ACTIFIO COPY DATA STORAGE

How It Works

An exploration of our technology, purpose-built to combat the copy data management problem at the core of the enterprise storage explosion.
# Table of Contents

- **Introduction** ................................................................. 3  
  - Current Data Management Approach .................................. 3  
  - Copies of Production Data Drives Data Growth.................. 3  

- **Actifio – Who we are** .................................................. 4  
  - **Actifio – a New Approach** ............................................ 4  
    - Redefining Copy Data Management .................................. 4  

- **How Copy Data Storage Works** ....................................... 5  
  - Virtualizing the Copy Operation ...................................... 6  
    - Discovery ............................................................... 6  
    - VMware™ vStorage APIs ............................................ 6  
    - Application APIs .................................................... 6  
    - Actifio Connectors .................................................. 6  
    - Other Applications ................................................. 7  

- **Virtualizing the Store Operation** .................................... 7  
  - Globally De-duplicating and Compressing Changed Data ....... 7  
  - Virtualizing the Storage Repository ................................. 7  

- **Virtualizing the Move Operation** .................................... 7  
  - Sync Replication ...................................................... 7  
  - Async Replication ..................................................... 7  
  - DeDup Async Replication ............................................. 7  
  - DeDup Backup Replication ........................................... 7  
  - Initial Data Seeding ................................................... 8  

- **Virtualizing the Restore/Recovery Operation** ..................... 8  
  - Mount ................................................................. 8  
  - Clone ................................................................. 8  
  - Restore ............................................................... 8  
  - File Level Recovery .................................................. 8  
  - Object Level Recovery ............................................... 8  
  - Bare Metal Recovery .................................................. 8  

- **About Actifio** ............................................................ 9  

- **Support Listing** .......................................................... 9  
  - Application and OS Integration via Command Line ............... 9  
  - Storage systems support ............................................. 9
Introduction

This document is intended to provide a more detailed understanding of Actifio’s advanced technology, and how Actifio’s approach to data management will enable IT professionals to stay ahead of the data explosion that is driving operational costs.

Current Data Management Approach

Data storage technology is undergoing some fundamental enhancements, such as the movement of production data storage from disk to SSD/Flash technology, storage arrays adapting to virtual machine environments, and backup moving from tape to disk. Despite these advances, its basic approach to the way data is managed within the enterprise hasn’t really changed in more than two decades.

The standard data management approach has been to deploy complex, expensive infrastructure silos—built with point tools for protection, disaster recovery, business continuity, test & development, analytics, compliance or other applications. Each point product has to create and manage redundant copies of data that also need to be stored and protected. Even with the introduction of thin provisioning, snap technology, de-duplication and compression, it does little to stem the growth of data. A radically different approach is needed to solve the data copy explosion problem; using the same tools with the same processes will never get ahead of data growth.

If you’re dealing with a storage explosion inside your enterprise, the copy data created by siloed data management applications is likely to blame for it.

Copies of Production Data Drives Data Growth

IDC estimates that in 2012, more than 60% of enterprise disk storage systems (DSS) capacity may have been made of copy data. Similarly, in 2012, copy data made up nearly 85% of hardware purchases and 65% of storage infrastructure software revenue*. In other words, 6 out of every 10 TB of disk storage systems capacity procured and deployed is allocated to copies of primary data. The explosion of copy data is due to the practice of creating separate copies of production data for Backup, Disaster Recovery, Test & Development, Data Analytics along with other IT data needs. The vast majority of shops deploy separate vendor specific tools, each performing the same four basic operations of Copy, Store, Move, and Restore of data. The result is an average of 13-120 duplicate copies of data and spending over 5 times more on infrastructure for managing copy data than the production data.

Most storage companies focus on increasing the speed and accessibility of production data. Actifio’s goal is to solve the bigger IT challenge of radically simplifying your copy data management environment.

THE STORAGE EXPLOSION

- **Copy Data** drives the storage explosion
- **Today:** Copy Data TCO is 5x production data
- **2013:** Copy Data Disk Cost passes production data

If you’re dealing with a storage explosion inside your enterprise, the copy data created by siloed data protection and availability applications is likely to blame for it.

