

ACTIFIO FOR AWS

Database Virtualization for Test Data Management in AWS

actifio
Radically Simple

ACTIFIO FOR AWS

Database Virtualization for Test Data Management in AWS

What Is Actifio?

Actifio virtualizes the data that's the lifeblood of business. We do it in more than 30 countries around the world. Our Virtual Data Pipeline™ technology enables businesses to protect, access, and move their data faster, more efficiently, and more simply by decoupling data from physical storage, much the same way a hypervisor decouples compute from physical servers. For enterprise-class backup modernization, self-serve instant data access, or service provider business transformation, Actifio is the first and only choice for radically simple copy data virtualization.

What Can It Do for you?

That depends who you are, and what problem you're trying to solve.

- Are you a **DBA** tired of mundane tasks like creating and masking database clones, compressing them, replicating them to AWS, then restoring databases in the cloud for your Dev, QA, and UAT testers?
- Are you an **enterprise architect** looking to automate test data management as part of a continuous integration process, so database copies are masked and spun up on test servers in AWS automatically?
- Are you a **QA manager** looking to improve quality and accelerate test cycles?
- Are you a **developer** or **UAT tester** looking for a self-serve way to spin up a masked copy of production data sets in AWS with automated data refreshes to accelerate testing cycles?
- Are you an **IT Ops Manager** tired of execs complaining about exploding storage costs, concerned about too little control over who's accessing copies of production data sets, and no data masking for test data on AWS?
- Are you an **application manager** who has adopted Agile or DevOps methodology and looking for ways to increase fidelity of test data?

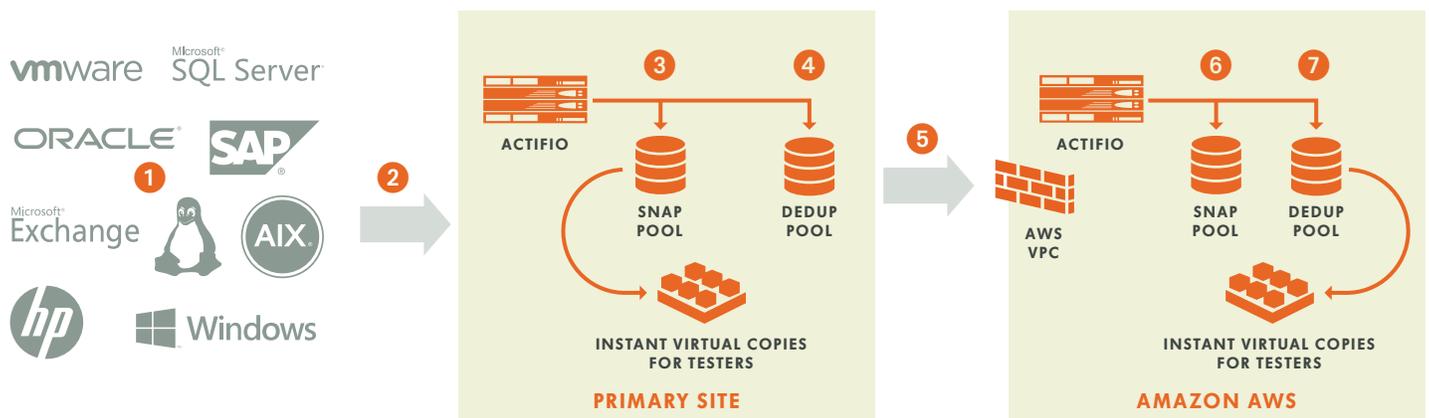
If the answer is yes to any of the above, this white paper will explain you how Actifio can help you achieve all of the above by leveraging Amazon Web Services.

Actifio enables DBAs and testers to provision virtual copies of multi TB DB in less than 5 minutes in AWS. To find out how, watch [this video](#) for Oracle, and [this video](#) for MS SQL

Using Actifio in AWS for test data management delivers:

1. **Speed**—Dev, QA, and UAT testers can provision multiple virtual copies of production data sets in a self-service manner and execute their test cases at an unprecedented speed in AWS. This allows the development team to develop and release more functionality with the same number of resources without increasing project timelines.
2. **Quality**—The quality of software improves exponentially because Dev, QA, UAT get to test on copies of production data sets instead of small, subset dummy data sets. This allows them to do scalability and performance tests early in development cycles.
3. **Control**—Enterprise IT is able to control Who can access What copies on Which servers in AWS for testing with role-based access control, audit trail, and automated data masking.
4. **Low TCO**—Actifio lowers the storage costs in AWS by up to 10x for large-scale test environments because of its copy data virtualization technology.
5. **Continuous Integration**—Automate end-to-end testing on virtual copies of test data, provisioned using Actifio REST APIs

FIGURE 1 Architectural Overview



Actifio Overview—How It Works

Actifio Virtual Data Pipeline (VDP) is a patented technology from Actifio that comes in two flavors: 1) Actifio Sky software appliance and 2) Actifio CDS physical appliance. Users can use Actifio Sky or CDS appliances on their premises; and Actifio Sky in AWS.

