WHITE PAPER

The Relevancy of Archiving for Business Transformation

By Hitachi Vantara

April 2018

Contents

Executive Summary

The center of the digital universe is data access. From a business perspective, leaving no data behind is critical to moving forward. Whether looking to develop solutions, programs, markets or revenues, your organization's ability to innovate will largely depend on data.

Leadership demands data-driven decision-making to deliver better business outcomes and assurances that the organization can meet regulatory moving targets. Lines of business want to harness customer data to target services, shorten go-to-market cycles and generate new revenue streams. End users need to access files anytime, anywhere and from any device to stay productive. IT is tasked with making it all happen, amid perpetual data growth, dynamic workloads and dwindling budgets.

Massive data growth, mostly file-based and unstructured, can be difficult to capture, store and access in all the ways that matter. Many organizations have held a "backup everything forever" mentality, which often leads to unknowable amounts of inactive or dark data stored across siloed technology stacks, including high-performance Tier 1 storage. In this era of all-the-time digital access, no one can afford to wait for data to be located and mined or deal with complex backup and recovery gyrations. Yet, data must be readily available and cost-efficient.

At Hitachi Vantara, we recommend a modern data archival and analysis approach to bring dark data to light or keep it from going dark in the first place. Having a simple, smart and agile archiving strategy means capitalizing on digital transformation benefits and cost efficiencies to achieve customer, compliance and revenue goals. It means reinvigorating your data to optimize business value. We can help you leave no good data behind.

Introduction

From internet of things (IoT) and cloud-based dynamics to digital transformation initiatives, data access is paramount to meeting your business challenges. Data continues to grow unabated and is not always stored or archived in the ways and places to support anywhere, anytime access and to actionize insights. We examine what is really at stake and how to strategically archive data to help you transform the business.

What's at Stake: Your Data

We already know data growth is a tremendous challenge. So is how you store it.

Do you have more than 50TB of file data scattered across your IT enterprise? Are your end users all over the map, working remotely from multiple bring-your-own-device (BYOD) and computing platforms and storing unprotected file data on laptops, tablets, phones and other user devices? Is your staff inundated with self-service and help desk requests for data recovery or discovery from a plethora of data sources: from email, web apps, cloud services and IoT devices? If your answer is yes, you are not alone.

Your organization's challenges may involve modernizing IT, adding new services or meeting expanded regulatory requirements. You may be focused on improving data availability, driving down storage costs, forecasting massive incoming data growth or repurposing inactive content. Whether you face existing IT issues or new business goals, time is not on your side.

Increasingly, departments want to do more with more: initiate more customer engagement activities, more innovative programs, more services. All demand more value from existing, expanding data.

The daily onslaught of new content and platforms and their variety, volume, velocity and complexity, is rapidly outpacing digital storage resources and budgets. Approximately 80% of new data is unstructured, usually sits on primary NAS storage and goes inactive or "dark" after 90 days of being parked there. At the other end of the spectrum, ad hoc or inadequate archiving can wreak havoc on IT agility, performance, utilization efficiencies and costs. At the same time, it can hinder opportunities to glean value from data or meet governance requirements.

Prevailing Governance

Regulations permeate every industry, at a time when data growth is unparalleled in quantity and complexity. Lately, we have seen paradigm shifts in the regulatory terrain, from new interpretations of the Privacy Act of 1974 in the U.S., revelations of widespread government surveillance, and the demise of Safe Harbor to the launch of EU GDPR (European Union General Data Protection Regulations) and myriad industry-specific governance requirements.

Governance and regulatory-related litigation have quickly saturated electronic information. For global companies or U.S. organizations looking to expand the customer base into global markets, game-changing rules, such as GDPR, will enforce significant privacy laws, regulatory fines and noncompliance litigation. The costs could pale in comparison to potential loss of consumer trust if there is wrongful acquisition or misuse of personal data.

Traditional NAS

The digital economy has reached a new threshold for enabling underused, old or forgotten data to be mined, repurposed and actioned for business. The promise for revenue and market share based on data-driven insights is staggering.

Most of the data pouring in, however, funnels to high-performance Tier 1 storage, which often includes multiple NAS devices. The problem with so much data growth is that you have to scale the NAS environment faster than expected, but NAS was not designed for petabyte scale and beyond. Maintaining traditional infrastructure requires labor-intensive management of file servers and NAS appliances for protection, performance, migrations and capacity planning.