---

* IDC Insight document, March 2013 (#238975)
Actifio – Who we are

Actifio is radically simple copy data management.

Our copy data storage lets businesses recover anything instantly, for up to 90 percent less. Actifio eliminates siloed data protection applications, virtualizing data management to deliver an application-centric, SLA-driven solution that decouples the management of data from storage, network and server infrastructure. Actifio has helped liberate IT organizations and service providers of all sizes from vendor lock-in and the management challenges associated with exploding data growth.

Actifio – a New Approach

Actifio virtualizes your copy data storage applications, consolidating them to enable you to recover anything instantly for up to 90% less.

Redefining Copy Data Management

In order to drive true change, Actifio set out to disrupt traditional rules of copy data management.

The first rule is to eliminate siloed systems by performing the same basic functions of: copy, store, move, and restore for all applications with heterogeneous storage using one simple solution. That means less software to support, fewer licenses to buy, no integration nightmares, and the elimination of operational complexity.

The second rule; don’t store something again if you already have it. Provide global de-duplication of all your copy data. This results in up to 70% less data to move across your network, 95% less data to store wherever you want it stored, and up to a 90% reduction in TCO.
FIGURE 2: CREATE SLA’S SIMPLY

The third rule is to design a solution that’s application centric. Actifio protects the data at the application level; this empowers customers to create different SLAs for each application. Actifio makes it economically feasible to have a copy of every application at different slices of time, enabling users to instantly recover applications from any such point of time.

The fourth rule is to provide flexibility by creating a fully functional storage system for your copy data that can be accessed instantly by directly mounting the data to any server, regardless of its size. Clone or restore data, as you would with conventional storage applications.

The final and most important rule is to make it radically simple to deploy, manage and integrate the solution to lower operational cost, improve RPO/RTO and solve your data management nightmare. The interface to Actifio is simple and intuitive; a few simple clicks and you’ve created SLAs and protected all of your applications, based on the importance to the business.

Actifio virtualizes your copy data storage applications, consolidating them to enable you to recover anything instantly for up to 90% less.

How Copy Data Storage Works

Actifio’s copy data storage can be deployed either within the SAN Fabric or “out-of-band,” over the network. Both methods have their advantages and provide customer options when deploying the Actifio solution. The heart of the Actifio solution is the Virtual Data Pipeline (VDP) technology. Its function is to virtualize production data copy management, eliminating redundancies and re-purposing the unique data for multiple data management applications. VDP efficiently captures a single copy or “gold copy” of changed data from the server and reuses the data for multiple uses, allowing the applications to directly access to copy data from Actifio without any data movement.

A single solution can now be deployed to replace one or all of your backup software, disaster recovery, business continuity or test and development tools. Also, this data source can now be used as the platform for search, compliance and analytics tools. The resulting simplicity of operations and reduction in infrastructure drives down the costs by up to 90% compared to legacy methods.

By introducing virtualization into data management, Actifio can deliver an application-centric, SLA-driven solution that decouples the management of copy data from storage, network and server infrastructure.
The next sections provide additional details for each of the copy, store, move and restore functions of the Virtual Data Pipeline Technology.

Virtualizing the Copy Operation

Actifio uses snapshots with changed-block-tracking technology to instantly capture application-consistent copies of data. This technology is the industry’s fastest, most efficient and scalable method of data capture, eliminating the traditional “backup window” while delivering data management SLAs previously impossible using traditional technology. Data copy can occur in multiple ways, depending on the user environment, leveraging technologies such as: VMware vStorage API’s™, Oracle RMAN™, Microsoft’s ™, as well as lightweight Actifio Connectors.

Discovery

In order to ease setup, Actifio performs a deep discovery of hosts, applications, volumes and file systems. VMware specific environments are discovered by querying vCenter for a list of ESX servers and the VM’s running on those servers.