Actifio's high-level architectural data flow is shown in **Figure 1**:

1. Actifio has a wide platform support for many physical and hypervisor platforms, databases such as Oracle and MS SQL, and ERP applications such as SAP, Oracle EBS, and Microsoft Dynamics.
2. Actifio Virtual Data Pipeline (VDP) does incremental forever ingestion at the block level. After the first ingest, Actifio ingests only changed blocks forever.
3. These changed blocks are stored in their native format in the "Snap Pool" as a "Golden Copy." Actifio VDP virtualizes the data in the Snap Pool so that users can provision multiple virtual DB copies instantly for testing purposes. Optionally, Actifio VDP automates data masking by leveraging user-supplied data masking scripts.
4. The changed blocks of data are then stored in "Dedup Pool." Here data is deduped across all applications to reduce the amount of data that needs to be replicated over WAN to AWS; and also reduce storage consumption so that this can be also used to retain data for multiple weeks or months for data restore purposes.
5. The deduped data is replicated efficiently with encryption over the WAN to Actifio Sky in AWS. Note that since the data is deduped, the same block of data on a different database or any other machine is never sent over WAN to Sky in AWS. The global data deduplication minimizes the bandwidth usage.
6. Dedup Data is stored in Dedup Pool in AWS, which consumes less storage. This is useful for testers to go back in time and spin up data sets for analytics or compliance testing purposes. Storage used for Dedup Pool is EBS (Elastic Block Storage) in AWS.
7. Data is rehydrated automatically from Dedup Pool to Snap Pool in AWS. Actifio VDP virtualizes the data in the Snap Pool (aka Golden Copy) so that users can provision multiple virtual DB copies instantly in on-demand EC2 instances for testing purposes. Storage used for Snap Pool is recommended to be EBS General purpose (SSD) volumes.

Dev, QA, and UAT testers can use the golden copy in Actifio Snap Pool to provision multiple virtual copies of production data sets, both in AWS and at their primary site. Enterprise users benefit more by leveraging Actifio virtual copies with on-demand EC2 instances in AWS because they don't have to invest, manage, and maintain infrastructure for dev-test at their primary site.

And the above architecture can also be leveraged by enterprises for backup, long-term data retention, and DR in AWS as well. Details for these use cases can be found in this white paper: [Actifio for AWS](#).

Features and Benefits

The next few sections describe the various features available for Dev and Ops.

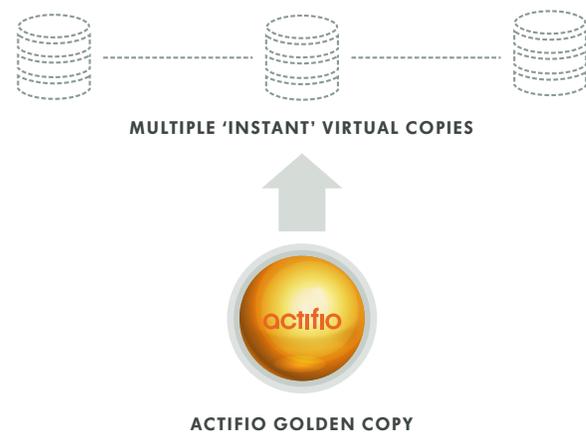
Instant Virtual Copies

Actifio Sky in AWS has the ability to provision multiple virtual copies of a data set instantly to multiple test servers simultaneously. These virtual copies are writable and scalable for unit testing, functionality testing, performance or load testing, and UAT testing. These virtual copies are storage efficient. For example if you have a 5 TB database and you create 10 virtual copies, it does not take any extra storage space. Yet each of the 10 copies seems like a full 10 TB copy. Only when testers change data while testing will incremental storage be consumed, storing only the changed data.

BENEFITS

1. Accelerate test data management
2. Reduce storage consumption by up to 10x

FIGURE 2 Instant Virtual Copies



Role-Based Access Control

Actifio administrators can specify WHICH users can access WHAT data sets on WHAT test servers in AWS. Reporting and audit logs list all user activities, including mount and un-mount operations and what data sets were accessed at what time.

