Rio-2 Hybrid Backup Server

A Revolution in Data Storage for Today's Enterprise

March 2018

BridgeSTOR®
The World's Storage Made Available

Notices

This white paper provides information about the Rio-2 Hybrid Backup Server as of the date of issue of the white paper. Processes and general practices are subject to change without notice. Readers of this white paper are responsible for making their own independent judgement and the uses of the Rio-2 product or other BridgeSTOR products referred to in this document. This white paper does not create any warranties, representations, contractual commitments, conditions or assurances from BridgeSTOR, its affiliates, suppliers or licensors.

Table of Contents

Introduction

Organizations of all sizes today are characterized by their "virtual" nature, defined as having many remote locations with many mobile workers and in an environment of constant change and with extreme cost restrictions. As a result, most of today's data storage is based on large storage arrays of spinning disks providing block and file storage to local users and servers. These expensive storage systems create isolated islands of storage that have proven to be costly to manage, backup, upgrade and maintain. A revolution is underway in which backup works far better when provided as a service by centralized, highly available, geographically distributed resources owned by both organizations and public service providers – a model that has come to be known as the Hybrid Storage Cloud.

Rio-2 Hybrid Storage Architecture

Hybrid Storage Cloud is fast becoming a corporate best practice because a Hybrid environment allows IT managers to use local storage resources combined with low cost Public Cloud Storage with a single view for all applications and users. Local disk may be used for caching heavily used applications while all data is conveniently sent asynchronously offsite to Public Cloud Storage for immediate disaster recovery. This is ideal for today's NAS or Backup Applications where the management of data is growing out of control. The Rio-2 Cloud Backup Server, a true hybrid solution, which includes advanced caching techniques that while not eliminating latency to the cloud will remove the latency from the local application and users. The local disk in the hybrid system is not only used as a storage area for the most accessed files allowing for fast and efficient local reads, but is also a staging area for transaction logs that are replayed by the hybrid systems backup threads which forces these threads to absorb the Internet latency. Once the cache fills up, the least accessed files will be removed from the cache, freeing up cache for new files and if a file is not in the cache it will be conveniently pulled from the Cloud Storage and re-added to the local disk cache
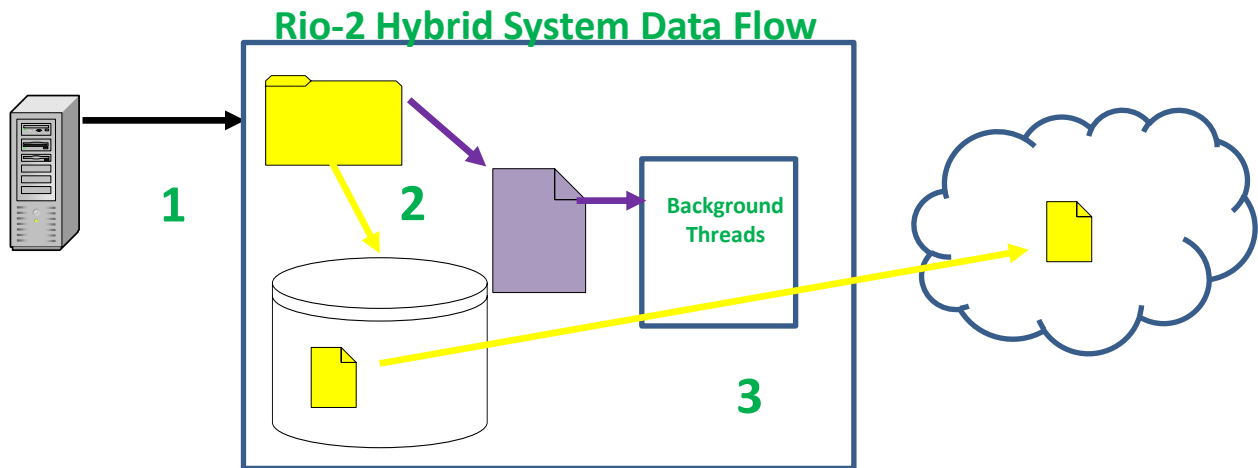
Rio-2 Hybrid System Data Flow

