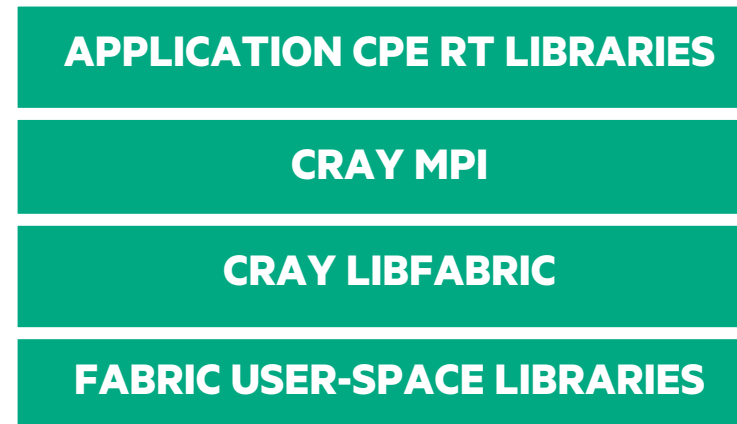
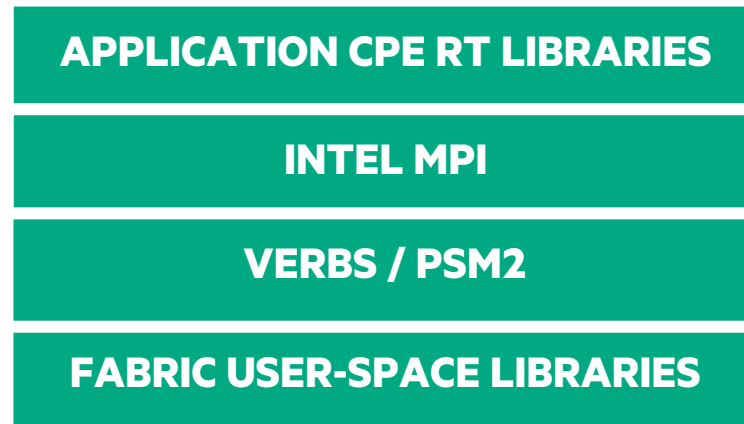
First Stage Big Container  
with all dev tools  
Cray MPI + Intel API



Second stage  
cherry pick rt files only

HPE HW Intel MPI Only

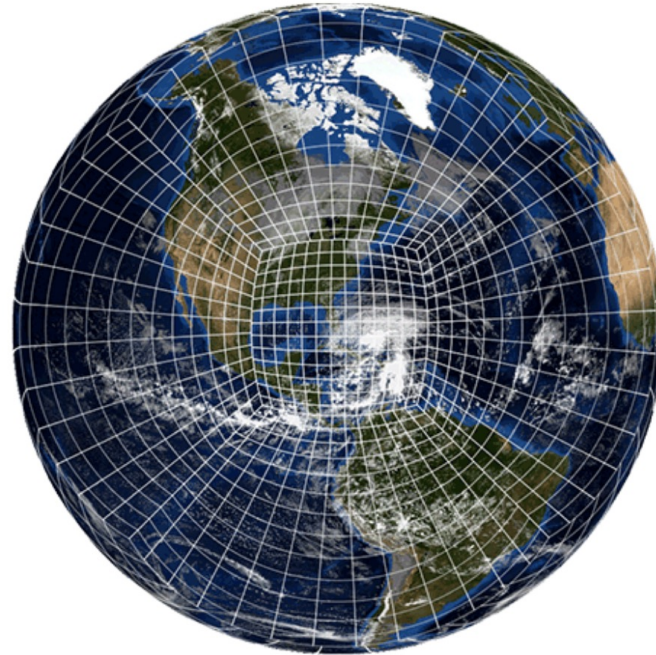
Cray MPI Only



# Applications: FV-3 and FVCOM

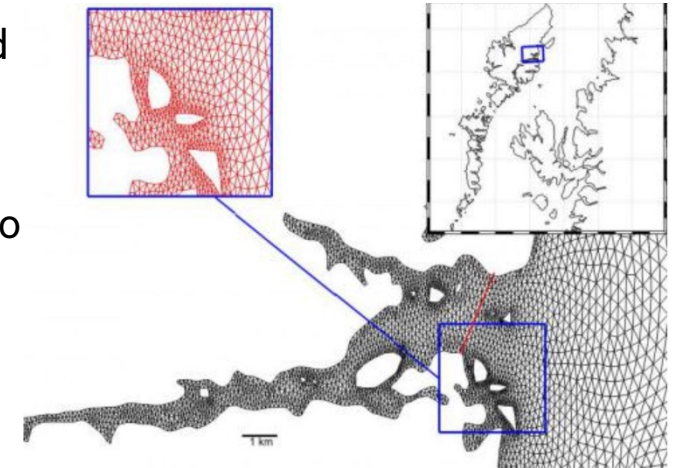
## FV-3:

