arcserve®

WHITE PAPER

Tape Backup Isn't Dead

As demand for data continues to skyrocket, high capacity storage mediums like magnetic tape are seeing a resurgence in value to IT (and Google™)

Summary

Tape backup isn't dead – far from it. For a 50 year old technology, magnetic tape storage is looking quite lively. As the backup media of choice for generations, tape looked like it would take a backseat to more popular disk- and cloud-based alternatives. But innovations in media density, performance, and reliability, coupled with a broader realization that data durability trumps just about every other form of business continuity problems, have earned tape a place in the hybrid data center. Real-world case studies at high-profile tech titans like Google have only helped to reinforce the belief that tape technology has a critical role to play in protecting corporate data assets.

For most organizations, the realization that data is critical to the ongoing functioning of every organization is not new, but the frequently short lived nature of today's data protection strategies often is. Data protection architectures increasingly rely on disk-based technologies, like VM snapshots, array-based replication, and cloud backups. High availability configurations using these approaches satisfy user demand for the fastest possible recovery time objective (RTO) with minimum disruption, but they also have weaknesses.

Tape backup is not a panacea, but as a tried and tested technology, it has a pivotal role to play in ensuring the durability of an organization's data. The reality of IT today is that no single architecture solves all problems. With combined physical, virtual, on-premises, colo, and the cloud playing a role in many IT infrastructure portfolios, a similar hybrid approach to data protection is needed to ensure the long-term health of corporate data assets.

The potential downside to a multi-faceted approach to data protection is the ongoing cost and complication of supporting a variety of vendor products. For small and mid-sized organizations whose IT resources are already stretched thin, this can be one backup product too far. Until now.

Table of Contents

Tape in the Data Center

Magnetic tape storage has been a feature of the data center since the very earliest days of computing. Reel-to-reel tapes that originally held a megabyte of information have evolved into today's high-capacity tape cartridges that can store multiple terabytes of data.

For today's "always-on" business applications with the most demanding RPOs and RTOs, high availability configurations provide the redundancy and swift failover needed to keep the organization up and running. However, a surprising number of IT systems continue to rely on conventional backup and recovery techniques that rely on tape or disk as a backup target.

Industry Standardization

One of the benefits of a technology with a long track record is industry standardization. Linear Tape-Open (LTO) and the common tape cartridge form-factor, Ultrium, have proven the most to be the most popular standards. The most recent iteration of the LTO standard, LTO-6, supports up to 2.5 TB of raw data per cartridge and transfer speeds of 160 MB per second. Tape-based data is more readily compressed than data on disk and vendors commonly quote two and three times the raw data capacity and transfer speeds for LTO-6.

The LTO standard is still evolving. LTO-8, anticipated in late 2017, will boost raw cartridge capacity to 12.8TB (32TB compressed) and transfer speeds to 472MB/sec. LTO-10, is anticipated to provide even more capacity, up to 48TB of raw uncompressed data, and transfer speeds over a terabyte per second.

Cost, Compatibility, and Durability

Industry standardization on the LTO specification has lead to widespread availability of cartridges and tape drive hardware from a broad range of manufacturers. This not only ensures a competitive marketplace that benefits cost-conscious IT buyers, but also offers compatibility with a broad range of third-party hardware and software solutions. Asking data protection vendors "Does it support tape?" is a given, and the answer is almost universally, "Yes."

Magnetic tape is an exceptionally durable medium, with manufacturers often quoting a 30-year lifespan. Bit Error Rate (BER), mean time between failure (MTBF) rate, and bit rot — the gradual decay of data stored on magnetic media—are all lower on tape when compared to disk. These attributes make tape the safest medium for long-term data storage.

Unlike spinning disks, data written to tape no longer consumes precious data center power resources. Combine these benefits with a modern tape library's exceptionally efficient use of data center floor space and the low dollar-per-terabyte cost of tape cartridges, and tape provides a highly cost-effective storage medium.

The Performance Choice

Counterintuitive as it may seem, tape can be the high-performance storage alternative. The Blue Waters supercomputer, operated by the National Center for Supercomputing Applications (NCSA) at the University of Illinois, is one of the fastest computers in the world and is tasked with a variety of computationally intensive applications, such as predicting the behavior of hurricanes, analyzing complex biological and engineered systems, and designing new materials. Blue Waters uses tape to store 380 petabytes of near-line data. While disk technology is superior for random access reads, tape often has the upper hand for sequential processing applications. Read/write rates of 61 GB/sec were one of the reasons NCSA chose tape for their supercomputer storage.

