

1. High Performance Computing

currently manages jobs across clusters with thousands of nodes, all of which PXE boot from a precrafted image. The LANL TOSS image is built with a lot of the software needed to conduct research. If a scientist needs certain libraries/software for their project, it must be approved and added to the image by HPC-SYS. Our research project explores an alternate, potentially more adaptable, workflow for running jobs using containers and object stores. If successful, our project presents a proof of concept that containers offer a means for scientists to have finer control over the environment in which their jobs run in.

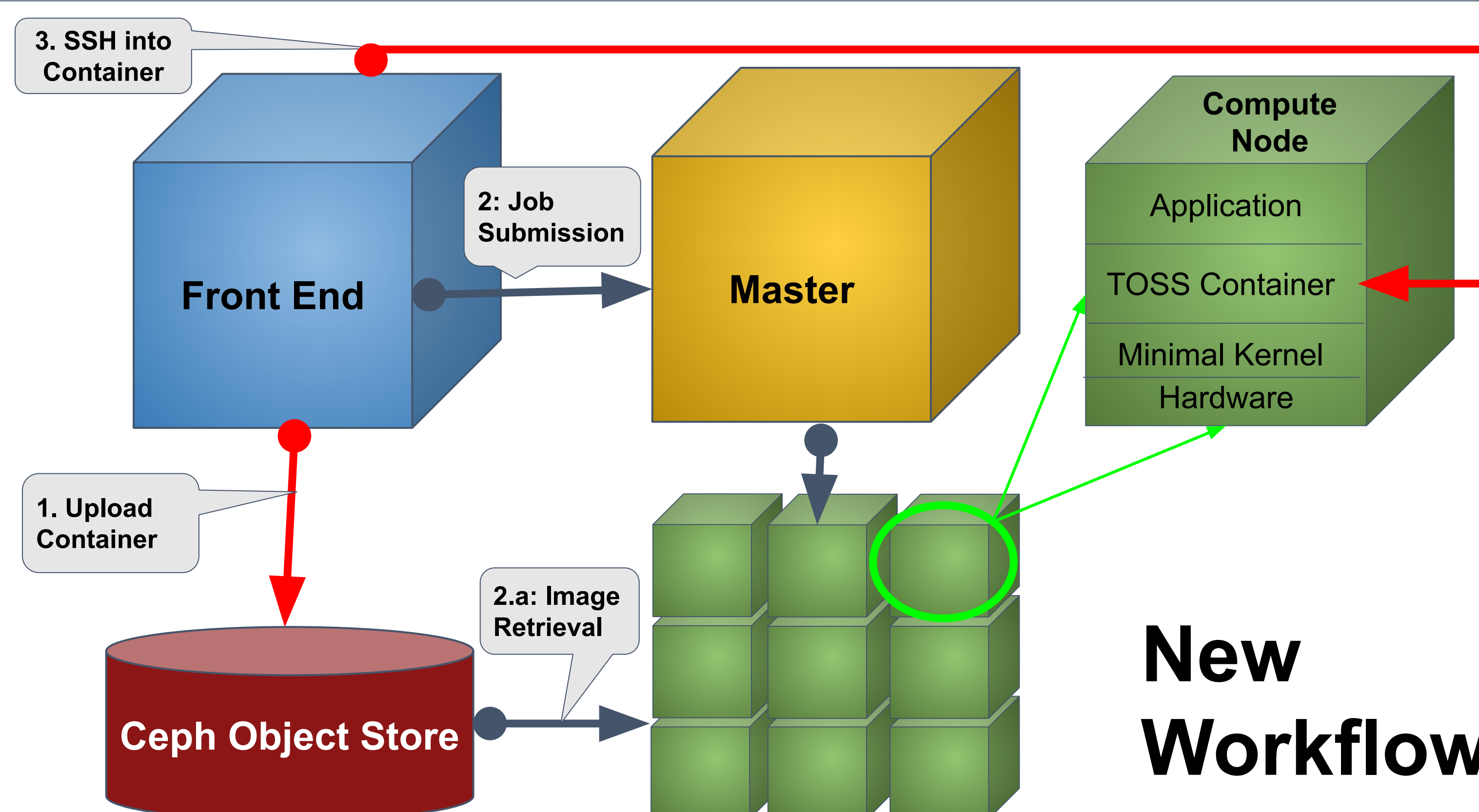
2. The Current Workflow

for running jobs on HPC systems at LANL follows a historic, yet rigid, process. The user submits a job from a front-end node using Slurm, and the job is distributed to the compute nodes. When the job finishes, Slurm combines output from all the nodes into an output file in the user's home directory, and everyone goes home mildly happy.

