

# Using containers in HPC for research workflows

**Mark Gray, Brian Skjerven and Marco De La Pierre**

*The Pawsey Supercomputing Centre is an unincorporated joint venture between*



*and proudly funded by*



# Pawsey Supercomputing Centre





# Improving research workflows

**What problems are we trying to solve?**

**Portability**

**Reproducibility & Provenance**

**Collaboration**

**Software dependencies**

**Ease of use**

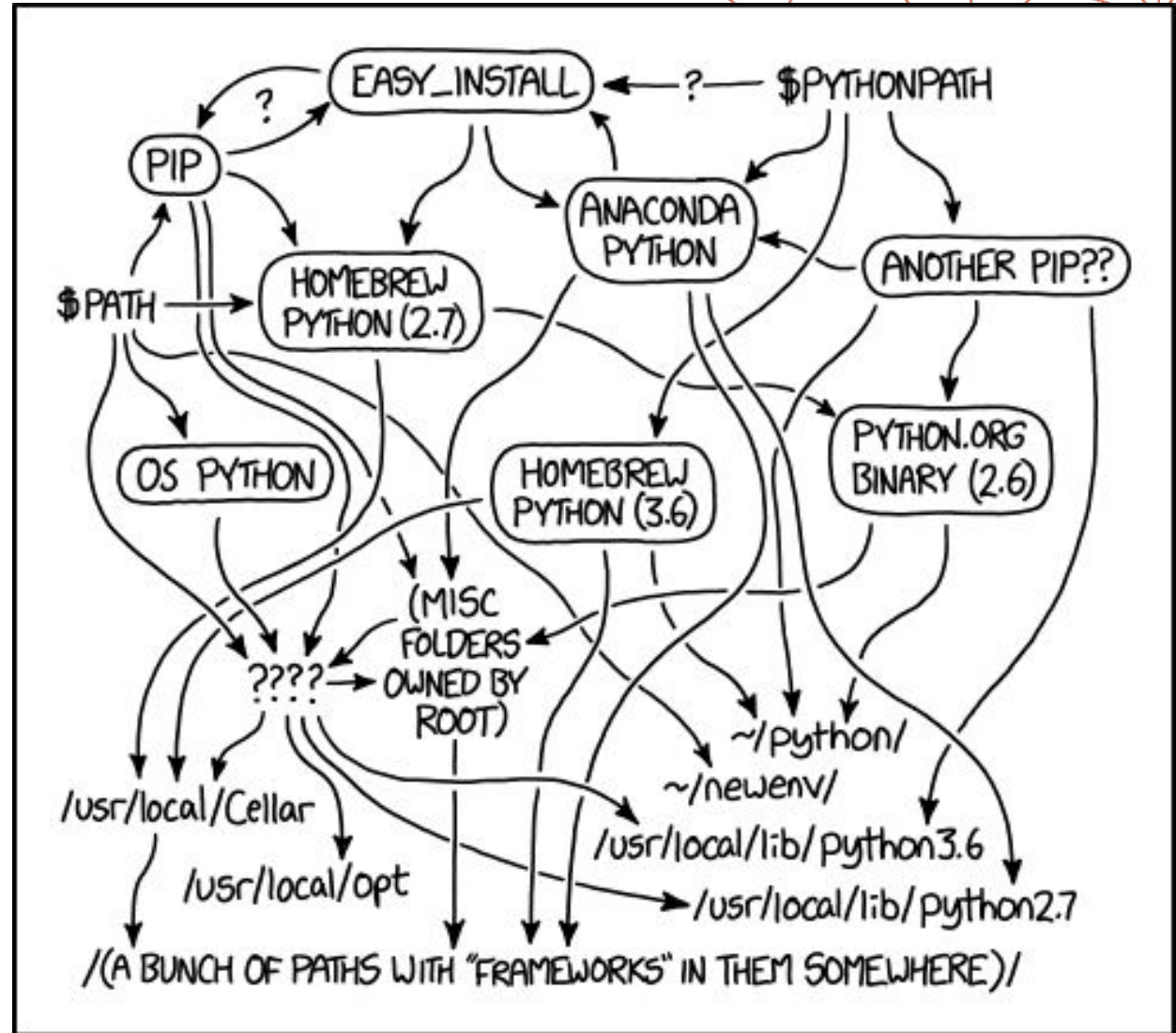
**Performance**

**Python and apps with heavy I/O**



# Python? I thought we like Python?

## We do, but ...



MY PYTHON ENVIRONMENT HAS BECOME SO DEGRADED  
THAT MY LAPTOP HAS BEEN DECLARED A SUPERFUND SITE.