Portability

Tape cartridges are the ideal medium for applications that require data to be securely transported offsite. Cloud storage service providers have made in-roads as a friction-free alternative to tape for offsite data storage, however the cloud is not appropriate for all applications and all organizations.

The volume of data that needs to be transported during backup operations is often a prohibiting factor in data protection schemes that use cloud. It is also not unusual for regulatory stipulations to specify that archives be stored in the country of origin. This can be difficult to guarantee when data is hosted in a third-party service provider's data center somewhere in the cloud. In addition, business continuity best practices may rule out dependence on a third-party service provider as part of the critical path of the organization's disaster recovery plans.

UDP appliances also fit the gateway category through the native capabilities to replicate data to private and public cloud services, MSPs or the upcoming Arcserve recovery as a service option. This is not just an evolution of what a data protection appliance is, but a true generational leap forward, a revolution in how you will be able to deploy data protection in your organization.

Good Enough for Google

In February 2011, Google's Gmail service had a very public outage . A software bug introduced in a new release of Gmail storage software resulted in users receiving login errors, experiencing empty Gmail mailboxes, and other Google Application problems. Gmail stores multiple copies of user's messages in multiple data centers and on tape backups. Due to the nature of the failure, Google engineers had to restore user mailboxes from offline tape.

After reviewing recovery from the Gmail outage, Google's Site Reliability group highlighted a number of lessons that had been gained from the experience:

No data loss, ever
Data is crucial to the functioning of every organization. This means data availability must be 100%. Most companies can survive periodic downtime for applications, but surviving data loss is much more problematic.

Backups are not the goal
Backups are essential, but it's the restore that is important. Put as much effort into backups as is necessary to guarantee that data can be restored as and when needed.

Redundancy is critical
Things fail all the time. Plan for failures by building redundancy into the system.

Diversity in everything
Maintaining a single backup copy is not going to help you restore data if the backup is corrupt. Build diversity into backup and recovery systems so that it is possible to recover the specific components that failed.

Prove it
Backups are useless if the restore doesn't work. Testing the recovery process under real-world conditions is essential.

Google's use of tape in their backup and recovery process for Gmail surprised many, but it highlights the versatility of the medium. Organizations large and small can learn from Google's experience of recovering from the Gmail outage, and apply those lessons to their own data protection and business continuity plans.

Simplifying Data Protection in the Hybrid Data Center

Security and risk management, which encompass various classes of data protection, are now among the top initiatives driving IT spending. Over the past several years, much work has been done to modernize and optimize IT workflows, but data protection has lagged behind. Dramatic growth in the amount of data used by applications, the cost of licensing a diverse set of data protection tools, and the administrative overhead of managing these tools are all combining to make modernization and simplification of data protection infrastructures a top priority.

Even the most modest organizations today have a hybrid data center that complicates data protection. On-premise physical and virtual infrastructures, off-premise collocated servers, and new cloud-based services each have unique backup and recovery considerations. This variety frequently results in a similar variety of data protection strategies that can include: disk-, tape-, and cloud-based backup; array-, application-, and database-level replication; virtual machine and storage array snapshots, and high availability configurations. Add to this mix dedicated appliances for backup and deduplication, multiple hardware, software, and service vendors to provide support, combined with the siloed management of tools based on inconsistent criteria, and you have a recipe for duplication, confusion, and almost guaranteed human error. This comment scenario says it all: An IT administrator takes on a bare metal restore who then hands it off to a virtualization administrator who recovers the hypervisor, and then passes the buck to an application and database administrator for recovery of data. Clearly, this isn't the most ideal process.

Recovery Comes First

Attempts to simplify the infrastructure must start with the basics: the need to recover data from loss or corruption. As Google discovered in their review of the Gmail outage, backup strategies must be driven by recovery needs. RPO and RTO are key metrics in this endeavor as they define the business case for recoverability. IT administrators use RPO and RTO to determine appropriate mechanisms, medium, and methods for data protection. Simplification is then a process of determining the most efficient unified set of tools that meet these needs.

D2D2T and D2D2C2T (Don't worry, we'll explain in a minute)

Disk-based backup has marginalized magnetic tape as a first line backup destination in the data protection arsenal. By comparison, recovery from a disk backup delivers far superior RPO and RTO performance. However, as organizations like Google have found, RPO and RTO do not tell the whole story, and the economic and durability advantages of tape mean that it still has a place in the hybrid data center.

