SC23
Denver, CO | i am hpc.

Perspectives and Experiences Supporting Containers for Research Computing at the Texas Advanced Computing Center

Erik Ferlanti, William J. Allen, Ernesto A. B. F. Lima, Yinzhi Wang, John M. Fonner

Texas Advanced Computing Center
The University of Texas at Austin
Austin, Texas, USA

THE AUTHORS

Erik Ferlanti

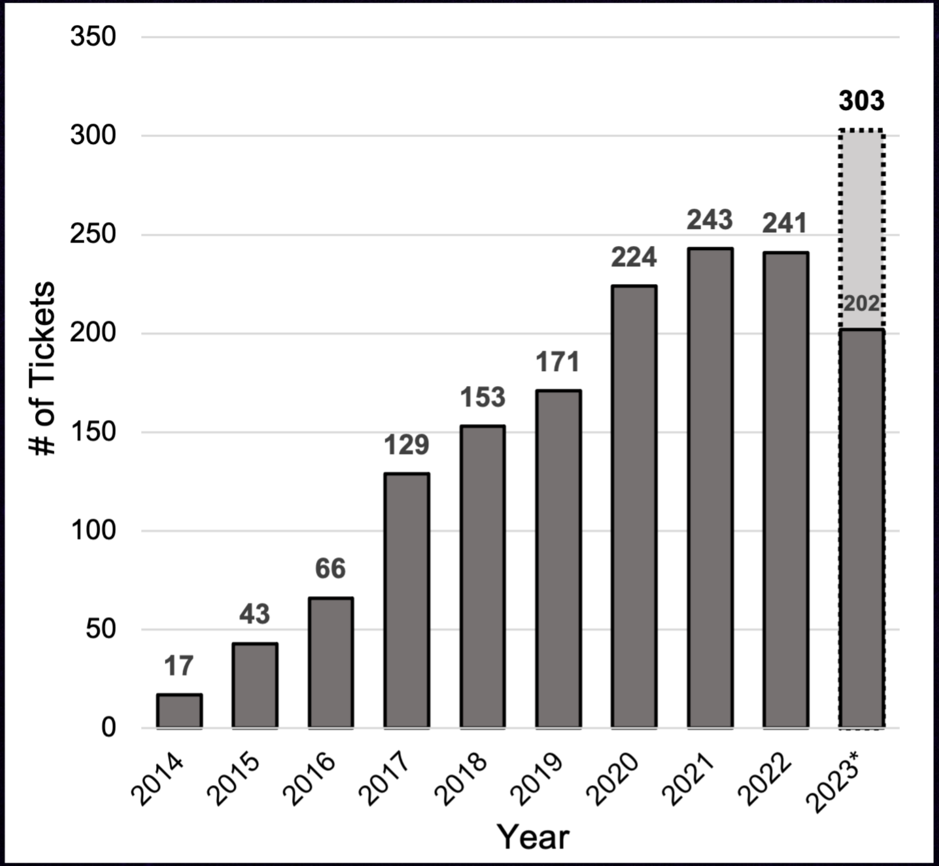
Joe Allen

Ernesto Lima

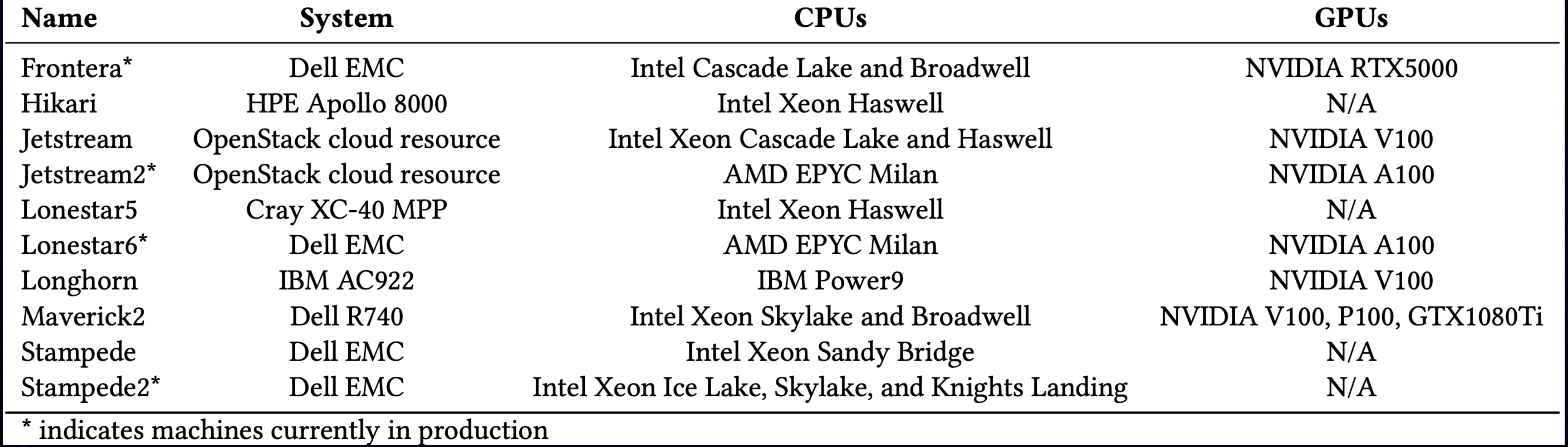
Ian Wang

CONTAINER INTEREST OVER THE LAST DECADE

- Limited to qualitative measurement
- System-wide module support started in 2016
- Most common ticket types:
  - help installing a software package
  - help pulling or running existing containers



TACC SYSTEMS WITH CONTAINER SUPPORT



Name	System	CPUs	GPUs
Frontera*	Dell EMC	Intel Cascade Lake and Broadwell	NVIDIA RTX5000
Hikari	HPE Apollo 8000	Intel Xeon Haswell	N/A
Jetstream	OpenStack cloud resource	Intel Xeon Cascade Lake and Haswell	NVIDIA V100
Jetstream2*	OpenStack cloud resource	AMD EPYC Milan	NVIDIA A100
Lonestar5	Cray XC-40 MPP	Intel Xeon Haswell	N/A
Lonestar6*	Dell EMC	AMD EPYC Milan	NVIDIA A100
Longhorn	IBM AC922	IBM Power9	NVIDIA V100
Maverick2	Dell R740	Intel Xeon Skylake and Broadwell	NVIDIA V100, P100, GTX1080Ti
Stampede	Dell EMC	Intel Xeon Sandy Bridge	N/A
Stampede2*	Dell EMC	Intel Xeon Ice Lake, Skylake, and Knights Landing	N/A

