

Riverbed and EMC Deliver Capacity-Optimized Cloud Storage for Backup, Recovery, Archiving, and DR

November 2010



Today's ongoing and rapidly-accelerating growth in data under management comes at the same time that organizations of all sizes are focused on cost reductions. How can customers balance these conflicting business drivers? How can IT departments keep up with the volume of new data—more and more of it critical and subject to compliance—and continue to meet data protection and availability service levels?

Data protection is a mature practice, and there are many existing, proven solutions that provide fast, reliable backup, typically to local disk, tapes, or virtual tape libraries. But traditional backup solutions are always limited by an IT organization's ability to accurately size, trend, source, budget for, secure and manage multiple disks, tapes, and libraries—often in different locations around the globe and often on different platforms. Moreover, backing up or restoring to a distant location puts a strain on expensive WAN links.

Many innovative organizations are turning to cloud storage to meet all of these demands. Cloud storage promises an elastic “pay-as-you-go” storage pool for backups and archives, which not only reduces expensive up-front backup disk and tape costs, but eliminates the need to maintain them off site. There are many proven cloud service providers delivering global access to cloud storage today, offering a variety of pricing and availability options.

Customers who are exploring cloud storage offerings to improve data protection—including backup, archiving, and disaster recovery operations—should start with a careful evaluation of each platform's reliability, availability and security features. Customers and service providers alike should then also evaluate how easily the cloud storage platform integrates into existing backup infrastructures, and whether it can deliver the performance required to meet increasingly demanding data protection SLAs.

What do you need to know about cloud storage?

Cloud storage is two concepts combined: a new approach for accessing storage via both public and private wide-area networks (WANs, the Internet) and a utility-driven IT model in which users pay only for resources they consume. Put simply, cloud storage is on-demand network access to a shared pool of storage resources, purchased and consumed on a pay-as-you-go basis, which are rapidly and easily provisioned and configured to meet each user's specific requirements.

A key benefit of cloud storage is the ability to consume storage only as you need it, automatically, from a self-managed pool that is optimized for remote and programmatic access. An ideal use case for companies exploring cloud storage is data backup and recovery. You shouldn't have to worry about having enough tape capacity, or predicting how much dedicated disk you will need for backups next month, next quarter, or next year. Nor should you incur the heavy capital costs of storage over-provisioning and off-site storage, as well as the administrative overhead of continually sourcing, configuring, and managing your backups.

S O L U T I O N P R O F I L E

Cloud storage delivers the right amount of storage resources, when you need them, at a much lower, more predictable cost.

What is driving the move to cloud storage?

When exploring any new approach to data protection, including storage in the cloud, common concerns arise:

- *How available is the storage platform, and will data be there when I need it?*
- *Is the security infrastructure in the cloud advanced enough to satisfy my recovery SLAs and compliance requirements?*
- *Can I rely on consistent performance for both backup and recovery operations?*
- *How much will I have to invest in training and reconfiguration of my existing backup infrastructure?*

Luckily, the last several years have seen major investments made in cloud technology innovations, and the emergence of many highly successful public cloud services offerings.

Cloud storage technologies have matured rapidly, driven by the needs of large-scale Internet properties that were designed from the ground up with security, massive scalability and very high availability in mind.

At the same time, customer and end-user experience and expectations have changed. More and more applications are built and delivered via a network (Internet, private WAN or LAN) with a browser-based interface. This in turn has led to more comfort decoupling the user experience from back-end services, and to the evolution of powerful new web-based, remote access architectures and protocols.

RESTful web services and SOAP are the primary methods for data transfer and manipulation in

the cloud, and enable developers to quickly deliver rich added-value functionality to users in a familiar way.

Finally, the growth in data of the last few years has been accompanied by an equally rapid growth in the number and types of mobile devices in use—including smart phones, netbooks and tablets—that need access to remote data from anywhere. This has further driven the demand for innovative HTTP-based data access and management solutions.

To meet this demand, EMC and Riverbed have combined the best of their development teams and expertise to create a powerful, flexible cloud-optimized **storage platform** plus a cloud-optimized **storage accelerator** to enable fast, secure, and efficient interface to the platform.