If, like many organizations, you have inactive data continuing to live on Tier 1 storage, you pay the price with:

- Increased sprawl based on numerous application storage silos.
- High administration and protection costs. The industry average is 80TB per FTE.
- Greater operating expenditures (opex), including power and cooling, data center footprint, technology refreshes.
- Storage inefficiency. Note that compression and single-instancing save space but degrade performance.

These challenges and trepidations hold companies back from reaching innovation goals and quickly add up, literally, to becoming the dominant cost categories within the total IT financial picture.

A Modern Archive: New Relevance for Stored Data

Archiving is so much more than your father's data retention "go-to." It's no longer that place to just park cold data. A modern renaissance of archiving is very much front and center in the wake of ubiquitous data growth, regulation and cybercrime. There is an urgency to it. On the spectrum from overwhelming to impossible, data accessibility can make or break how you succeed or fail to achieve these business-critical issues.

The imperative is clear: Transform the business with data or risk everything. If you want to protect, grow and innovate the business, then you must transform your data storage, too.

So, what if you could exploit data fully for greater competitive advantage, revenue streams and business continuity while increasing IT efficiencies, lowering costs and meeting the dizzying number of retention requirements? What if you could store data from multiple application workloads on a singular storage platform and then use analytic tools to understand that data up front and generate business intelligence and insights?

Future Proof for Compliance and Regulatory Perpetuity

What if your future looked like this: nimble visibility, governance and control of data that you can securely move to, from and between on-premises and cloud environments? This would be accomplished using automated, cost-efficient and intelligent policies, automation based on metadata, workflows and user-defined policies, which support a proactive, mobile, compliant, data-driven business environment?

Now, perhaps more than ever, IT must exact precision with regards to retention, discovery, monitoring, deletion and perpetuity of data. Building a secure and breathable governance and compliance strategy requires efficacy, control, visibility and agility throughout the data life cycle.

A single, unified data platform can help you balance policy-based data retention and disposal objectives with sustainable availability of trusted data for authorized users at the right time.

Transform Data Into Actionable Business Value

Solving file-based problems requires modern, flexible architecture capable of more than storing and archiving data. Status quo file services do not scale seamlessly and are not economically feasible or technologically sustainable using basic functionality of legacy NAS platforms. Coalescing that data stashed across Microsoft Windows, NFS and vintage NAS systems, across remote and branch offices, and on myriad BYOD devices into a centralized, usable repository is an excellent way to ignite potential.

The C-suite wants to know that files services are capable of meeting business needs while whittling away complexity, cost and risk. Lines of business want to be assured that all data can be curated to help boost new services and revenues. At the end of every day, the goal is to make every bit of your organization's data consistently insightful, actionable and auditable.

An Intelligent Archiving Solution

In today's dynamic digital enterprise, data access is almost always business-critical. You want to keep data active and at the center of the business, for both traditional and transformational objectives. Intelligent archiving is a method for modernizing, activating and innovating throughout the data life cycle while reducing overall storage costs and complexities. Think: online access of all of your data, all of the time.

The Solution

The foundation for intelligent archiving is a scalable, smarter and secure data platform. Hitachi Content Platform (HCP) portfolio is exactly that – and so much more.

Backup and recovery operations mandate more simplicity, less risk. No problem. You want to transform data into the valuable assets your organization demands. We can help you get there right away.

IDC expects the overall File and Object Based Storage (FOBS) market to grow at 5.8% CAGR for 2016–2021 — higher than the enterprise storage systems market, which is growing at 2.7% CAGR for the same period. IDC expects the FOBS market to reach $28.3 billion in revenue in 2021.

Source: IDC – Worldwide File- and Object-Based Storage Forecast, 2017–2021, September 2017

