

Data Storage in the Cloud – Can you Afford Not To?

EXECUTIVE SUMMARY

Storing data in the cloud using a Whitewater™ cloud storage gateway from Riverbed Technology overcomes what is becoming a serious challenge for IT departments: how to manage the vast, and ever-growing, amount of data that must be protected. Whitewater eliminates concerns about data security, data transmission speeds over the Internet, and data availability, while providing great flexibility and a favorable return on investment. Moving data to the cloud replaces the high costs of tape and disk storage systems with a pay-as-you-go cloud storage tier. This paper makes the business case for cloud storage, outlining where capital and operational costs can be eliminated or avoided by using the cloud for backup and archive storage.

Cloud storage offers advantages

Cloud computing has a wide variety of definitions today. Simply put, in this paper, cloud computing refers to accessing servers or storage over the Internet for the purposes of running applications or storing data. This analysis looks at the viability of using the cloud for data storage, specifically pay-as-you-go data storage services from providers such as Amazon, AT&T, Nirvanix and EMC Atmos, including use for backup, recovery, replication and disaster recovery (DR) operations. (Some would call this as a “public” cloud since these providers open their services to all.)

As a new delivery model for IT services, cloud computing is still being evaluated by most CIOs. Its primary advantage is the ability to increase IT capacity without investments in new infrastructure, personnel and software licenses. Common concerns include the fact that proprietary, critical data leaves the premises. This is addressed in *Cloud storage concerns, and how they can be addressed*, below.

For data storage, cloud computing offers a safe, attractive and affordable solution to what is becoming a serious challenge for IT departments: how to manage the vast and ever-growing amount of data that must be backed up and maintained. Even organizations with an effective data storage process currently in place may find these processes overwhelmed by data volumes in the future. Riverbed customers report that their data increases by 40 to 60 percent annually, putting their storage systems on the fast track to full capacity and obsolescence.

According to Riverbed customers, some of the advantages gained by using cloud storage include:

- Pay-as-you go capacity costs
- Data accessibility from anywhere
- Automatic off-site copy of data for DR
- Unprecedented flexibility and scalability
- Low cost due to economies of scale of cloud storage providers
- High availability due to the redundancy built into cloud storage

Cloud storage concerns, and how they can be addressed

Four key concerns dominate views today on storing an organization’s data in the public cloud. Of these, security is the greatest. Data must be safe during transmission to and from the cloud, as well as while it resides within the cloud. Slow data transmission speeds across the WAN are another concern. Moving terabytes of raw data to and from the cloud can be very time consuming, making it a significant problem. Third, data must be available when needed. Data inaccessibility can have a very significant business impact. Finally, the concern around the return on investment (ROI) that will be gained by use of cloud storage which is especially acute during times of shrinking IT budgets.

Riverbed recently introduced cloud storage gateway technology that addresses each of these cloud storage concerns, making the public cloud a feasible storage tier for most organizations. Riverbed® Whitewater® cloud storage gateways, available in both hardware and virtual configurations, require no change to a company’s existing backup infrastructure. On the front-end, Whitewater appliances present themselves as CIFS shares or NFS mounts to the existing backup software and appears as a standard disk target to the backup software. At the back-end, they have built-in support for APIs from Amazon Web Services S3, AT&T Synaptic Storage as a Service Microsoft Windows Azure, Nirvanix Cloud Storage Network, Rackspace Cloud Files, clouds based on EMC Atmos or OpenStack (Swift) object storage. These compatibilities allow Whitewater to be immediately deployed in most IT environments.

Whitewater cloud storage gateways address public cloud storage concerns in the following ways:

Data security. Whitewater cloud storage gateways encrypt data in-flight to the cloud with SSL v3 and at rest using 256-bit AES encryption. This dual-level encryption ensures that any data moved into cloud storage will not be compromised and creates end-to-end security. Whitewater gateways leverage an innovative key management system that allows organizations to carefully manage data security. Encryption keys are kept safe by IT management.

Transmission speed across the WAN. Whitewater cloud storage gateways leverage existing Riverbed data acceleration technology, including the industry-proven WAN optimization and deduplication functionality in the Steelhead® product line, to speed the transfer of data to and from the cloud. Whitewater appliances receive data from the backup application, deduplicate it, compress and encrypt it, store it in a local disk cache, translate it to cloud protocols, and then send to the cloud, minimizing transmission time to and from the cloud. The deduplication technology (inline, variable-length, byte-level) reduces the amount of data sent to the cloud by 10 to 30 times, on average. In addition, because cloud storage providers charge based on the amount of storage capacity, deduplication reduces the cost of using it as a storage tier.