The Cloud Storage Platform: EMC's Atmos

EMC delivers cloud-optimized storage—ready whenever you need it, wherever you need it. EMC pioneered the cloud storage market with Atmos, which was purpose-built for cloud-scale data storage and is based on a global network of highly efficient and interconnected storage devices.

The Atmos architecture is designed for massive scalability, both within and across locations, to support the growing cross-industry demand for elastic enterprise object storage (Fig. 1).

EMC has also focused from the beginning on management simplicity and comprehensive policy support, enabling customers to easily define and manage the movement of their most critical data, as well as to specify how and where that data is stored.

Atmos delivers a broad range of immediate benefits to cloud storage service providers and, in turn, to their end-user business customers:

S O L U T I O N P R O F I L E

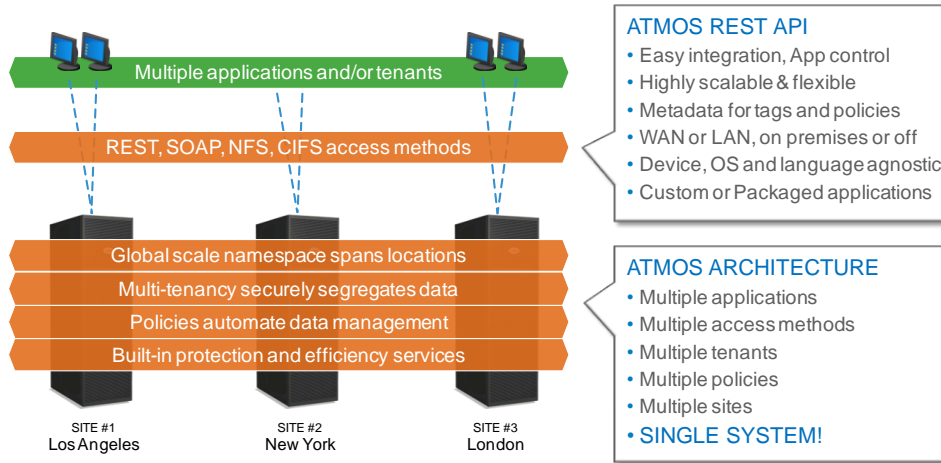


Fig. 1: Key Features of EMC's Atmos Cloud Storage Platform

- **Policy-based, Automated Management:**

Atmos automatically distributes content based on business policies specified by an IT administrator, giving the administrator full control over how and when data is replicated, and where it should reside.

End-to-end automation is included: auto-configuration, auto-management, and auto-healing. Policies also control a wide range of all-in-one data services, including versioning, compression, deduplication, and spin-down.

- **Flexible, Intuitive Management:**

Atmos presents data objects in the cloud via a single, unified namespace, which enables universal access regardless of location, reduces application development complexity, and lowers operational overhead. GUI, command line and API management interfaces are available.

- **Cloud-Optimized Data Protection:**

Atmos includes a full suite of advanced content-protection methods, specifically designed for large-scale distributed object storage. Atmos' GeoProtect feature gives the administrator fine-grained controls to balance protection levels, performance, and storage costs for all data residing in the cloud.

GeoProtect includes automatic replication to multiple geographic locations as well as "erasure coding," which further improves availability and accessibility by dividing storage objects into multiple segments and also distributing these segments across geographies.

- **Secure Multi-Tenancy:**

Atmos was designed from first principles for very high levels of data security—both at rest and while in transit—so that multiple applications and users can securely share a common infrastructure.

Every user and application space is partitioned, creating a trusted multi-tenant platform. This is especially important for service providers who wish to offer secure environments for multiple customers, and also benefits larger enterprises that wish to logically separate data within their own Atmos clouds.

- **Broad Protocol Support:**

Atmos is ideal for today's distributed web application developers, providing flexible web services access (REST/SOAP) for web-scale applications and file access (CIFS, NFS, IFS) for traditional applications.

S O L U T I O N P R O F I L E

**The Cloud Storage Accelerator:
Riverbed's Whitewater**

Cloud-optimized object storage is an ideal solution for today's data protection challenges, but only if customers can securely and efficiently *access* these storage resources. In our experience, any storage solution that doesn't **complement** existing data protection technologies and procedures will **complicate** or **replace** them.