- Atmospheric dynamical Core for weather and climate models.
- Finite-volume on cubed-sphere capable of hydrostatic and non-hydrostatic simulations
- Currently used in the US operational forecast model.



## FVCOM:

- finite-volume, prognostic, unstructured-grid, and free-surface.
- coastal ocean circulation model developed by UMASD-WHOI joint efforts.
- ideally suited for many coastal and interdisciplinary scientific applications.





# Workload description / context

---

## Compilers:

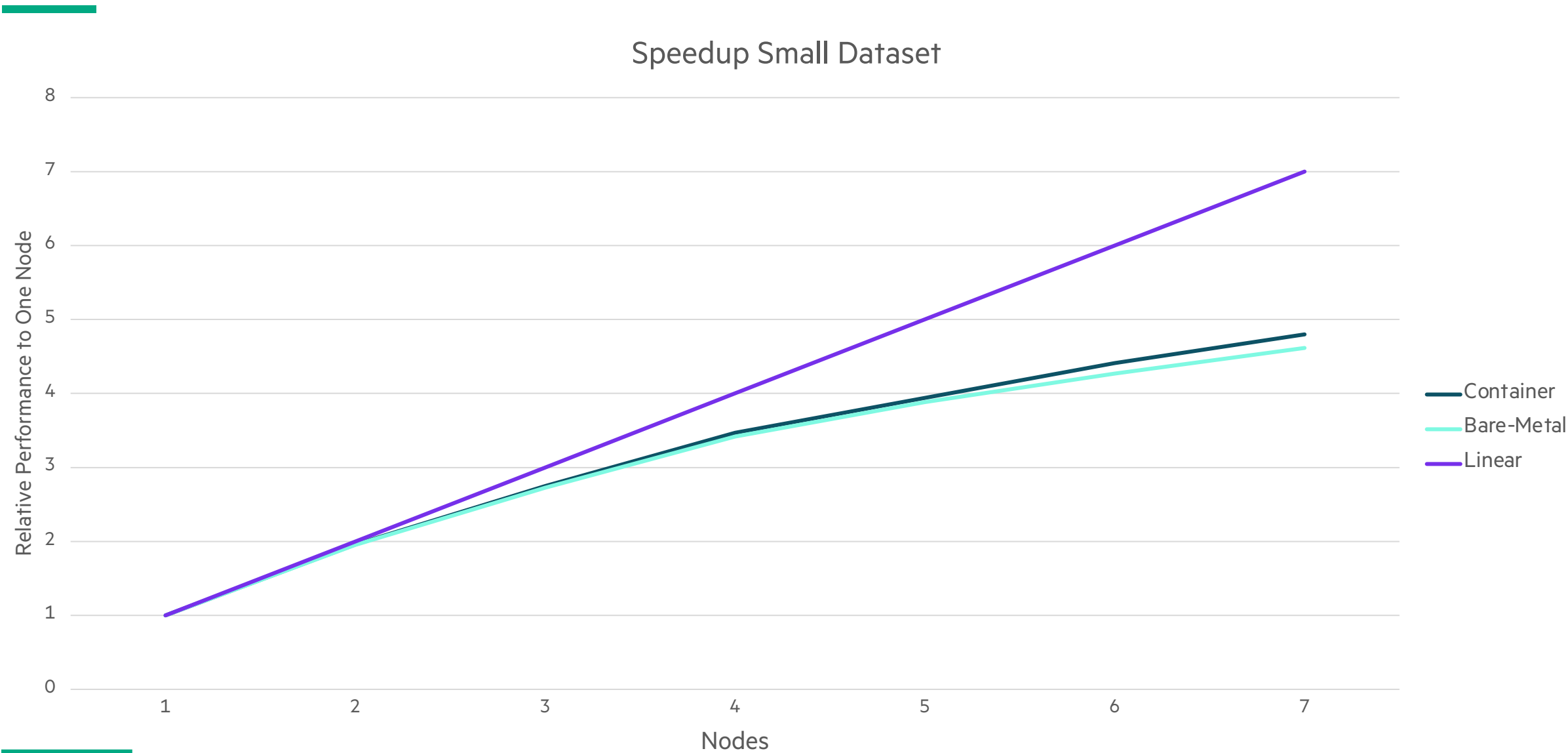
- Intel® oneAPI, ifort and icc.

