Smarter cloud storage solutions

A special eGuide from IBM

e book

www.technologydecisions.com.au

technology Decisions
IT leadership & innovation
www.technologydecisions.com.au

WESTWICK-FARROW MEDIA

This special eGuide comprises three informative white papers that will assist you in comparing cloud storage solutions in both private and public environments.

From user case studies to product and solution comparisons, the information in this eGuide will assist you in making the right decisions to help save money and increase the effectiveness of your cloud storage architecture.

From real-time compression to dynamic provisioning and performance optimisation, these IBM white papers look at the constantly evolving requirements for storage systems operating in cloud environments.

We trust you find this eGuide of value.

Best regards

Geoff Hird
Publisher – *Technology Decisions*
www.TechnologyDecisions.com.au

BUILDING A SMART STORAGE CLOUD WITH IBM SMARTCLOUD STORAGE ACCESS

JANUARY 2013

Many organizations find that their large networked storage deployments are siloed or dedicated to specific applications, making it difficult to dynamically provision, optimize performance, or share expensive resources. Transitioning to cloud-based storage provides an attractive solution and many public cloud storage services are readily available. However, some lack the processing power, data transfer speeds or high service levels that active enterprise data centers require. And even when a public cloud storage service offers the service levels that a corporation needs, the organization must choose between contracting outside to a public cloud provider and being able to leverage expensive investments already made in on-premise storage systems. The problem is that IT does not want to give up the advantages of operating scale-out NAS or unified SAN/NAS storage but would like to take advantage of the cloud's efficiency, agility and increased simplicity.

The answer to these organizations is IBM SmartCloud Storage Access, which lets organizations turn on-premise storage systems including IBM Scale Out Network Attached Storage (SONAS) and IBM Storwize V7000 Unified Storage into powerful private clouds. The cloud-based storage services created by IBM SmartCloud Storage Access combines the scale-out and unified features of the underlying storage systems into highly flexible and manageable cloud-based storage.

Overall, IBM SmartCloud Storage solutions enable business to easily consume or deliver storage services via managed, hosted, public, and private clouds. IBM SmartCloud Storage Access is a key member of IBM's SmartCloud Storage offerings for private clouds. It enables a business to build a storage cloud service that provides simple and powerful cloud functionality without the usual layers of complexity. This Solution Profile will briefly describe IBM SmartCloud Storage, and examine IBM SmartCloud Storage Access as the enabling feature set for a powerful and simplified private cloud. The IBM-enabled cloud offers strong benefits at three levels: building a highly scalable, secure and flexible cloud; running the hosted cloud for continued benefits to the corporation; and extending real benefits to the private cloud end users.

IBM SmartCloud

IBM SmartCloud is a huge IBM portfolio covering a variety of IBM products, services and partnerships that cohesively integrate into IBM-enabled cloud offerings. Broad IBM SmartCloud portfolio groupings include Foundation (private and hybrid clouds), Infrastructure as a Service (IaaS), Platform as a Service (PaaS), Software as a Service (SaaS) and IBM's SmartCloud consulting and implementation arm.

Very briefly, IBM SmartCloud is an entry point to different cloud-enabled computing structures including private clouds, cloud services and cloud-enabled applications. This includes physical hardware like the IBM SONAS and the Storwize V7000 Unified storage systems, both of which have

87 Elm Street, Suite 900 · Hopkinton, MA 01748 · T: 508.435.2556 · F: 508.435.2557

native virtualization support for the IBM virtualization layer, and software stack solutions like Websphere that support application infrastructure. IBM Tivoli can provide the management middleware for data protection, provisioning automation, and cost tracking tools. The final layer provides workload integration tools that can integrate the IBM cloud framework with business applications.

IBM SMARTCLOUD STORAGE

Underlying many of the IBM SmartCloud solutions is IBM SmartCloud Storage, which provides cloud-optimized storage services for storage, compute, backup, and archive clouds.

IBM SmartCloud Storage Enables:	Building	Running	Consuming
Storage Cloud	**IBM SmartCloud Storage Access**, IBM Active Cloud Engine, storage and data services for the cloud	SmartCloud for managed and hosted services	Object storage with SmartCloud Enterprise
Compute Cloud	IBM SmartCloud Virtual Storage Center, TSM for Virtual Environments, Flex System V7000 Storage Node , storage and data services for cloud	SmartCloud for managed and hosted services	Persistent storage using SmartCloud Enterprise
Backup Cloud	TSM Suite for Unified Recovery with Front Safe Portal, storage and data services for cloud	SmartCloud Managed Backup	Object storage with SmartCloud Enterprise
Archive Cloud	TSM Suite for Unified Recovery with Front Safe Portal, storage and data services for cloud	SmartCloud Content Management	Object storage with SmartCloud Enterprise

Figure 1 – IBM SmartCloud Storage components for building, running, or consuming cloud storage services

IBM SmartCloud Storage is built on IBM hardware and software with available expert consulting services to build private, public and hybrid clouds. Management and hosting services enable flexible and policy-driven management while end-users enjoy rapid access, self-provisioning, and user-level reporting.

IBM SmartCloud Storage Access

Drilling down further into the Storage Cloud model, we see IBM SmartCloud Storage Access as a fundamental building block for leveraging IBM storage infrastructure like IBM SONAS and Storwize V7000 Unified into highly manageable, flexible and robust private clouds.

IBM SmartCloud Storage Access provides a web-based interface that is easy to both deploy and use to define, manage, and provision cloud storage. Immediately, it layers self-service capacity

87 Elm Street, Suite 900 · Hopkinton, MA 01748 · T: 508.435.2556 · F: 508.435.2557

provisioning, monitoring and reporting on top of and integrated with storage management (Tivoli Storage Productivity Center) and assigned storage arrays.

IBM SmartCloud Storage Access facilitates building private clouds using new or existing IBM storage systems. IBM's SmartCloud approach greatly reduces the usual risk, cost and complexity of building private clouds out of an all-new top-to-bottom stack of previously unintegrated components. The all-IBM infrastructure guarantees interoperability between storage and storage management making the private cloud highly manageable and offering highly secure IT controls for security, governance, capacity, and allocation.

