Hitachi Data Systems File and Content Solutions Portfolio

Silverton Consulting, Inc.
StorInt™ Briefing

Normal

Today, file data is expanding without limit. SOme suggest that this data will grow 40X over the next decade, at which time ~80% of all company data wil be unstructured. Such unrestrained growth becomes an unmanageable burden un its own right. Further, regulatory compliance and/or E-discovery may necessitate even more rigorous guardianship over an organization's data.

Nevertheless, the principal advantage of file data is its ease of use. Therein lies the dilemma. How does IT maintain unstructured data's versatility while at the same time providing better management and control over this data throughout its lifetime wherever it lies in the enterprise?

The answer to this data conundrum must come from a portfolio of file and content solutions, which adds to traditional file storage infrastructure, capabilites for archive, compliance and search, in a tightly integrated set of offerings. The Htiachi Data Systems (HDS) file and content solutions product family includes both file and object storage solutions as well as advanced storage services that can be joined together to resolve such contrasting requirements.

Unstructured data in the enterprise

Enterprises are struggling today with an explosion of data. Equally troublesome to IT staff is that more and more of this data exists outside the data center, in potentially hundreds of remote offices, thousands of mobile devices and even cloud service providers. Gaining control over all this information has become an almost impossible task.

To support such unruly data growth, IT typically deploys additional storage in separate 'silos'across workloads or lines of business. This leads to storage infrastructure sprawl, increased management complexity and significant inefficiencies in utlilization.

On top of this, corporate backup window appear to be shrinking. The issues are multi-faceted but three are most prominent:1) aal of that growing mass of unstructured data represents corporate IP which must be backed up;2) much of that file data, possibly 70% is seldom changed after an initial flurry of activity and 3)

Many files are duplicated of other data. Hence, backup streams lengthen as file data proliferates, protecting the same data over and over again, leading to tighter backup window and increased costs.

In addition, many organizations have to operate under regulatory compliance, which mandates adherance to specific requitements for applicable data. For example, in the US, healthcare companies must keep medical records for 7 years or the life of the patient, European data protection acts require secure storage and access for personal information, even restricting where data can reside and similar regulatory environments exist in other regions. Regulations such as these just add to the burden of enterprise data.

However, compliance is not the only reason for data persistence. E-discovery requests can examine all data in a corporation looking for any shred of evidence germane to some litigation. Such activity could easily duplicate or place an expiration/ deletion hold on significant amounts of data just on same minor relevance to legal proceedings.

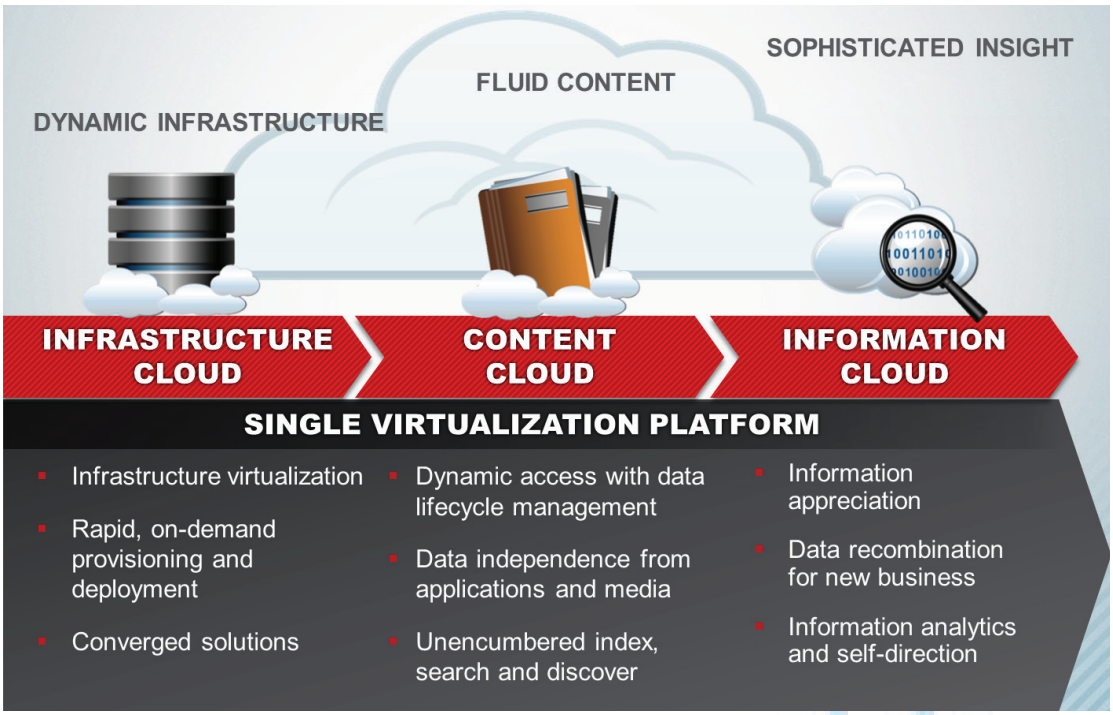
Big data analytics also have an enormous appetit for an organization's data, in whatever form it takes. Leaving this data scattered across silos complicates analysis and may result in segments of missing or misplaced data that could easily lead to incorrect analyses and flawed conclusions. To further complicate matters, traditional unstructured storage solutions aren't built to support advanced analytics which may require another copy of the data to be stored in the analytics layer, causing yet another silo of data adding to the already strained storage environment.

