Singularity Containers

Containers for …
HPC, analytics, machine learning, reproducible and trusted computing

http://sylabs.io

DAVE GODLOVE

- Software Engineer, Sylabs Inc.
- Previous:
  - Computational Biologist, National Institutes of Health, High Performance Computing Center
  - Postdoctoral Fellow, National Institute of Mental Health

GREGORY M. KURTZER

- CEO, Sylabs Inc.
- Senior Architect, RStor Inc.

- Previous:
  - HPC Systems Architect, LBNL (~20 years)

- Open Source Work:
  - Founder and project lead: Warewulf (2001)
  - Founder: Centos Linux (2003)
  - Founder and project lead: Singularity (2015)

Institut Pasteur

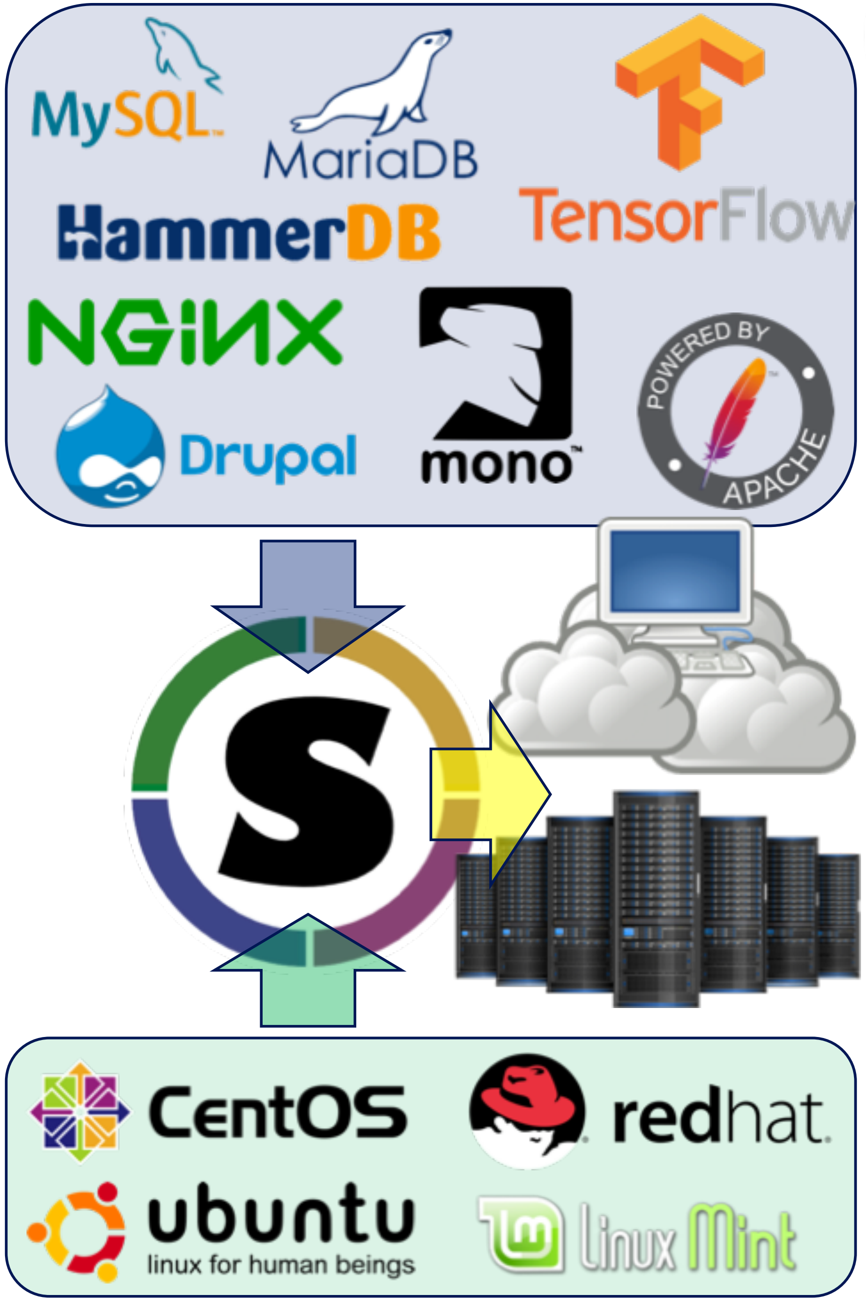
“Singularity + Warewulf + Centos: Winning combo!”