VMware™ vStorage APIs

Actifio leverages VMware snapshots and a feature of VMware’s vStorage APIs for Data Protection called Changed Block Tracking™ (CBT). This enables Actifio to capture changes from the production environment using true incremental-forever architecture. Changed Block Tracking lets Actifio copy only the blocks that have changed since the last data collection on to the Actifio system.

Actifio is fully integrated with VMware vSphere™, beginning with version 4.0. Actifio communicates directly with the vCenter management server and initiates the data collection process by taking a VMware snapshot of a customer’s Virtual Machine. VMware VM snapshots provide Actifio with an application consistent view of the blocks on the virtual disk(s) inside of a VM.

Application APIs

Several enterprise class applications and platforms provide advanced interfaces for better manageability of copy data. Examples include Microsoft Windows VSS and Oracle RMAN. Actifio directly interfaces with these API’s to capture an application-consistent snapshot of the data and import only the changed blocks into Actifio. This provides the most efficient level of data capture at the application level.

Actifio Connectors

Actifio leverages host connectors for physical servers in order to obtain application consistency during the precise moment when the snapshot is taken. The Actifio connector is a lightweight piece of
software that is installed on the server to provide tighter integration with various applications.

Actifio Connectors are also used to include file systems mounted on servers that have DAS, SAN or NAS devices as storage. This means that both file and block copy data on those servers can take advantage of Actifio’s advanced copy data management capabilities.

Other Applications
For applications and operating systems that are home grown or lack built-in interfaces for quiescing, Actifio’s internal snapshot technology can quickly capture crash consistent views of the application data. In addition pre- and post-snapshot scripts can be run to ensure that the data is consistent. Actifio leverages advanced Copy On Write (COW) snapshot technology for physical server environments such as Microsoft Windows™, Unix, Linux, and IBM i Servers™. Once the copy data is saved using Actifio snapshot technology, customers can leverage Actifio’s advanced copy data management features such as single image, global deduplication, compression, remote replication and instant restore.

Virtualizing the Store Operation

Globally De-duplicating and Compressing Changed Data
Once changed blocks are imported into Actifio, the next step is to globally de-duplicate and compress the changed data. Newly imported data is de-duplicated across all physical and virtual servers managed by Actifio’s copy data storage and are compressed prior to being written to disk. Global data deduplication and compression provide the following benefits:

- Globally de-duplicated and compressed data can be economically stored on disk for long periods of time
- The de-duplicated data pool provides a second, fully independent copy of the data for protection against physical problems with the production storage environment
- Deduplication and compression optimizes the data set for transport between sites

Virtualizing the Storage Repository
In addition to reducing the storage footprint by over 10X, Actifio enables customers to use any storage device. Customers can now have their SLAs dictate the type of storage they use, rather than be constrained by their storage vendor. Many users have re-purposed their existing storage for storing data copies or have opted for lower cost storage devices. This capability further lowers the overall storage costs by over 50%.

Actifio lets you dictate the storage to be used for a given application—not the storage vendor—using a simple but powerful SLA.

Virtualizing the Move Operation
Data movement is the single largest inhibitor to efficient data management in a distributed environment. Actifio delivers the industry’s most robust, efficient, secure, and scalable data movement technology that not only drives down the overall network usage by over 95%, but also eliminates the need for a dedicated WAN accelerator/optimizer. Actifio offers four types of storage replication: Sync, Async, and DeDup Async replications are used for DR and BC purposes, creating remote copies of application data that are available for instant recovery operations. DeDup Backup replication is used for backup purposes to store remote copies of application data for long-term retention. Actifio Sync and Async are similar to the synchronous and asynchronous replication offered by traditional storage arrays. Both leverage a Fibre Channel connection between customer sites for efficient data movement. All methods described use encrypted communication in data replication.

Actifio storage replication does not require storage array vendor licenses as data is sent from one Actifio system to another. Replication is heterogeneous, from any supported array to any supported array (Tier 1 to Tier 2 and/or Vendor A to Vendor B) and fully integrated with VMware vCenter Site Recovery Manager™.