## Libraries:

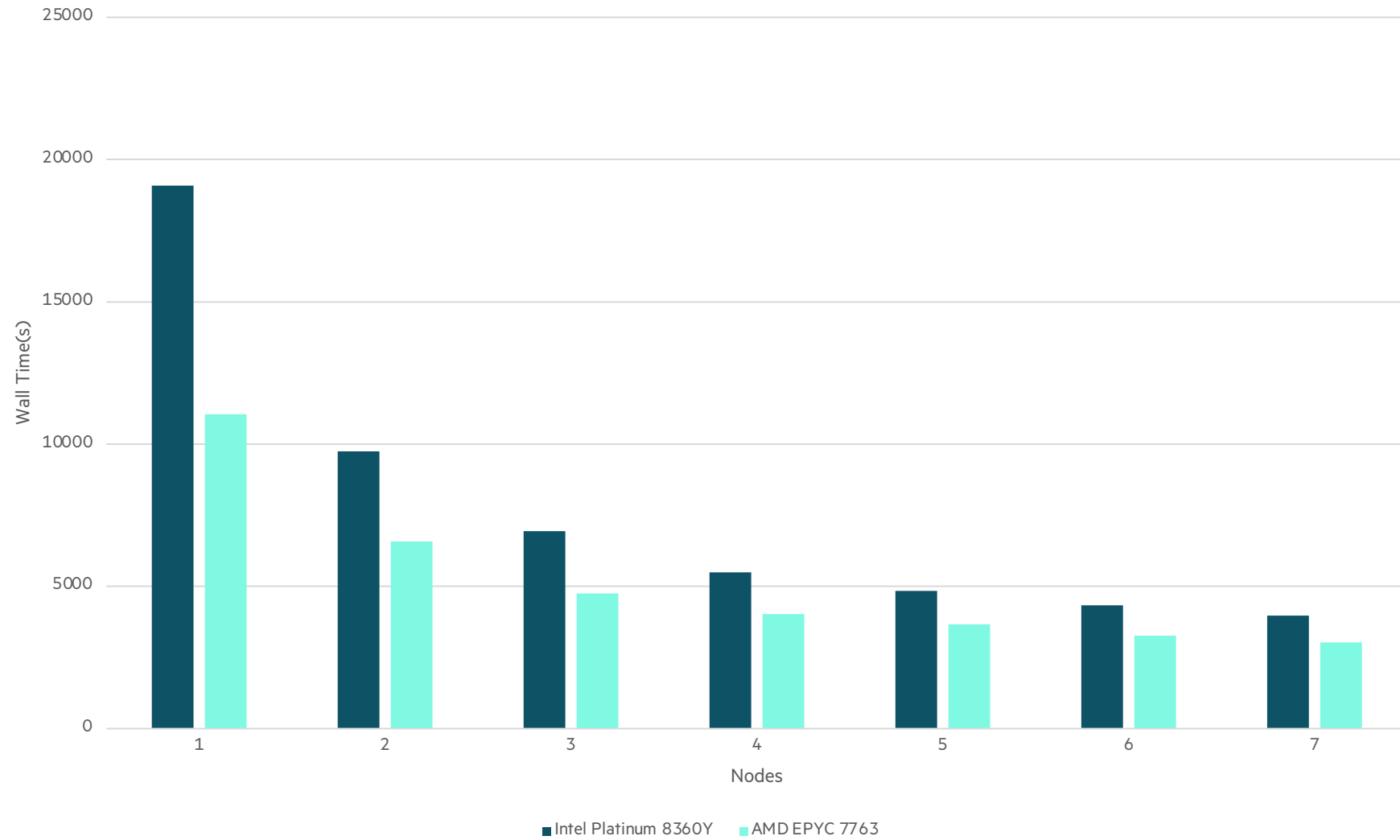
- NetCDF: Both NetCDF-C and NetCDF-Fortran (built with OneAPI)
- MPI: OneAPI
- Julian: for calendar calculation.
- Proj4: used for coordinate conversion.
- Petsc: the toolkit of scientific computation. This library is required when using semi-implicit scheme, data assimilation, non-hydrostatic, or SWAVE module.
- ESMF: This library is required when online nesting or WRF-FVCOM coupling is applied.