Introduction to Containers and Singularity

CONTAINERS 101

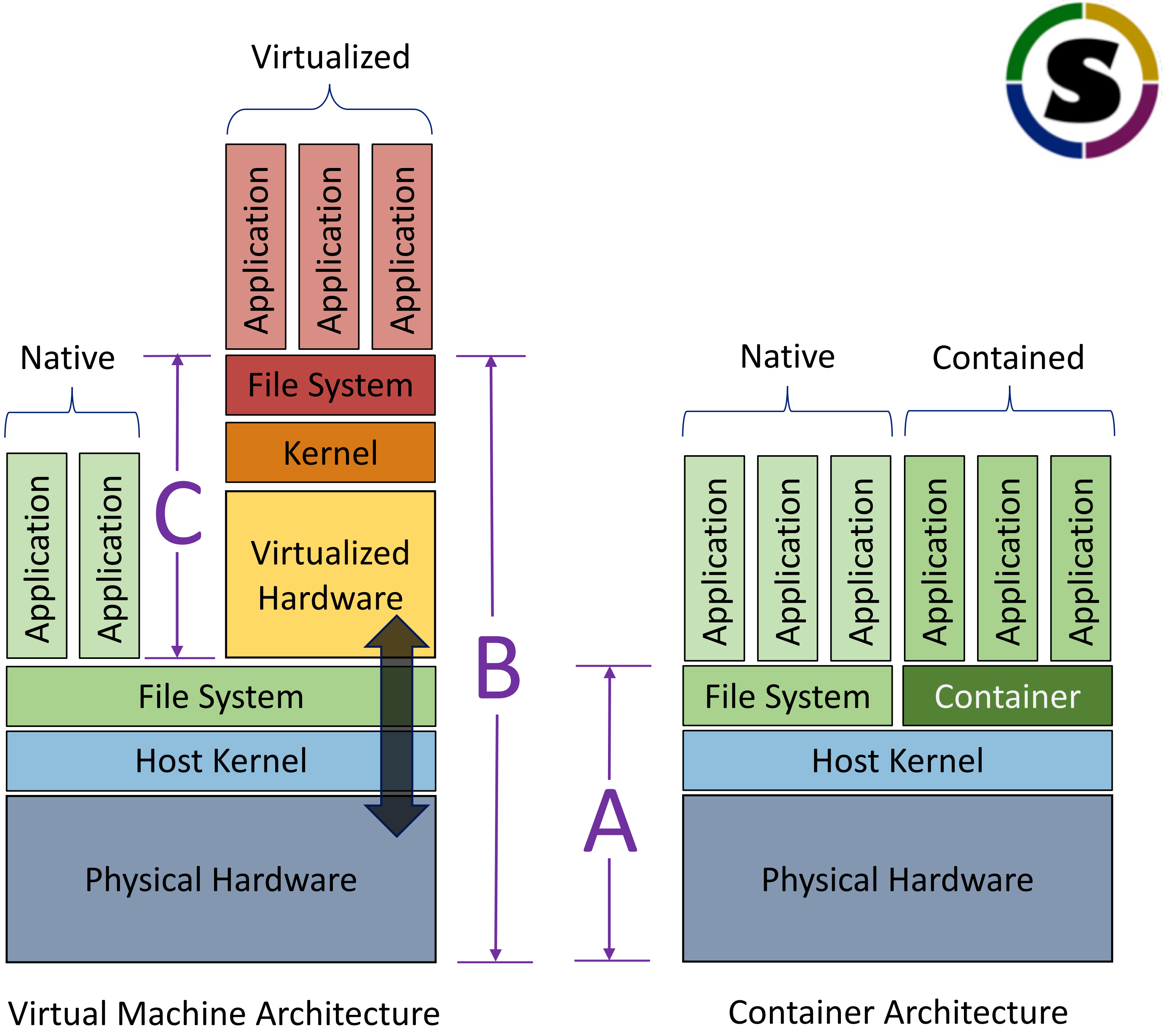
- Containers are encapsulations of operating system environments
- Includes all applications, libraries, configs and dependencies
- Workflows are completely self contained and portable
- Easy to share, distribute, validate and reproduce

- Multiple implementations exist
- Singularity differentiates itself in its architecture and container format



COMPARISON: VIRTUAL MACHINES

A is always less then B due to C



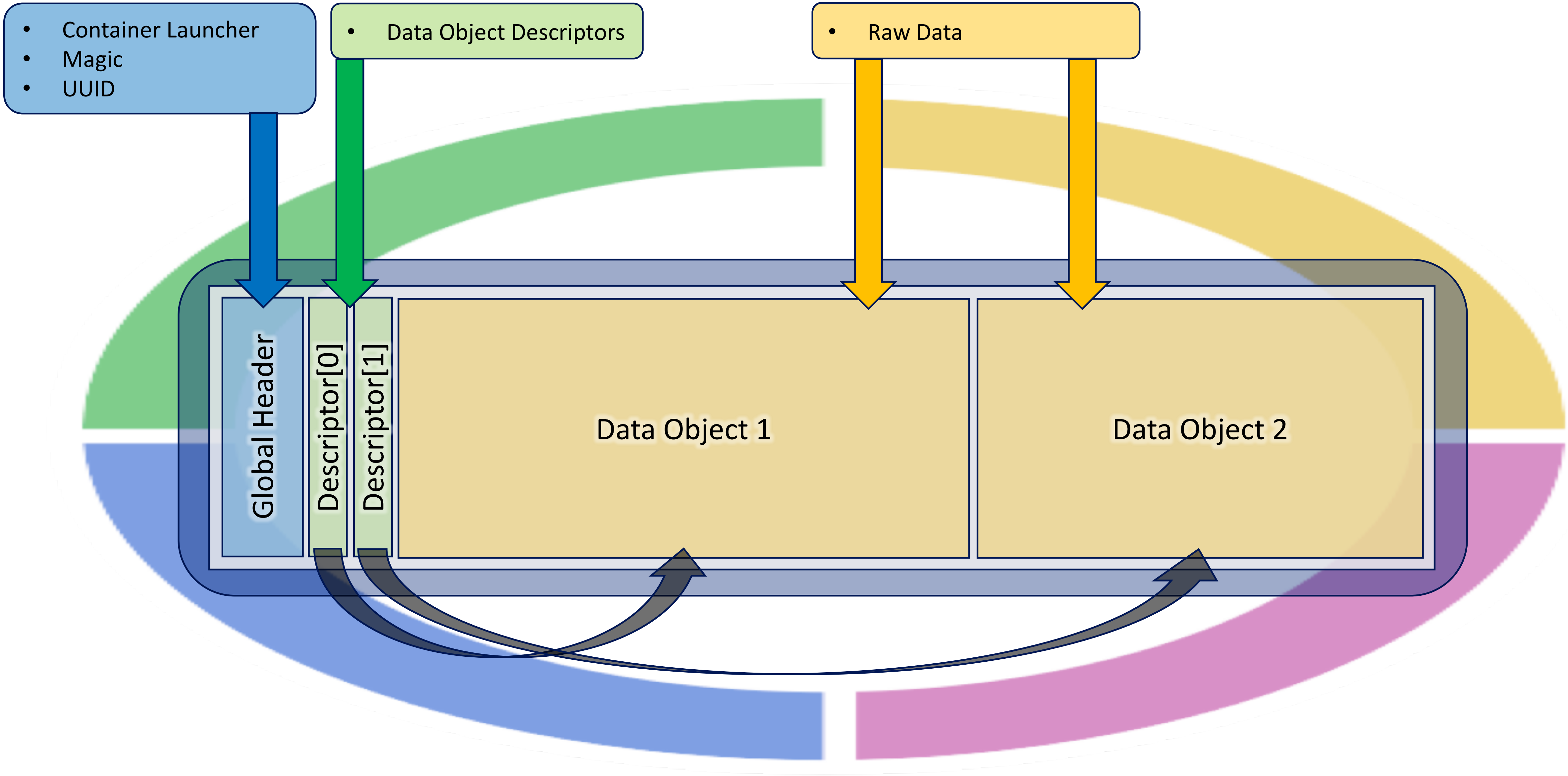
What makes Singularity different?

UNIQUE FEATURES OF SINGULARITY?