HOW IBM SMARTCLOUD STORAGE ACCESS MAKES IT EASY

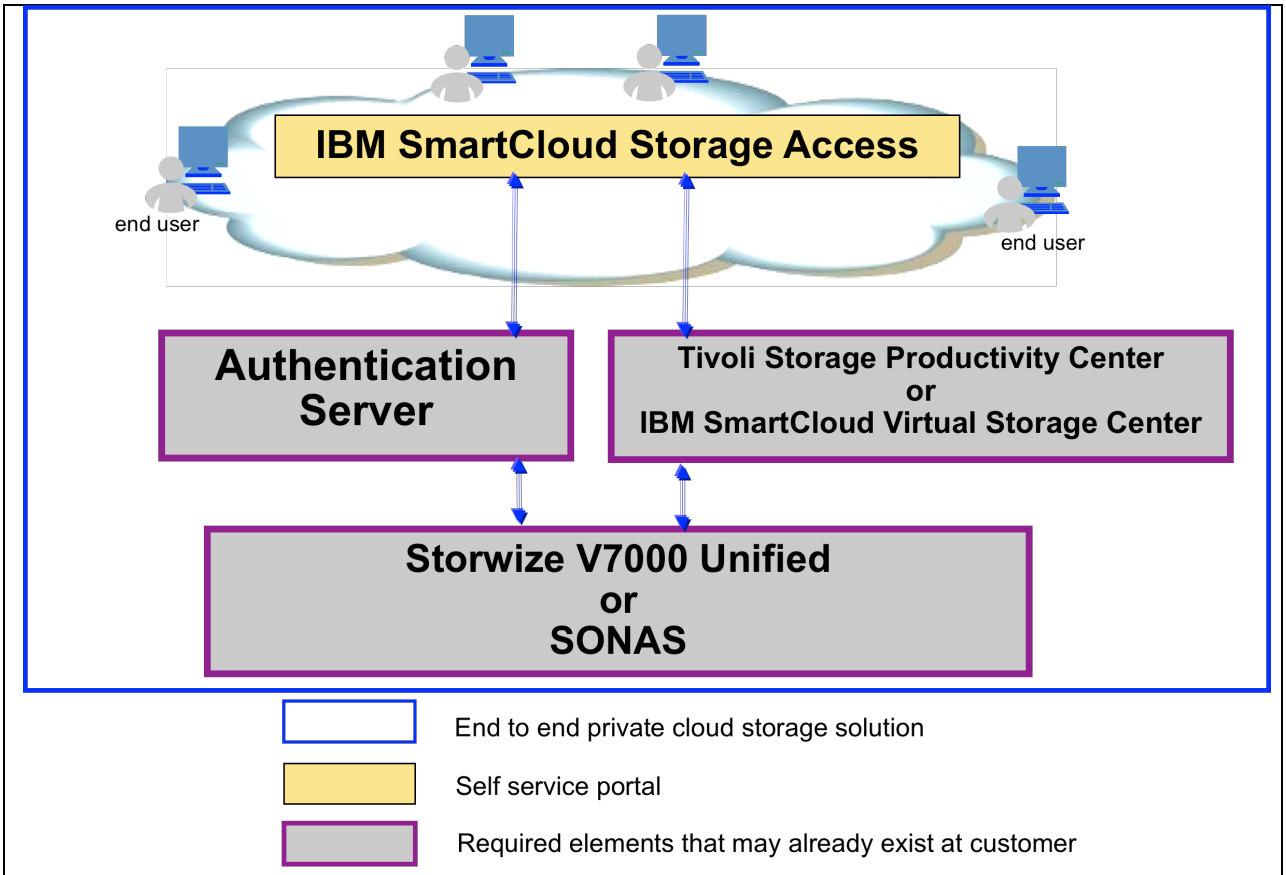


Fig.1: IBM SmartCloud Storage Access infrastructure

- **Secure Access.** IBM SmartCloud Storage Access leverages existing authentication services to determine access at every level from the storage systems to the cloud provisioning. Role-based controls enable cloud owners to grant or deny various access levels as needed for individuals or groups. End-users may log into the portal to request storage provisioning and sharing and to access usage reports. Administrators may use the portal to manage assigned levels from the entire cloud to a regional portion and may schedule automatic operations or establish policy-driven triggers.

- **Simplified Provisioning and Access.** Users can request and access storage without having to know low-level infrastructure architectures or specific device types. This enables IT to deliver storage to end users while abstracting out internal operational details. End users can easily find

87 Elm Street, Suite 900 • Hopkinton, MA 01748 • T: 508.435.2556 • F: 508.435.2557

their virtualized storage without worrying about storage locations and pathnames, and can access their data from multiple endpoints including desktops and laptops.

- **Easy On-Demand Provisioning.** Administrators can set up policy-driven approval processes for users. For example, administrators may choose to allow end users to directly self-provision their own storage within defined boundaries. IT's job is made easier as IBM SmartCloud Storage Access lets end-users provision capacity on demand subject to IT permissions.
- **Complete Reporting.** Both administrators and end-users may run dashboard reports for queries on historical usage. The reports are customizable by service role: administrators can view usage data for the end users or for the whole department or per storage resource pool, while end users can view their own usage data.
- **Definable Service Levels.** IT can define a service catalog to help balance performance and optimize capacity utilization across multiple tiers of storage infrastructure. Users can more easily provision appropriately matched capacity for their needs (and budget). IBM Tivoli Productivity Center (TPC) will automatically conduct storage management activities based on defined service levels.

Building Productive and Efficient Private Clouds

IT shops faced with the challenge of building out a private storage services cloud will find that IBM SmartCloud Storage Access 1) delivers the functional and usability requirements demanded by both operators and end-users, and 2) nets valuable organizational cost and competitiveness benefits.

IT EFFICIENCY

IT can spend a lot of time and effort provisioning storage on a manual, per-application or per-user request basis. This ongoing storage workload can decrease dramatically by implementing a storage services cloud with IBM SmartCloud Storage Access. IT now has the ability to support self-provisioning within role-based boundaries out of a pre-defined storage services catalog, which also enforces automatic business policies.

These same capabilities efficiently align expensive storage with business needs, which optimizes both performance and capacity. And by serving virtualized storage as a service instead of lining up dedicated devices, IT effectively aggregates demand and can more readily handle individual application/user volatility. In addition, IBM SmartCloud Storage Access simplifies storage space planning and enables better strategic options for future hybrid or public storage adoption.

CLIENT AGILITY

There is obvious value to end-users in reduced time to provision new storage (actual reductions could be 20:1 or more depending on existing practices). There is also value in consistently being able to access multiple grades or tiers of storage. With IBM SmartCloud Storage Access, users can align their actual needs with offered services to meet performance, budget, integrity, DR, and other requirements.

Agile storage requires that IT can de-provision and/or re-align storage on-demand. Yet dedicated infrastructure is often either over- or under-provisioned, with few options to correcting, iterating, growing, or shrinking allocations. IBM SmartCloud Storage Access provides virtualized storage services that can be realigned elastically and dynamically as needed, along with the insightful reporting necessary to making both tactical and strategic provisioning decisions.

87 Elm Street, Suite 900 · Hopkinton, MA 01748 · T: 508.435.2556 · F: 508.435.2557

COST VISIBILITY

Increasing IT efficiency and client agility contribute to increased business productivity. Upstream cloud storage services enable more agile development, while downstream employees, partners, and clients all benefit as applications run faster and smoother in production operations. These benefits, along with improvements in storage administrator overhead, storage provisioning time, and effective storage utilization, all provide cost visibility and justification for private cloud storage services.

Taneja Group Opinion

IBM has put a lot of effort into organizing and integrating its vast portfolio of IBM SmartCloud solutions. Key to IBM's approach is to make sure that their client enterprises can access cloud benefits in whichever way proves best for the client, from building internal private cloud services to running services on managed or hosted clouds to simply better consuming cloud services. Cloud storage services sit in the middle of most enterprise cloud designs and underpin everything from backup to mission-critical applications, so IBM's SmartCloud Storage solutions need to be especially robust in operations and simple to adopt.

IBM SmartCloud Storage Access meets these criteria by providing the self-service provisioning portal. It simplifies private cloud deployment by readily extending existing infrastructure, management, and staffing skillsets right into private cloud service delivery. With little configuration, current IBM SONAS and Storwize V7000 Unified arrays can be transformed into high-functioning self-service storage clouds. We expect that IT organizations with IBM storage will find that transitioning from manual storage allocation practices to offering storage as a service will experience overwhelming cost, agility, and efficiency benefits with little or no transition cost or pain.

We think that the ease and simplicity of deploying IBM SmartCloud Storage with IBM SmartCloud Storage Access is exceptionally attractive, and should be considered by all IT organizations looking to build cloud storage services. IBM manages very large storage clouds with their own portfolio, both public and managed/hosted for the largest of enterprises, and can be trusted to intimately understand and ultimately deliver rock-solid cloud solutions.

---

87 Elm Street, Suite 900 · Hopkinton, MA 01748 · T: 508.435.2556 · F: 508.435.2557

THE CLIPPER GROUP
Navigator™
Navigating Information Technology Horizons℠
Published Since 1993 Report #TCG2012013LIR Updated October 2, 2012

Significantly Improving Storage Efficiency — IBM Delivers Real-time Compression on Unified Storage

Analyst: David Reine

Management Summary

When you go into the supermarket do you ever wonder how many different products are on the shelves? Do you wonder how they determine what goes on each shelf? All of the shelves are shared among many products. *What gets placed on the bottom shelf? What goes on the top?* One week your favorite item might be at eye level, and the next time there may be a sale item in that same spot. *How do they determine the value (such as return per foot of frontage) of each shelf?* The supermarket often positions high-margin (impulse buying) items close to related items; for example, premium buns, wraps and condiments at the deli counter, trying to get the greatest return from that positioning. The overall goal is to get the greatest return from all of the store's shelf space both premium and hard-to-reach, so this is an ever-changing proposition, with grocery stockers making regular adjustments to maximize the latest strategy and promotions. Maximizing sale of more profitable items is an important measure of *effectiveness*.

