Datrium Solution for Jenkins

Highlights

- Lowers storage requirements
- Reduces overall runtimes
- Preserves all artifacts
- Improves virtual resource utilization
- Simplifies management

The Challenge

Jenkins is a valuable tool for automating, building, testing, and deploying code, especially when coupled with virtualized infrastructure. However, Jenkins users face some challenges when trying to build, test, and archive large projects.

1. Jenkins architecture requires that artifact data go through a master node when using slaves, which becomes a significant bottleneck.

2. Data is frequently copied twice for each build and every test cycle, first to the master, and then to the NFS mount point.

3. Large file archiving is not natively supported.

Workflows that use static slaves with one or more executors may lead to leakage of state between concurrent or previously executed builds, causing incorrect results. These challenges are depicted in Figure 1 below.

The Solution

Datrium DVX, with zero-copy and near-instant clones, can be used with transient Jenkins worker VMs and Containers to deliver an optimal solution. The basic workflow first creates a worker VM by cloning a Jenkins slave template. Next, an upstream virtual disk (vdisk) or persistent volume (PV) is cloned and attached to the worker.
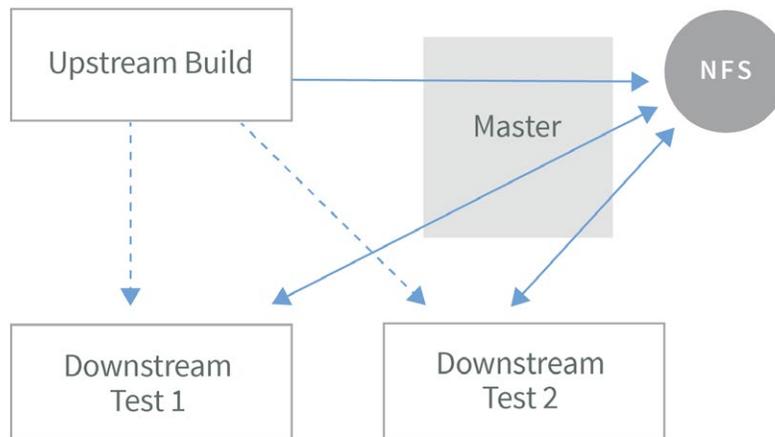

Figure 1 – Bottlenecks: NFS and Jenkins Master

Once the build completes, the vdisk or PV gets detached, and the worker is destroyed. Consider a typical process that builds the product in a top-level job, which then triggers many downstream jobs, each responsible for testing different areas of the build. The cloned vdisk or PV provides the upstream artifacts without transferring any data, and each detached vdisk or PV becomes a complete copy of the build and test artifacts. This new process is depicted in Figure 2.

Then, artifacts are made available via a hyperlink on the build page, exposing both an NFS mount point and HTTP URL for browsing, and because the vdisk or PV is named according to the job name and build number, lookup and retention management is straightforward.
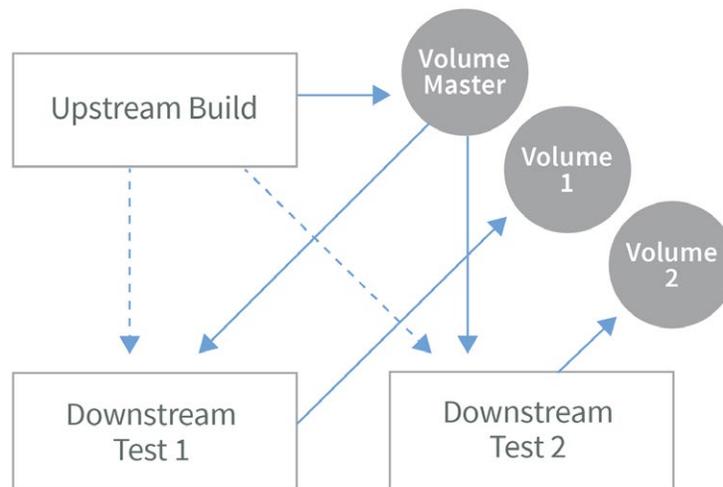
Figure 2 – Optimized Jenkins Workflow

Key Benefits

Efficient Build and Test

Advanced snapshot and cloning features support efficient data handling for build artifacts and job archives. The downstream cloned vdisk or PV is a snapshot of the base vdisk or PV, and it consumes no additional space until written to. All writes are both compressed and deduped, further reducing space consumption. The zero-copy clones and data reduction techniques typically result in space savings of 20x over the traditional local storage and NFS based workflows, and data transfer times are all but eliminated, reducing run times by 30% or more.

Clean Build Environment and Flexible Workers

Fast and efficient clones support ephemeral, template-based, workers – with many benefits. Because the worker is cloned from a base template and powered-on before each build, it provides a known initial state without any residual state from previous builds. Spinning up a new worker for each build also makes it easy to specify CPU and memory configurations, according to job requirements. Finally, expanding the size of a cloned vdisk or PV, if needed, is painless. The DVX workflow provides complete control of each worker and individual build environments.

Comprehensive Artifacts

DVX enables cloned vdisks or PVs to be efficiently stored and easily managed. Cloning vdisks or PVs for downstream builds substantially improves the overall build and test workflow. However, the resulting vdisk or PV also serves as a complete archive of the build itself. The vdisk or PV can be cloned again instantly for analysis or ad-hoc testing; plus, exposing artifacts as a clone of the original vdisk or PV ensures that the original artifacts are preserved.

Distributed Resource Scheduling

DVX cloning techniques allow worker VMs and Containers to be initially placed on any Host or later migrated within the cluster without penalty. VMware distributed resource scheduling (DRS) can be enabled and configured in numerous ways to optimize overall resource utilization.

Conclusion

Some of the most challenging aspects when using Jenkins are efficient artifact handling and slave management. DVX helps Jenkins users to address production challenges using advanced cloning and data reduction technology.