# Containers: a crash course

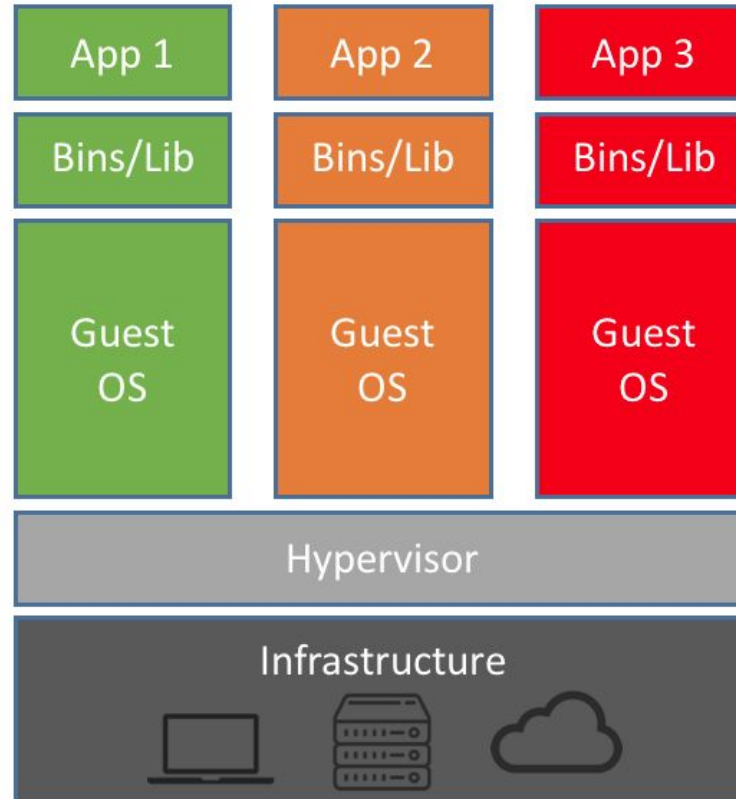
## Containers

- **Virtualized OS**
- **Use host's kernel (and sometimes other stuff)**
- **Kernel is shared with other containers**
- **Light weight (only what you put in)**

## VMs

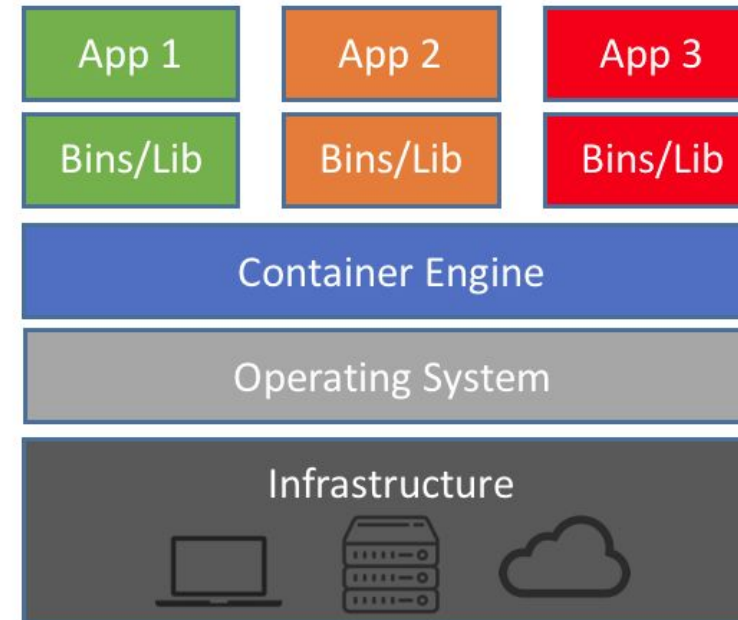
- **Virtualized hardware**
- **VMs have their own kernel**
- **Hypervisor manages sharing hardware between VMs**
- **Full blown OS, drivers, kernel, etc.**