However, maximizing the use of shelf space is another important goal, and is a measure of *efficiency*. Any empty space is a lost opportunity. Thus, stockers move the cans and boxes around to keep the shelves as full as possible. Packaging with a smaller footprint increases the density, so that is one way of getting more product into the same space. Additionally, some food companies offer freeze-dried items or condensed items, such as instant coffee, soup-in-a-cup, condensed milk, concentrated juice, and powdered drinks. All you have to do is add water to the contents to reconstitute it. This also improves shelf efficiency plus there is less weight and volume to transport. As a good rule of thumb – the better the space efficiency, the better is the opportunity to maximize the return on the shelf space. In order to make a profit, the supermarket has to balance the goals of effectiveness and efficiency. Sometimes, the efficiency and the effectiveness are at odds on a cost per unit basis. However, if you can improve the efficiency without affecting the effectiveness, then it is a winning proposition.

Trying to fit more "product" into a given amount of space becomes even more significant as we transition from our personal grocery shopping to our professional IT lives, especially as we try to figure out what to do with all of the information that needs to be stored. When it comes to storing data, we find the same concepts of effectiveness and efficiency. Often, these are at odds with each other. Usually, when you make something *more effective* (think "higher qualities of service"), you do so by spending more per unit of storage. Conversely, when you make storage *more efficient* (costs less per unit of storage), you often do so by lessening the quality of service in one or more ways. **What we all seek is that elusive deal where we don't affect the quality of service, yet are able to store a lot more (without spending a lot more).**

This is exactly what IBM has done with its latest improvement to its *Storwize V7000 Unified*

IN THIS ISSUE

mid-range storage system. By adding a new feature called *Real-time Compression*, you can store more (probably, a lot more) without having to buy any more hardware and without any impact on storage operations. The cost of this is a modest software licensing charge. To learn more about Real-time Compression, please read on.

The Quest for Storage Efficiency

The long-term storage of data deals with information that needs to be preserved and protected as a result of application requirements or enterprise, industry, or even governmental regulations and standards. The value of data to the business, however, may change over time, along with the urgency to access it, usually in terms of "how fast" or "how often".

- *How can we store it economically, yet maintain rapid availability when necessary?*
- *How can we reduce the amount of physical storage required to preserve and protect a given amount of data?*
- *How can we deploy storage more effectively in terms of different tiers of storage, yet also deploy it more efficiently in order to reduce the total cost of ownership (TCO)?*

First, we can separate this pile of bits and bytes into two basic categories: data that is "likely needed" and "less likely needed", loosely corresponding to *active data* and *archived data*. Active data may be needed at any moment and any significant delay (think "a second or two, at the maximum") will slow down the pace of business. While archived data may be needed at any moment, the expectations for instantaneous delivery are few, because it is realized that it is just one of thousands or millions (or more) of data that has been archived, usually in a manner that places cost minimization ahead of near-instantaneous retrieval speed.

Of course, this isn't a binary world most of the time, as there may be many levels of quality of service (platinum, gold, silver, etc.) and cost (from very high to very low). Optimally placing data at the best storage tier (where quality of service matches cost of delivery) is the name of this game. Whether you do this through automation or manual administration, let's assume that it gets done and done well.

While IT management may give you a good pat on the back, more than likely it will be followed by that stinger of a question "What have you done for me lately?" This, of course is driven by the rapid growth of data outstripping the near stagnant storage budget. Wouldn't you like to be able to pull something out of your professional bag of tricks, besides just saying that you are waiting for the next refresh of technology to lower your average costs? Well, now you can. Exactly how you will do this depends on what storage now is sitting on your floor.

Introduction to IBM Storwize V7000 Unified

If you already have an *IBM Storwize V7000 Unified* storage system, you are way ahead in your quest to make something seemingly magical happen. But, let's assume that you haven't done that yet, and explain what Storwize V7000 has been delivering since it was announced in 2010. If you are familiar, just skip to the section "New to Storwize V7000 Unified" on the next page.

Unlike some vendors, IBM has a multiplicity of storage solutions for the data center. This includes the *DS3000*[1] family for the entry-level data center, up to the IBM *System Storage DS8000*[2], *XIV Storage System*[3], and the *Scale-Out Network Attached Storage* (*SONAS*)[4] for the largest enterprises. Most importantly for the mid-sized business or segregated outpost of a larger one, IBM has a variety of mid-range storage solutions, in the $25K to $250K range, including the *Storwize V7000 and Storwize V7000 storage systems*. IBM's customers do have many options and this allows them to choose one or more solutions that best fits *their* storage requirements rather than trying to force fit a single, "universal" solution into their data center, unless that is what is desired.

In an era of rampant data growth (with more to store), there are increasing requirements for

[1] See **The Clipper Group Navigator** entitled *Delivering Enterprise Features for the Mid-Range – IBM Introduces DS3500 with TPC MRE*, dated June 7, 2010, at http://www.clipper.com/research/TCG2010027.pdf.

[2] See **The Clipper Group Navigator** entitled *Maximizing the Business Value of Information and Lowering Energy Consumption with IBM's DS8700*, dated April 13, 2010, at http://www.clipper.com/research/TCG2010018.pdf.

[3] See **The Clipper Group Navigator** entitled *XIV Gen 3 – IBM Lowers TCO and Raises Performance and Functionality*, dated August 5, 2011, and available at http://www.clipper.com/research/TCG2011028.pdf.

[4] See **The Clipper Group Navigator** entitled *IBM SONAS – A New Kind of Infrastructure to Meet Large-Scale Storage Challenges*, dated March 15, 2010, at http://www.clipper.com/research/TCG2010010.pdf.

scalability at each tier. Simply put, an array selected must be able to scale to handle growing business requirements. However, the ability to scale is not the only requirement. Today, more so than ever before, enterprises of all sizes are looking for the storage management functionality once only found in the most expensive, high-end systems.

In an attempt to meet both the scalability and functional requirements, IBM Storwize V7000 was introduced in October 2010 as a multi-tiered, scalable mid-range storage solution with enterprise-class capabilities and outstanding performance.[5] It had outstanding power, flexibility, ease-of-use and functionality, in many ways similar to that of IBM's DS8000 enterprise-level systems. It had the following features and capabilities.

- **SSD devices** with very high IOPS – for database applications requiring high performance;
- **High-performance SAS** drives – to satisfy mission-critical Tier-1 application requirements;
- **High-capacity nearline SAS** drives –to fulfill the rapidly growing needs of Tier-2 business-critical applications, such as email and other web-facing requirements;
- **Easy-to-use graphical user interface** for storage administrators;
- **Easy Tier** – for automatic and dynamic migration between tiers; and
- **Virtualization** – to improve storage utilization rates, in much the same way that data centers have consolidated and virtualized their server infrastructure to improve the TCO of the IT environment.

Six months later, IBM added two-way clustering, 10Gbps iSCSI, additional drives, and improved *VMware* management. In 4Q11, IBM added unified block and file storage, greater flexibility for remote mirroring, enhanced *Tivoli* support, and even more new drive options.

It can have a range of storage capabilities (i.e., multiple tiers to satisfy multiple needs like active and archived data, to name just two), uses automation to move data between the tiers, is easy to administer, handles both file- and block-oriented access (that's what makes it "unified"), can virtualize legacy arrays behind it, and is priced competitively for the mid-range. While that sounds like everything that you might want (and the list of capabilities is much longer, see Exhibit 1, above), remember that your boss is going to ask, sooner or later, for another round of improvements.

