

# Containers and Reproducibility in Computational Science

**Reproducibility** is the degree to which scientists are able to achieve the same results using the same methodology.

Computational science experiments in general suffer from **poor** reproducibility. Complex dependencies make it very difficult for one researcher to recreate the software environment used in another researcher's experiment.

Enter the **software container**: an encapsulation of a software environment and all its dependencies into a single sharable unit. Containers allow researchers to share their experiments' software environments with effectively perfect fidelity.

Computational scientists have previously **avoided use of containers in HPC systems** due to the fact that users are able to gain **root access** to the host server from inside the container, which introduces intolerable security risks.

***Our aim was to develop a system that allows for the secure building and rebuilding of software containers in HPC systems. We also aimed to automate the entire container build process from version control to deployment.***

## Key Software

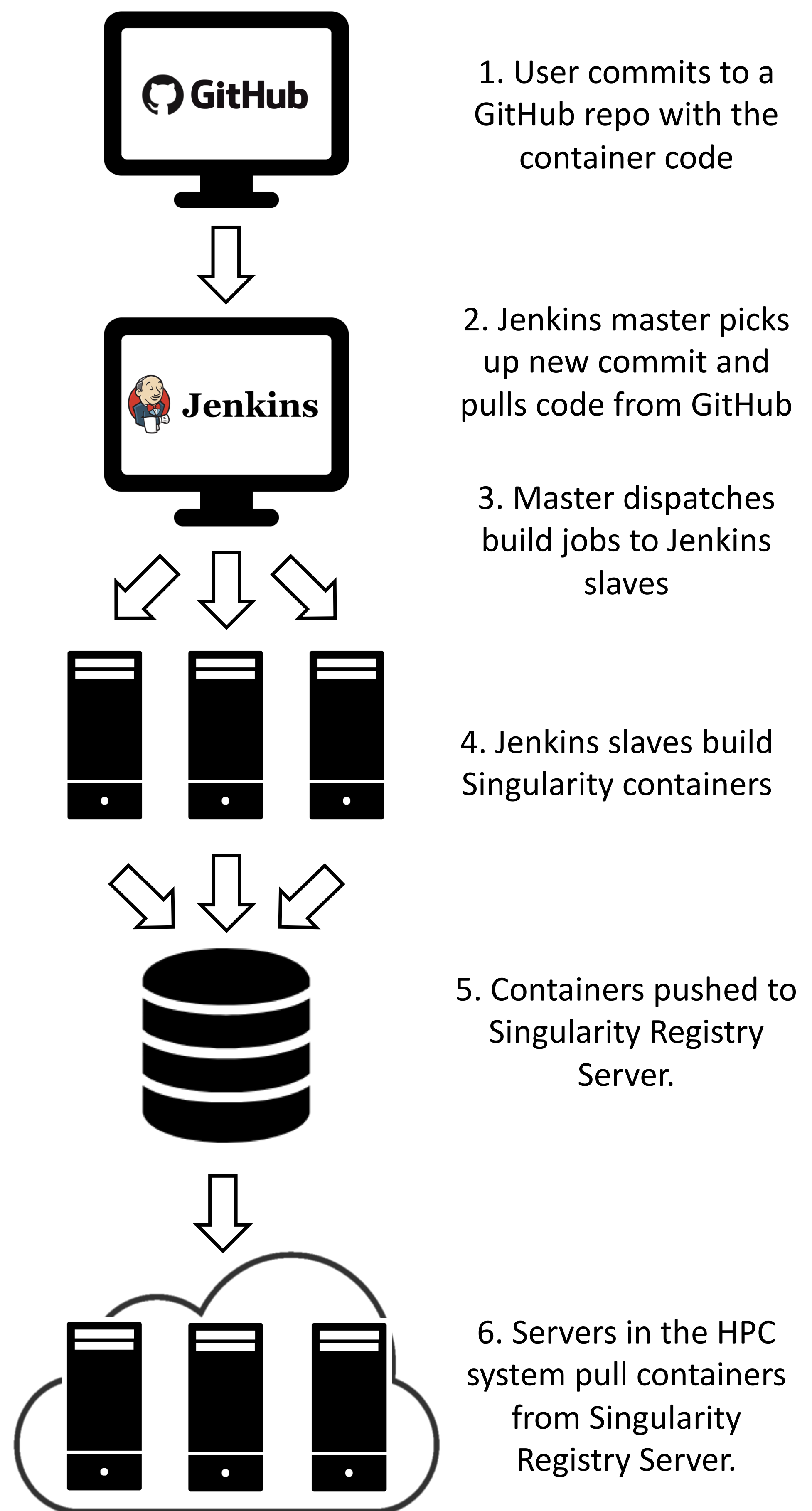


**Singularity**  
containers for the  
secure use of  
containers in HPC  
systems



**Jenkins**, an automation  
server used for the  
continuous integration,  
delivery and  
deployment of  
containers

## System Implementation



## Project Outcomes

- ✓ Our system was able to automatically build and deploy containers safely in an HPC system
- ✓ We encapsulated our entire system in Docker containers, making it highly portable and shareable.
- ✓ With some additional work, our system has the potential to help improve the reproducibility of computational science experiments.



## Authors

Robert Perrott PRR0B006@myuct.ac.za  
Rowan Hodson HDSROW001@myuct.ac.za

## Supervisor

Robert Simmonds  
simmonds@cs.uct.ac.za

University of Cape Town, Department of Computer Science