BENEFITS

1. Better security with role based access control
2. Auditability of who did what operations

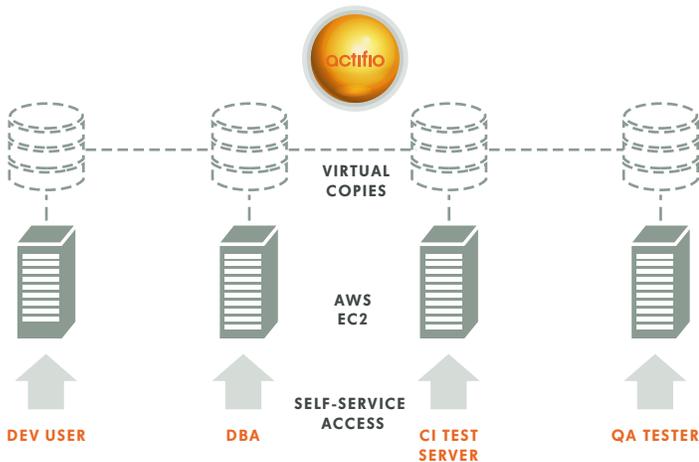
Self Service Access

Once the Actifio administrator has setup Role-Based Access Controls (RBAC), Dev, QA, UAT and DBAs can provision virtual copies on their test servers in a self-service fashion. This eliminates complex, dragged out processes involving tickets, approvals, sys admins, storage admins and many manual processes that often take weeks. See **Figure 3**.

BENEFITS

1. Accelerate unit testing, QA, & UAT testing
2. Develop and Release applications faster
3. Automation eliminates boring tasks and manual processes
4. Automated refresh and masking eliminates boring tasks for Ops

FIGURE 3 Self-Service Access



Automated Data Refresh and Masking

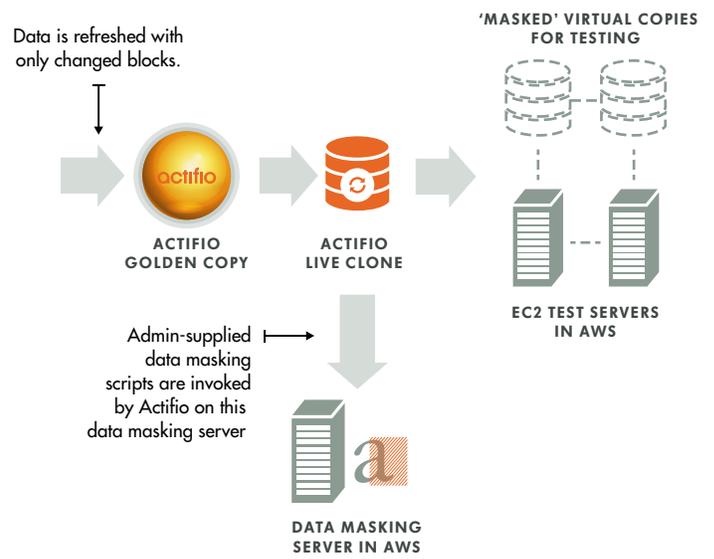
Operations can specify how often to refresh data - once a day, once a week etc. When a data refresh cycle starts, Actifio VDP detects changed blocks and replicates only changed blocks (after dedup) to Actifio Sky in AWS. This keeps the bandwidth requirements very low, so data refresh is automated and efficient. See **Figure 4**.

Operations can specify, in Actifio, a data masking script and a data-masking server in AWS. Actifio Sky automates the process of taking a live clone of golden copy, mounting it on data masking server, invoking the data masking script, and, upon completion of data masking, making the masked copy available for testers. Dev, QA, UAT testers can provision virtual copies of ONLY masked clones.

BENEFITS

1. Data refresh with "only changed blocks" reduces bandwidth needs.
2. Achieve security and compliance with data masking.
3. Automated refresh and masking eliminates boring tasks for Ops

FIGURE 4 Automated Data Refresh and Masking



Test Data Versioning and Test Data Promotion

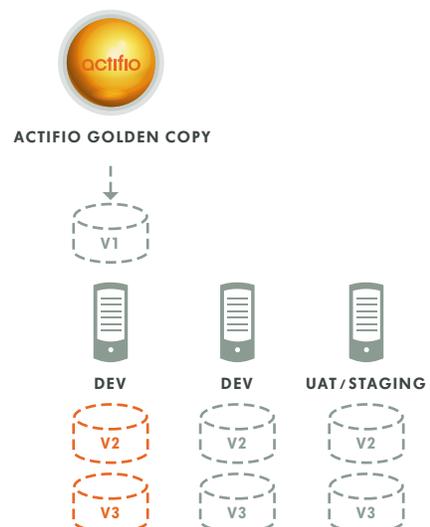
As shown in the picture above, users can label a virtual copy of production data set as, for example, V1. They can also take a snapshot of V1 and label it as V2. Users can then make DB schema changes in V2 and create further virtual copies of V2 and promote this version to other test servers.