Exhibit 1 — Storwize V7000 and Storwize V7000 Unified Management Components and Capabilities

- Virtualized unified storage capable of enjoining external arrays of many types
- IBM *FlashCopy Manager*
- *Tivoli Storage Productivity Center*
- *Tivoli Storage Manager*
- *Tivoli Storage Manager FastBack*
- IBM *Systems Director*
- IBM *Active Cloud Engine*
- Block and file replication
- Asynchronous remote block mirroring
- Thin provisioning
- Storage performance monitoring
- Capacity planning and forecasting
- Real-time Compression [new]

Source: IBM

New to Storwize V7000 Unified

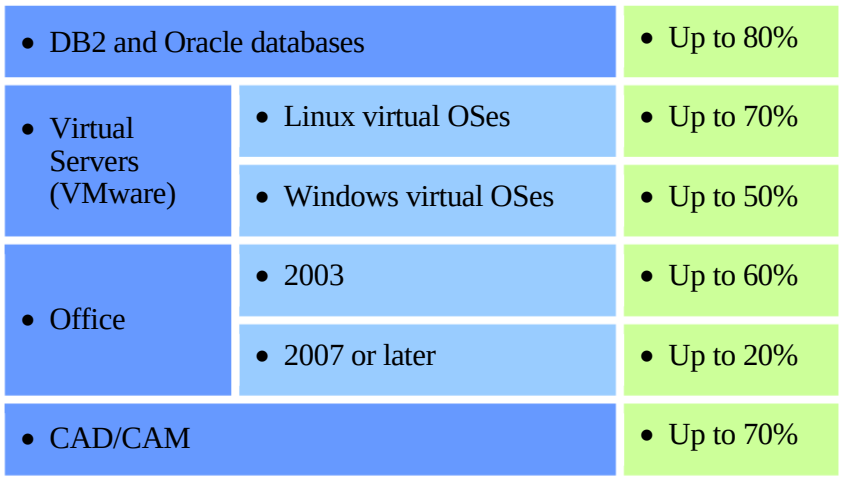
IBM has continued to improve the functionality offered in Storwize V7000, which is part of IBM's *Smarter Storage Strategy*. And now, **IBM has announced *Real-time Compression* for Storwize V7000 Unified. This is the surprise that you need to pull out of your bag of tricks.** There are other additions concurrent with compression. These include four-way clustering for even more scalability, support for FCoE, VMware support for file access, and even more software interoperability. These features are added to other innovations from the full spectrum of IBM storage products, including an innovative GUI from XIV, *Easy Tier* from the DS8000, and the *Active Cloud Engine* from SONAS. These all represent capabilities that distinguish Storwize V7000 Unified from its competitors. **While all of these are important, what you really want (right now) is the special effect of Real-time Compression.**

Real-time Compression

Now, with Real-time Compression, the data center can store more data without having to buy any more storage, simply by obtaining a license to turn on RTC, an expenditure with a quick

[5] See **The Clipper Group Navigator** entitled *IBM Brings Enterprise Functionality to Mid-Range Storage* dated October 7, 2010, and available at http://www.clipper.com/research/TCG2010047.pdf.

Exhibit 2 — Components and Capabilities



Source: IBM

ROI. There is no additional hardware to buy and nothing to change; simply a license to procure.[6] If you already have Storwize V7000 or Storwize V7000 Unified, simply obtain a license and improve your physical storage capacity. If you have a block-based array other than Storwize V7000, consider putting it behind Storwize V7000 or the IBM SAN Volume Controller (which also can do Real-time Compression), enabling the IT staff to achieve compression on the non-Storwize-V7000 array and getting the advantages of common management and virtualization, plus Real-time Compression. Truly, this a triple benefit.

Storwize V7000 Unified takes advantage of IBM innovations introduced on its larger platforms, to enable the data center to store even more active and less active data, more efficiently, in the same unified space. **Real-time Compression works behind the scenes to compress and decompress your data, dynamically, in real-time, making compression, literally, a "no-brainer". No compromises need to be made between "hot", active data, and less active data.**

How Real-time Compression Works

Real-time Compression uses a time-tested and patented *Random Access Compression Engine* (*RACE*), used previously in IBM Real-time Compression Appliances. This technology, according to IBM, allows you to store up to five times more data for DB2 and Oracle databases, with significant savings for other kinds of data, too (as shown in Exhibit 2, above.) **By using storage more efficiently, the data center can slow the increase in capacity and reduce acquisition costs, rack space, energy consumption, and software costs for additional functions.** IBM has even made available a *Compresstimator* tool to evaluate the expected benefits in your environment. Your mileage will vary, depending on your mix and volumes of data, as you would expect.

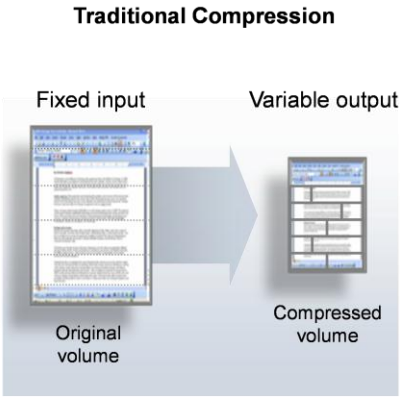
Real-time Compression is an innovative, simplified compression that is fully integrated into Storwize V7000 and Storwize V7000 Unified to further improve their storage efficiency with a delivery compression of 50% or better. **It is a high-performance implementation that has been designed to support active, primary workloads**, in support of SSDs and high-performance SAS devices, **and also for less active data**, using high-capacity nearline SAS drives. **In other words, it won't slow down what you now are doing; it will just do it with less physical storage space.** (This sounds like magic but is no illusion!)

The inclusion of this dynamic compression capability reduces the acquisition budget, saves rack space, reduces energy costs, and minimizes licensing costs for software with charges based upon total physical storage capacity. In short, compression will help the data center freeze storage growth, or at the least, enable the enterprise to slow down additional storage purchases. If additional acquisitions are still required, a new Storwize V7000 Unified system will provide the data center with more new capacity at a lower

[6] Of course, if you don't already have Storwize V7000, you'll have to get one, but with the space-saving benefits of compression, it shouldn't be hard to convince your boss.

Exhibit 3 — Traditional Compression

- Traditional compression engines take a fixed input and produce a variable output depending on compressibility
- Some or all data written is stored uncompressed and only compressed later
  - Reduces efficiency because additional storage required
- Compression ratios depend on chunk size used
  - Big chunks can have poor I/O performance
  - Small chunks offer lower compression ratios
  - Compression typically degenerates over time creating fragmentation requiring more garbage collection
  - Poorer I/O performance



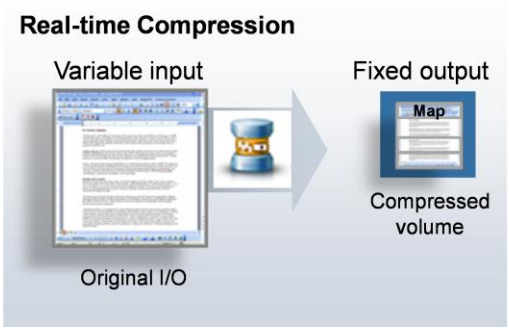
Source: IBM

cost per TB. The previous version of Storwize V7000 (without compression) already was an attractively priced mid-sized offering. Even with the added licensing list price of $9,000 per shelf of disk, Storwize V7000 Unified is now an even better bargain on a TCO/TB basis. Other vendors, such as EMC and NetApp, provide compression as well. Traditional compression engines take a fixed input and produce a variable output depending on compressibility (See Exhibit 3, above). Active data is first stored uncompressed and compressed later through a pre-scheduled post process activity (i.e., it is not compressed in real-time, as it is first being stored).