Gaining control over unstructued data

For these reasons, Hitachi has designed their integrated fie and content services portfolio to help enterprises better manage this burgeoning collection of data. It is the integration of NAS, object storage, indexing and search that sets HDS apart. The ability to move data among virualized storage technolgy in an automated manner improves organization of, and access to, content, reduces the cost and complexity of storing, preserving and prorecting data and prepares the storage environment for the coming of big data analytics. While HDS integrated solutions and services make it possible for IT to completely transform their unstructured data environment, for most data centers, this is a journey of multiple steps.

The first step is to start with an infrastructure cloud, a more flexible and responsive virtualzed storage environment to house all company data moving away from difficult to scale, traditional storage silos. The infrastructure cloud provides the base level of services that primes an organization to consolidate unstructured data and take initial steps to gain control over its growth and proliferation.

Hitachi Data Systems File and Content Services Portfolio



Adding object storage as well as index, search and discoverability of data across all applications or devices, takes IT organizations to the content cloud. With content cloud, customers can start to reduce management, storage and access costs while at the same time automating information lifecycle. Next, by making use of an object storage repository where content can be preserved and protected, IT staff can start to use system and custom metadata to apply retention and other automated policies to manage these data elements.

With dynamic infra tructure and fluid content in place, analytic tools can more easily mine- and discover trends to produce useful insight. This is the information cloud, where vital business insight can be produced from analyzing all of an enterprise's data. The information-cloud is not a single service or solution and there may be many intermediate steps along the way that can help data center staff to gradually realize its capabilities.

Gaining control over unstructued data

With dynamic infra tructure and fluid content in place, analytic tools can more easily mine- and discover trends to produce useful insight. This is the information cloud, where vital business insight can be produced from analyzing all of an enterprise's data. The information-cloud is not a single service or solution and there may be many intermediate steps along the way that can help data center staff to gradually realize its capabilities.

HDS offers integrated file storage, archive and search solutions that can help customers realize the infrastructure and content cloud of today while taking steps towards the information cloud of tomorrow. The solution is a combination of:

- **Hitachi Network Attached Storage (HNAS)** gateway provides file protocol services and uses Hitachi Unified Storage or Hitachi Virtual Storage Platform (VSP) for back end storage.

- **Hitachi Unified Storage (HUS)** supplies both block and file storage interfaces in one integrated system.
- **Hitachi Unified Storage VM (HUS VM)** provides an enterprise class block and file storage solution in a single, integrated system.
- **Hitachi Content Platform (HCP)** offers an intelligent object store to preserve, protect and access data and uses Hitachi Unified Storage, HUS VM or Virtual Storage Platform (VSP) for backend storage.
- **HitachiData Ingestor (HDI)** operates in conjunction with the HCP to offer a bottomess and backup free file serving solution for remote office/branch office environments.
- **Hitachi Data Discovery for Microsoft® SharePoint® (HDD-MS)** environments allows SharePoint data to be archived out of its collections database that offers a number of unique customer benefits.
- **Hitachi Data Discovery Suite (HDDS)** delivers federated search and indexing across a wide assortment of file storage.

HDS NAS platforms for file data consolidation

HDS offers three product groups specifically designed to help enterprises to consolidate unstructured data, namely HUS storage, HUS VM storage and HNAS gateways. All threeproducts use the same advanced storage features for file data, based on the BlueArc® SiliconFS, hardware accelerated file system. SiliconFS delivers modern capabilities with superior scalability and file performance.

These three systems offer a great solution to the storage proliferation that accompanies enterprise file growth. Most organizations purchase storage equipment that meets current and projected future needs but quickly runs out of capacity, performance or connectivity. This causes additional systems to be purchased, which rapidly multiplycreating anadministrative nightmare. By introducing HNAS, HUS or HUS VM storage into the environment, file data can be consolidated into fewer more capable systems, which reduce storage sprawl, provide simpler data management and as such, better control file growth.[1][2][3]

HDS NAS platforms for file data consolidation

1 Please see http://www.hds.com/products/storage-systems/hitachi-unified-storage-100-family.html for more information
2 Please see http://www.hds.com/products/storage-systems/hitachi-unified-storage-vm.html
3 Please see http://www.hds.com/products/file-and-content/network-attached-storage/ for more information

HCP is an intelligent object storage solution that can be standalone or connect to otherHUS, HUS VM or VSP storage. HCP can automatically apply lifecycle policy management to preserve, protect and access data throughout an enterprise.[4]

As such, HCP insures that its data is always preserved, safely and perpetually, within the content repository. Thus, once data is placed under HCP control, it never has to bebacked up again, reducing backup time And processing requirements as wellas freeing up storage capacity.