In their post-mortem of the Gmail outage, Google identified diversity and redundancy as two critical components of a successful data protection architecture. By incorporating magnetic tape as a tier in the overall data protection scheme, sources of backup data are diversified across multiple media. If a disk-based backup source becomes unrecoverable, a redundant copy on tape can be used in its place. The longevity, durability, and cost-effective nature of tape media makes it a great medium for failsafe copies that you hope never need to be used, but are there when necessary.

Among the wide variety of new backup infrastructure acronyms, disk-to-disk-to-tape (D2D2T) and disk-to-disk-to-cloud-to-tape (D2D2C2T) are becoming more popular, indicating a renewed interest in tape as a data protection medium. These configurations give administrators the speed and flexibility of disk when responding to applications with demanding RPO and RTO needs, and the added security of tape-based data, should all other sources fail.

Unified Data Protection and the Hybrid Data Center

Hybrid data center architectures place enough stress and strain on the IT infrastructure without adding the additional risk of incompatible data protection solutions. Yet this is exactly what organizations face when deploying a combination of legacy tools and newer single-purpose approaches to data protection. The lack of integration inevitably creates silos that add unwanted management, increase overhead costs in the infrastructure, and complicate recovery - effectively working against the objective of reducing risks to operational data.

Arcserve® Unified Data Protection (UDP) is a single, unified data protection solution that provides the flexibility to support a wide variety physical, virtual, and cloud IT platforms, a diverse range of application RPO and RTO requirements, and an array of backup media, including tape.

Arcserve UDP integrates disktodisk backup, tape backup, replication, high availability, and global deduplication in a highly scalable, single architecture. With agentless deep integration to support a variety of hypervisors and a modern task-based approach to administration, Arcserve UDP is able to automate complex repetitive tasks and provides all data protection and recovery from a single pane of glass. Assured Recovery™, a unique capability of Arcserve UDP, gives administrators complete confidence in their data protection architecture by automating risk-free testing of disaster recovery scenarios, without the need for end-user downtime.

Arcserve UDP and Magnetic Tape

With its enterprise-level functionality, Arcserve UDP gives small and mid-sized organizations a singe tool for their hybrid data centers. The solution offers bare metal recovery (BMR), local and remote standby, instant VM recovery, push-button failover and failback, and granular recovery from any backup media, including tape.

Unlike point solution approaches to data protection, Arcserve UDP's support for magnetic tape is not a last minute add-on. With a 25 year history in providing data protection for organizations and enterprises of all sizes, Arcserve has deep experience supporting tape, with tape-awareness built into the core of the technology. This is critical when integrating tape with modern data protection technologies like global deduplication, incremental backup, and VM snapshots.

Arcserve UDP 8000 Appliance Series

The Arcserve® UDP 8000 Appliance series provides a "set and forget" backup and recovery solution in a cost-effective, purpose-build data protection appliance. The hardware easily integrates into existing data protection schemes and offers cloud-native capabilities, unmatched ease of deployment and use, and output to magnetic tape.

Arcserve UDP 8000 Appliance supports global deduplication, multi-site replication, tape backup, and automated data recovery. Further, it seamlessly integrates with Arcserve UDP software to provide a distributed data protection architecture. Global deduplication technology actively reduces backup and bandwidth costs by combining patented, industry-leading deduplication, coupled with incremental backup technology to reduce disk, tape, and network capacity needs. Together, these capabilities increase operational agility and simplify disaster recovery.

Conclusion

As demand for data continues to skyrocket, high capacity storage mediums like magnetic tape are seeing a resurgence in value to IT. Modern approaches to data protection that enable fast failover and RPO and RTO in seconds demand a more responsive medium, but for many other business applications that require longevity and durability of backup media, tape still has a valid role to play in the architecture.

As Google found during the Gmail outage, tape-based data is important for your overall data protection architecture. Enabling tape to take its place as a data protection medium, however, requires software and hardware solutions capable of using tape. Today's intense focus on virtualization and the cloud has lead to many solutions focusing entirely on fast RPO and RTO, while ignoring tape altogether. This restricts the flexibility of administrators in designing sustainable and workable approaches to minimizing data loss and corruption.

Arcserve UDP is a modern approach to data protection built with 25 years of expertise in backup and recovery, high availability, and disaster recovery. It offers a unique, unified approach with all of the tools an IT administrator needs from a single web-based console, as well as the flexibility to use them however they see fit. If tape has a role to play in any data protection infrastructure, Arcserve is there to help.

arcserve®

For more information on Arcserve, **please visit arcserve.com**

[1]Google Apps Incident Report: Gmail Outage – February 27, 2011
https://static.googleusercontent.com/media/www.google.com/en//appsstatus/ir/nfed4uv2f8xby99.pdf