# Containers vs. VMs



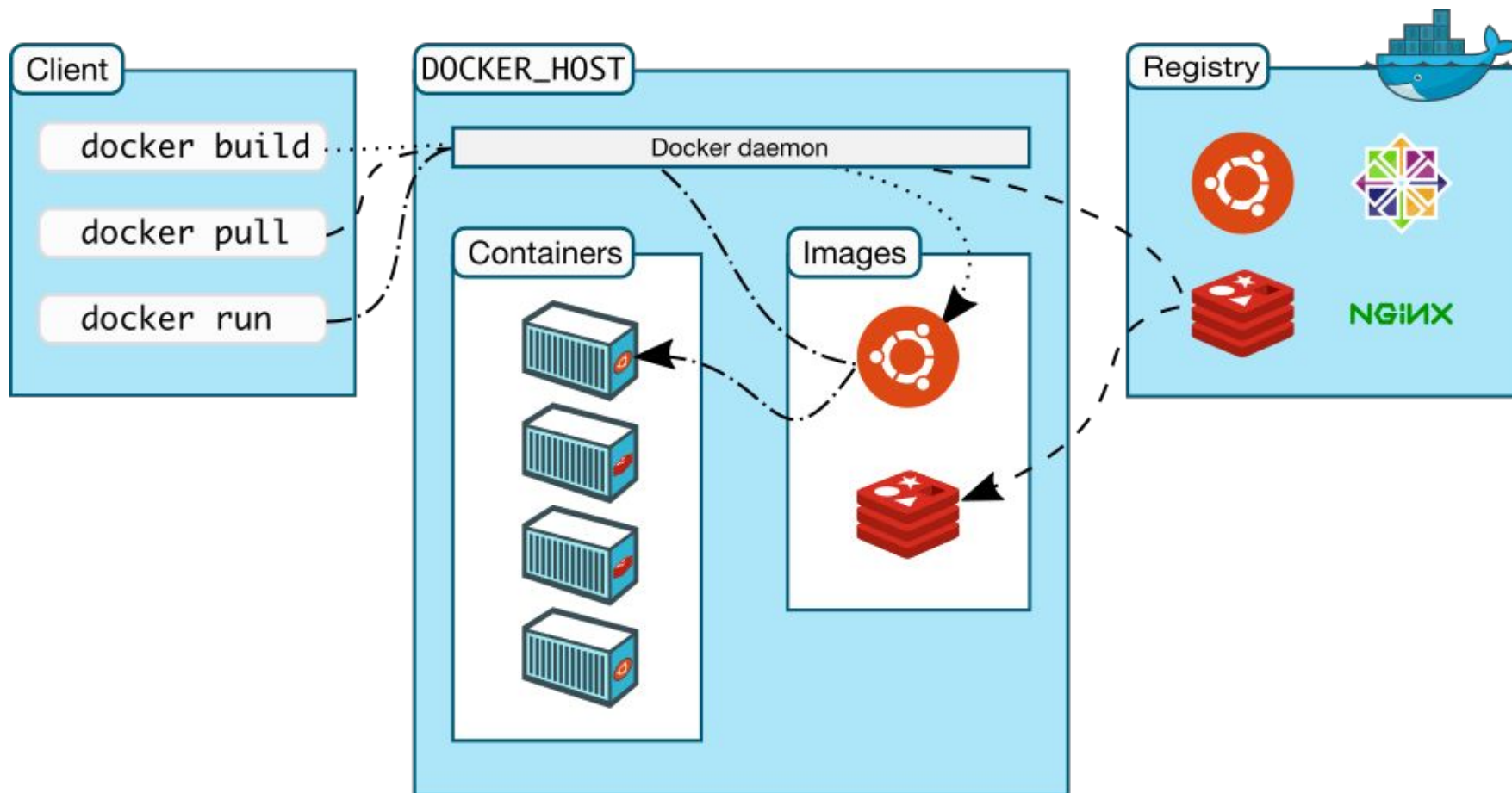
Machine Virtualization

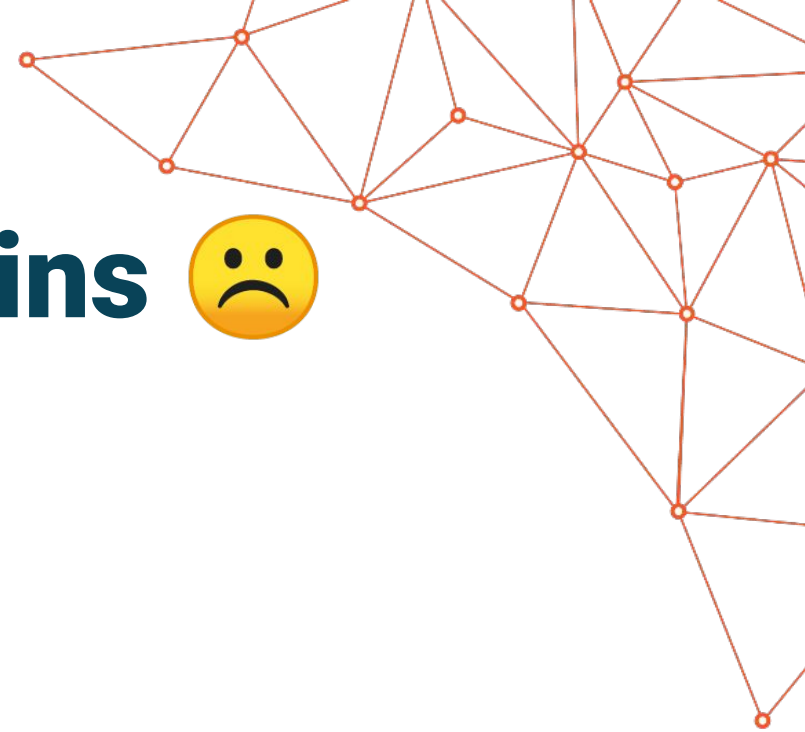
- **Containers are isolated**
- **Share OS and sometimes bins or libs**



Containers

# Docker Concepts





# **Containers are easy, yay!**

## **Not so much for HPC sysadmins**

- **Why not just run Docker?**
- **A few issues:**
  - **Security – Docker assumes root privileges**
  - **Batch Systems – Doesn't play nice with schedulers**
  - **Underlying system – Usually requires an up-to-date kernel**



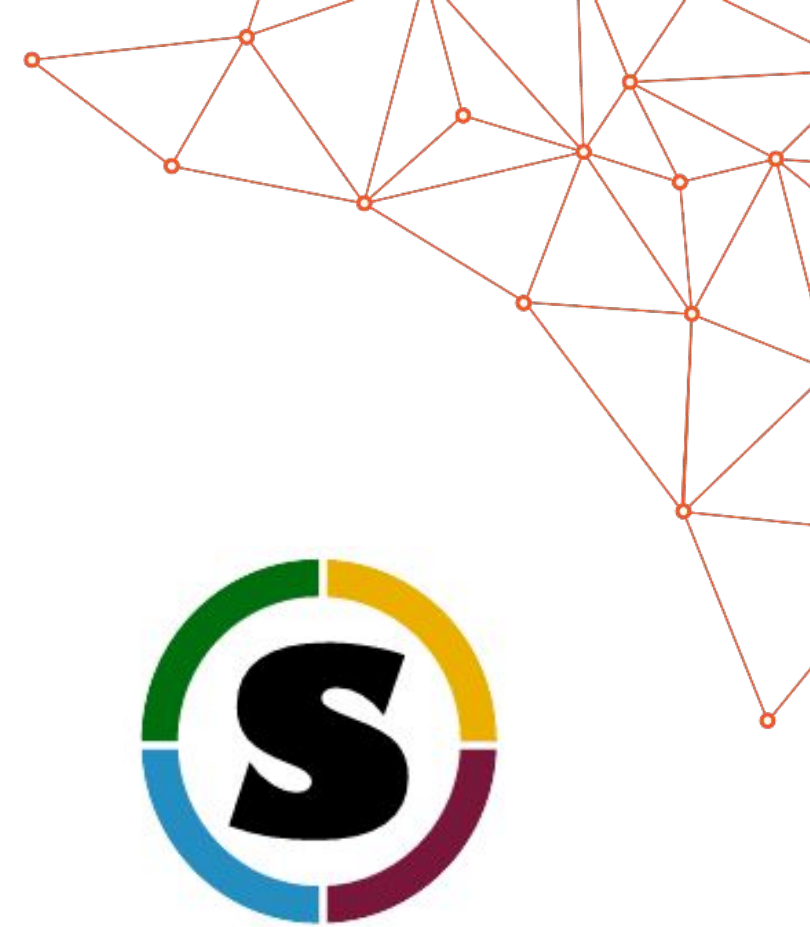
# Shifter

- **Open-source project developed by NERSC**
  - Cray & CSCS
- **“Docker-like” interface on a Cray**
- **Easy to use**
  - Users create images locally (or use Pawsey provided)
  - Upload to Dockerhub and pull to Pawsey
  - Minimal modification to workflows
- **GPU & MPI support**



