

# TechChannel

## How to Enhance Your Data Storage Strategy

Virtual tape best practices and tools for  
maximizing performance; how data temperature  
influences business storage decisions

SPONSORED BY



---

**3** Driving Business Storage Decisions With Data Temperature

---

**7** The Rise of Edge Infrastructure Storage

---

**9** Virtual Tape Performance Best Practices and Tools

---

**14** How Air-Gap Backup and Tape-Based Technology Protect Against Cybercrime

## Monitor and Measure Your Data Center



Running an efficient data center requires constant monitoring and measurement. This means understanding everything from storage best practices and tools for maximizing performance—to knowing your data's temperature in order to make accurate business decisions.

Data storage costs, for example, differ based on speed, availability and enterprise data services. Data temperature often determines the percentage of budget that gets devoted to storing different types of data. Once you create process maps based on the data's temperature, you can better manage both the data and the IT around the data.

Virtual tape performance is also dependent on several factors: total system configuration, Tape Volume Cache capacity, number of channels, read/write ratio and data characteristics, and more.

This TechChannel e-book features insight from IBM's Eric Herzog on hot, warm and cold data—and why data temperature should drive business storage decisions. Lastly, Kyndryl's John Shuman and Patrick Stanard explain which virtual tape best practices and tools you can leverage to maximize performance.

**Keelia Estrada Moeller, Senior Editor**

# Driving Business Storage Decisions With Data Temperature

IBM's Eric Herzog on hot, warm and cold data and why data temperature should drive business storage decisions

BY JENNIFER GOFORTH GREGORY

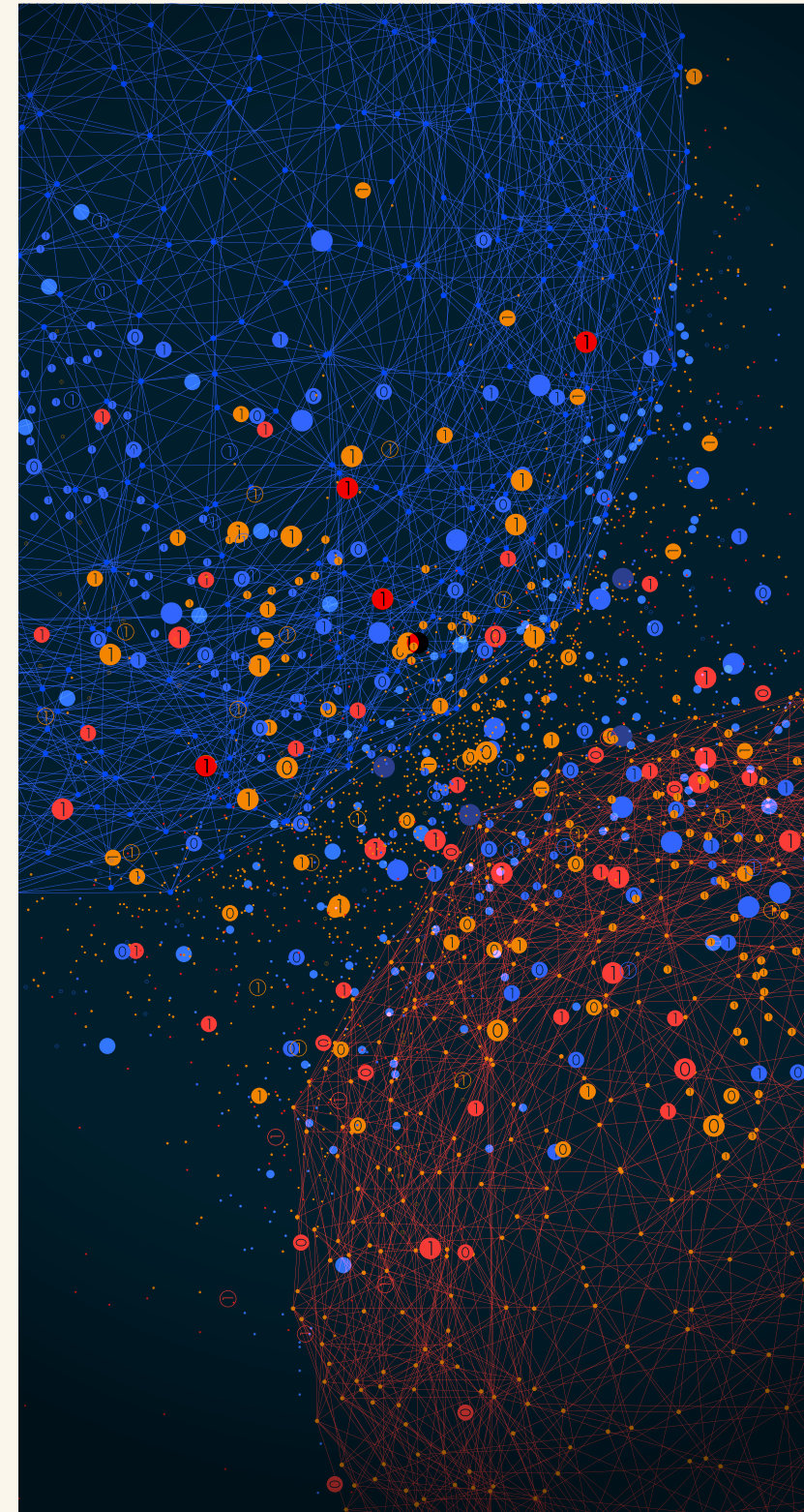
**Y**our company's most valuable asset, by far, is its data. Your patents, customer records, designs, logistics, supply chain and more are all examples of company data. But the types of data are not equal in their value (or temperature). Some data is critical to the daily operation of the company, while other data may never be needed again after collection. Because of the resources required to protect and manage data, organizations must know the temperature of their data to make accurate business decisions.