Data availability. Because Whitewater cloud storage gateways are often implemented as part of a DR solution, it is critical that companies can access data as quickly as possible, wherever it resides. Whitewater gateways use a local disk cache to store the most recent backups. Because most restores are from recent backups, typical data retrieval is from the local cache. In the uncommon case where the backup data needed is no longer within the local cache, the Whitewater appliance will supplement whatever data is no longer in the cache by pulling the required data from the cloud. The Whitewater appliance minimizes the amount of cloud-based data needed in order to satisfy this type of restore.

The business case for cloud storage

Compared to tape or disk backup, storing data in the public cloud has been shown to reduce costs and backup time in a number of ways. Listed below are some examples, along with estimates of avoided costs gathered from Whitewater users:

Tape backup (assuming LTO-4 tape with 120 MB/sec throughput and 800 GB capacity)

- LTO-4 drives (\$10,000/drive)
- LTO-4 tape library (\$30,000 each)
- LTO-4 tape (\$30/tape)
- Tape maintenance (18% to 23% of total cost of tape system)
- Off-site tape vaulting (\$200 - \$500 per /month)
- Power and cooling costs

Disk backup

- Storage array (\$5,000 to \$150,000, often \$1,000 per TB for enterprise class storage)
- Storage array maintenance (9% to 18% of the price of the array)
- Power and cooling costs

The “set and forget” nature of the Whitewater cloud storage gateways compares favorably to the on-going labor involved with tape and disk storage. If labor cost is estimated at \$80,000 per man-year, and with the assumption of one person needed for every 50 terabytes of stored data, labor cost savings from using Whitewater from labor alone can be substantial.

Other advantages can be harder to quantify, but still increase the ROI from using Whitewater cloud storage gateways. Because backups are faster than tape due to the data being written to an appliance (rather than a tape drive), more frequent backups are possible, making the recovery point closer to the current time. Another advantage is that writing data to the Whitewater gateway generally improves the number of successful backup jobs by 5 to 10 percent by eliminating a number of tape-related issues that often cause backups to fail, saving labor and minimizing risk for the organization by reducing the amount of time a server is unprotected.

Another area where Whitewater has the potential for significant cost savings is legal compliance. Audits and legal action typically put a hold on all data, which means that companies using tape backup can no longer reuse these tapes. It is not unheard of that data can be on legal hold for several years, possibly representing tens of thousands of dollars worth of tape. During a legal action, using a Whitewater gateway and storing the data in the cloud replaces the enormous cost of tape infrastructure and vaulting with the per-use fee of the cloud storage provider.

For companies currently using disk backup, switching to Whitewater allows them to reclaim disk-based storage for other uses and avoid buying disk to forward provision in anticipation of future needs.

As part of a DR strategy, companies will often use a replication application to replicate their data to a second site and also back up to tape to provide recovery in the event that the replicated data is erased through human or software errors. With a Whitewater gateway writing data to the cloud, the tape system is unnecessary; the cloud is the replication backup.

Results from Riverbed customers

International Justice Mission

International Justice Mission is a human rights agency that secures justice for victims of slavery, sexual exploitation and other forms of violent oppression. The agency's mission encompasses victim relief, perpetrator accountability, victim aftercare, and structural transformation.

IJM upgraded from a tape backup system to a Whitewater cloud storage gateway as part of the agency's transfer of critical systems from its Washington D.C. headquarters to a lights-out data center. "One of the holdups was that we needed a backup system that we could manage without having someone physically present, or having to hire additional staff," explains Chris McPeek, manager of global systems and infrastructure at IJM.

"Whitewater lets us do backups and restores without anyone on site," McPeek says. "And it saves us from having to hire a backup engineer to manage the system."

After getting Whitewater up and running – "one of the easiest installs ever," according to McPeek – IJM was able to eliminate all of the labor associated with the previous tape backup system, such as rotating the tapes in the catalog and moving them off-site. "Whitewater lets us do backups and restores without anyone on site," McPeek says. "And it saves us from having to hire a backup engineer to manage the system." *Costs avoided: salary for an entry-level engineer and tape transport and vaulting.*

Another significant cost savings involves the agency's 15 field offices. Prior to deploying the Whitewater gateway, McPeek was planning to acquire tape backup systems or network attached storage (NAS) devices for each of those sites. Instead, a backup server running Symantec Backup Exec pulls data in from all of the field offices and writes it to the Whitewater gateway, which then sends it to the cloud. "Essentially we just point our backups at the Whitewater appliance and let it rip, and off it goes out into the Internet," he adds. *Costs avoided: backup systems for 15 offices.*

IJM's data consists, in part, of crucial information related to the prosecutions of perpetrators. It must be retained for as long as necessary until the cases go to trial. The volume of data at IJM grows by about 40 to 60 percent per year. Currently the amount of data being backed up runs about 4.5 terabytes. It would not have taken too long before the previous tape backup system was overwhelmed. By going with Whitewater, which scales up easily, the company will avoid large capital expenses related to addition tape backup equipment. *Costs avoided: tape system upgrade.*

For IJM, whose mission is saving lives, all these avoided costs represent more than a healthier balance sheet. Says McPeek, "We can almost measure per person how many lives are saved because of having Riverbed technology such as Whitewater in our network."