HCP provides a breadth of industry-standard protocols, through which archiving applications can automatically move data residing on any tier of storage to a fully integrated, secure and highly available archive platform. With agentless data management and high-density architecture, HCP acts as an archive data orchestrator and delivers remarkably low total cost of ownership. Our intelligent archiving platform eliminates storage silos and enables you to consolidate data on a single platform, so you can rapidly repurpose data and drive efficiencies across the whole storage infrastructure. HCP supports any size data environment, from 4TB to 4EB+ of data and billions of objects and files.
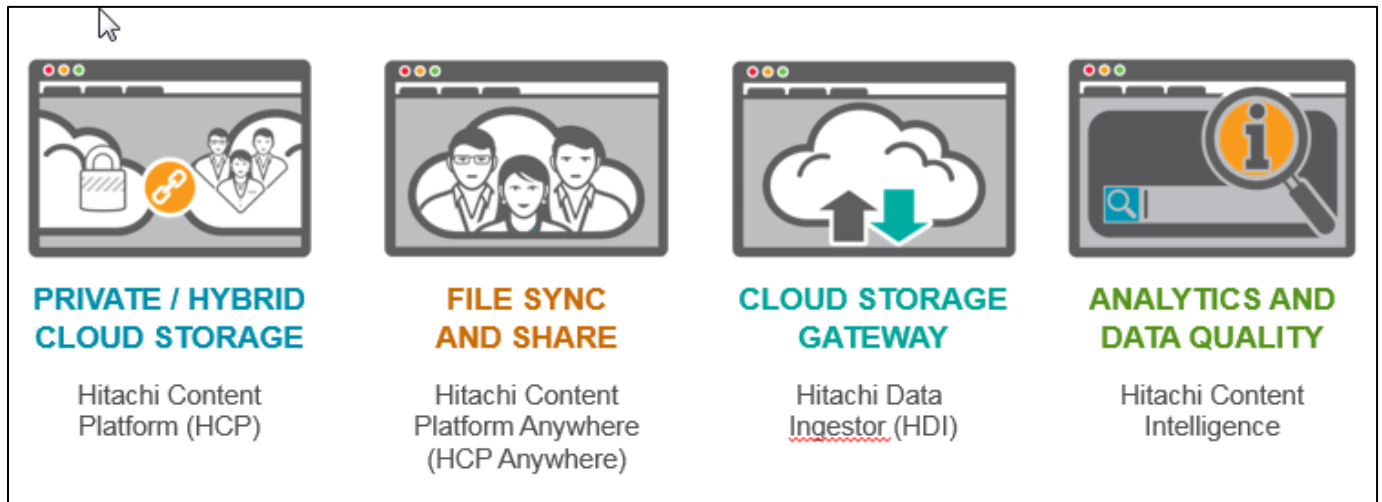
The Portfolio

The HCP portfolio is a software-defined, cloud-ready platform based on object storage architecture. With industry-leading capacity and performance, this portfolio connects and centralizes data generated by people and things to deliver:

- Data governance and control.
- Data mobility and access.
- Data analytics and insights.
- Data infrastructure for timeless agility.

Our tightly integrated HCP portfolio is specifically designed to:

- Modernize existing tools and infrastructure to shrink costs, boost efficiency and improve operations.
- Activate data with policy-driven attributes and metadata, which provides context for automation, exploration, mobility and governance.
- Innovate with analytic tools that help you find, qualify and cleanse data to derive actionable insights across all your data.

Figure 1. Hitachi Content Platform Portfolio

PRIVATE / HYBRID CLOUD STORAGE	FILE SYNC AND SHARE	CLOUD STORAGE GATEWAY	ANALYTICS AND DATA QUALITY
Hitachi Content Platform (HCP)	Hitachi Content Platform Anywhere (HCP Anywhere)	Hitachi Data Ingestor (HDI)	Hitachi Content Intelligence

The Strategy

Now is the time for a strategic plan that harmonizes ongoing data growth and protects your investments: no rip and replaced required. Being smart starts with intelligent storage tiering. As you know, tiered storage infrastructure can increase efficiency and automate file management. Typically, data is categorized via tiers:

- Tier 0 for ultra-high performance.
- Tier 1 for active, mission-critical or revenue-generating data.
- Tier 2 for backup and recovery, reference and sensitive information.
- Tier 3 for fixed content, archives and long-term retention.

Modern archiving starts with data movement. We recommend offloading less-accessed data from your Tier 1 NAS storage to Tier 2 object storage using HCP or Hitachi Data Ingestor. Placeholders or "stubs" stand in for the files you move from Tier 1 to HCP. Whenever an end-user clicks on a stub, the file is readily retrieved from HCP and opened. HCP uses backup-less functionality, so the only files requiring traditional backup are those still in your NAS environment. With less data being backed up, your backups will complete faster. By moving inactive data from the backup stream to more suitable, secure storage, you win back time.

This strategy enables you to:

- Optimize NAS infrastructure while reducing future NAS device purchases.
- Streamline backup operations, shorten backup and recovery operations.
- Provide faster access to business information.
- Increase efficiency and useful life of existing infrastructure.
- Reserve primary storage for frequently accessed, high-performance data.
- Lower costs for backup software, licensing and management resources.
- Improve server utilization, storage capacity and data duplication issues.