"The key is understanding the temperature of your data, which includes the value, how it fits from a corporate perspective, and any legal/compliance issues. Once you understand where your data is and what it does, you can then optimize your availability, cost, performance and cyberresilience around the storage that supports your data," says Eric Herzog, VP of Business Development and Evangelism and VP of Global Storage Channels, IBM Storage Division.

Herzog says many organizations overlook seemingly small pieces of data that can be much more important than expected, such as a SKU number on a pair of pink socks. If the SKU is incorrect, distribution may send the socks to a store in California instead of the New York City location that needs them, leaving the NYC store totally out of their customers' favorite pink socks. The SKU seems to be a small detail, but it's extremely valuable because it is the identification used to move the product through the supply chain and, eventually, results in revenue. And the results can be far-reaching. If these socks are not available when a customer wants to buy them, the company not only loses the initial sale of the socks—but also, potentially, a loyal customer.

## Data Temperature Drives Many Business Decisions

When determining data temperature, Herzog says organizations must consider two things: frequency of access and value to the company. A customer database in a retail business is accessed every 30 seconds or less and a customer database at a B2B company may only be accessed four or five times a day, but both databases are critical to their company's operation. If you only look at amount of access, it's easy to overlook potentially hot data, which leads to negative business consequences in terms of protection.





“By correctly identifying the temperature of your data, you can make sure valuable data is always preserved from threats—from earthquakes to ransomware. Next, you can map your IT process properly based on the heat of the data to most efficiently use manpower resources and processes. Because protection and processes are expensive, especially for Fortune 2000 companies, knowing the temperature helps you properly profile the data for CAPEX, OPEX and IT process management,” says Herzog.

Data storage costs differ based on speed, availability and enterprise data services, and the temperature often determines the percentage of budget that gets devoted to storing different types of data. Data that is hot and very active should be prioritized on the fastest and most highly available storage. All data, however, doesn't need this level of performance. When you understand your data's temperature, you can make the