Traditional compression ratios depend upon the size of each "chunk"; while larger chunks can have a poor I/O performance and higher

Exhibit 4 — IBM's "Compression Without Compromise"
Variable Sized Compression Enhances Performance and Efficiency

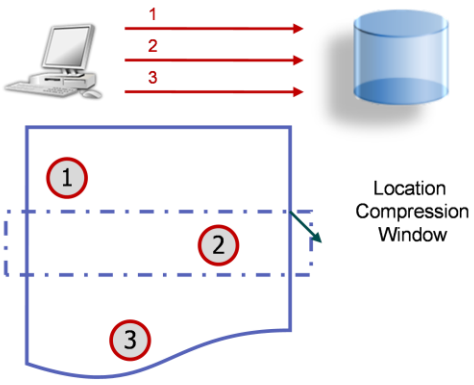
- RACE flips this approach, taking a variable data stream and producing a "fixed" output
  - Compressed volume has a consistent logical layout
  - Temporal locality: data that's accessed together is compressed together
  - Variable sized input chunks get better compression
  - Less data is compressed and decompressed
  - Fewer disk I/Os
  - Better performance



Traditional Compression	IBM Real-time Compression
1 MB Read	0 MB Read
1 MB Decompress	0 MB Decompress
100 Byte Update	0 Byte Update
1 MB Compress	100 Byte Compress
1 MB Write	<100 Byte Write
Total I/O: 2 MB	**Total I/O: <100 Bytes**

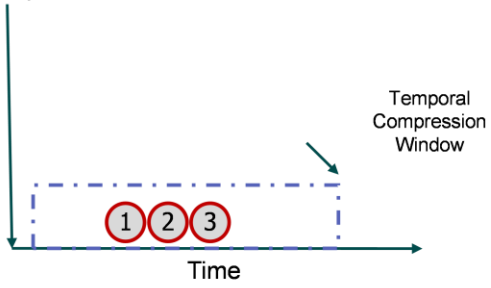
Source: IBM

Exhibit 5 — IBM's "Compression Without Compromise" Explained
Designed for Real-Time Random Access to Active Data



- Applications make multiple updates to data
- Traditional compression uses fixed-sized chunks and compresses each update based on its physical location on a volume
- In this example, **three** separate compression actions

- RACE compression acts on data that is written around the same time ("temporal locality") not according to physical location
- Temporal locality is more related to real system operation
  - Applications may make related updates to different parts of a volume
- RACE takes advantage of the structure of the data and its relations
- In this example, only **one** compression action
- Better compression and decompression efficiency and performance



Source: IBM

ratios; smaller chunks offer lower compression ratios but better performance. In fact, traditional compression methods typically degenerate over time, creating fragmentation requiring even more administrative clean-up. On the other hand, IBM's Real-time Compression is a dynamic procedure, operating immediately, in an unobtrusive, easy-to-manage process.

What makes IBM's implementation superior?

Using RACE, IBM provides compression for both active and inactive data, not simply compressing secondary copies of the data, as do some compression schemes. **Real-time Compression always is active, there is no need to schedule a post-process compression task, freeing up administrative personnel to be doing more productive work.** Real-time Compression provides a high performance engine to support workloads that other solutions cannot, significantly expanding the amount and type of data that is available to be compressed. RACE takes a variable-length data stream and produces a fixed output, significantly improving storage efficiency. (See Exhibit 4, at the bottom of the previous page.)

RACE compressed data has a consistent, logical layout, with data that's accessed together being compressed together. With fixed-sized chunks to store, Storwize V7000 Unified can get better compression, with less data being compressed and decompressed when accessing a specific piece. As a result, there are fewer disk I/Os and, thus, better performance. (See Exhibit 5, above.)

Real-time Compression eliminates the requirement to reserve additional buffers for uncompressed data in transit. It supports all Storwize V7000 and Storwize V7000 Unified devices, including internal or external virtualized storage, enhancing the value of the data center's existing storage assets. Real-time Compression has been integrated nicely into the Storwize V7000 GUI to enhance its usability, giving IT staff more information about the compression performance and savings being achieved.

Existing Storwize V7000 volumes can be compressed easily with volume mirroring, eliminating unused space during the conversion. In terms of scalability, Real-time Compression currently supports up to 200 compressed volumes for each Storwize V7000 control enclosure.[7]

[7] IBM has indicated that the Storwize V7000 will support more than 200 volumes in the near future.

Conclusion

Why should your data center deploy Storwize V7000 or Storwize V7000 Unified with Real-time Compression? From an economic standpoint, these systems provide the enterprise with the financial benefits it needs to meet the growing storage requirements within the limits of a constrained budget. **If you already own a block-only Storwize V7000 or Storwize V7000 Unified, all you need is a license for Real-time Compression and you can take advantage of more physical storage without significant effect on your budget. If you own other block storage, you can put that storage behind a new Storwize V7000 Unified system and receive the same, space-saving benefits.**

Real-time Compression provides the data center with the flexibility to increase capacity and maintain performance, for all kinds of data, including your most active data. *With the same storage budget*, the data center can expect to see a reduction in per usable TB storage costs for most common configurations in the order of 30-to-40%, according to IBM. *To maintain existing usable capacity*, the data center can expect to see a reduction in cost per usable TB of approximately 25-to-35 percent. By applying compression to existing storage, the data center might be able to double usable capacity for a very modest incremental investment.

If your data center needs to compress both block and NAS storage, Storwize V7000 Unified provides the single platform that you need to simplify deployment. If you are suffering from too much data with an insufficient budget, Storwize V7000 Unified with Real-time Compression may well be the solution you need. It will be like getting more shelf space without enlarging the grocery store and without paying the going price for additional shelves. Isn't that the kind of surprise you want to pull out of your bag of tricks the next time you are asked? Check it out!

About The Clipper Group, Inc.

The Clipper Group, Inc., now in its twentieth year, is an independent publishing and consulting firm specializing in acquisition decisions and strategic advice regarding complex, enterprise-class information technologies. Our team of industry professionals averages more than 25 years of real-world experience. A team of staff consultants augments our capabilities, with significant experience across a broad spectrum of applications and environments.

- ***The Clipper Group can be reached at 781-235-0085 and found on the web at www.clipper.com.***

About the Author

David Reine **is a Senior Contributing Analyst for The Clipper Group.** Mr. Reine specializes in enterprise servers, storage, and software, strategic business solutions, and trends in open systems architectures. In 2002, he joined The Clipper Group after three decades in server and storage product marketing and program management for Groupe Bull, Zenith Data Systems, and Honeywell Information Systems. Mr. Reine earned a Bachelor of Arts degree from Tufts University, and an MBA from Northeastern University.

- ***Reach David Reine via e-mail at dave.reine@clipper.com or at 781-235-0085 Ext. 123. (Please dial "123" when you hear the automated attendant.)***

Regarding Trademarks and Service Marks

The Clipper Group Navigator, **The Clipper Group Explorer**, **The Clipper Group Observer**, **The Clipper Group** *Captain's Log*, **The Clipper Group Voyager**, *Clipper* **Notes**, and *"clipper.com"* are trademarks of The Clipper Group, Inc., and the clipper ship drawings, *"Navigating Information Technology Horizons"*, and *"teraproductivity"* are service marks of The Clipper Group, Inc. The Clipper Group, Inc., reserves all rights regarding its trademarks and service marks. All other trademarks, etc., belong to their respective owners.

Disclosures

Officers and/or employees of The Clipper Group may own as individuals, directly or indirectly, shares in one or more companies discussed in this bulletin. Company policy prohibits any officer or employee from holding more than one percent of the outstanding shares of any company covered by The Clipper Group. The Clipper Group, Inc., has no such equity holdings.