Sync Replication
Synchronous replication can be guaranteed between customer sites up to 300KM apart.

Async Replication
Asynchronous replication has no distance limitation, and will send data over the WAN as fast as network bandwidth allows.

DeDup Async Replication
The third type of primary storage replication, called Dedup Async, is an industry first that is uniquely delivered by Actifio. This is the industry’s most efficient way of moving data, driving down bandwidth requirements by over 10X and enabling Big Data applications to be effectively managed. Dedup Async replication provides asynchronous data replication, but uses globally deduplicated and compressed data movement over the network. This approach uses a fraction of the bandwidth required with traditional replication technologies. Combined with the ability to instantly access data, Dedup Async enables disaster recovery protection for a much broader set of applications where traditional storage replication technologies are cost-prohibitive.

DeDup Backup Replication
Customers who require off-site backups for long-term archival or compliance purposes can leverage Actifio’s proprietary deduplication-aware replication protocol. This replication enables the transmission of only the globally unique blocks that are needed in the remote Actifio system, ensuring the most efficient data movement possible.
Furthermore, the Dedup Backup replication and Dedup Async replication services utilize the same data stream for even more efficiencies. With traditional technologies, a customer will require bandwidth for primary storage replication as well as backup replication. With Actifio, these two services share the same bandwidth. For example, when the primary storage blocks are replicated by Dedup-Async replication, those blocks do not need to be resent when Dedup Backup replication starts some time later, since those blocks are already at the DR site. The end result is that both primary storage and backups can be replicated sharing the same globally deduplicated and compressed data stream.

Initial Data Seeding
This feature is used to establish additional sites for DR or bringing new applications online at a remote site. Customers can choose from a variety of techniques, from shipping appliances for local data copy, to using removable media for smaller data sets. In all cases, this feature eliminates the time, expense and bandwidth of seeding the primary data over the WAN.

Virtualizing the Restore / Recovery Operation
Actifio delivers the unique capability of enabling applications to directly use point-in-time copy data, without the need for a traditional “restore” operation. This is because Actifio is an intelligent storage subsystem that can instantly create a “view” of the application data from any point in time, and allow applications to access it efficiently via a Fibre Channel or iSCSI interface, just as if accessing a traditional storage system.

Application data form any point-in-time copy can be accessed on any system connected to the Actifio copy data storage system. Common use cases are to recover a VM after a software issue, to retrieve a file that was deleted accidentally, or to use a virtual copy of a production data set for test and development. There are three different methods than can be used to access the data stored on Actifio copy data storage: Mount, Clone and Restore.

Mount
The mount function is the most frequently used data access method, as it directly leverages the virtual copies of data stored in Actifio. Actifio stores a copy of the production data and the changes over time. Since Actifio already has the data and can service IO directly, there is no need to copy the data anywhere. Virtual copies of the data can be mounted instantly on any system in the environment using efficient block-level iSCSI and Fibre Channel protocol.

By eliminating the data movement from the process, data sets of any size can be accessed instantly on any server in the environment, virtual or physical.

Clone
The clone function is used to create an independent copy of a data set for any number of reasons. The most common uses are: development and test automation, data audit for compliance, data warehousing, e-discovery and user acceptance testing. Virtual server or physical server data sets can be copied from any application consistent point in the system to a separate storage location anywhere in the customer environment.

Restore
The restore function will effectively revert the production data to look exactly as it did at the time of the data collection point. This is the least frequently used data access method but it is very important to have it when it’s necessary. Typical use cases for restore would be to recover an entire server or application to a valid state after a massive data corruption or storage array failure.

File Level Recovery
While all datasets can be accessed very quickly at the VM level or at the volume level, Actifio provides a file level recovery feature for Windows systems. File level recovery allows the administrator to browse the data set directly from the Actifio GUI to retrieve individual files or directories from any NTFS volumes that are managed by Actifio.

Object Level Recovery
Actifio also provides object level recovery for Microsoft SharePoint™ and Exchange™ systems. Individual objects can be retrieved or restored back to production from within the application content database including emails, documents, mailboxes, contacts, etc.