Recognizing this challenge, Riverbed has applied its expertise in WAN optimization technology to the access, security and performance challenges unique to cloud storage.

WAN optimization from Riverbed has been battle-tested and proven to significantly reduce the amount of data on the wire, optimize the performance of the TCP protocol, and accelerate the broadest number of key business applications of any solution of its kind.

In particular, Riverbed's solutions are especially well-suited to accelerating long-distance data protection operations, including backup and recovery, replication, and DR.

Building on the comprehensive, multi-tiered WAN optimization features in its Steelhead product line, Riverbed has now developed an appliance specifically for cloud storage acceleration: **Whitewater**.

Similar to how Steelhead overcame WAN protocol inefficiencies for enterprise WAN-delivered applications, Whitewater attacks the challenges of cost-effectively accessing, storing and protecting data in the cloud.

Whitewater is essentially **cloud storage acceleration in a box**. Customers can easily leverage inexpensive, pay-as-you-go cloud storage services, like those based on Atmos, for off-site storage of backup data, and enjoy simple configuration, faster data transfers, and built-in comprehensive security features.

Moreover, implementation is non-disruptive: customers can use the same data protection infrastructure they use today, the same protocols, and the same schedules and procedures. The only significant operational impact will be the time and effort saved by no longer having to manage or recover from off-site disk and tapes.

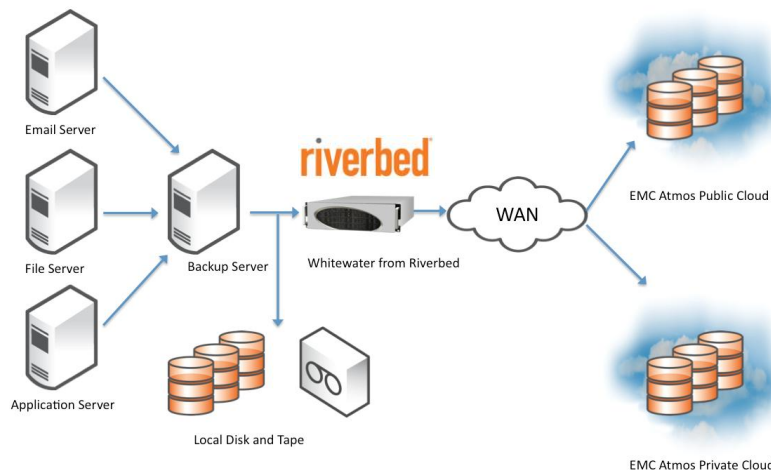


Figure 2: Atmos/Whitewater Architecture Diagram. The Whitewater physical or virtual appliance resides in the customer's datacenter and is accessed in the same way as any disk backup target from existing backup applications.

S O L U T I O N P R O F I L E

Fast, Secure, and Efficient

With Whitewater, the cloud becomes a secure and reliable backup target, accessed like any other storage target at a customer's site. Now, however, that backup target can be sized up or down on demand, capacity-optimized and accelerated when Whitewater is paired with a cloud storage service built on Atmos.

Unlike Steelhead, Whitewater is a single-ended solution: only one appliance is needed at the customer site to accelerate backups and recovery to/from the cloud. The Whitewater appliance is available in both physical and virtual editions, and delivers:

- **End-to-end security for off-site data:** All backup data written to Whitewater is fully encrypted while in transit via SSL v3 and while at rest in the cloud with 256-bit AES. Whitewater manages the encryption keys, which can be transferred to any other physical or virtual appliance for recovery of a customer's current cloud data—anywhere.
- **Accelerated network performance for faster recovery:** Recent backup data sets are stored on Whitewater before they are replicated to the cloud, providing reliable local data recovery that remains available through transient cloud connectivity problems.