If you have cold data that is not expected to be accessed but must be retained, another component of your tiered strategy could be to move this content to public cloud. HCP is an adaptive hybrid cloud solution, which supports data placement to major public cloud storage services, including Amazon S2, Microsoft Azure and Google Cloud

Storage. HCP also maintains security by encrypting data prior to transferring it. You retain the encryption keys and are able to easily and securely move and control content to and between one or more public clouds as you need.

This strategy enables you to:

■ Automate data tiering and archiving using customized management policies.

■ Move to pubic cloud or on-premises storage using simple, user-defined policies.

■ Respond to immediate needs and scale on demand for capacity.

■ Manage more storage with lower costs and less resources.

> "The HCP portfolio of products turbocharges your hybrid cloud strategy with valuable capabilities, including centralized visibility and control, multiprotocol and mobile device support, integrated search and analytics, encryption at rest at no additional charge, and adaptive cloud tiering, which enables HCP to tier content to one or more public cloud providers."
>
> -- Brian Garrett, VP of ESG Lab[1]

A more transformative scenario is to modernize the entire storage environment by migrating off NAS completely. In this case, you would move all that file data to HCP object storage, making it a Tier 1 storage solution. We call this "Files Services Redefined." While providing an excellent foundation for streamlining data movement and capitalizing on storage cost efficiencies, this approach requires rethinking or moving away from NAS environment processes.

This strategy enables you to:

■ Replace expensive, limited NAS and file server technology with a single scalable, efficient platform.

■ Capitalize on HCP's built-in data automation, protection, governance and metadata-based search and analytics capabilities.

■ Improve productivity and efficiency of mobile users by enabling data access from any location with any device.

■ Add the ability to search and gain control of user data (corporate policy initiatives) for review and monitoring to lower costs, address compliance concerns and reduce corporate risk.

> "IDC expects that the archival market will continue to see demand from increasingly data-driven businesses that will continue to find ways to store data in a cost-effective manner — and at the same time reserve the option to run analytics on it at a moment's notice."
>
> — *Laura DuBois, Program VP,*
> *Storage and GRC Infrastructure*

The Difference

Our proven strategy and Hitachi Content Platform portfolio offer a solution like no other. We provide you with the technologies, tools and services to identify, analyze and exploit all of your data. The HCP intelligent archiving solution equips you with new data-centric capabilities to unclog growth challenges, storage inefficiencies and regulatory escalation.

■ **Unlimited capacity scale and stellar performance.** Scale seamlessly and deploy as software-defined storage or as an appliance. HCP is equally proficient at serving billions of small to large objects.

- **Extensive application ISV ecosystem.** Experience easy integration and interoperability of third-party independent software vendor (ISV) applications with HCP, including archiving applications, such as NICE, Splunk, FileNet and Veritas Enterprise Vault, which leverage HCP as a storage back end.

- **Multiprotocol and heterogeneous access.** Support legacy applications using NFS, CIFS, SMTP or WebDAV, and modern applications using HTTP/REST, S3, OpenStack Swift or WebDAV. Data ingested via any one protocol is fully accessible by every other protocol.

- **Compliance storage modes.** Satisfy regulatory requirements that require immutable, undeletable "write once, read many" (WORM) storage, guaranteed authenticity or proving chain of custody, and retention for a set period or forever.

- **Construct hybrid storage pools.** Use HCP adaptive cloud tiering functionality to automate data movement to, from and across your choice of leading public cloud service providers, including Amazon S3, Google Cloud Storage, Microsoft Azure, or any other S3-enabled cloud service.

- **Policy-based automation.** Define cradle-to-grave data management plans that govern an object's protection class, access speed and disposition policies.

- **Sophisticated metadata search and analytics.** Create and modify custom metadata at any time during an object's life cycle. Multiple authors can have separate sections of custom metadata. Use the HCP's built-in metadata query engine, or enable data categorization, natural language searches and data refinements using Hitachi Content Intelligence.

Read the Spin Master Case Study

DOWNLOAD

"What got us here will not get us there. We continue to evolve and now have a foundation that works, that we can layer upon without having to rip and redo. We have agility to deploy new initiatives quickly. The breadth and depth of Hitachi technologies and services are what get us out of fire-fighting mode and into enablement, so we can foster revenue generation. At the end of the day, it's really about putting big smiles on kids' faces."

— Pravine Balkaran

Global Head of IT, Spin Master

Moving Forward

For more information on how the HCP portfolio can meet your unique challenges, please visit www.HitachiVantara.com

Hitachi Vantara