* indicates machines currently in production

APPTAINER CONFIGURATION

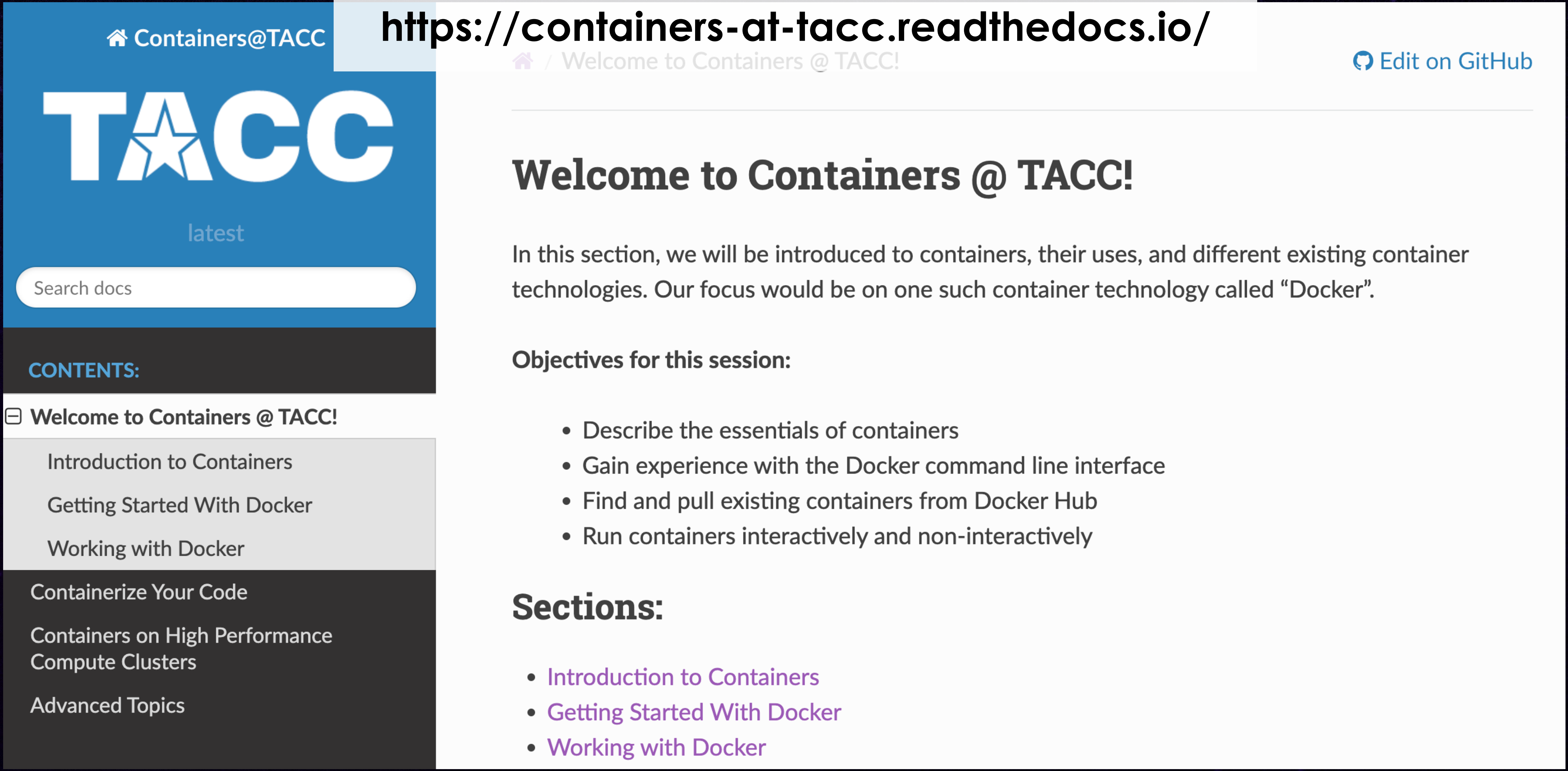
- Modules named “tacc-singularity” and “tacc-apptainer” to infer site customization
    - Both modules currently load Apptainer
- Image cache stored in user’s /work directory by default
- /home, /work, and /scratch mounted by default using Underlay
- Set User ID features are currently disabled for an evolving set of reasons
- Containers currently not allowed on login nodes

CONTAINERS AS MODULES

- "module load biocontainers" currently makes available an additional **3,247** bioinformatics apps (many with multiple versions)
- Modules generated using RGC (https://github.com/TACC/rgc)
- Executables are exposed as bash functions to match user expectations

```
$ type bwa
bwa is a function
bwa ()
{
    singularity exec ${BIOCONTAINER_DIR}/biocontainers/bwa/bwa-0.7.17--
pl5.22.0_2.sif bwa $@
}
```

https://containers-at-tacc.readthedocs.io/


Containers@TACC
TACC
latest
Search docs
CONTENTS:
Welcome to Containers @ TACC!
Introduction to Containers
Getting Started With Docker
Working with Docker
Containerize Your Code
Containers on High Performance Compute Clusters
Advanced Topics
/ Welcome to Containers @ TACC!
Edit on GitHub

Welcome to Containers @ TACC!