- **Built specifically to support HPC/Science:** Singularity was built by demand, requests, threats and bribes by researchers, scientists, and computational users
- **Single file based container format:** verifiable via checksum and cryptographic signatures ensuring reproducible and validated software environments during runtime
- **Extreme mobility:** using standard tools (rsync, scp, GridFTP, NFS, etc.)
- **Controls compliant:** images can be easily archived and managed as any other data
- **Compatible:** with complicated architectures (e.g. HPC, Machine Learning, Cloud, etc.)
- **Security model:** designed to support untrusted users running untrusted containers *(rather then trusted users running trusted containers)*

The Singularity Image Format (SIF)

SINGULARITY IMAGE FORMAT (SIF)

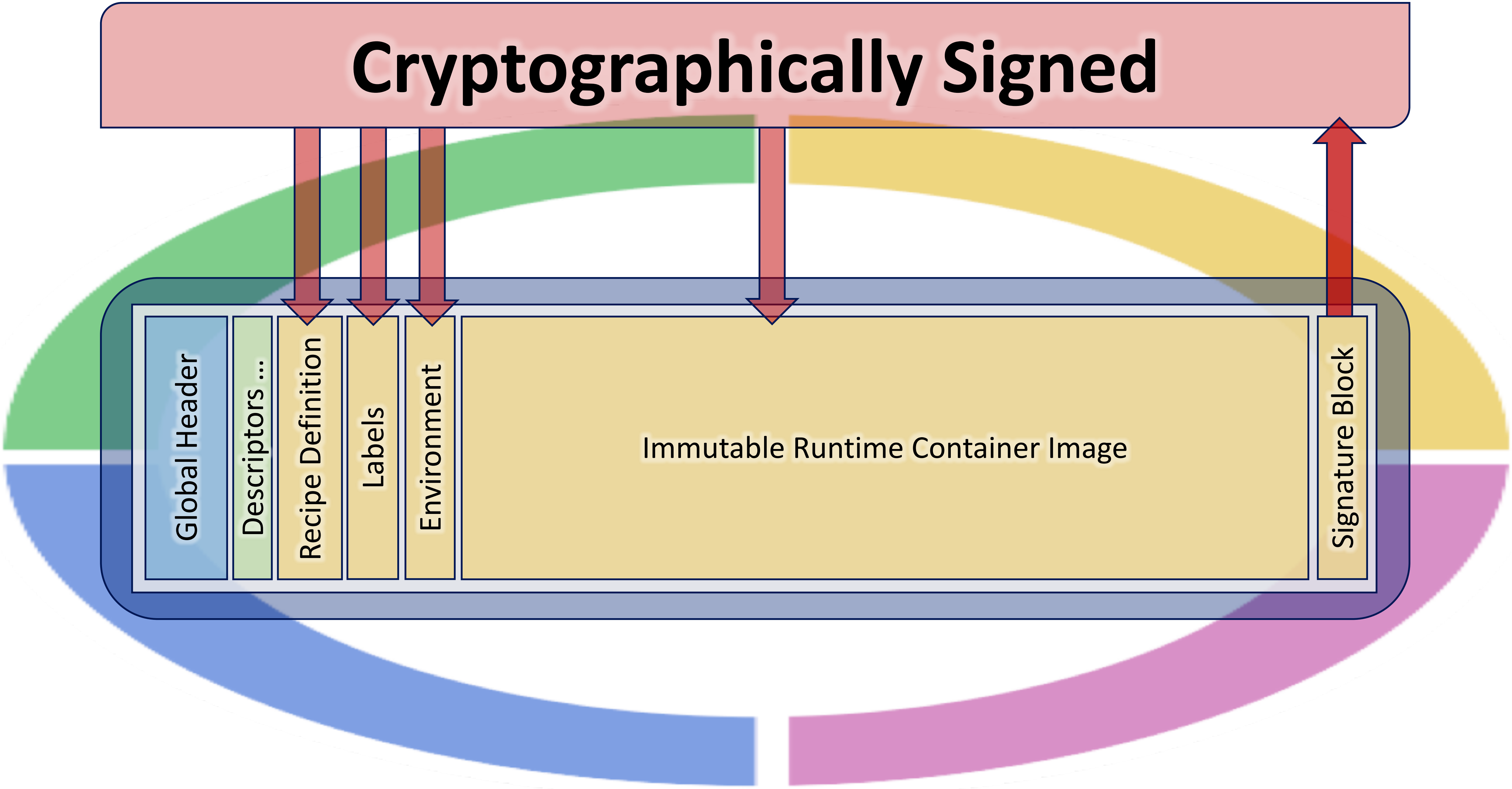


- Container Launcher
- Magic
- UUID

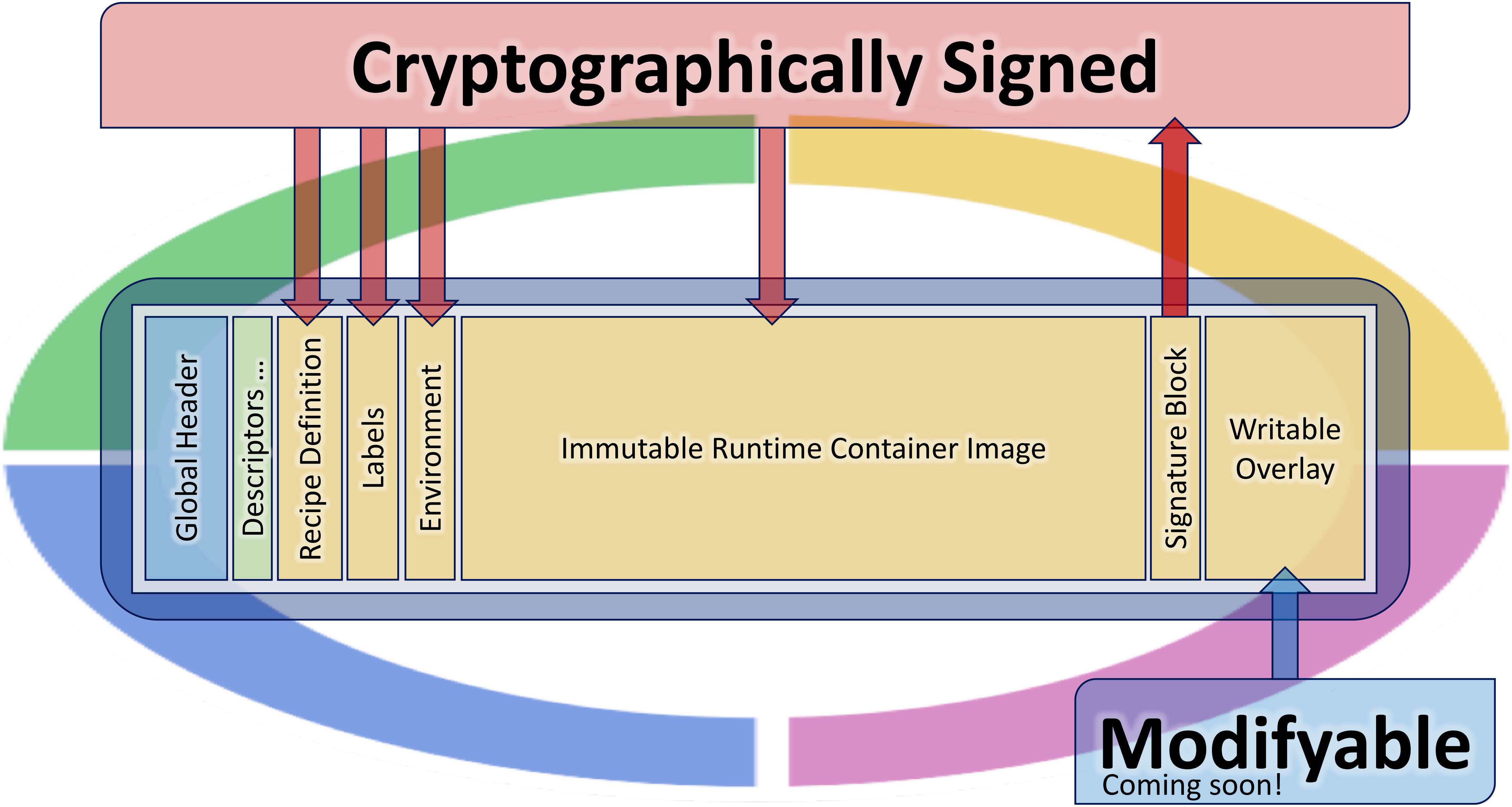
- Data Object Descriptors

- Raw Data

SIF EXAMPLE: TRUSTED, SIGNED CONTAINERS



SIF EXAMPLE: TRUSTED, EVOLVING CONTAINERS



SINGULARITY IMAGE FILE (SIF): FEATURE LIST