Figure 1: Rio-2 Hybrid System Data Flow

Here is the data flow of the Rio-2 Hybrid shown in Figure 1:

1. Rio-2 Clients which may be NFS, SMB or iSCSI VTL send their data directly to the Rio-2 System. (Yellow Folder Above).

2. The Rio-2 front end processing will write the data to disk and update the local transaction log.

3. Background Threads in the Rio-2 will then read the transaction log and send the file into the cloud storage

Backup and the Cloud

The basic tool for accessing and managing Cloud storage environments has been a traditional web browser, proprietary web applications or Backup Software Cloud Connectors. Backup Software has recently added the capability of sending data into the cloud but these methods tend to be slow, proprietary or require 3rd party vendors. Cloud latency is always an issue as backup traditionally transfers large chunks of data which reduces the backup windows. Finally, what about all those old tapes that are laying around or sitting in vaults supplied by third party vendors. With the price of Cloud Storage falling it can be the new 3rd party tape storage vault. Currently Backup technologies may send data to disk or tape and it's up to the Administrator to choose which is better for their unique situation. A true Cloud Hybrid Backup product must be designed to support both methods.

The Secret Sauce

Every Cloud Application communicates with Cloud Storage using the Amazon S3 REST (REpresentational State Transfer) protocol which has become the de-facto standard. A true Hybrid Cloud Solution must work equally well with relatively low performance Public Cloud Storage and/or with the high performance Object Storage of a Private Cloud. From a Backup Application perspective, a solution should be economical enough to be used by applications calling for low cost Cloud Storage for archiving, while also having performance characteristics strong enough for the extraordinary demands of a full backup. Finally, this solution should easily slide into existing enterprise environments with little or no impact to the enterprise environment.

A true Hybrid Backup Cloud Solution should also provide important complementary features:

1) Highly secure with in-line data encryption

2) A global view of cloud storage with file sharing capabilities so that users may share files no matter where users or data are located

3) Data reduction compression when appropriate for optimal use of network bandwidth and storage capacity.

4) High performance caching to reduce latency to cloud storage.

The ability to provide full enterprise functionality and performance while also being as transparent as possible is difficult to solve for any cloud solution and absolutely separates a true enterprise product from the rest of the field.

File to Object Mapping

In the default installation, Rio-2 will always map individual files to Cloud Storage objects. Rio-2 has optional file compression and if selected will break files into smaller objects that will be saved in a proprietary format in the Cloud Storage.

Rio-2 requires two buckets requiring the same Access Key and Secret Keys. The first bucket is used for file system metadata, such as dates and times, modified date and Windows ACL's. The second bucket is used for storing objects in their native format. The bucket format will be an identical path to the local file system representation.

The Rio-2 Cloud Backup Server

BridgeSTOR's **Rio-2 Cloud Backup Server** is a Hybrid Storage Appliance that provides access to a broad range of Public Cloud Storage or Private Cloud Object Storage by appearing as an industry standard NAS device or a Virtual Tape Library (VTL) over iSCSI. When a Global File System is required to share backup data across multiple locations are required, the optional **Global View Manager** runs as a separate virtual machine locally or in the Cloud – providing global file accessibility, secure file sharing and other enterprise features for a virtually unlimited number of Rio-2 Cloud Backup Servers.

The Rio-2 VTL support allows administrators to extend their existing tape strategy to back up to the cloud or to easily transfer existing tapes to the cloud. All recently backup files or tapes will be kept in local storage supported by read and write access at near local disk speeds while reducing backup windows. Rio-2 backup threads will then transfer the data in the background completing the Hybrid Strategy. Based on a least accessed policy these local files will be retained until they are no longer required and will be removed from the local cache. If a file is required that is not currently in the cache Rio-2 will seamless recover the file from the cloud. For tape images, only the individual file(s) will be pulled from the Cloud Storage. There is no reason to pull the entire tape.