# Singularity

- **Similar to Shifter**
- **Developed at Lawrence Berkley Lab**
- **Own image format (but can run Docker images)**
- **MPI & GPU support**



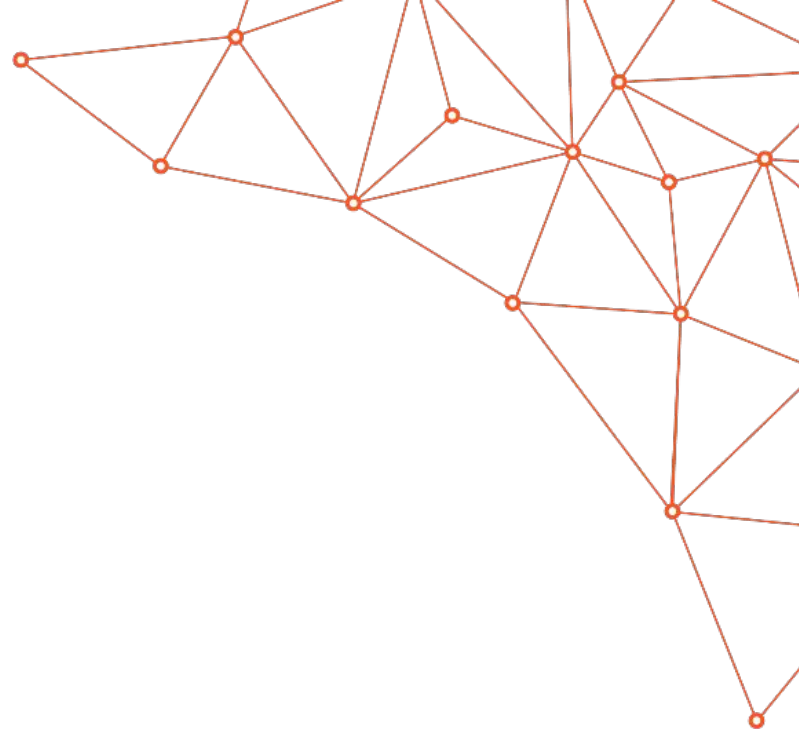
# MPI Performance

- **OSU Benchmarks**

- **Bandwidth**
- **Latency (MPI/MPI+OpenMP)**
- **Bi-directional Bandwidth**
- **Collectives**
  - **All Gather latency**
  - **All Reduce latency**

- **Runs**

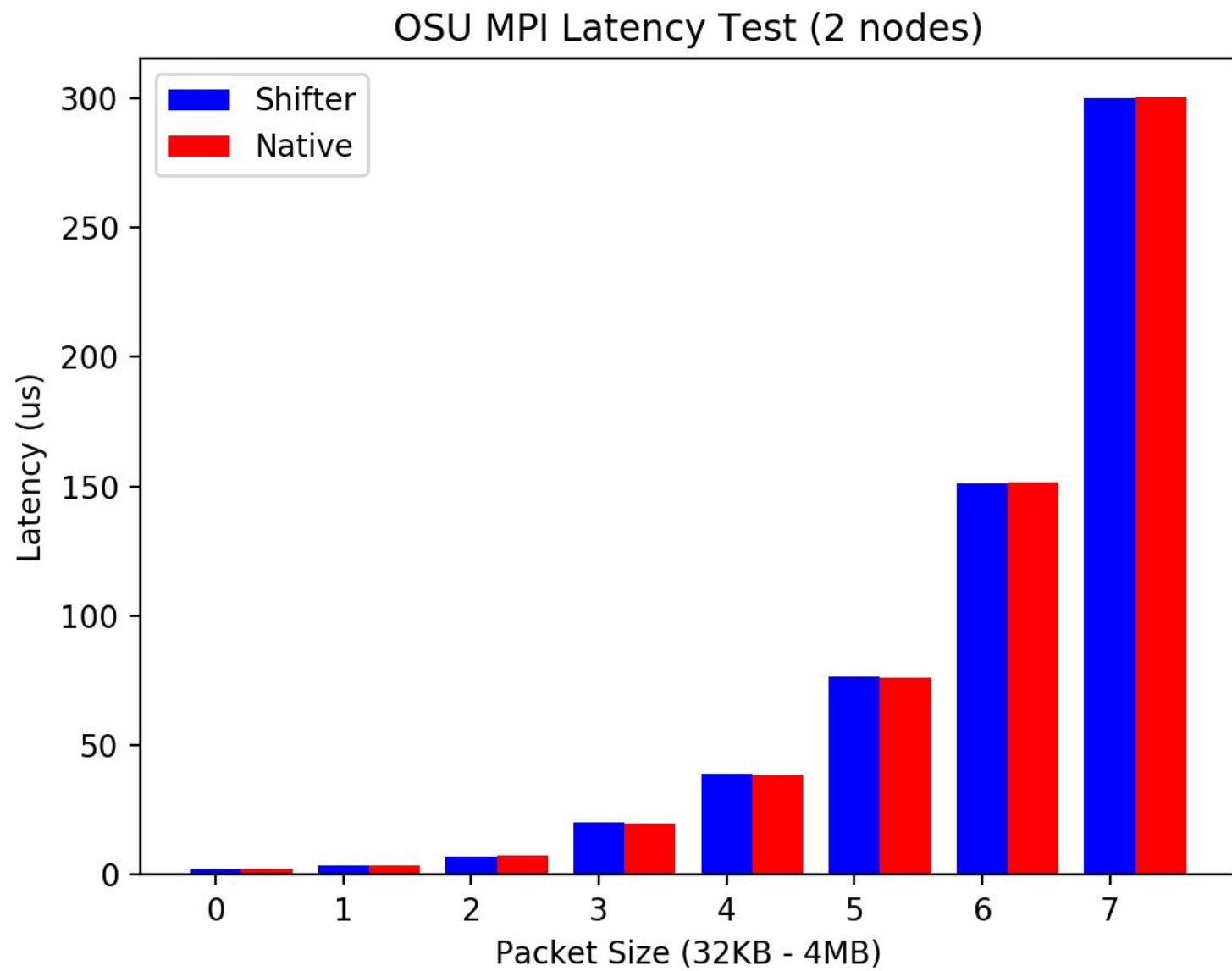
- **Shifter (MPICH/Cray MPI, GCC 5, Python 2)**
- **Native (GCC 5, Cray MPI)**