## **Shouldn't You Trust the Pioneers In z/OS® Storage Management With Your Most Important IT Asset?**

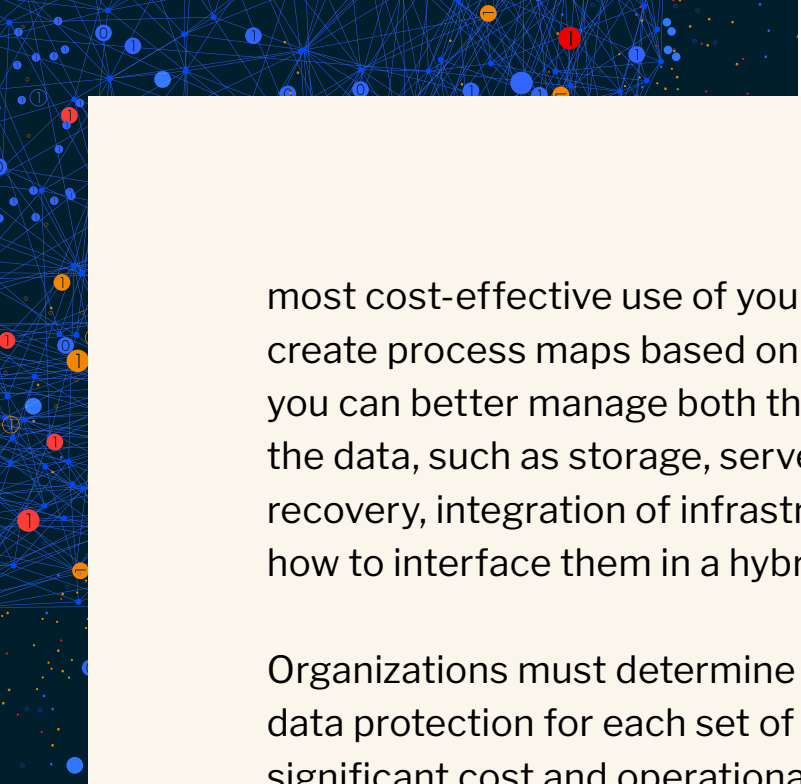
For nearly 40 years now, DTS Software has been recognized worldwide as the leader in enterprise storage management automation software. Our integrated suite of software products helps our customers to monitor, manage, and control their storage subsystems for optimal output.

**Take us for a complimentary test drive!**

Let us show you what we can do for you. DTS offers free analyses and recommendations of your ACS Routines, DTS Policy Rules, or Storage Allocation Automation with ACC Monarch. Contact us today to get started.

**DTS**  
**SOFTWARE, LLC**

919-833-8426 • [www.dtssoftware.com](http://www.dtssoftware.com)



most cost-effective use of your physical assets. Once you create process maps based on the data's temperature, you can better manage both the data and the IT around the data, such as storage, servers, networking, disaster recovery, integration of infrastructure with security and how to interface them in a hybrid cloud configuration.

Organizations must determine the appropriate level of data protection for each set of data, which can have significant cost and operational consequences. When considering data protection, Herzog says organizations must determine how often each data set needs to be backed up and what tier backup each data set falls into. Performing automated backups that meet your business needs starts by understanding the characters of each data set.

## Using Technology to Manage Storage Based on Temperature

In addition to backing up data, data protection means protecting data from cyberattacks, such as malware, ransomware and data breaches. Herzog says that once you know the temperature of your data, you can decide if it's a fit for IBM Safeguarded Copy, which allows you to rapidly recover data after an attack. Because you need additional storage for this technology, he says most organizations

don't need to do Safeguarded Copy on the entire data set, but instead only make frequent copies of hot and critical data.

The temperature of data often changes, as do instances of data with different temperatures being stored together. Herzog says that technology such as the AI-based tool Easy Tier automatically determines the temperature of the data and then moves it seamlessly to appropriate storage based on the temperature. For example, certain financial data may only be hot at the end and beginning of a quarter. When the data cools down, the system moves it to a lower-performing tier and then moves it back to a higher-performing tier when its temperature increases. This process optimizes the underlying storage infrastructure to optimize your budget and cost space.

## Creating a Process for Regularly Checking Data Temperature

Because the importance of data changes and evolves over time, organizations must regularly review data temperature. For example, the data for many physical assets of a large company (chairs, desks, whiteboards, etc.) is going to be the same for many years. However, some items will be added, such as new technology. Other items will be removed—such as fax machines, which were



previously a staple of most offices and are now almost non-existent.

While it's easy to focus on the specific data items as they evolve, organizations must also keep their eye on their data strategy and how it relates to data temperature as their business needs evolve. Storage is critical for cybersecurity strategy, both in terms of saving and money and keeping critical data safe.

