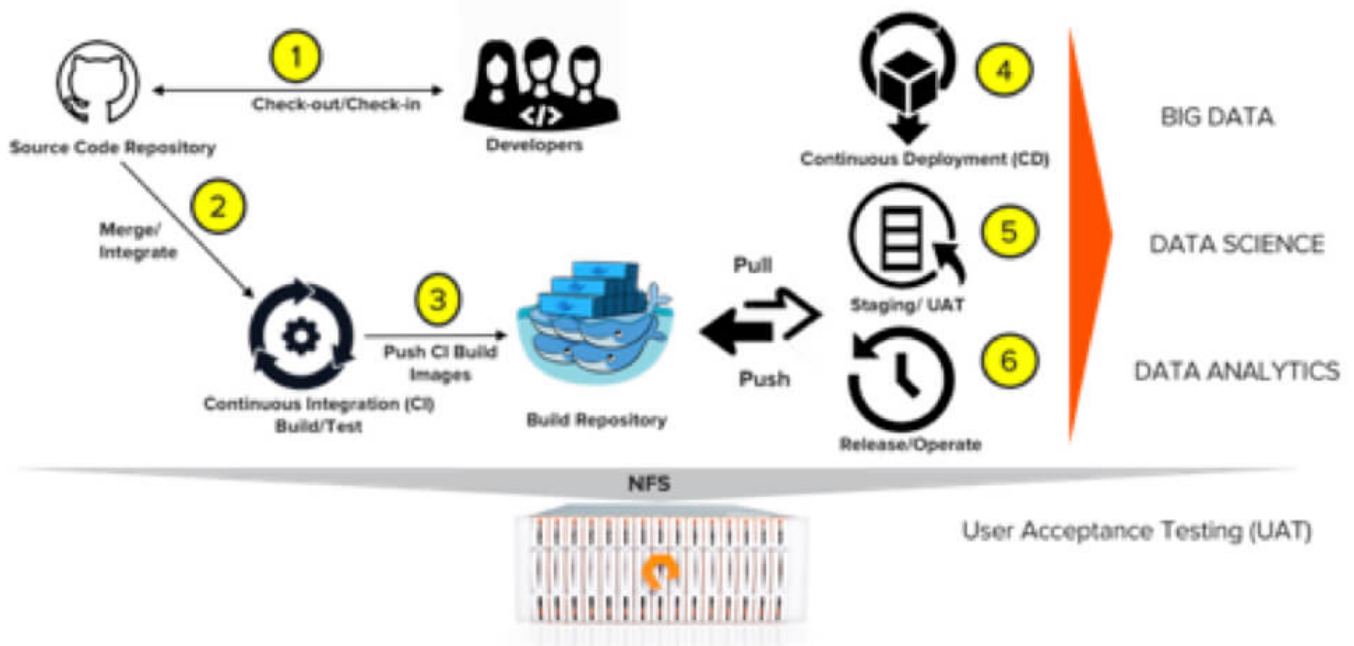


Why Your Docker Images Should Jump Ship to Flash?

Organizations looking to modernize and automate application development and delivery with containers and other Dev/Ops tools find that streamlining workflows and centralizing data is key. Provisioning and consolidating heterogeneous workloads seamlessly with a data hub such as Pure FlashBlade can provide linear performance scalability, data availability, and protection during application development, testing, and delivery. In this blog, we'll explain how private Docker Registry can co-exist with other software development workloads on FlashBlade, including data pipelines for Data Analytics, AI/ML.



What's the Big Deal About CI/CD?

The Continuous Integration (CI) and Continuous Delivery (CD) process is important both for developing web and mobile apps and for connecting to Big Data and AI/ML data pipelines. DevOps methodologies not only enable agile software development - they also provide agile techniques in AI/ML and Big Data analytics environments to build, test, and deploy applications as part of the software delivery process.