Rio-2 NAS Support

Rio-2 exports existing NAS protocols SMB and NFS for backup administrators that would prefer to write files to the Cloud or to Object Storage in the same manner as they write to traditional local NAS storage. Simply map a drive letter to the Rio-2 over Windows or NFS, point the backup software to the share, and start sending the data to the cloud. It's that easy. The Rio-2 Hybrid nature will keep the data local for instant recoveries while all data is conveniently sent to the cloud for offsite storage.

Rio-2 VTL Support

Rio-2 also exports an iSCSI VTL simulating an HP library and LTO tape drives over iSCSI. The administrator may set the number of slots required for the backup application reducing costs associated with slot count. Tapes are stored on disk by barcode location. A Windows share may be used to view and manage the or the tape images in the Cloud. The Rio-2 GUI acts as a library door. The GUI is used to move tapes in and out of the library. There is even automatic refill if a backup policy removed the tape from the library,

a blank tape will be inserted in the previous location. The backup software may easily copy or send backup data to the tapes. The tapes also support compression so data will be sent across the wire compressed and saved in a reduced footprint.

Rio-2 Native Veeam Support

Rio-2 has been built on Centos7 and as such looks like any other Linux Server to Veeam allowing Veeam to treat Rio-2 as a "Linux Repository". Veeam will load its backup agent in Rio-2 allowing Veeam to send backup data directly to Rio-2 bypassing SMB or NFS. This approach allows for high speed ingest into Rio-2 where the data will be sent to the Rio-2 caching layer where background threads will send the backup images to the cloud. Veeam has no idea it's actually writing data to the cloud.

Rio-2's High Security

Security is the number one concern of IT managers with regard to the Cloud, and it's no wonder, since using a Public Cloud facility such as Amazon's AWS S3 to store your company files demands the assurance of high security.   BridgeSTOR's Rio-2 with ProtectedCloud™ Security fills this need with a whole family of features that work together for the protection of your data.

Rio-2's comprehensive enterprise feature set is based on BridgeSTOR's own **Cloud Storage File System (CSFS™)**, the product of five years of development, which encrypts files at the sub-file (block) level, and therefore reading the information contained in a file requires all the blocks to be reassembled into their original sequence when a file is retrieved. Thus, each data block, even if it is simply clear text, will not typically contain any meaningful information content.

CSFS also uses highly secure AES 256-bit encryption with XTS extensions which was initially designed for block storage. This means that Rio-2 combines a user-defined encryption key with a BridgeSTOR-defined XTS algorithm for each block processed. This insures that two blocks even with identical data will look different after being encrypted, for further protection against hackers.

Integration with Microsoft Active Directory

Microsoft Active Directory (AD) is commonly used for user authentication across enterprises, and Rio-2 is easily integrated with AD. This means that the same access control policies for users and groups that are already in place may be applied to Hybrid Cloud Storage access via Rio-2.

Combining all the above methods with single file instance and compressed data (more on Rio-2 data reduction in the next section) in the Cloud makes data processed by Rio-2 nearly impossible to be read by unauthorized persons either while in flight, or while stored in the Cloud. All data is encrypted and decrypted only at the authorized Rio-2 site(s), so data is secure while being transported over the wire or stored at rest in the Cloud.

Rio-2 Global File System

A true enterprise Cloud Storage Solution should provide a global view of files with secure, global access and sharing plus other important features at the file system level. BridgeSTOR's Cloud Storage File System (CSFS) creates a native POSIX file system interface that includes the capabilities to create, mount, delete and export CSFS volumes. CSFS back-end technology communicates to Public or Private Cloud Storage over a REST interface. REST was originally developed for Amazon S3 Cloud Storage and is a de-facto standard in the industry. Not only is REST currently used by Amazon and other Cloud vendors, it has also been adopted by Private Cloud Object Storage vendors.