- **Cryptographic signatures and verification:** The Singularity image format supports both region checksums as well as signing for your verification pleasures

- **Faster access to container meta-data:** Container meta-data is now part of the image file but outside the container

- **Support for multiple system partitions**: A single container image can *contain* multiple container regions and/or a writable overlay

- **Support for Checkpoint Restart:** Internal support for checkpoint-restarting for mobility of state

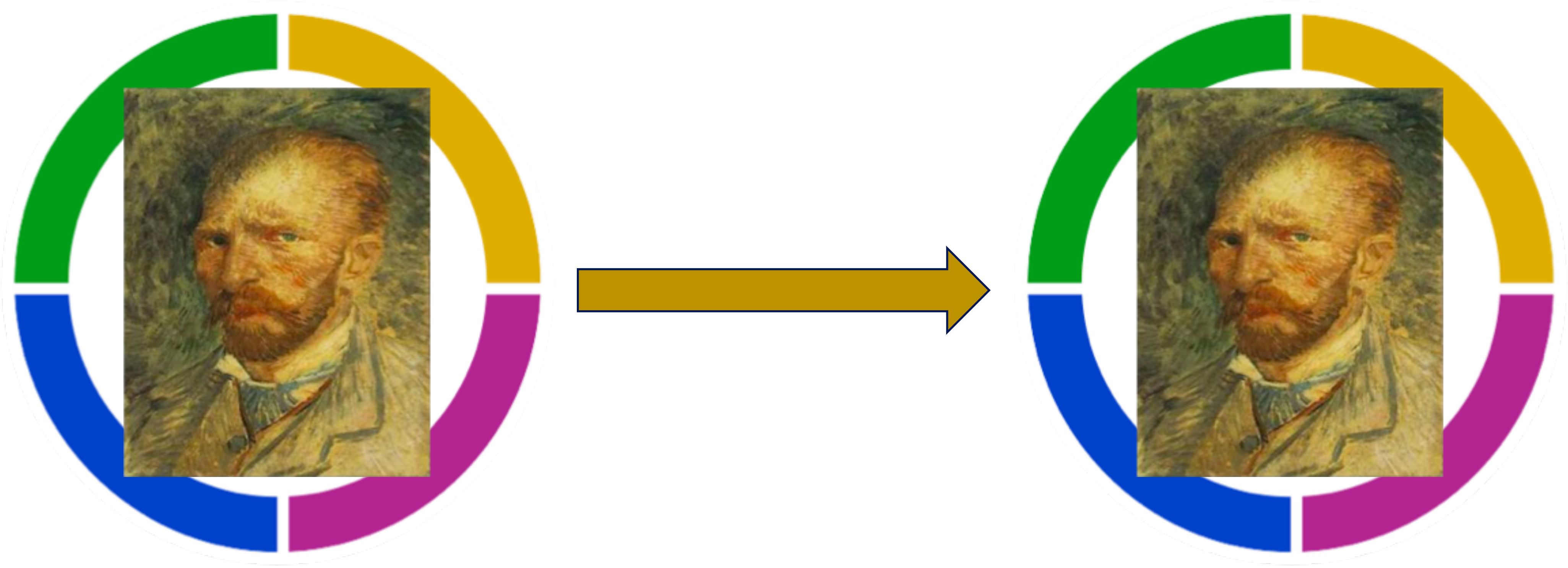
Reproducibility and Portability

IS THIS "GOOD ENOUGH"?

=?

WHAT IS YOUR REPRODUCIBILITY PLAN?

BIT FOR BIT SOFTWARE REPRODUCIBILITY



SHA:
5f09a35a642a68c467bf230f5e5ea3218e4177a0

SHA:
5f09a35a642a68c467bf230f5e5ea3218e4177a0

University of Michigan

“Singularity is a fabulous tool for providing forward and backward software compatibility on clusters and for reproducibility”

ENTERPRISE

Microsoft Azure

amazon web services™

Google Cloud Platform

SINGULARITY ENABLES EXTREME MOBILITY

SCP, SFTP, GridFTP/Globus, Rsync, NFS, Lustre, Object Stores, etc..