In addition, HCP intelligent object storage incorporates extensivemetadata that can be used in conjunction with its pol icy management to automate data lifecycle for compliance and other regulatory requirements. For instance, medical records can be left un-altered for an appropriate time until its expired after which it can be deleted. Also, HCP policies canbe used that will encrypt data and secure access to it,in order to comply with data protection regulations.

Integration between HNAS and HCP as well as HDI and HCP provides AS environments with backup-free ndemingly bottomless storage without compromising users' and applications' access to data. This approach helps eliminate the aforementioned silo of Infrastructure by regularly moving static content to a centralized, self-protecting, multi-tenant repository that provides the necessary segregation of data without the need for physically separate systems.

Furthermore, a number of 3rd party applications that provide industry specific data management and access have been developed to connect directly to the HCP content repository. For example, specialized applications focused on medical imagery, records management, and email archives to name just a few, are available that interface directly with the HCP. Such support for myriad 3rd party and custom applications as well as typical protocols such as HTTP/REST, NFS, CIFS, SMTP and others, coupled with its built in multi-tenancy, make HCP a key component of the content cloud. The HCP enables all data to reside in a single system, integrated and managed in a unified manner with data that can freely move among media types, applications and geographically dispersed or even mobile users.

HDI distributed data caching for HCP

HDI functions as an edge based file server for remote and branch offices (ROBO) that integrates seamlessly with an HCP backend. HDI can cache files locally for ROBO quick access while at the same time replicating this data to a core HCP content repository to preserve, protect and provide access over its lifetime. Also, HDI can supply automated storage tiering, where files are migrated from HDI to HCP when they no longer need to be accessed. As a result, ROBO file data can be placed under HCP storage management reducing storage costs and backup requirements.

---

4 Please see http://www.hds.com/products/file-and-content/content-platform/ for more information.

Significantly while still preserving access to the data when needed.[5] This gives those remote and branch offices backup-free and seemingly bottomless storage with local IO Performance and ensures access to a consistent set of data as well as providing for their unique needs at any given site.

Hitachi Data Discovery for Microsoft with HCP, HNAS, HUS or HUS VM

HDS supplies the Hitachi Data Discovery for Microsoft® SharePoint® connector to archive SharePoint 2010 data to Hitachi HCP, HNAS, HUS or HUS VM storage systems. Hitachi Data Discovery for Microsoft® SharePoint® integrates with SharePoint single or multiple server farms so that data or folders can be migrated to Hitachi file or fixed content storage rather than consuming space in constrained collection databases. As such, archived SharePoint data are replaced with pointers or stub files that preserve access paths but minimize collection capacity requirements.[6]

Moreover by migrating such data, SharePoint collections shrink to only include active data that significantly reduces database backup time and server load. In addition, by bringing down collection size, customers may be able to lessen SharePoint and SQL Server licensing costs that can represent substantial expenditures for large enterprises.

HDDS federated search

HDDS provides a single search engine for many enterprise NAS storage systems but also offers deeper integration with HCP for more comprehensive object metadata searching. HDDS operates with HCP, HNAS, HUS, HUS VM, NetApp® Filers, and/or Microsoft Windows Storage Server 2008 (WSS) to provide full-text plus metadata indexing and search of an enterprise's unstructured data.[7]

As a result, with one command, operator's can search every file in an enterprise. Such federated search capabilities can be crucial to the success of an E-discovery request or other information query that needs to examine all files across a data center.

In addition to its federated search, HDDS can be used to migrate files from HNAS, HUS, HUS VM or WSS to HCP providing intelligent storage tiering for file data. Any keyword that can be searched on can just as easily be used as criteria for migration. Thus, HDDS in conjunction with HCP canreduce backup requirements, produce better governance and control the cost of e-discovery for any enterprise file data.

---

5 Please see http://www.hds.com/products/file-and-content/data-ingestor.html for more information.

6 Please see http://www.hds.com/products/file-and-content/data-discovery-for-ms-sharepoint.html for more information.

7 Please see http://www.hds.com/products/file-and-content/data-discovery-suite.html for more information

Summary

In short, the HDS file and content solutions portfolio spans a wide range of capabilities in an integrated solution that brings together NAS, object storage and search. While the individual offerings deliver compelling value on their own, the biggest ROI is realized when they work together as part of an infrastructure, content or information cloud where the unruliness of enterprise file data can almost be tamed, if not harnessed to generate new insights. It may be just an initial step, but with today's HDS file and content solutions product family the promise of the information cloud can finally start to be realized.

Silverton Consulting, Inc. is a Storage, Strategy & Systems consulting services company, based in the USA offering products and services to