After publication of a bulletin on *clipper.com*, The Clipper Group offers all vendors and users the opportunity to license its publications for a fee, since linking to Clipper's web pages, posting of Clipper documents on other's websites, and printing of hard-copy reprints is not allowed without payment of related fee(s). Less than half of our publications are licensed in this way. In addition, analysts regularly receive briefings from many vendors. Occasionally, Clipper analysts' travel and/or lodging expenses and/or conference fees have been subsidized by a vendor, in order to participate in briefings. The Clipper Group does not charge any professional fees to participate in these information-gathering events. In addition, some vendors sometime provide binders, USB drives containing presentations, and other conference-related paraphernalia to Clipper's analysts.

Regarding the Information in this Issue

The Clipper Group believes the information included in this report to be accurate. Data has been received from a variety of sources, which we believe to be reliable, including manufacturers, distributors, or users of the products discussed herein. The Clipper Group, Inc., cannot be held responsible for any consequential damages resulting from the application of information or opinions contained in this report.

WHITE PAPER

Eliminate Shortcomings in Your Cloud Architecture with Smarter Storage

Sponsored by: IBM Corporation

Laura DuBois

January 2013

Global Headquarters: 5 Speen Street Framingham, MA 01701 USA P.508.872.8200 F.508.935.4015 www.idc.com

EXECUTIVE SUMMARY

Enterprises are deploying private and hosted clouds to respond faster to business demands for new applications, to improve service levels for given workloads, and to reduce cost across the datacenter. The enterprise storage architecture deployed is paramount to ensuring agility, performance, and reliability of a private or hosted cloud. Integration of storage infrastructure across hypervisors, cloud orchestration software, and open source architectures is required. Storage architectures for private and hosted clouds must support a range of mixed application workloads and the dynamic placement of data on the right storage tier. These essential storage functions enable faster service response times, improved service-level agreements (SLAs), and greater agility, performance, and reliability for private and hosted clouds.

IBM has a differentiated storage portfolio, including XIV, Scale Out NAS (SONAS), SAN Volume Controller (SVC), Tivoli Storage Manager (TSM), Tivoli Storage Productivity Center (TPC), and SmartCloud solutions. These offerings enable firms to optimize their storage systems in order to achieve their private and hosted cloud objectives. This paper examines the benefits of and barriers to private cloud deployments and identifies the storage challenges enterprises face with these deployment models. It also discusses IBM's robust storage solutions with a focus on the state-of-the-art XIV storage architecture. XIV's ease-of-use, reliability, and performance attributes eliminate storage vulnerabilities that firms face when implementing private or hosted clouds. The offering is being used in some of the largest private and hosted clouds today. IDC spoke with several IBM customers using XIV in private and hosted cloud deployments. Their comments are included in this paper.

SITUATION OVERVIEW

As a result of business pressures and economic realities facing IT organizations today, technology executives need to continuously implement more efficient ways to run IT services for their companies. Increased reliance on IT systems to run business operations has heightened the need for reliability and performance across the datacenter. The consumerization of IT and the speed of global business have brought about demand for self-service, personalization, and instant access to IT. Acceleration in the delivery of products and services dictates faster availability of IT services and infrastructure. Business units and executives expect IT to quickly respond to new application and project requests while meeting service-level requirements. These factors are driving IT organizations to rethink the way infrastructure is provisioned, allocated, and managed for applications and lines of business. This forces IT executives

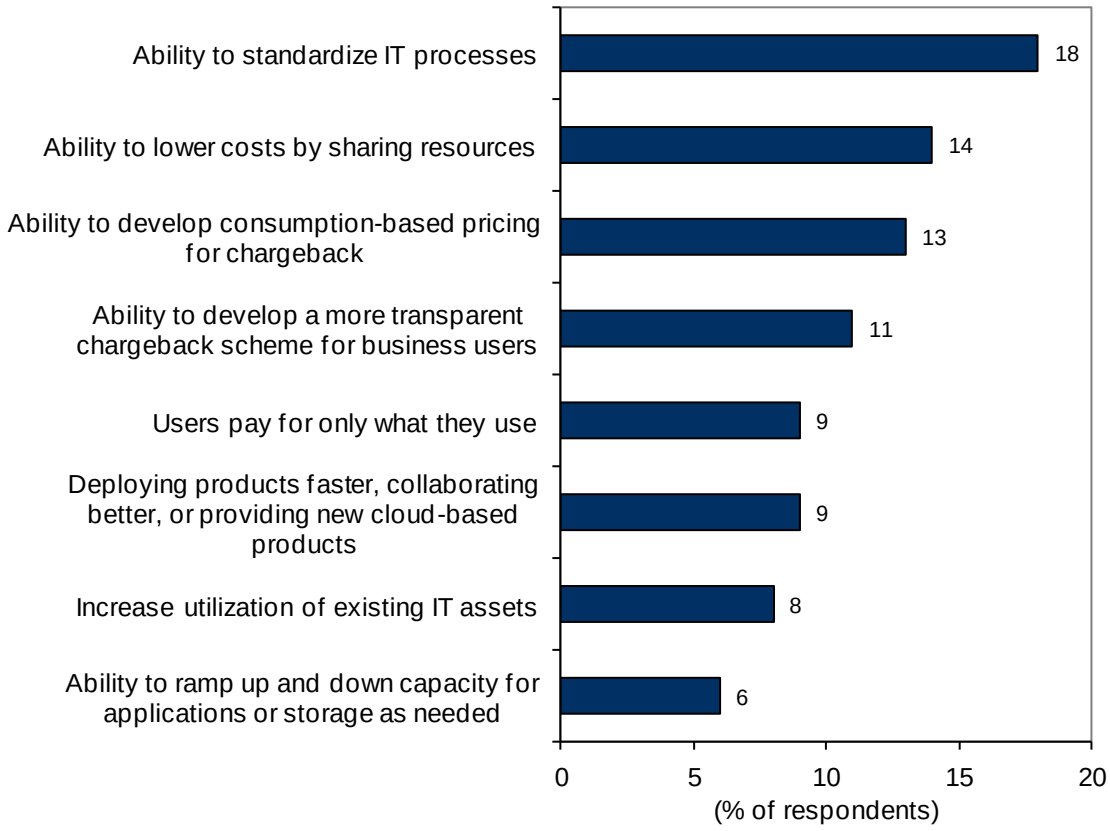
to implement cloud projects that leverage shared infrastructure and the necessary automation and management controls to run IT as a service.

An IDC cloud study of 800 firms conducted in June 2012 cited cost savings and faster time to market as the two most critical benefits of cloud. However, cloud attributes including user self-service, standardization of IT services, elastic scaling, metering/chargeback, and flexible migration between public and private clouds were also important. Among firms considering or deploying cloud services, concerns with public cloud include lack of interoperability standards, data/system portability, and lack of migration choice/flexibility. However, private cloud infrastructure, both on premise and hosted in a virtual private cloud or across a dedicated infrastructure, can address these concerns. As Figure 1 shows, datacenters are considering and/or deploying private or hosted clouds for three leading reasons: to standardize IT processes, to lower costs by sharing resources, and to develop consumption-based pricing for chargeback.

FIGURE 1

Drivers for Use of Private Cloud

Q. Which of the following best describes the primary reasons why you are using — or planning to use — a private cloud?



n = 800

Note: Responses of 4% or less were removed.

Source: IDC's *CloudTrack Survey*, 2012

Private Cloud and Hosted Services Cloud Infrastructure

In principle, a private cloud is implemented by a single company but has the same infrastructure attributes of elasticity, shared infrastructure, secure multitenancy, self-service provisioning, and so on as a public cloud. A private cloud can be managed internally or by a third party and hosted internally (within the company's datacenter) or externally. However, obstacles of security, service reliability, and internal IT staff knowledge required to reorient IT operations to private cloud are concerns with private cloud. Security concerns stem from leveraging shared infrastructure and broader access to critical IT services. Private cloud features, including stringent role-based permissions, integration of service catalogs, secure multitenancy, two-factor authentication, and encryption can address security concerns. Limited internal staff capabilities can be met with training/certifications and/or the deployment of a private hosted cloud that leverages service provider skills. A system that is easy to learn and use gives firms an advantage in terms of staff proficiency. Private cloud service uptime, response time, user downtime, and reliability SLAs are satisfied, in part, with the right underlying storage infrastructure.