Single file containers means that containers are easy to manage

Open Science Grid

XSEDE
Extreme Science and Engineering Discovery Environment

PENGUIN COMPUTING

COMPUTE

HPC/EPC Compatibility

HPC SPECIFIC TECHNOLOGIES

- **Low latency, very high throughput networking**: usually InfiniBand for x86

- **Parallel File systems:** highly parallel read/writes across many nodes (thousands of processes accessing the same file), Lustre or GPFS

- **Message Passing Interface/MPI:** low level, highly optimized internode communication library, Open MPI, Intel MPI, MPICH, MVAPICH2, etc…

- **Resource Managers/Schedulers:** the scheduler is responsible for job priorities and preemption and the RM executes jobs on nodes

- **Graphical Processing Units/GPUs:** massively parallel computations at the chip level

EPC (ENTERPRISE PERFORMANCE COMPUTING)

- Enterprise users interested in AI, Deep Learning, compute drive analytics, and IOT increasingly demand HPC-like resources.
- Singularity’s features also make it the perfect container solution for this new type of “Enterprise Performance Computing” (EPC)

HPC/EPC COMPATIBLE SECURITY MODEL

- **Base security assumption:** Untrusted users running untrusted containers
- **Limit user’s potential security contexts:** We can not allow users to escalate to root, even in containers that they control (and know the root password to)
- **Allow user’s access to data they own:** And limit access to data that they don’t own

Washington University in St.Louis

“Users have been asking for containers for years, but I've always resisted. Singularity addressed the majority of my complaints and couldn't have been easier to install”

HPC/EPC COMPATIBLE CONTAINER ARCHITECTURE

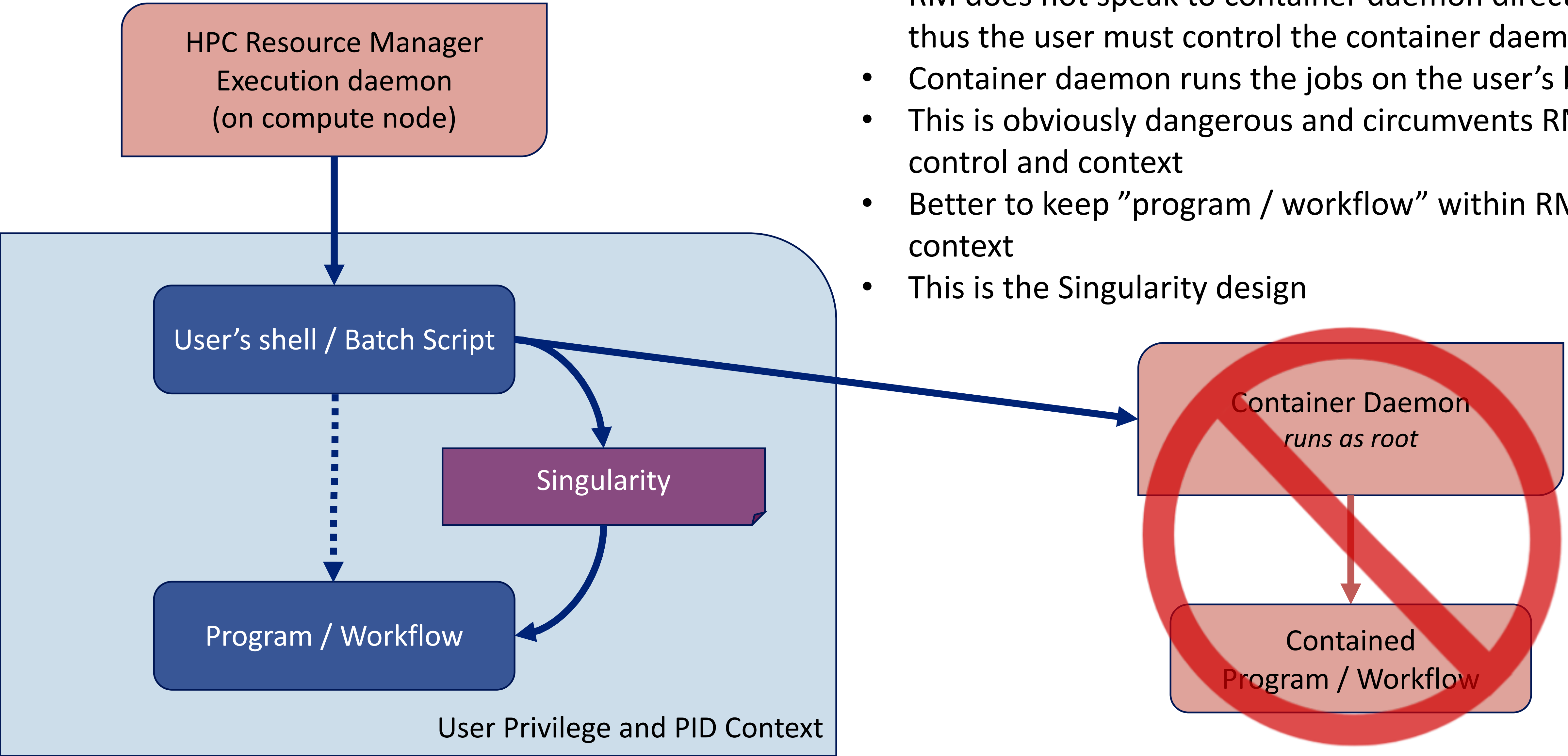
- **MPI support:** MPI jobs are also easily supported using a hybrid model
- **GPU:** Users are themselves within a container, and thanks to the Singularity security model, they can not escalate. This means we can share the GPU device into the container.
- **Resource managers:** Container processes are decedents of the RM (rather then a root owned container daemon which the RM has no direct control of)
- **Single file container images:** This image format is highly optimized for parallel FS

ARL

Army Research Laboratory

“Singularity is the best option among the big three considerations for HPC”