“Data storage is going to remain a chief part of organizations for the long-term future. By building the right data and cyberresilience into your storage strategy, you can ensure that your company is both protected and positioned to increase revenue,” says Herzog.

# The Rise of Edge Infrastructure Storage

BY EVELYN HOOVER

Storing edge data is one of the reasons IBM announced **IBM Spectrum Fusion** back in April 2021, which seamlessly spans edge, core and cloud. Designed to simplify data accessibility and availability, Spectrum Fusion is the first container-native solution designed on Red Hat OpenShift. The new offering fuses IBM's general parallel file system technology and its data protection software.

“IBM sees that containers are where the world is going,” says Eric Herzog, vice president, Business Development and Evangelism and Vice President of Global Storage Channels, IBM Storage Division.

It's common for workloads and applications to create multiple copies of the same data set—one for the edge users, one for the core users and one for the cloud users. Spectrum Fusion offers a streamlined solution with a single copy of data.

“By having a single copy, we give customers much better CapEx and OpEx, because they only have one copy,” Herzog says. “If you’ve got 20 PB and you make four copies of it, you have to buy 80 PB. You have to manage 80 PB and all of the aspects of OpEx, like power, cooling, rack space.”

Spectrum Fusion wasn’t the only announcement Big Blue made this past April. Updates to the **IBM Elastic Storage System** (ESS) family were also unveiled. The revamped ESS 5000 supports 10% more density with a total capacity of 15.2 PB. The new ESS 3200 offers double the read performance of its predecessor, the ESS 3000. The all-flash systems are a strong play for AI, analytics and big data.

The ESS updates are important for a few reasons, Herzog says. First, the new ESS family automatically joins the existing file system in the global native space, saving time and manpower. Second, the performance and capacity scales linearly, which helps clients save on CapEx.

Herzog views these latest storage announcements as testament to IBM’s leadership in the container-native storage, which he says will be important to clients interested in AI, big data and analytics.

[Learn more about IBM storage advancements in the full article](#)



# Virtual Tape Performance Best Practices and Tools

Running an efficient data center requires constant monitoring and measurement. Learn virtual tape best practices and tools you can leverage to maximize performance

BY JOHN SHUMAN AND PATRICK STANARD

**T**homas Monson once said, “When performance is measured, performance improves. When performance is measured and reported, the rate of improvement accelerates.” Tape performance is one of the myriads of topics that is reported on for data centers. Reporting always leads to better performance as changes and tuning can occur.

## Virtual Tape Performance Factors

Virtual tape performance, for instance, depends on several factors, such as total system configuration,