If users decide to make further changes off V2, they can take a snapshot of V2, and label it as V3, as shown in picture. Users can then further create virtual copies off V3 and promote them to the rest of the test servers. See **Figure 5**.

BENEFITS

1. Ensure all testing happens off new schema on a copy of production
2. Enable QA to take snapshots of their test environment and hand the snapshot copy to developers for RCA, debugging

FIGURE 5 Test Versioning and Promotion

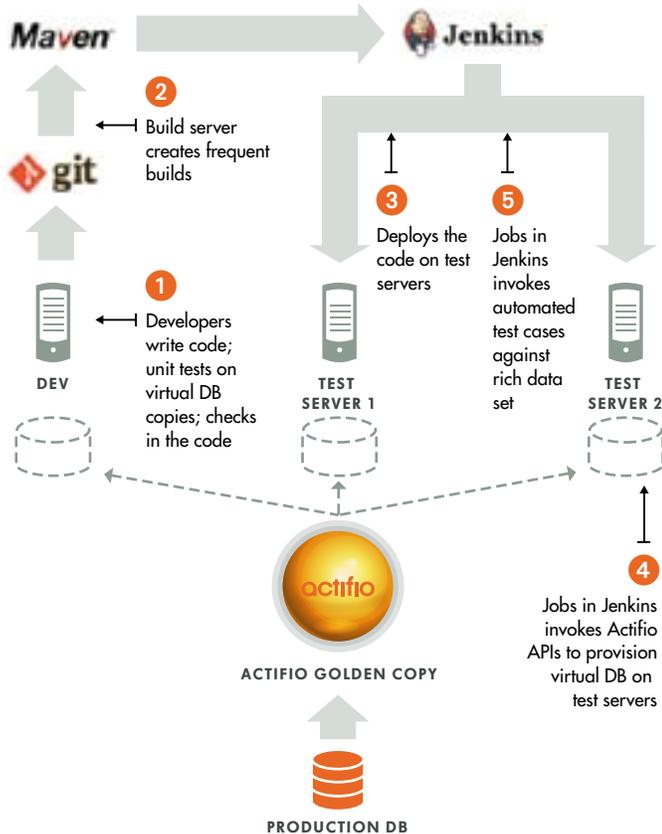


API Integration

All the features described above are available through APIs. So, users can choose to integrate to their existing portals or existing Continuous Integration tools to invoke Actifio REST APIs to provision virtual copies.

One such example of continuous integration is shown in **Figure 6**.

FIGURE 6 API Integration



Summary

The benefits of Actifio are significant across all of the stakeholders in your test data management model:

For Dev, QA, UAT users...

1. Perform unit testing on (virtual copies of) multi TB data sets
2. Automate Build tests on virtual copies
3. Run performance and scalability tests earlier to catch bugs sooner
4. Avoid last minute 'design impacting' bugs product release delays
5. Improve business impact and team morale by delivering more functionality on time and on budget

For DBAs

1. Eliminate boring manual tasks such as DB cloning, DB refresh, data masking, bringing DB online on test servers
2. Leverage virtual copies as a playground for analytics and experimenting with optimized queries for Product Managers and Marketers
3. Release more functionality faster and differentiate against competition
4. Deliver on promised roadmap on time to existing customers
5. Capture larger market share faster with constant, high-quality releases

For IT Managers / Directors / CIOs

1. Lower TCO by reducing storage consumption by up to 10x
2. Leverage subscription model by using Actifio in AWS
3. Leverage Actifio for Business Resiliency and Mobility as well and further reduce TCO

Still not convinced? Try it yourself, FREE.

1. 30-day free trial of Actifio in AWS is available [here](#)
2. Watch 5 minute Videos of virtual DB provisioning [here](#) for Oracle and [here](#) for MS SQL
3. For addition information, visit Actifio at actifio.com or email us at info@actifio.com or call us at **855-886-8997**