DEEPER LOOK INTO RESOURCE MANAGEMENT

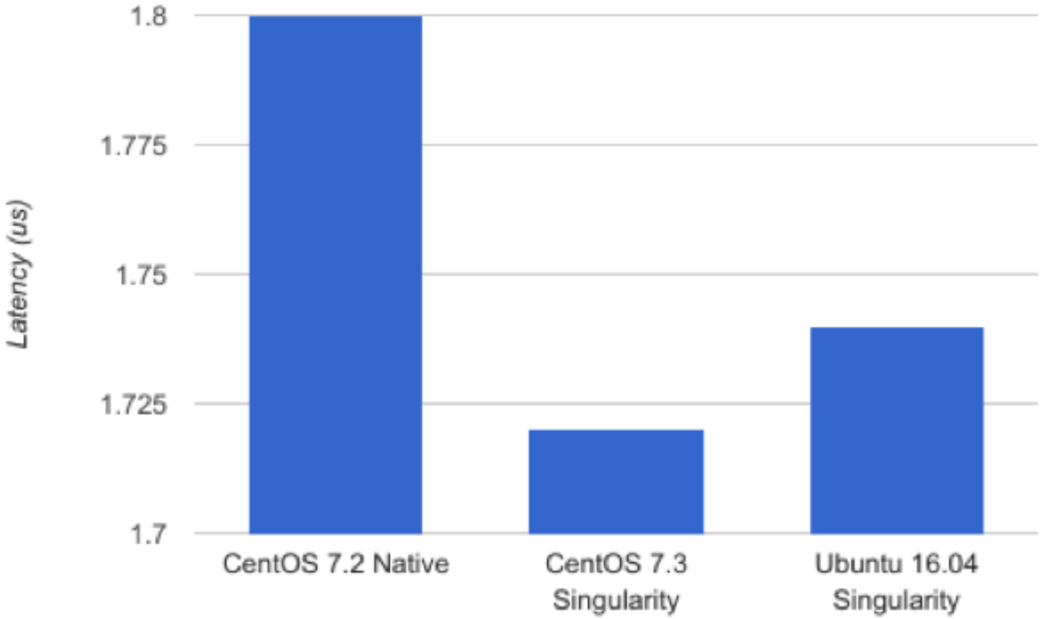


- RM does not speak to container daemon directly thus the user must control the container daemon
- Container daemon runs the jobs on the user's behalf
- This is obviously dangerous and circumvents RM control and context
- Better to keep ”program / workflow” within RM context
- This is the Singularity design

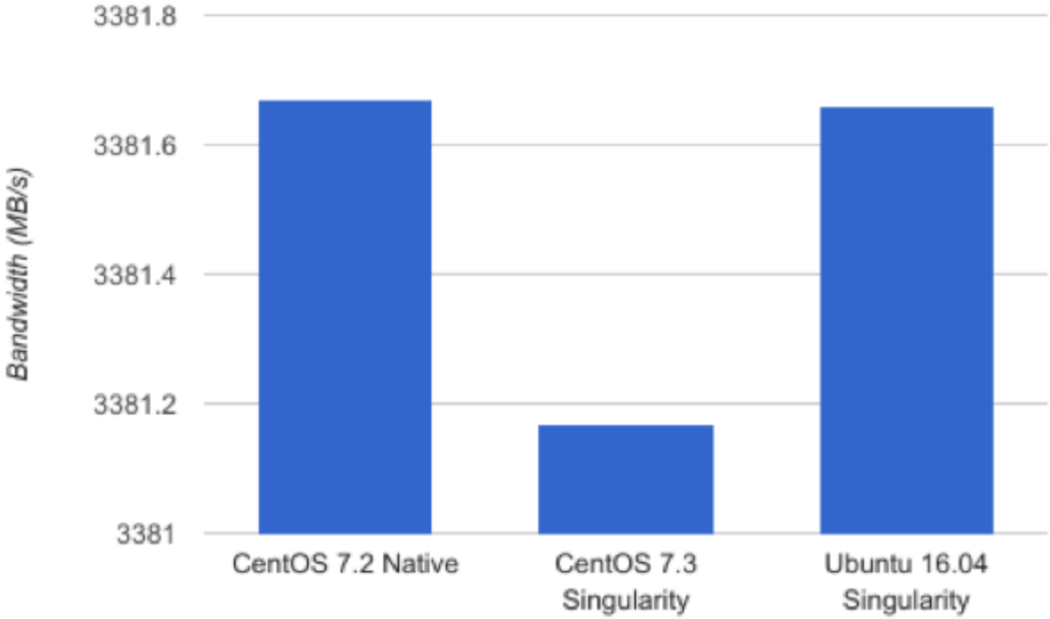
Performance

CONTAINERIZED MPI LATENCY COMPARISON

MPI OSU Latency Test msgsize = 1 byte



MPI OSU Bandwidth Test msgsize = 4MB



Open MPI 2.0.1 with OSU Micro Benchmarks 5.3.2

RUTGERS

“Used (Singularity) to get around a GLIBC version requirement for binary distribution of the NCI GDC download tool on CentOS 6”

THE UNIVERSITY OF
CHICAGO

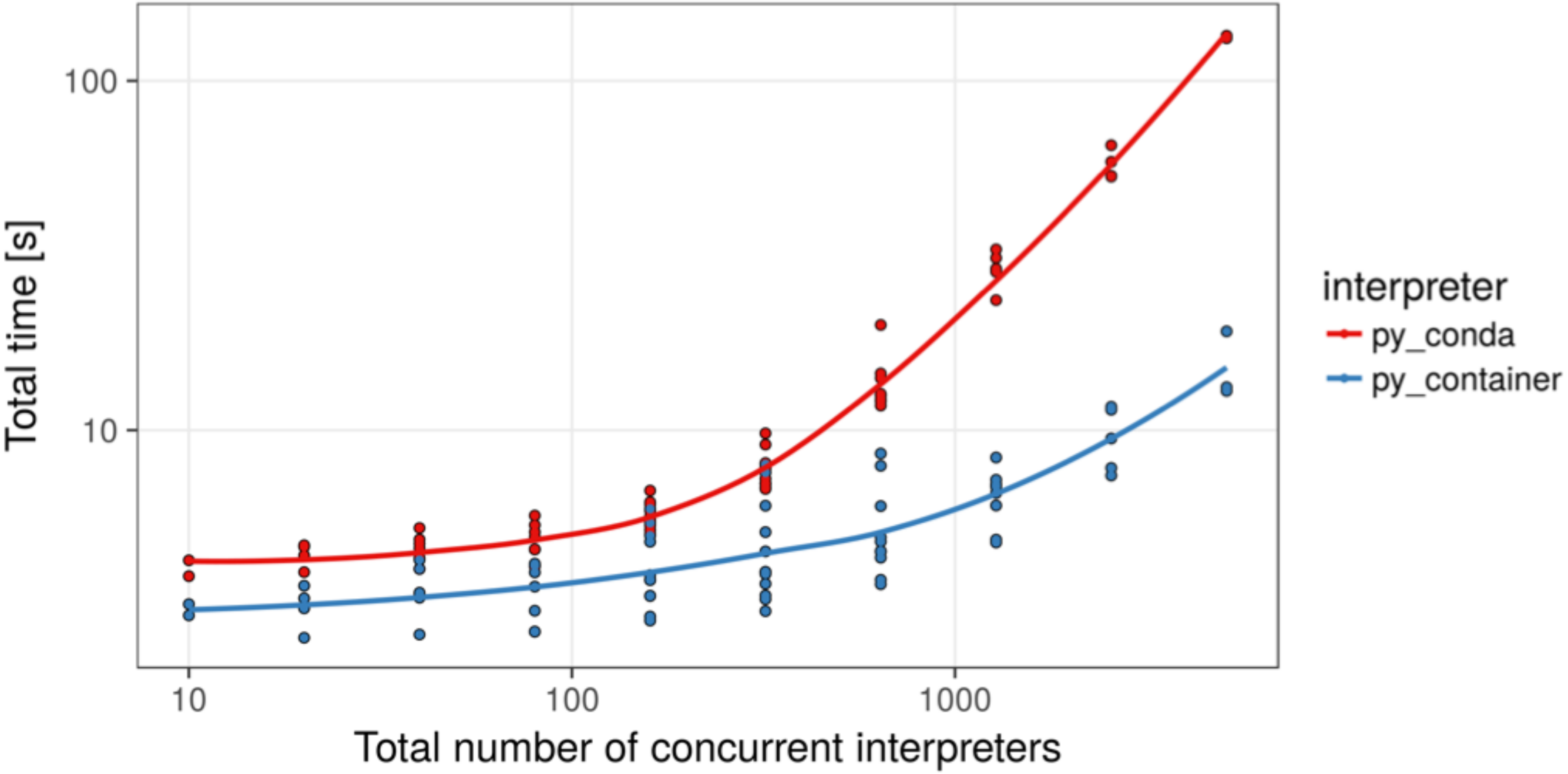
“Singularity allowed us to use software that was otherwise impossible to install under SL6, such as TensorFlow”