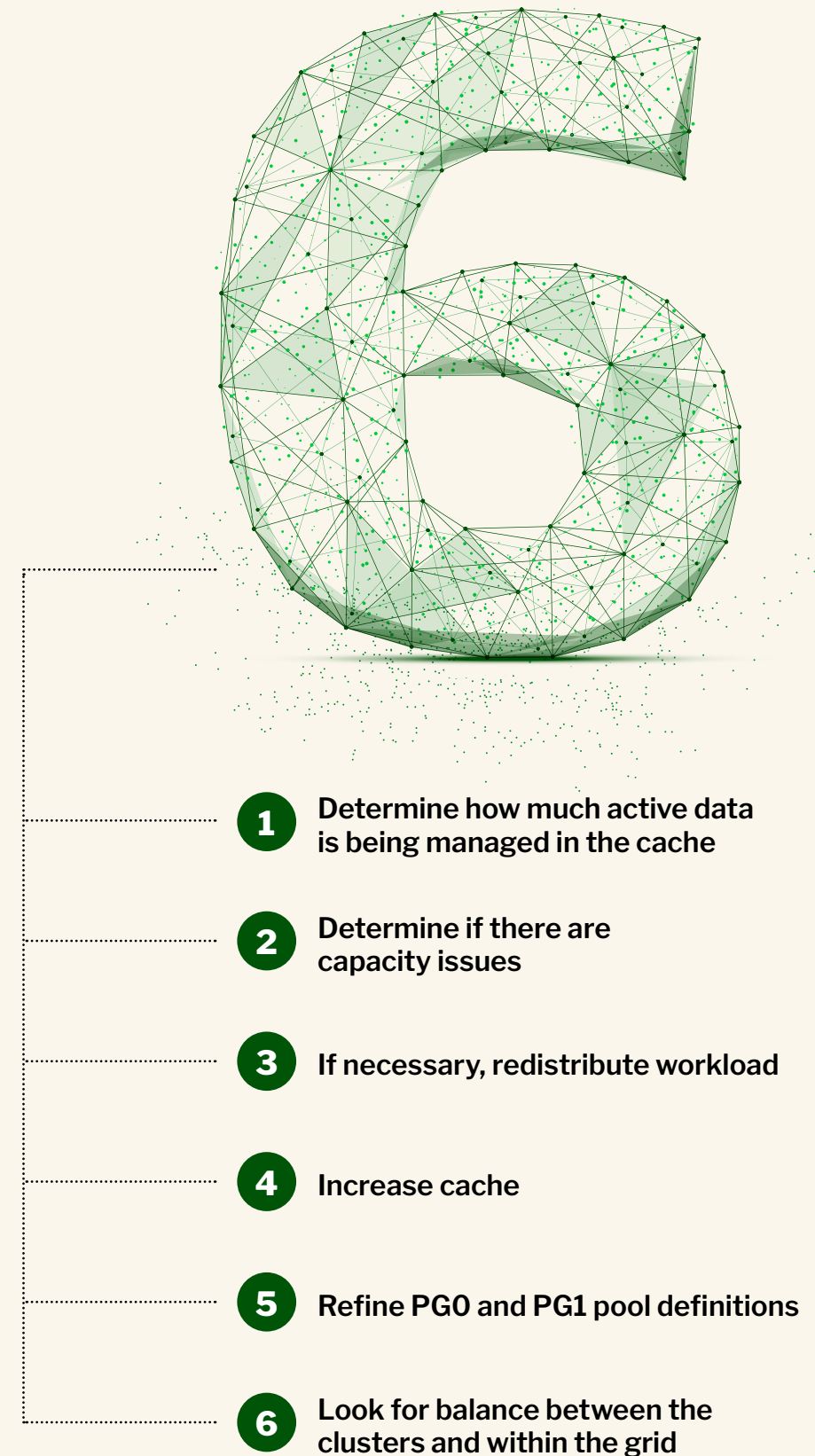
Tape Volume Cache capacity, the number of physical tape drives that are available to the TS7700, the number of channels, the read/write ratio and data characteristics, such as block size and the mount patterns.

In general, tape performance and capacity planning tools rely on Bulk Volume Information Retrieval (BVIR) data about local volumes in the TS7700. This includes volume status, cache content, physical to logical volume mapping, physical media pools, physical volume status and copy audit.

# Performance Best Practices

The IBM Washington Systems Center recommends TS7700 performance best practices that approach analysis from the TS7700 individual cluster level as well as the grid level. Start with 90-day trending to determine the general trends and then drill down on details at the weekly or daily level if necessary. Here are some performance best practices you should follow:

- 1. Determine how much active data is being managed in the cache** and optional external physical tape libraries (i.e., TS4500, back-end tape). This includes binary gigabytes (GiB) transferred and active logical mount times. If there are any significant or unexpected spikes, then move to the weekly and daily data.
- 2.** If the subsystem has back-end tape, then it's necessary to **determine if there are capacity issues** that result in poor performance. Report on key indicators like recall mount time, mount hit percentage, time-in-cache and physical mounts per day. If recall mount time is too high then maybe there are too many simultaneous requests, or the cache hit ratio is low (<80%).
- 3. If necessary, redistribute workload** to reduce mount requests
- 4. Increase cache** to allow the data to remain in residence longer



**5. Refine PG0 and PG1 pool definitions**—as data that's in PG0 moves off cache to tape quickly, while data sent to the PG1 pool is more likely to be recalled and is designed to keep the data in cache longer. Make sure data is being sent to the correct pool.