As this research highlights, firms starting to implement a private cloud infrastructure often face issues in the infrastructure and storage domains, which can stymie the overall functionality and benefits of a private cloud offering. Storage vulnerabilities in private and hosted clouds are identified in the following section.

Storage Challenges with Private and Hosted Cloud Deployments

Deployment of a private or hosted cloud requires a range of technology components. Compute and storage virtualization are table stakes. Also required are self-service provisioning, automated provisioning, workflow orchestration, SLA management, monitoring, metering, and chargeback. Achieving these capabilities requires the integration of storage management functions with hypervisors and cloud orchestration tools. Lack of such integration is the single biggest storage challenge in private and hosted cloud deployments.

Automation for Standardized Storage Services

With virtualization, the time to provision new systems has been dramatically reduced. However, separate workflows and provisioning processes are still required to carve out and present logical unit numbers (LUNs) to the hypervisor. This creates separate and often disjointed workflows that can introduce delays and errors. For a private cloud deployment, a single provisioning action and workflow process should provision the virtual machine (VM), networking, and physical storage. This requires that a set of storage services (or a storage services catalog) be presented, which allows an IT generalist to easily select the right storage policies based on a given VM or application profile. Firms deploying private clouds at scale require automation via cloud management frameworks, such as OpenStack.

Enterprise Storage Complexity

Configuring storage for a private cloud deployment can be anything but efficient and automated. Administrators must select RAID groups, set up multipathing and host configurations, determine thin or fat volumes, and configure snapshots and/or replication.

Depending on the storage used, painful LUN resizing or expansion processes may be required. If more storage capacity is required, a rebalancing project may be needed to manually balance usage of the storage infrastructure. Once storage is configured, ongoing tasks such as distribution planning, load balancing, performance monitoring, and capacity or performance upgrades must be conducted. This specialized skill in managing storage inhibits the ability of general-purpose IT staff to provision storage for their own assets.

> The issues we had were four different types of storage: It became harder and harder to manage the data. To get TBs free to link them to a new server was very difficult. Storage configuration was tricky. We had trouble extending the current storage capacity — being able to leverage storage that was underutilized while expanding storage in other places; doing this without bringing down any systems was hard."
> — Executive Director, Educational Institution, Germany

Insight in Storage Resources Among Tenants

Metering, chargeback, and showback are fundamental benefits to leveraging shared storage infrastructure across multiple departments or organizations. IT organizations as well as lines of business seek visibility into how storage resources are being consumed and perhaps ultimately to pay for only the resources they consume. However, it remains very difficult for a cloud team to get an accurate picture of storage resource allocation, usage, and cost according to "logical" organizational or departmental boundaries. Leveraging private cloud infrastructure across different IT groups, business units, and subsidiaries requires a level of granularity in resource tracking and utilization.

Standardizing Services, Satisfying Business-Driven SLAs

IT professionals seek to provide faster time to infrastructure deployment, more rapid response to new business projects, and the ability to scale infrastructure dynamically to maximum capacity periods. The challenge is meeting business-driven time-to-market requirements while having only weeks to get infrastructure services up and running. To meet specific SLAs, IT organizations are defining standard services (or what one firm called "common patterns") within the infrastructure.

The ability to meet documented SLAs on performance, resiliency, availability, and recovery is heavily dependent upon the storage infrastructure. Cloud SLAs mandate nondisruptive maintenance windows, including microcode and hardware upgrades. All components in a private or hosted cloud storage system must be able to be added without bringing down user and application access or introducing performance degradation. Performance requirements vary based on the storage tier, but performance response times are measured in milliseconds, even during overhead processes such as RAID rebuilds. Storage performance needs to keep pace with the downstream demands that cloud I/O and virtualization place on storage processing. Hotspots can be prevented with the right storage architecture. When hotspot conditions are encountered, the system must be able to dynamically respond, ensuring that processing is balanced across applications and users.

Dynamic Scale, Mixed Workloads, and Performance and Capacity Optimization

While improvements in CPU processor speed and computing power have been realized, many workloads with higher IOPS and/or lower latency requirements become spindle bound. Server virtualization and client virtualization exacerbate the performance problems even further by increasing the amount of and randomization of I/O. The use of techniques such as short-stroking drives, adding more spindles, and/or introducing SSDs to deal with performance constraints is common. However, private cloud services seek not only to deliver IT efficiency and faster response times but to do so in a cost-effective manner while maximizing utilization of infrastructure.

Private and hosted clouds seek to deploy storage architectures that can support a range of workloads. With traditional storage deployments, the storage vendor wants to optimize the system for the workload, databases, Exchange, etc. With a cloud offering, the cloud architect or provider likely will not have insight into how the customer will be using the storage. Thus, manual tuning for a specific workload is not a viable option. The customer will ask for storage and for an SLA, and that's it. However, not many storage vendors can deal with mixed workloads.

> We can't just buy a general storage box. We need 'specialized hardware,' and what that means is that this is SLA driven. I need hardware that can mix any workload and be able to support a certain SLA for our performance tier. — Storage Architect, Leading eCommerce Company, United States

IBM STORAGE SOLUTIONS OPTIMIZE CLOUD DEPLOYMENTS

IBM has a differentiated portfolio of storage solutions including scale-out, block-and-file storage systems, storage management software, data protection and recovery offerings, and cloud management being deployed in private and hosted cloud architectures today. At the center of IBM's private and hosted cloud strategy is the IBM XIV storage system, which can be deployed standalone or behind an IBM SVC virtualization engine or SONAS gateway. XIV's scale-out storage architecture offers storage automation, ease-of-use, performance, and reliability attributes that are required for private and hosted cloud deployments. XIV also integrates with TSM, TPC, and SmartCloud storage offerings to provide a comprehensive data protection, storage management, and cloud management strategy. Details of the integration between XIV and IBM's other private and hosted cloud storage offerings are included in the next section. First, we provide an overview of the XIV architecture and its distinguishing features.

IBM XIV AND HOW IT ELIMINATES STORAGE VULNERABILITIES WITH PRIVATE AND HOSTED CLOUD

From IDC's perspective, the following features distinguish XIV from other block storage architectures available in the industry:

- A distributed grid-based storage approach with massive parallelism

- Unique data distribution architecture/algorithms
- Performance and reliability attributes
- Virtualization, cloud, and open source integration
- Easy-to-use GUI and provisioning simplicity
- Modular design in which each module has its own multicore CPU, RAM cache, SSD as cache (optional), and commodity HDD capacity, serving as a single elastic data resource
- No tiers to manage, which eliminates complexities and offers strong performance gains over traditional RAID

The XIV architecture is virtualized from the ground up, with the ability for massive parallelism to gain enterprise performance from commodity components in a single (drive) tier of storage. The XIV management model is differentiated by the ease-of-use and storage management features the architecture affords.

Another factor differentiating the XIV approach from market alternatives is the inclusion of data protection functions such as snapshots and replication services natively with the array. The XIV system architecture is not based on internal storage tiers; instead, the whole system, comprising both HDD and SSD, is presented as a single tier to applications and users. Thus, in effect, more drives are working for each application, which offers massive parallelism and performance advantages.