Although Cloud Storage and Object Storage is highly scalable, it has traditionally been limited in its ability to work with file structures and their metadata. In many cases, for example, it's not even possible to easily move files around or rename directories. CSFS overcomes this by creating a Global Name Space that is completely independent of underlying storage, and when using the Global View Manager, users have a fully global view of their Cloud Storage.

To accomplish this, CSFS separates the metadata from the physical data location while keeping the two combined as a single object. This significant architectural achievement allows metadata to be managed separately from the physical data while both are maintained in Cloud Storage. The Global View Manager maintains a global name space view that can be accessed by all Rio-2 Cloud Backup Servers. The user may view all files and directories without actually accessing the Cloud Storage platform. However, for disaster recovery purposes, CSFS also duplicates all metadata inside the Cloud Storage. This allows CSFS to rebuild all metadata in case of a local disaster.

Rio-2 Data Reduction

In many applications of Cloud Storage, the use of data reduction via single file instance and compression provides both a reduction in network bandwidth requirements and in storage capacity requirements.   But when performance considerations make data reduction undesirable, the administrator may turn it off.

CSFS single file instance operates at the sub-file level. As its name implies, all files are decomposed into segments – chunks, clusters or blocks - that are examined for redundancy versus previously stored information. In CSFS sub-file technology, blocks of data are "fingerprinted" using a hashing algorithm that produces a unique, "shorthand" identifier for each data block. These unique fingerprints, along with the blocks of data that produced them are indexed, optionally compressed and encrypted and then retained. Duplicate copies of data that have previously been fingerprinted are duplicates, leaving only a single instance of each unique data block along with its corresponding fingerprint. The fingerprints along with their corresponding full data representations are stored in a compressed (and optionally encrypted) form.

Finally, based on the data being processed, CSFS can also significantly reduce storage utilization for some instances of unstructured data, using file-level deduplication, typically referred to as single instance storage and using content addressable deduplication. The latter incorporates optimizations for virtual machine image data reduction.

Global Single File Instance

CSFS will perform single file instance against all the files in the global name space to which it is exposed. That collection of data can originate in multiple locations worldwide. Therefore, when the Global View Manager is in use, duplicate files are only stored once globally for highest possible capacity optimization.

Smoothing the Transition to the Cloud

A true enterprise Cloud Solution should be capable of helping administrators make a smooth transition to the Cloud. The Rio-2 Cloud Backup Server does this not only by allowing many legacy storage systems to be accessed via Rio-2, but enabling the movement of files from those systems to Public or Private Cloud Storage in the background with no interruption of user access.

Managing Cloud Storage as if it's Local

A true enterprise Cloud solution is transparent enough to require little management once set up and tuned if necessary. But it should provide visibility of critical aspects of the system and make it possible to manage both Public and Private Cloud Storage with the same set of tools.

The Rio-2 Management GUI (graphical user interface) uses your Internet browser to provide visibility of important aspects of your Cloud Storage systems and when you detect the need for more storage space in the Cloud, Rio-2's use of storage virtualization makes adding more cache capacity an easy and painless process. Rio-2 Thin Provisioning techniques also provide assurance that you can add more physical storage to a Rio-2 local cache at any time that it is needed. Thus, the Rio-2 user has an ever expanding, virtually unlimited volume or share capacity without purchasing an excessive amount at the start.

Raising the Bar in Primary Cloud Storage Access

BridgeSTOR's Rio-2 Cloud Backup Server was not the first Cloud backup product to the market, but BridgeSTOR developers took their time and applied the lessons learned by many of those who were introduced in earlier stages of the transition to the Cloud. The result is a true enterprise Cloud Solution with many exclusive features that has set a new standard for Cloud Storage access.

Those considering joining the already well proven advantages of the Storage Hybrid Cloud revolution, or those in the throes of that transition, should carefully consider their options in light of the attributes of a true enterprise hybrid Cloud Backup Server as discussed in this paper.