Mid-size Company

The systems administrator at this company was just about to request approval of a 4x increase in physical data storage capacity when he learned about Whitewater. The company was using a tape backup system at the time, and the growing amount of data being backed up was going to exceed the system's limit. Increasing the capacity was going to cost several hundred thousand dollars.

That would be added to the acquisition and on-going costs of the tape system which, for this company, included: 48 tapes/week at a cost of \$30 each, the 10-drive tape library (\$100,000 for the ten drives and \$30,000 to \$40,000 for the library), off-site tape

storage (hundreds of dollars per month), and approximately one-half of one full-time employee's salary to manage the process. Other costs included the off-site provider's fee to deliver a tape for a restore (\$400 to \$1,000, depending on how quickly the company needed the tape) and new tape drives for those that needed to be replaced (\$10,000/drive/every few years) due to constant use.

Choosing to deploy Whitewater and cloud storage instead of upgrading the tape system was "a real game changer for us," says the systems administrator. It took him less than half an hour to set up the Whitewater appliance which, unlike the tape system, requires no on-going labor to manage. The deduplication performed by the Whitewater appliance greatly reduces the amount of data that goes into cloud storage. The systems administrator reports an initial deduplication factor of approximately five, meaning that of the 150 terabytes of data backed up at that time, his company only paid to store about 30 terabytes in the cloud. They expect their deduplication rate should keep improving as the Whitewater appliance sees the same data over and over, so that 30 terabytes should grow very slowly over time.

Choosing to deploy Whitewater and cloud storage instead of upgrading the tape system was "a real game changer for us," says the systems administrator. It took him less than half an hour to set up the Whitewater appliance which, unlike the tape system, requires no on-going labor to manage. The deduplication performed by the Whitewater appliance greatly reduces the amount of data that goes into cloud storage.

Pump Solutions Group

Pump Solutions Group (PSG) is a global manufacturer of pumps with facilities in the U.S., Germany, China, India, and France. When PSG needed to unify the disparate networks of its subsidiaries and build a common DR framework for the organization, the company became an early adopter of a Whitewater, which the company deployed originally in its Wilden Pump subsidiary based in California.

Before Whitewater was installed, the backup process at Wilden Pump began with disk-to-disk replication using either Dell EqualLogic or an HP SAN for disk storage. Data from the disk was then copied to tape. Jeff Rountree, global network manager at PSG, estimates the costs of the tape backup system at approximately \$30,000 for the initial investment (tape libraries and tape media) and \$5,600 in annual maintenance and upkeep. Part of the annual maintenance and upkeep included approximately two hours per week of someone's time for managing the tapes. Those two hours a week have been freed up for higher value work. *Cost savings: \$5600 per year plus 100 hours annually.*

Rountree has found other areas where the Whitewater appliance saves money. One is in the amount the company pays to its cloud storage provider, AT&T, each month. Because Whitewater deduplicates and compresses the backup data before sending it to AT&T, the cost of cloud storage is much less than it would otherwise be. Explains Rountree: "The amount of data we back up is around 1.2 terabytes, but what we put out to cloud ends up being only 100 gigabytes." At the rate they are paying for cloud storage (25 cents/gigabyte) that represents a savings of \$275/month. *Annual savings: \$3,300.*

Another savings relates to the company's SANs. Using SANs for long-term storage is expensive, and now that Wilden Pump is storing data in the cloud, the company does not need to purchase additional SANs. "EqualLogic is fairly inefficient when doing replication," Rountree says. "Space is reserved on the target EqualLogic for the source volume on the source EqualLogic, so even if there is only 80 gigabytes, for example, used on a 500-gigabyte source volume, 500 gigabytes are reserved on the target EqualLogic that cannot be used." In an example he related, 97 percent of delegated replication space on the SAN was empty but could not be used for anything else. "Considering an EqualLogic would cost \$30,000 to \$75,000 or more, purchasing a SAN for long-term storage is a significant investment, and it does not protect data as well as Whitewater does with off-site replication," he adds. *Costs avoided: \$30,000 to \$75,000.*

"I can do the same number now and do them faster, and we are thinking about doing full backups more frequently. With Whitewater, more backup work gets done in less time."