One of the most common challenges for end-users, like data scientists, is to get a consistent platform with the right operating system, libraries, and tools to test during research and training. It's no small task, for instance, for researchers and data engineers to create different prototypes and test purpose-built data pipelines for Kafka deployments in production.

Docker Images, Docker Registry, and DockerHub

A Docker Registry is the most common form of repository for storing and hosting Docker images, particularly the public Docker Registry, DockerHub. However, storing proprietary code that needs to remain private in a public setting presents certain challenges, as most of the publicly-hosted repositories do not

have the ability to integrate with common development and test tools. DockerHub allows one free private repository. Additional cost applies if multiple repositories are required on DockerHub.

Containers like Docker allow data science teams to provision the required platform quickly, and then build new Docker images. The generated Docker images are saved on persistent storage – such as FlashBlade – so the images are protected and available at all times. These images are shared between teams and lines of business for different use cases. The underlying Docker containers are required to scale performance and capacity as needed, with Docker image sizes ranging from a few MBs to north of 1GB as more and more layers are added to the base image.

Private Docker Registries Provide the Best of Both Worlds

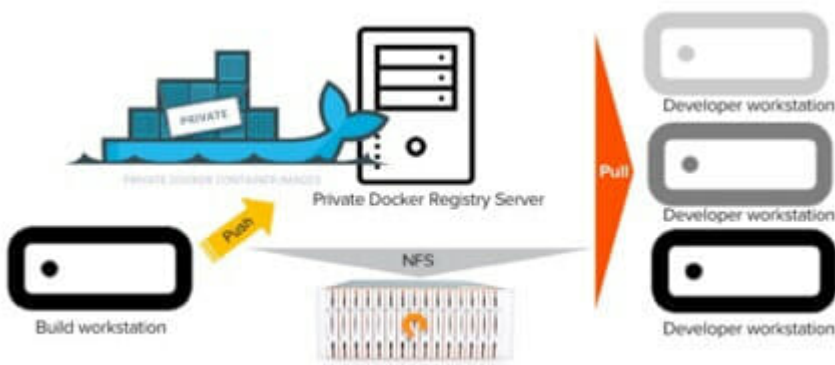
Private Docker Registry version 2, configured on FlashBlade over NFS, delivers cost-optimized performance and data reduction alongside easy capacity scaling for hosting and distributing Docker images. Moreover, changes and updates from different teams to Docker images remain consistent with this configuration.

Why Flash?

At first glance, flash may not be the least expensive storage option. But there are a few additional considerations that more than justify the cost of using flash. For instance, a unified flash-based data hub like FlashBlade reduces build and test cycles by up to 33% during the CI process, in which Docker images are pushed to a private Docker Registry. The Docker images are then pulled from the private Docker Registry in the CD process for A/B testing or Blue-Green deployments. In mid to large enterprises, push/pull operations on Docker images to and from a private Docker Registry can run into the 100s or 1000s per day. The all-flash power of FlashBlade handles this volume easily – with no performance impact, inline data reduction included.

Furthermore, as build and test cycles are reduced during the development and deployment process, applications can be consumed faster in other connected data pipelines, like Data Analytics and AI/ML. Using FlashBlade as a data hub to linearly scale performance and capacity for heterogeneous workloads not only enables data scientists and analysts to accelerate their workflows and pipelines – it drives the elimination of data silos.

As shown in the image above, there may be single or multiple build servers in a development environment where the build pipelines can be configured to push the new/updated Docker images to the private Docker Registry. End-users can manually or automatically (using Ansible playbooks) store and access Docker images from different development, test, and production ([Kubernetes](#) and/or Docker Swarm) clusters and collaborate across teams.



Ready to Jump Ship to Flash?

To summarize, a Dregistry can be hosted in the public cloud or in private data centers and cloud environments. Privately hosted registries offer more control and tighter integration with different development and deployment tools of choice. To drive it all, a data hub like FlashBlade delivers persistent, high-performance, data reduction-enabled all-flash storage that's cost-efficient and easily scalable.

Get the details on configuring private registries on [FlashBlade](#).