**6. Look for balance between the clusters and within the grid.** One cluster may have more virtual devices active and flows between the clusters can be tuned. Note that the TS7700 subsystem includes a series of self-regulating mechanisms that optimize resource usage. Limited resources such as CPU, cache bandwidth, cache size, host channel bandwidth, grid network bandwidth and physical drives need to be shared by all tasks within the system. The TS7700 subsystem throttles these resources when reaching its limits. For example, the deferred copy throttle regulates outgoing deferred copies to other clusters to maintain host throughput.

The incoming copy queue represents the amount of data waiting to be copied to a cluster. Data written to a cluster in one location can be copied to other clusters in a grid to achieve uninterrupted data access. Copies can be directed to specific clusters and the frequency of copy operations occurring. Each cluster maintains its own list of copies to acquire, then satisfies that list by requesting copies from other clusters in the grid according to queue priority. TS7700 prioritizes tasks in the order of immediate copies, recalls, copy export, host I/O, reclaims, pre-migration and lastly deferred copies.

## Improving Performance With Compression

One of the quickest ways to improve performance is to reduce the amount of data being copied through compression. Prior to the TS7700 release 4.1.2, compression was achieved through the default compression built into the FICON adapters connecting the TS7700.

Release 4.1.2 introduced LZ4 and ZSTD compression on top of the already available FICON compression. Each of these compression options have different performance characteristics versus effectiveness. LZ4 delivers moderate improvements in compression while consuming a small amount of TS7700 processing cycles. ZSTD, on the other hand, delivers greater improvements in compression storage savings while consuming more server processing power.

IBM ran tests using 18 separate workloads to compare the compression ratios of the new compression options with FICON. According to IBM, the results were as follows:

- LZ4 had an average increase in compression ratio of +32% compared to FICON
- ZSTD had an average increase in compression ratio of +105% compared to FICON



This was a benchmark test; customer results may vary, and all workloads will have different results.

Compression becomes a powerful tool for the reduction of the Recovery Point Objective time. As an example, consider an application that writes 300 GB of data. The host has a compression rate of about 3:1 and compresses the 300 Gb to 100 GB. ZSTD doubles the compression to 6:1 and now the data being sent over the network is only 50 GB, which could dramatically reduce RPO.

## Tape Tools

Tape tools like IBM's VEHSTATS and **IntelliMagic's Vision**, turn this data into spreadsheets, charts, and graphics. VEHSTATS is designed to produce reports based on statistics generated as the result of BVIR requests. The **IBM Techdocs public URL** points to the location where the tools can be downloaded.



## TechChannel Webinars

### Free, Expert-Led Education; Live or On-Demand

Want to keep up on the latest IT trends? Wondering what your peers are doing? Need a solution for a business challenge? With many in-person conferences and expos on hold, there's no better way than complimentary, expert-led TechChannel Webinars to get the latest on all things tech. Attend now and interact with the technologists; ask the questions that matter most to you.

View all live and  
on-demand webinars



**Running an efficient data center requires constant monitoring and measurement of all assets. There are many levers that influence efficiency in a tape subsystem, and the environment continues to grow.**

Other related documentation and tools can be found [here](#). Start by reading the Updates.txt files for changes. A document that describes each field in detail, the Vehstats Decoder v2.4, is available as well as the BVIRHSTx collection and the VEHSTxx reporting tools.

## **Constant Monitoring and Measurement Is Key**

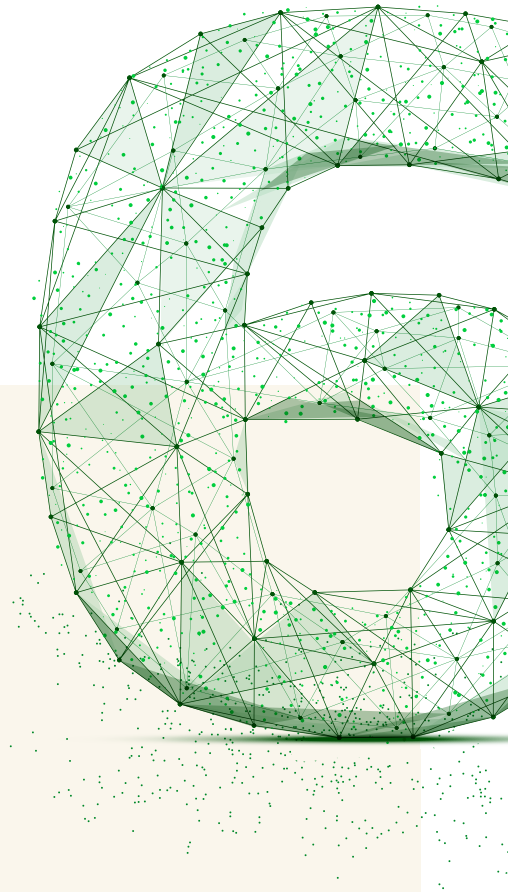
Running an efficient data center requires constant monitoring and measurement of all assets. There are many levers that influence efficiency in a tape subsystem, and the environment continues to grow. For