3. The New Workflow

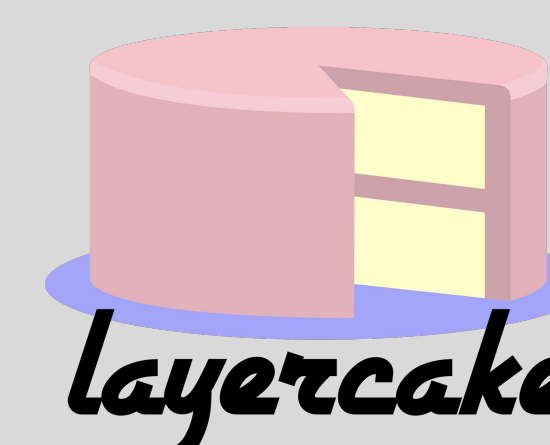
for running jobs is the same as before for those not using containers. If users want to bring their own container, they simply need to:

1. Upload the container image to Ceph Object Store
2. During job submission, specify flags in the srun/salloc command:
 - a. `--ch-image`, the name of their container
 - b. `--ch-mount`, a comma-separated list of devices and file-systems to mount to the container.
3. When SSHing, users are dropped into the container namespace instead of the bare node



TEXAS A&M
UNIVERSITY

Donald Elrod
Texas A&M University
donald.elrod.96@gmail.com

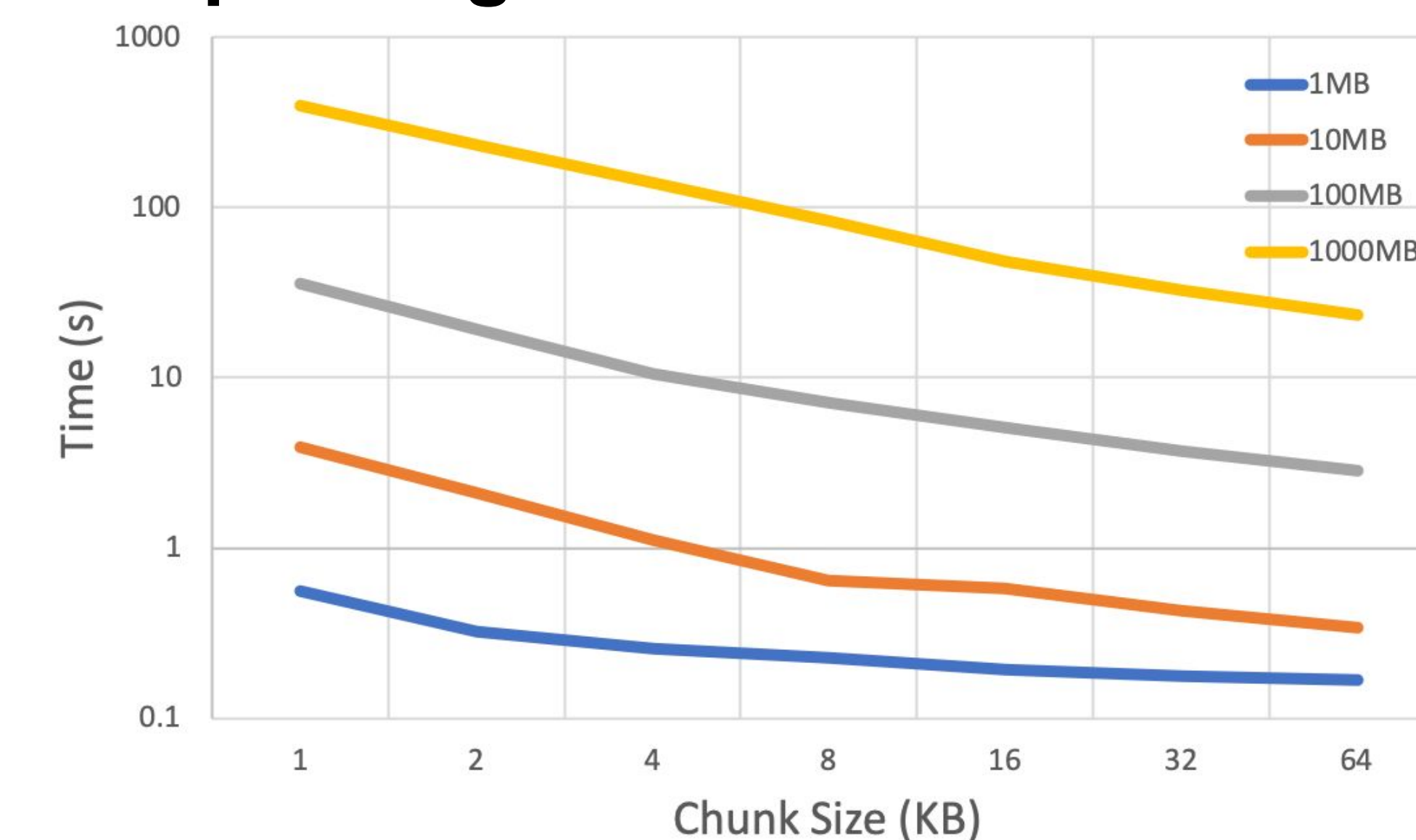


HPC can EVOLVE:
CONTAINERS allow scientists to
bring their own
SOFTWARE-STACK,
be more **PRODUCTIVE,** and
ENHANCE SECURITY
SEAMLESS | TRANSPARENT

4. Benefits of the New Workflow:

- For users:
 - Flexible software stack
 - Ability to take advantage of the latest frameworks and programming languages
 - Allows more productivity
 - Keeps HPC current with today's software and framework ecosystem
- For admins:
 - Jobs are unprivileged
 - Users have (essentially) root privilege inside their containers
 - Users are dropped into their job's namespace while SSHing, providing a thin extra security layer

5. Ceph Image Download Performance



6. Future Work:

1. Experiment with RADOS block device support
2. Test load-balancing across OSD's for container distribution
3. Test jobs using new workflow
4. Project hardening
5. Auto encapsulation
6. Incorporate Ceph image retrieval fully into the Charliecloud-Slurm plugin instead of python2

Acknowledgements:
Reid Priedhorsky
Catherine Hinton