With Whitewater appliances, backups run faster, so full backups can be done more frequently. And it is now possible to perform multiple backups simultaneously. Rountree is evaluating how the company's backup practices will change now that they have Whitewater. "We used to limit how many backups we did at one time," he says. "I can do the same number now and do them faster, and we are thinking about doing full backups more frequently. With Whitewater, more backup work gets done in less time."

For Rountree, who knows that the ultimate purpose of backed-up data is to be able to restore it when necessary, Whitewater gives the company better protection for its data. "A huge advantage of Whitewater is how quickly we can restore data now," he says. "With tape, you can lose part of a week's data depending on when you last backed up, which means you would lose the most

recent information. With Whitewater, a full restore can be done within an hour or two, and can be done from any location". (*Value: the cost of lost data.*)

Whitewater Cloud Storage Gateway Overview

Riverbed Whitewater cloud storage gateways allow organizations to eliminate tape, reduce costs, streamline administration and improve disaster recovery readiness by leveraging the scalability, flexibility and pay-for-use pricing of cloud storage.

Disaster recovery capability is greatly improved by secure off-site storage within the public cloud without the expense of redundant hardware, power, cooling and hosting costs. Backup data stored in the cloud is available anywhere with an internet connection to provide fast, flexible recovery should the unexpected occur.

Whitewater safeguards data during transmission and at rest using secure socket layer V3 (SSLv3) technology and strong AES 256 bit encryption. This dual layer of security ensures that sensitive data is protected end-to-end.

Leveraging Riverbed's industry leading deduplication, compression and optimization technologies, Whitewater gateways shrink data set sizes by 10 to 30x on average, substantially reducing cloud storage costs while accelerating data transfers. This massive reduction in data sets allows a single gateway to replace more than a petabyte of data stored on tape.

Whitewater gateways secure, optimize and accelerate backup and recovery to the cloud without changes to the existing infrastructure. Whitewater acts as a broker to connect any of the following data protection applications to the following cloud storage providers:

Data Protection Applications

CA ARCserve, Commvault Simpana, EMC Networker, HP DataProtector, IBM Tivoli Storage Manager, Quest vRanger, Symantec NetBackup and Backup Exec, Veeam Backup and Replication

Cloud Storage Providers

Amazon Web Services S3, AT&T Synaptic Storage as a Service, Microsoft Windows Azure, Nirvanix Cloud Storage Network, Rackspace Cloud Files, clouds based on EMC Atmos or OpenStack (Swift) object storage.

Conclusion

The public cloud can offer a safe, attractive and affordable solution for the storage of backup data sets. Concerns around security, transmission speed, and accessibility exist but each can be overcome by the application of technologies that exist today from industry leading vendors such as Riverbed. From interviews with Riverbed Whitewater cloud storage gateway users, key benefits such as reduced system management requirements, avoidance of new tape and storage array purchases and the ability to safely store data off site provide a substantial return on investment. Cost savings continue to accrue by large reductions in data set size by use of deduplication and the resulting reduction in cloud storage costs. Customers leveraging Whitewater gateways are pleased that they have moved their backup data sets from tape technologies to the public cloud, for from their experience, they can no longer afford not to.

About Riverbed

Riverbed delivers performance for the globally connected enterprise. With Riverbed, enterprises can successfully and intelligently implement strategic initiatives such as virtualization, consolidation, cloud computing, and disaster recovery without fear of compromising performance. By giving enterprises the platform they need to understand, optimize and consolidate their IT, Riverbed helps enterprises to build a fast, fluid and dynamic IT architecture that aligns with the business needs of the organization. Additional information about Riverbed (NASDAQ: RVBD) is available at www.riverbed.com.



Riverbed Technology, Inc.
199 Fremont Street
San Francisco, CA 94105
Tel: (415) 247-8800
www.riverbed.com

Riverbed Technology Ltd.
One Thames Valley
Wokingham Road, Level 2
Bracknell. RG42 1NG
United Kingdom
Tel: +44 1344 401900

Riverbed Technology Pte. Ltd.
391A Orchard Road #22-06/10
Ngee Ann City Tower A
Singapore 238873
Tel: +65 6508-7400

Riverbed Technology K.K.
Shiba-Koen Plaza Building 9F
3-6-9, Shiba, Minato-ku
Tokyo, Japan 105-0014
Tel: +81 3 5419 1990