Recovery from the cloud is faster via Whitewater's best-in-class deduplication and WAN optimization features, which dramatically reduce the time to recover anything, from a single file to a volume or more. In addition, QoS capabilities in the appliance can throttle the bandwidth consumed by a backup job if need be.
- **Lower off-site storage requirements:** Whitewater utilizes Riverbed's patented data deduplication technology to compare new backup sets and deltas to data already written

to the cloud, and only writes new data, saving significant bandwidth to/from the cloud and requiring less cloud storage.

Deduplication is *byte-level*, which is significantly more granular than block-level and yields much higher deduplication ratios—and therefore far greater network efficiency.

- **Easy integration into existing backup infrastructures:** Whitewater looks like any other local disk target to existing backup software (e.g., Symantec NetBackup/Backup Exec or IBM Tivoli Storage Manager) and can be accessed via CIFS or NFS protocols.

At the cloud side, Whitewater supports all the important cloud access APIs, including REST, EMC Atmos, and HTTP, so there is no need to recode existing backup software to communicate with the cloud.

How Whitewater Works

Implementing cloud backup storage with Whitewater and Atmos is simple and straightforward:

1. Without modifying your current backup infrastructure, add a Whitewater physical or virtual appliance as a new backup target;
2. Backups directed to Whitewater are automatically deduplicated at the byte level, reducing data storage requirements by 10x to 30x on average;
3. Whitewater's local disk retains copies of recent backups, providing a source for LAN-based recovery of recent data;
4. Deduplicated backups are written to the Atmos cloud using an object-based REST API plus Riverbed WAN optimization technology, reducing WAN bandwidth requirements and shrinking backup and recovery times;

S O L U T I O N P R O F I L E

5. Encryption keys are managed by the customer on the Whitewater appliance, enabling recovery from the cloud to any Whitewater appliance, anywhere.

Whitewater ships as a no-cost virtual appliance with a raw local disk capacity of 2TB and in two

physical appliance models that scale up to 16TB of local disk.

Depending on the type of data and backup frequency, customers can expect 10-30x data reduction via deduplication, and a corresponding reduction in cloud storage requirements.

Taneja Group Opinion

Riverbed and EMC have each pioneered many enterprise WAN optimization and storage technologies, and both companies enjoy strong reputations as innovators. Cloud storage is poised to fundamentally change the way businesses of all sizes think about storage capacity (utilization, trending, and purchasing) and how to best protect critical data at the most predictable and efficient cost point. In our view, these innovators have targeted the right customer challenge, and together can deliver a new way of thinking about data protection.

By ensuring compatibility between the Atmos cloud storage platform and the Whitewater cloud storage accelerator, Riverbed and EMC have tapped into the most important driver of cloud computing: the desire for clients to move away from CapEx-heavy upfront investment in IT resources, and toward a service-based, more predictable OpEx model. And no IT capital expense is growing faster in most organizations than storage.

Atmos by itself is highly reliable, available, secure, and globally distributed. With policy-based automation, Atmos greatly simplifies data management operations while giving the IT administrator complete control. Atmos is proven in the market, and gives cloud service providers an ideal platform to maximize their investments in infrastructure by providing new and innovative high-performance backup and archiving services to their customers.

By adding Whitewater as the front-end access point to Atmos, the end-user benefits are multiplied further: existing backup processes require no modification (no “forklift” upgrades) because Whitewater speaks their language already. Whitewater delivers unmatched byte-level deduplication, full encryption, and a virtual edition that allows users to restore data to nearly any server, in any location.

Customers will enjoy cloud backup and replication at LAN-like performance that is as reliable and secure as local disk or tape (if not more so)—at a lower cost. Service providers will be able to offer new storage solutions without being limited by high conversion costs or inefficient network access. In our opinion, the Riverbed-EMC solution goes a long way to tearing down any remaining barriers to the adoption of cloud storage.

NOTICE: The information and/or product recommendations herein are based upon public information and sources and may include personal opinions of the Taneja Group and others, which we believe to be accurate and reliable. However, as market conditions change and are not within our control, the information and recommendations are provided without warranty of any kind. All product names mentioned herein are the trademarks of their respective owners. The Taneja Group assumes no responsibility or liability for any damages whatsoever (including incidental, consequential or otherwise), caused by your use of, or reliance upon, the information and recommendations herein, nor for any inadvertent errors that may appear in this document.