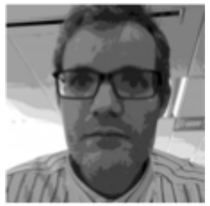
SINGULARITY INVOCATION PERFORMANCE OVER SHARED STORAGE

Objectives:

1. Measure scaling of python startup and import speed with increasing numbers of concurrent python interpreters
2. Compare scaling of a standard python installation with an identical containerized installation (Singularity).

*Note: Underlying file system is NFS, max jobs was 5120 over 320 nodes, graph is **logarithmic** on both axis.*





DR. WOLFGANG RESCH

HTTPS://GITHUB.COM/WRESCH/PYTHON_IMPORT_PROBLEM

Use Cases

nextflow

- Nextflow is a workflow management language for data-driven computational pipelines
- Nextflow uses Singularity to deploy large-scale distributed scientific workflows
- Commonly used in genomics pipelines
- Supports both HPC cluster and cloud based resources in a portable manner
- Used by:
  - Center for Genomic Regulation (CRG)
  - Pasteur Institute (France)
  - SciLifeLab (Sweden)
  - Sanger Institute (UK)

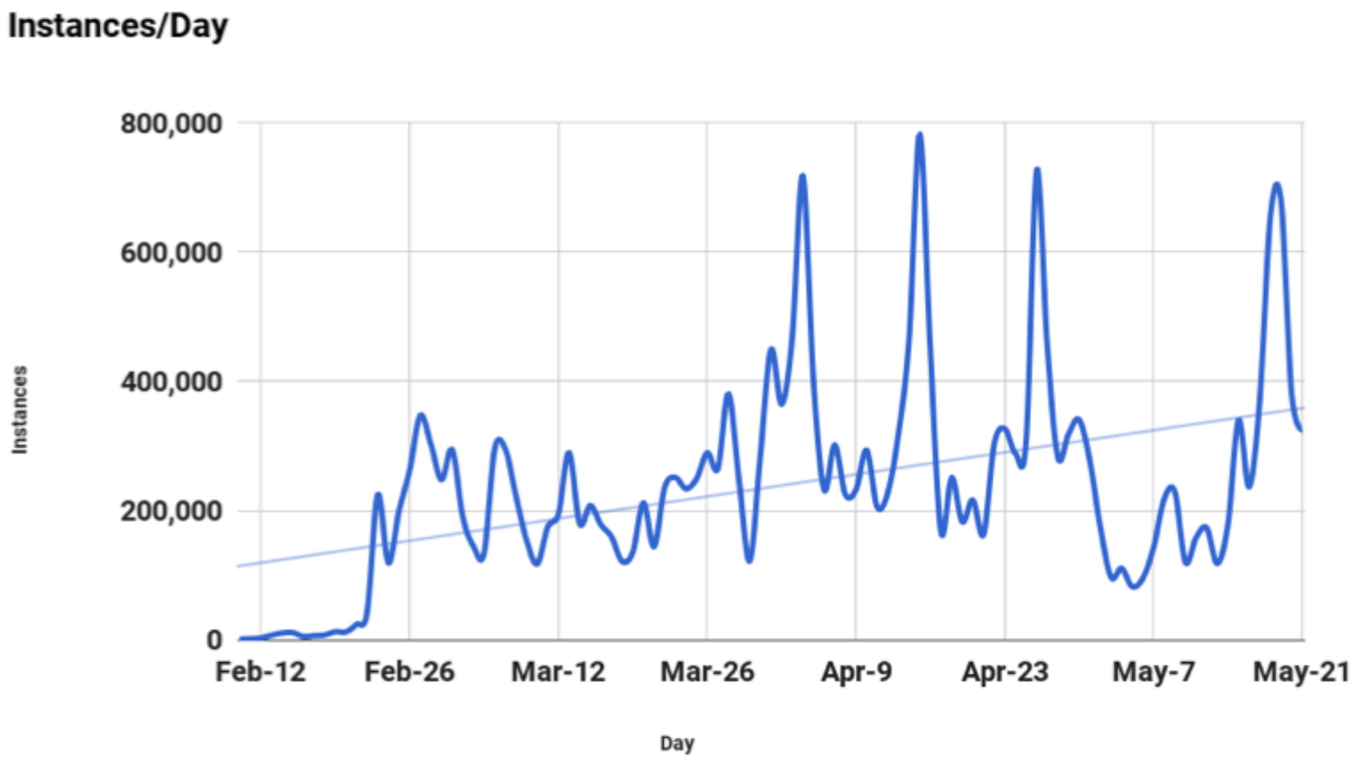
- Among standard HPC use cases…
- Researchers are using iPython Notebooks via JupyterHub
- iPython JupyterHub kernels were deployed in Singularity containers
- Once the container is deployed via JupyterHub, the job runs within the container while maintaining access to local node resources
- This is a multi-user environment so Docker is a non-starter