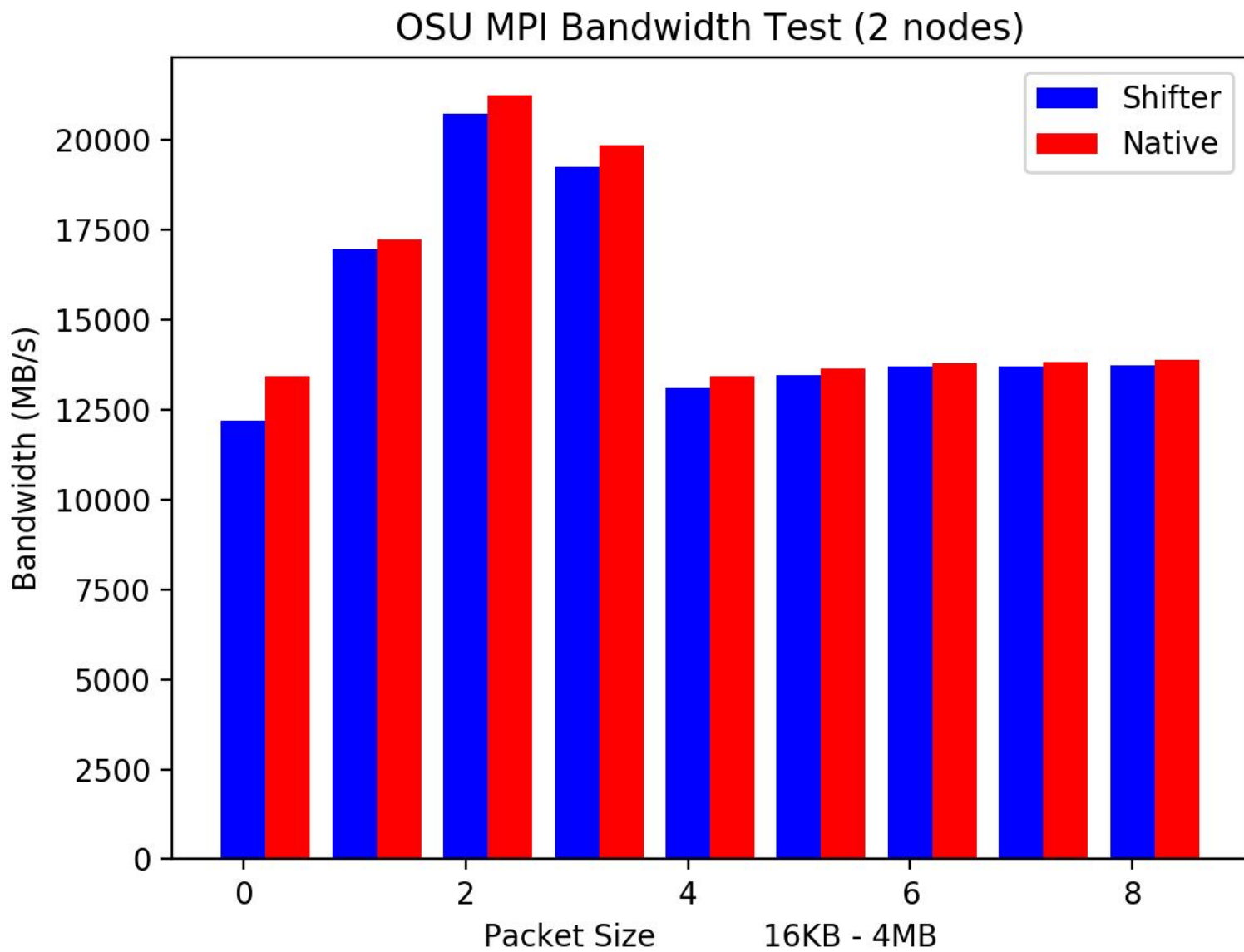


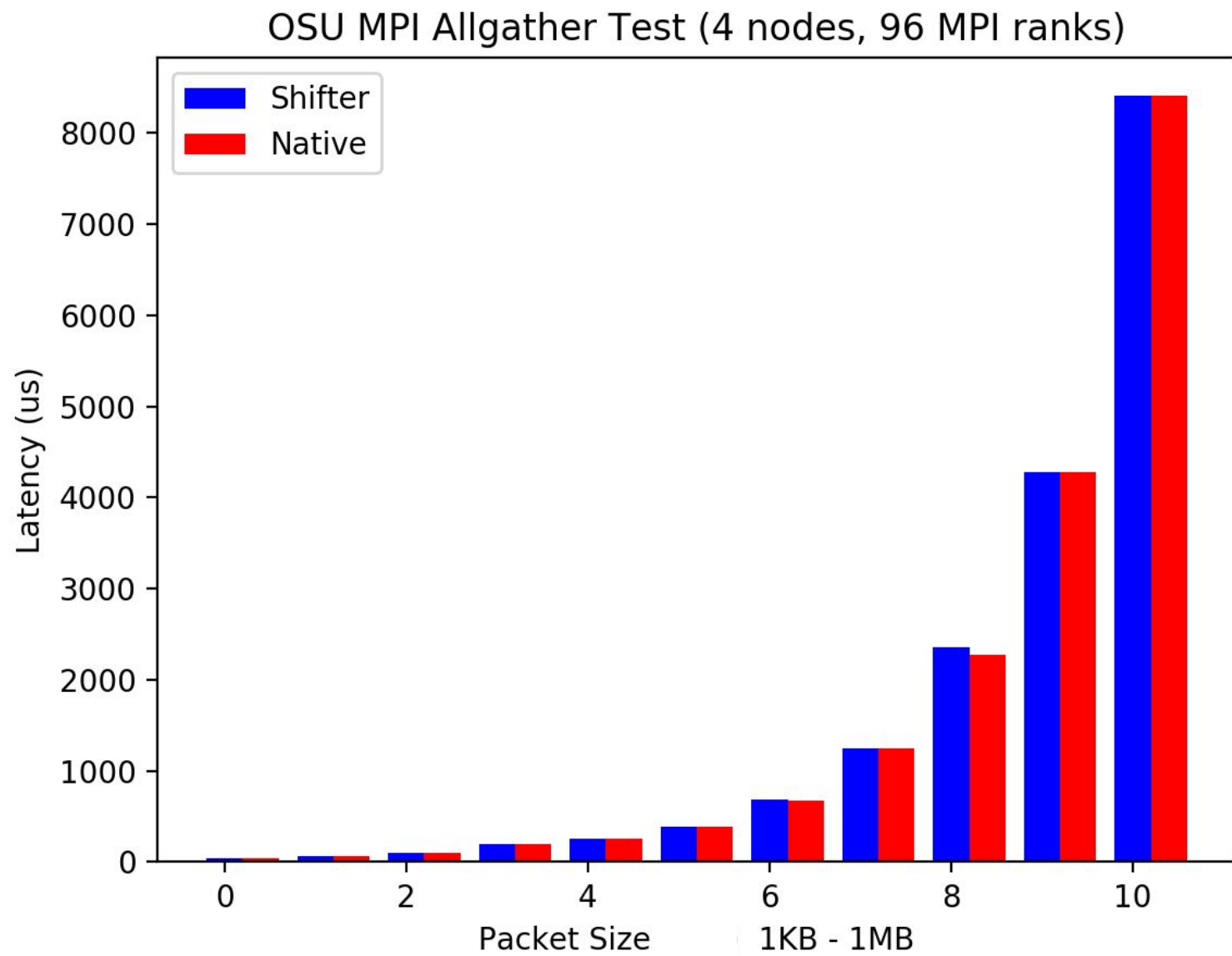
# Native MPI Libraries

- **Shifter images can use Cray MPI libraries and Aries interconnect**
- **Image built with stock MPICH**
  - **ABI compatibility with Cray MPI**
  - **Update Idconfig to look at Magnus paths**