In this section, we will be introduced to containers, their uses, and different existing container technologies. Our focus would be on one such container technology called “Docker”.

Objectives for this session:

- Describe the essentials of containers
- Gain experience with the Docker command line interface
- Find and pull existing containers from Docker Hub
- Run containers interactively and non-interactively

Sections:

- Introduction to Containers
- Getting Started With Docker
- Working with Docker

TRAINING APPROACH

- Material hosted on readthedocs.io to support self-guided learning
- Use the Docker toolchain for container creation
- Encourage using the local laptop as the development environment when possible
- Provide MPI- and GPU-enabled base images compatible with TACC systems
- Cover fully-worked hands-on examples throughout
- 3-6 hours of content repackaged as appropriate

BASE IMAGES

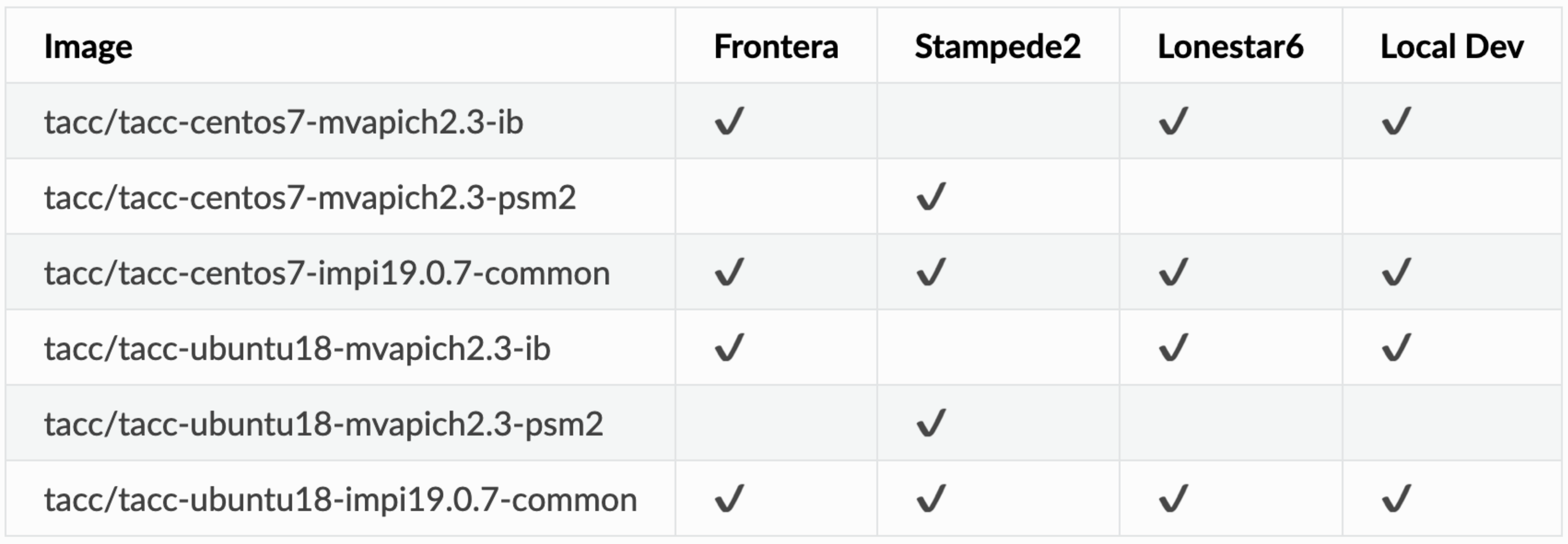
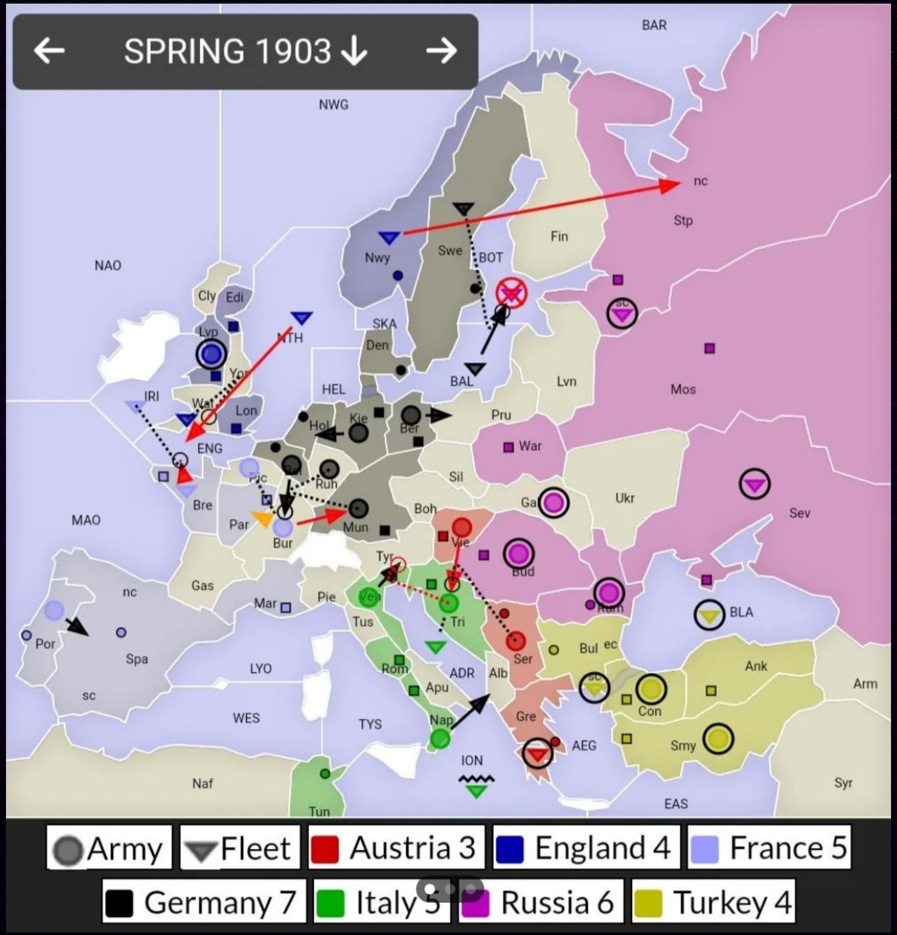