- ALICE jobs are packaged into Singularity containers
- Jobs are executed via Singularity through a modified SLURM script
- At any given moment in time, there are about 2000 Singularity containers active on the system

GSI Green Cube
6 stories tall
30,000 sqft
12 MegaWatts
PUE = 1.07 (world record)

Open Science Grid

- The OSG uses Singularity to provide a consistent runtime environment across heterogeneous resources worldwide
- Container images are distributed via CVMFS to all sites
- About half a million jobs are run through Singularity per day



OAK RIDGE
National Laboratory

Titan @ Oak Ridge Leadership Computing Facility

Adam Simpson (front) and Matt Belhorn (back), high-performance computing user support specialists at the OLCF, use the Singularity application to develop containers that will allow newer systems to run deep learning packages.

https://www.olcf.ornl.gov/2017/05/09/containers-provide-access-to-deep-learning-frameworks/

NIH National Institutes of Health
Turning Discovery Into Health

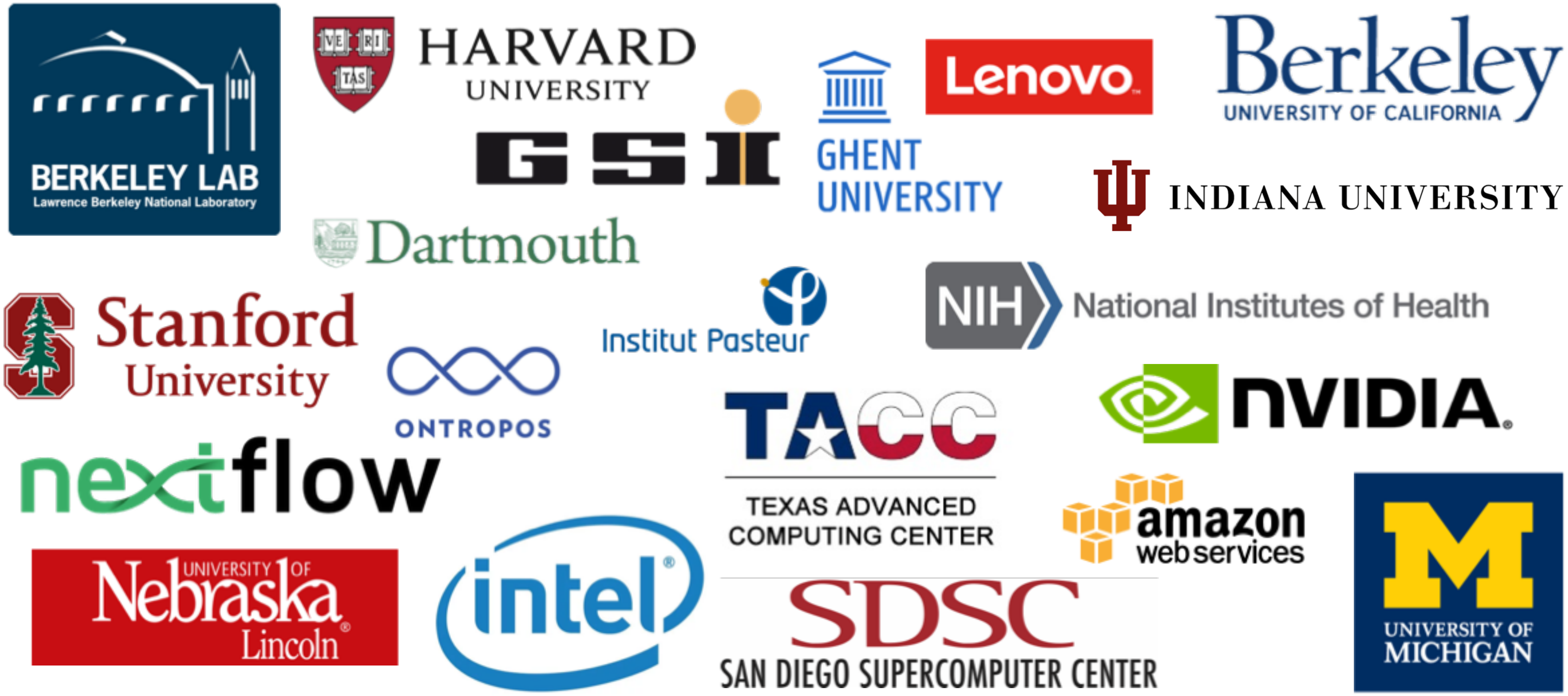
- The NIH uses Singularity to provide programs like TensorFlow and OpenCV3 which are difficult or impossible to run with their current operating system
- With Singularity they can create ”portable reproducible data analysis pipelines”
- Some applications have been installed into Singularity containers and used as standalone programs via environment modules for the users (the users don’t even know they are running within a container!)

Coming Soon

SOME ITEMS ON OUR ROADMAP

- Virtual container boot within instances
- Evolving signed containers
- Optional non-SetUID execution modes (Linux capabilities)
- OCI (Open Container Initiative) compliance
- Native Kubernetes support
- Performance profiling of contained applications
- Checkpoint / Restart of containers
- OS X and Windows support

SINGULARITY: CONTRIBUTORS


BERKELEY LAB
Lawrence Berkeley National Laboratory
VE RI TAS
HARVARD UNIVERSITY
GSI
GHENT UNIVERSITY
Lenovo
Berkeley UNIVERSITY OF CALIFORNIA
INDIANA UNIVERSITY
Dartmouth
Stanford University
ONTROPOS
Institut Pasteur
NIH National Institutes of Health
NVIDIA
nextflow
TACC TEXAS ADVANCED COMPUTING CENTER
amazon web services
UNIVERSITY OF Nebraska Lincoln
intel
SDSC SAN DIEGO SUPERCOMPUTER CENTER
M UNIVERSITY OF MICHIGAN

THANKS!

- http://www.sylabs.io/ The brand new home of Sylabs Inc
- http://singularity.lbl.gov/ Help docs for getting started
- https://singularity-hub.org/ CI build service for Singularity containers!
- https://singularity-tutorial.github.io/ Hands on Singularity tutorial for beginners.