- **At runtime Shifter mounts Magnus libraries into container**









# Back to Python (and Containers)

- **Python now much more scalable**
  - Radio astronomers rejoice
- **Ease of package install**
- **Pawsey provides a base Docker images for Python**

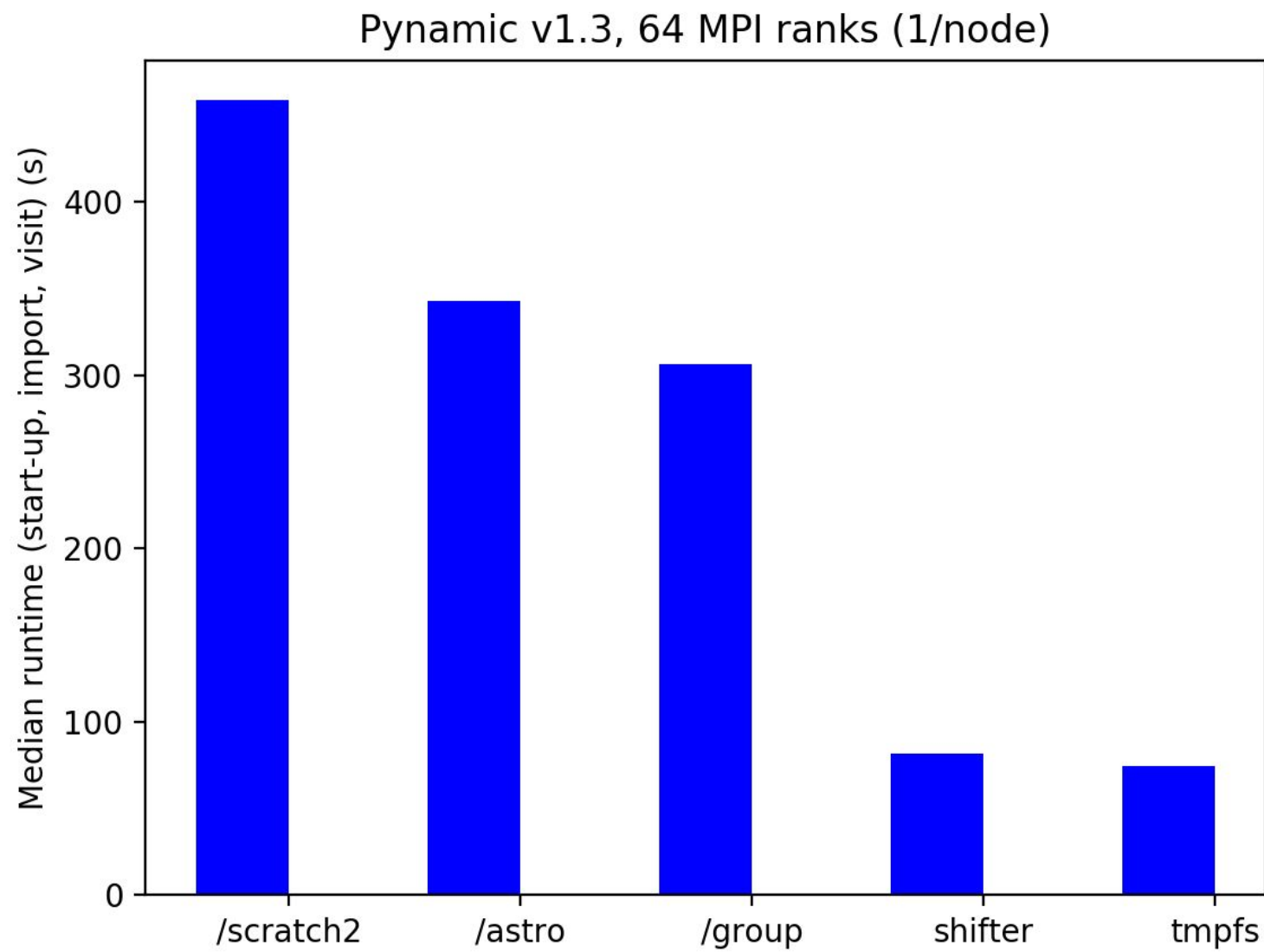
<https://hub.docker.com/r/pawsey/hpc-python/>