Image	Frontera	Stampede2	Lonestar6	Local Dev
tacc/tacc-centos7-mvapich2.3-ib	✓		✓	✓
tacc/tacc-centos7-mvapich2.3-psm2		✓		
tacc/tacc-centos7-impi19.0.7-common	✓	✓	✓	✓
tacc/tacc-ubuntu18-mvapich2.3-ib	✓		✓	✓
tacc/tacc-ubuntu18-mvapich2.3-psm2		✓		
tacc/tacc-ubuntu18-impi19.0.7-common	✓	✓	✓	✓

EXAMPLE USE CASES

- Jupyter Notebook environment that is consistent between Cloud and HPC
- VS Code Server container on Longhorn
  - Based on https://github.com/coder/code-server
- Tightly controlled ML development environment running external code
- Isolated environments for AI agents on Frontera and Lonestar6 playing Diplomacy
- Tapis and Abaco APIs all built around containers



CONCLUSIONS AND FUTURE PERSPECTIVES

- Containers are a hammer that makes other problems look like a nail.
  - Reduce the friction of broad software support
  - Provide software continuity between users, between laptop and HPC, and across production systems
- Education and training are critical!

- Regular evaluation of container platforms
- Support for building containers on TACC systems
- Kubernetes offerings
- Multi-architecture builds will increase in importance

THANK YOU!

Questions? Comments? Feedback? Suggestions?

John Fonner

jfonner@tacc.utexas.edu

(also on Slack)