instance, TS7700 started supporting object storage with Transparent Cloud Tiering in TS7700 release 4.2. With this level of complexity, lack of measurement quickly leads to increased cost. Like other peripherals to the mainframe, performance increases translate directly to the bottom line.

---

**JOHN SHUMAN** is the executive chief mainframe architect for Kyndryl.

**PATRICK STANARD** is the executive U.S. chief mainframe architect for Kyndryl.



# How Air-Gap Backup and Tape-Based Technology Protect Against Cybercrime

BY JOHN SHUMAN AND PATRICK STANARD

**T**oday, many corporations and organizations employ a backup and recovery strategy known as 3-2-1. This has been around for a long time and is a very effective approach. When the methodology of air gap is added to this backup strategy, a copy of that data is not remotely accessible thus creating the concept of air gapping.

As old school as magnetic tape might be, it's really one of the preferred methods as far as one of the copies for the 3-2-1 strategy because it's reliable. Studies have shown that magnetic tape can last for up to 30 years where hard drives on the average are roughly five years. Other reasons are scalability, cost effectiveness and longevity. IBM is a leading tape storage vendor in the industry and has over 2x the market capture of any vendor. The air-gap backup topic has brought tape-based technology back into focus as a solid solution to protect against cybercrime. Tape is all about "offline" storage.

The air-gap solution is created when a tape cartridge is secured in offline storage as this is a medium that is not connected to the internet or network in any way. A hacker cannot corrupt or steal what they cannot access. Thanks to longstanding tape technologies, the capabilities of software-defined storage give worldwide enterprises solid tools to define against cybercrime.

The overall objective of air-gap backup is to expand the scope of data protection to ensure that both cybersecurity and business continuity are protected. Air-gap backup should be a part of any high availability and disaster recovery solution. Ideally, air-gap backup provides immutability, isolation as well as granularity of the data protected.

[Learn more about air-gap backup solutions in the full article](#)

This e-book was published by



953 Westgate Drive, Suite 107, St. Paul, MN 55114 // (612) 339-7571

staff list

**Senior Content Strategist:** Claire O'Leary

**Senior Editor:** Keelia Estrada Moeller

**Art Director:** Kelsey Hanscom

**Project Manager:** Noelle Heaslip

**Audience Development Director:** Linda Holm

**Publishing Director:** Mari Adamson-Bray  
(612) 336-9241 // mbray@techchannel.com

**Account Executive:** Darryl Rowell  
(612) 313-1781 // drowell@techchannel.com

© Copyright 2021 by MSPC, a division of MSP Communications. This e-book could contain technical inaccuracies or typographical errors. Also, illustrations shown herein may show prototype equipment. Your system configuration may vary slightly. This e-book may contain small programs that are furnished by MSPC as simple examples to provide an illustration. These examples have not been tested under all conditions. MSPC, therefore, cannot guarantee or imply reliability, serviceability or function of these programs. All programs contained herein are provided to you “AS IS.” IMPLIED WARRANTIES OF MERCHANTABILITY, NON-INFRINGEMENT AND FITNESS FOR A PARTICULAR PURPOSE ARE EXPRESSLY DISCLAIMED.

All customer examples cited represent the results achieved by some customers. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

The articles in this e-book represent the views of the authors and are not necessarily those of MSPC or TechChannel.



# TechChannel

TechChannel.com is home to a variety of content to help you get started on your security journey.

[Learn more about storage](#)

SPONSORED BY