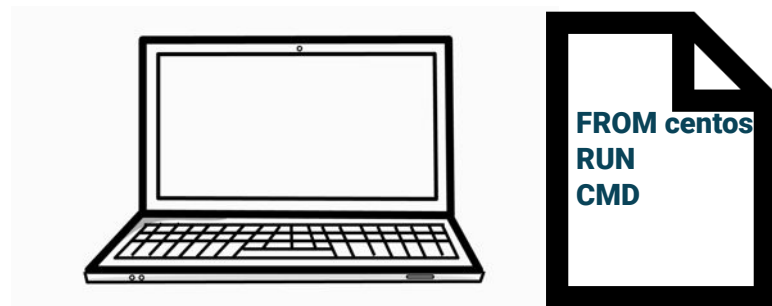
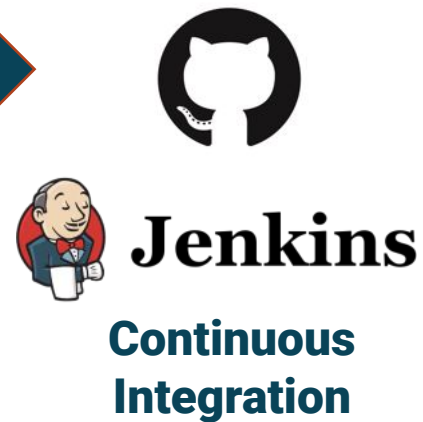
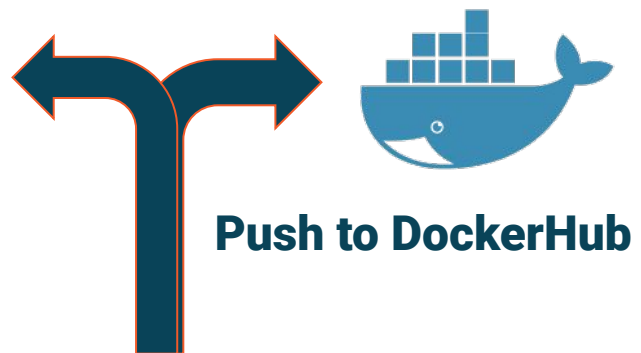
# Dynamic

- **Synthetic Python code**
  - **Test dynamic library loading in Python applications**
  - **Built on MPI**
- **Generates shared objects**
  - **Creates C source files**
  - **Each lib has multiple Python callable functions (avg. number for randomness)**
  - **Cross library function calls**



# A real live workflow

Share w/collaborators



Develop container locally



Pull to Pawsey

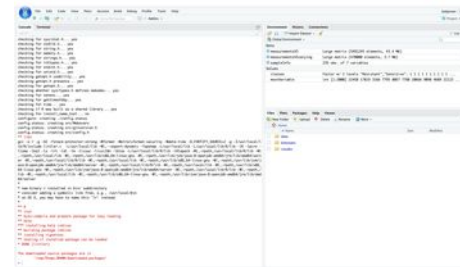


Private  
Registry

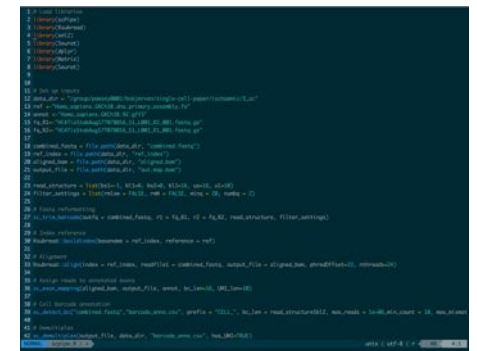
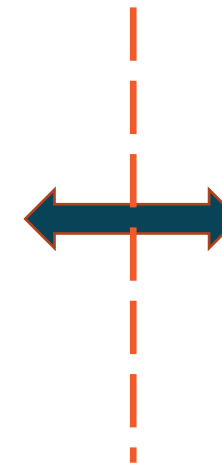
# Case Study

## Telethon Kids Institute / UWA

- **Single Cell workflows**
- **Combination of Software**
  - R (Bioconductor, CRAN, Git)
  - Others (e.g. Cell Ranger)
- **Issue of Scale**
  - Workflow is “bursty”
- **Collaboration**
  - Closed filesystems at Pawsey
  - Containerised file sync