# FVCOM Performance: Container vs. Bare-metal

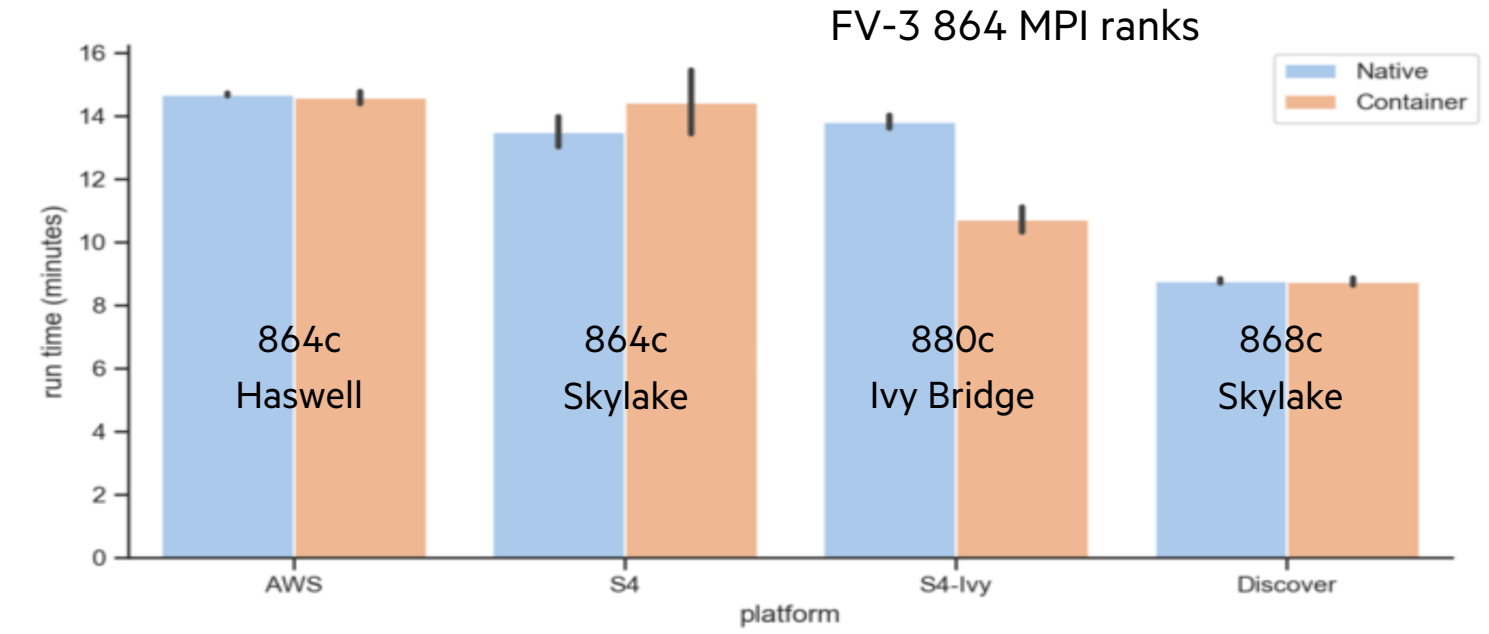


# FVCOM Container Performance: Milan vs. IceLake (AVX-2 binaries)



# FV-3 Performance – Previous work

Containers can perform better because of portable optimal settings.



HPC platforms:

- AWS – 24/c5n/EFA
- S4 -27/EDR/DFR compute, University of Wisconsin’s (SSEC)
- S4-Ivy – 44 compute/FDR
- Discover 31/FDR NASA: NCCS Discover supercomputer.

Miesch, M., N. Acks, T. Auligne, D. Hahn, S. Herbener, D. Holdaway; S. Nolin, K. Peck, J. Stroik, and Y. Tremolet, 10.2 - High-Performance "Supercontainers" for Earth System Prediction, Virtual Presentation, 101st Annual Meeting of the American Meteorological Society, 1/14/21.  
<https://ams.confex.com/ams/101ANNUAL/prelim.cgi/Paper/382433>



## Conclusions

---

- ❑ Reproducibility, portability, and agility
- ❑ Marketplace
- ❑ Curated HPE container
- ❑ Intel MPI has great portability, need bind for SS11



# Thank you

Paulo de Souza, paulo.souza@hpe.com  
Usama Anber, usama.anber@hpe.com