Bare Metal Recovery
Actifio provides Bare Metal Recovery (BMR) optionally to deliver protection of physical system state of Windows and Linux machines, in order to get servers quickly back up and running in event of failure (RTO). BMR provides the ability to protect and restore both physical system state and data based on SLA, restoring to both similar or dissimilar hardware
About Actifio

Actifio™ pioneered the industry’s first storage system optimized for managing copies of production data, eliminating redundant silos of IT infrastructure and data management applications.

Actifio copy data storage is based on patented Virtual Data Pipeline™ (VDP) technology, delivering dramatically enhanced business availability by eliminating backup and restore windows, and by creating virtual point-in-time copies of data on-demand for use by any business application. Actifio provides the most efficient way to manage data growth while solving your biggest IT challenges around data management.

Support Listing

Fully Automated Application Integration
- VMware ESX 4.x, 5.x
- Microsoft Exchange 2007, 2010 and 2013
- Microsoft® Exchange DAG
- Microsoft® Sharepoint 2007, 2010 and 2013
- Red Hat Enterprise Linux®, Advanced Server 5.x, 6.x
- SAP BRTools
- SUSE Linux Enterprise Server 11
- Oracle 10g, 11g

Application and OS Integration via Command Line
- Novell NetWare V6.5
- Sun Solaris 8, 9, 10
- IBM AIX V4.3.3
- IBM AIX SLTM V6.1, V5.2, V6.3
- IBM AIX V6.1
- IBM DS2
- IBM iSeries
- IBM z/VSE V4.2
- IBM PowerVM Virtual I/O Server 1.2, 1.3, 1.4, 1.5
- HP-UX 11.0, 11i, V1, V2, V3
- Red Hat Enterprise Linux®, Advanced Server 2.1, 3.0, 4.0, 5.0
- SUSE Linux Enterprise Server 8, 9, 10
- HP Tru64 5.1A, 5.1B
- HP OpenVMS 7.2-2, 8.2, 8.3
- SGI Irix 6.5.28, Altix SLES 9
- IBM i series Gateways
- NetApp V-Series
- ONStor Clustered NAS Gateway

Storage Systems Support
Specific models of the following storage systems:
- Compellent SC800
- EMC Symmetrix DMX and 8000-series models
- EMC CLARiiON CX-series models and FC4700
- IBM Storwize V3700
- IBM Storwize V7000
- IBM TotalStorage Enterprise Storage Server®, IBM System Storage DS3000, DS4000TM, DS5000, DS6000TM, DS8000, N series
- IBM XIV® Storage System
- Hitachi Data Systems HUS
- Hitachi Data Systems Thunder, Lightning, TagmaStore, AMS, WMS, Universal Storage Platform
- Huawei storage
- Sun StorEdge systems, Sun StorageTek systems, FlexLine 200
- Hewlett Packard MA8000, EMA12000, EMA16000, EVA family, MSA family, XP family, 3PAR
- NetApp FAS, E-Series
- Bull StoreWay
- Fujitsu Eternus
- NEC iStorage
- Nexsan storage
- Pillar Axiom
- Storage virtualized behind IBM SVC
- LSI Engenio 2600
- Violin Memory

Host Multipathing Software
- IBM System Storage Multipath Subsystem Device Driver (SDD)
- Symantec/Veritas Volume Manager 3.5 MP3, 4.0, 4.1, 4.3, 5.0
- PVLinks for HP-UX
- MPxIO for Windows® and IBM AIX®
- MPxIO for Solaris
- Native NetWare multipathing driver
- Native VMware multipathing driver for VMware ESX 4.X and later
- Native multipathing drivers for OpenVMS, Tru64, SGI Irix

Support for SAN Switches—Selected Models from the Following Suppliers
- Brocade
- McDATA
- Cisco
- CNT

*For the latest listing of what Actifio supports, please contact your local Actifio representative or mail Actifio info@actifio.com

For more information on how Actifio can radically simplify your data management, visit www.actifio.com.