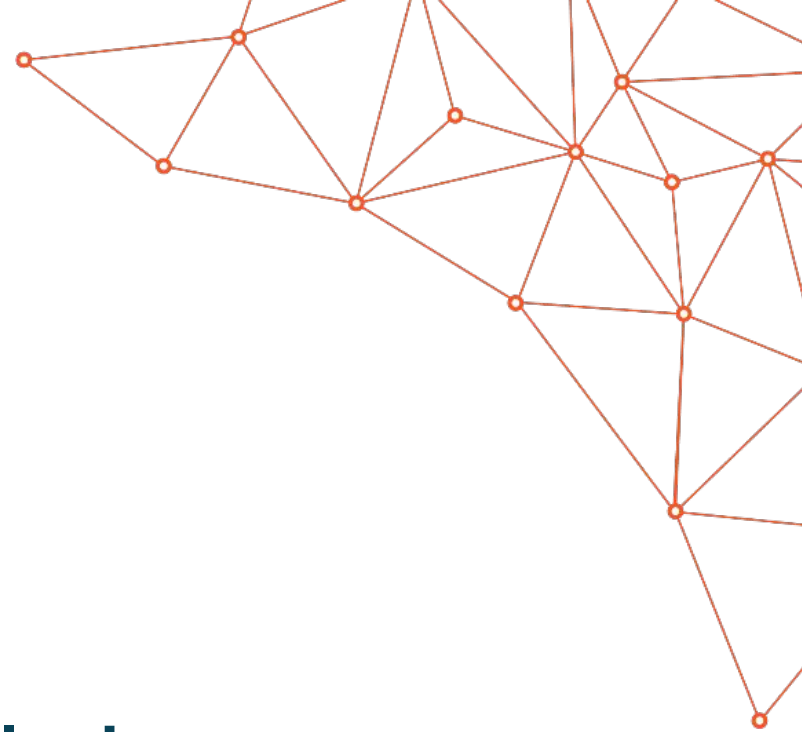


**RStudio Container  
(Local/Cloud)**



**RStudio Container + R Script  
(Pawsey)**





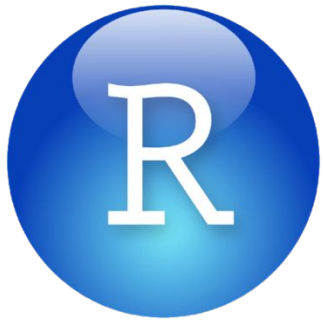
# Single Cell Gen. pipeline

- **One container for Cellranger**
- **One container for R downstream analysis**
  - **Seurat / SingleR**
  - **Straightforward setup of RStudio server on Pawsey Cloud**
  - **Monolithic approach**
    - **Same R container used on HPC (Rscript) and Pawsey Cloud (RStudio)**
- **Docker use case: has host-container directory mapping**
  - **Puzzling for beginners**
  - **To facilitate use: define and use aliases and variables**

# Containers in use at Pawsey

BIOCONDA®

 Keras



 Bioconductor  
OPEN SOURCE SOFTWARE FOR BIOINFORMATICS

 PyTorch

10X  
GENOMICS



Open  FOAM

 TensorFlow

**...and adding more  
each day**

# Challenges and future plans...

**So, a funny thing happened on the way to  
whole-centre container workflow rollout...**

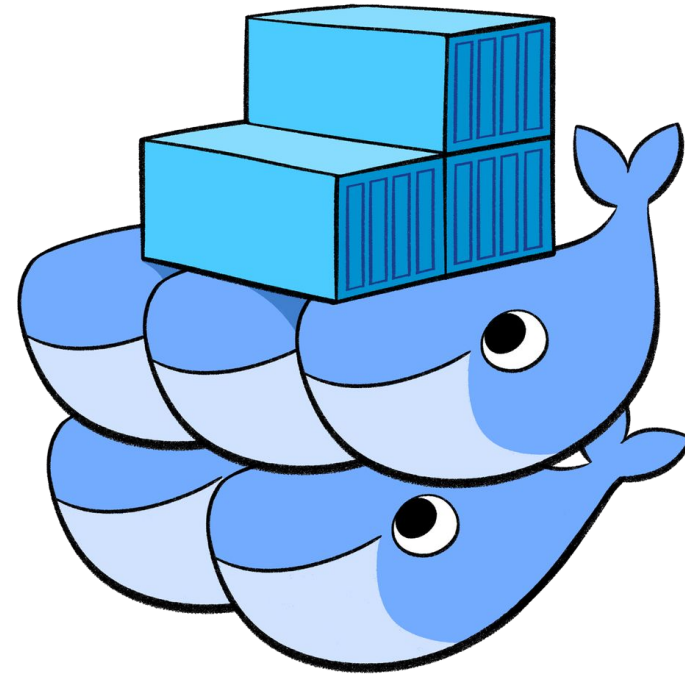


# Orchestration: talking to each other

## Kubernetes



## Docker Swarm








# Containers: external and internal

**User  
workflows**





# Questions?



HOME ABOUT US ▾ SYSTEMS & SERVICES ▾ SCIENCE SHOWCASE ▾ NEWS & EVENTS ▾ SUPPORT USER PORTAL JOBS CONTACT PAWSEY FRIENDS

## Online Training: Overview of Containers in HPC

📅 21 February 2019  
🕒 9:30am - 11:00am  
📍 ,

**Thursday 21<sup>st</sup> February 2019, 9:30am – 11:00am AWST**  
**Please join us for our Webinar Session (90 minutes) :**

- » Brief introduction to containers
- » Overview of container options at Pawsey
- » Example workflows and benchmarks using containers (hands-on)

**Who should attend?**

<https://pawsey.org.au/event/>



# Sample Jobscript

```
#!/bin/bash
#SBATCH --nodes=1
#SBATCH --partition=workq

module load shifter

srun -n 24 shifter run pawsey/hpc-python:latest python
PythonScript.py args
```