XIV Architecture

The XIV distributed architecture comprises a series of modules that are interconnected and communicate over an internal, redundant InfiniBand or Ethernet network. Each module consists of its own multicore CPU, RAM cache, SSD as cache (optional), and HDD capacity, which all work in parallel. This parallelism is important because the XIV system distributes storage processing among the nodes or modules. Thus, as the number of storage modules increases, so does the amount of parallelism available in the system. The XIV system leverages these modules to present a single large, elastic data store to the application network. A flexible, distributed cache allows the XIV system to leverage large slots for reads while managing a smaller slot size, resulting in a superior cache hit ratio and, consequently, better performance. Aggressive prefetching is enabled by the large cache-to-disk bandwidth available within each module, together with the extremely large aggregate module interconnectivity bandwidth that is available on the XIV backplane.

As data enters the XIV system, it is randomly and evenly distributed across all modules and disks in the system through XIV's data protection scheme. XIV stores and distributes data by breaking it down into 1MB chunks called "partitions," each mirrored for redundancy to another module, ensuring no single point of failure. The system distributes all the partitions automatically and uniformly across all the disks by means of a pseudo-random distribution algorithm. This approach to resiliency offers performance gains over traditional RAID rebuild times while eliminating the overhead in initial RAID configuration. The XIV system automatically distributes the application load across all modules evenly,

putting the combined power of all modules at the service of all requests concurrently. When the number of functional disks or modules changes, the system is rebalanced to ensure optimal data layout and, consequently, optimal use of system resources at all times — without the need for manual intervention or capacity rebalancing.

With the XIV architecture, the storage system does not need the tuning or changes that other enterprise systems require. When changes to an application or its I/O pattern occur, XIV automatically rebalances resource utilization across the system. Having a single drive type and pool within the storage system negates the need for storage tiering complexities. To ensure application performance in the storage pool, XIV can set quality-of-service (QoS) characteristics through Performance Classes, limiting IOPS and MBps throughput per host. This eliminates the need for storage administrators to monitor and tune storage for performance over time. The addition of disk modules is accompanied by a corresponding increase in processing power, cache, storage capacity, and connectivity. According to IBM internal research, for OLTP-type workloads, SSD caching provides up to 4.5 times the performance without much extra cost. The system monitors its components continuously, reacting to existing and potential issues by activating self-healing as needed and returning to full redundancy rapidly and without human intervention. As part of this proactive approach, the system uses disk-resident diagnostics to predict potential disk failures. It accesses all disk drives and areas to assess health status and increases protection levels by retiring suspect disks before they fail and rebuilds their data while a redundant version is still available. Further, XIV storage is server agnostic, working well with IBM hosts and other vendor systems.

XIV Integration Within the Cloud Ecosystem

Hypervisors form the underlying platform for a private or hosted cloud. Datacenters with private and hosted clouds must support a range of hypervisors. XIV supports all the major virtualization platforms, including VMware vSphere, Microsoft Hyper-V, Xen, and KVM. Many leading private and hosted clouds are using open source technology and standardizing on OpenStack. As a result, integration with the OpenStack Cinder API or other cloud frameworks such as VMware vCloud and IBM cloud management tools is essential. In OpenStack environments, which currently use only the iSCSI protocol, XIV enables high performance by means of high-bandwidth 10GbE connectivity. XIV's support for this API differentiates the offering from other storage arrays, although they will follow in time.

For cloud management, XIV can be integrated with the IBM SmartCloud storage family. Additionally, XIV can serve as the high-performance storage architecture behind IBM SmartCloud Enterprise+ deployment. XIV supports VMware vStorage APIs, including VAAI and VASA, and also offers plug-ins to virtualization management frameworks, including vSphere clients as well as support for vCloud Director and Hyper-V System Center, among others. Microcode within XIV supports Microsoft System Center Virtual Machine Manager (SCVMM) 2012 out of the box. The IBM plug-in for vCenter (IBM Storage Management Console for VMware vCenter) combines the benefits of the XIV architecture with storage management functions that can be performed from the vCenter client. Storage functions such as volume creation, autodiscovery, capacity reporting, and monitoring can be performed. This integration with virtualization is essential for private and hosted cloud deployments because it enables greater visibility into and management of storage by a broader set of IT administrators.

XIV has an adapter for VMware vCenter Site Recovery Manager (SRM) that helps support business continuity solutions by synching storage failover with SRM failover. IBM SVC also contributes virtualization advantages by supporting VAAI and SRM and offering visibility through the IBM vCenter plug-in. For overarching cloud management, enterprises can leverage the IBM SmartCloud storage family, integrating XIV storage, SONAS, and SVC. XIV and SVC system integration with OpenStack Cinder/Nova-Volume is an IBM strategy for addressing the growing trend to adopt cloud management frameworks.

> XIV is ideal for cloud delivery of services. We use all of the tools within the XIV device to support cloud strategies, including virtualization and thin provisioning, so we can add capacity very quickly and scale up very easily. It's very, very elastic — it's very, very scalable — all of the things that cloud defines in terms of reactivity, and it has responsiveness to requirements. XIV is designed to support those sorts of business requirements. — Phil Clark, Channel and Partnership Director, niu Solutions, United Kingdom

IBM SONAS and SVC and XIV Integration

Storage for private and hosted cloud infrastructure from IBM includes IBM XIV, SONAS, and SVC. SONAS is IBM's scale-out NAS offering, ideally suited for applications ranging from high-performance analytics and rich media to computer-aided design, oil and gas exploration, and genomic sequencing that require SONAS' large addressable namespace and high-performance scalable storage. SONAS is being used in private cloud deployments in financial services, research, and educational institutions to provide petabytes of storage under a single management interface. XIV supports integration with the SONAS gateway, which enables administrators to dedicate some or all of the available storage capacity in the XIV storage system for network file serving. This enables XIV to provide a unified storage approach, with Fibre Channel, iSCSI, and NFS/CIFS protocol support, and extends the value of IBM in private and hosted cloud deployments.

IBM SVC is an innovative block-based storage virtualization platform that virtualizes and combines storage capacity from different storage architectures into a single resource pool. An SVC system is configured as a series of LU nodes and deployed in pairs that sit between application servers (hosts) and the storage arrays. The SVC system presents itself as a SCSI target to the hosts and as a SCSI initiator to the storage arrays. The SCSI LUNs that the hosts see are virtual disks or volumes, which are allocated by SVC from its virtualized storage pool. Administrators can manage a range of storage systems with a common approach and also optimize capacity utilization, which makes provisioning easier and enables nondisruptive data migration and copy services between disparate storage products. SVC is being used in private clouds today for data migration and ongoing virtualization of storage. An XIV system or multiple XIV systems can sit behind an SVC cluster to offer a scale-out block storage architecture, as well as coexist with other heterogeneous storage.

IBM TSM, TSM for Unified Recovery, and FlashCopy Manager and XIV Integration

For optimal protection and recovery of data and systems in a cloud, IBM offers TSM, TSM for Unified Recovery, and FlashCopy Manager. TSM is an enterprise-class protection and recovery solution for a broad range of platforms, applications, and virtualized infrastructure in both centralized and distributed locations. As more virtualized environments use array-based snapshots and clones to enable fast protection and recovery of applications and images, FlashCopy Manager offers a centralized, robust application-consistent hardware-based local and remote snapshot protection approach. TSM for Unified Recovery is a suite that includes all IBM's protection and recovery offerings — TSM, FlashCopy Manager, and FastBack — for end-to-end protection and recovery from a centralized management framework. For protection and recovery, XIV comes standard with its own snapshot and replication functions. However, XIV snapshots can also be centrally managed using FlashCopy Manager. For centralized backup, XIV data can be protected in an application-consistent manner using TSM, which supports tape, disk, and deduplication target appliances such as ProtecTier.

IBM TPC and SmartCloud and XIV Integration

XIV systems integrate with the IBM SmartCloud storage family to enable the automation of cloud storage services, including provisioning, monitoring, and access in a private or hosted cloud deployment. IBM TPC is the company's management solution for virtualized, heterogeneous server and storage environments, offering storage analytics, reporting, and control. TPC is the management framework to control all IBM storage systems, including XIV. Management of multiple XIV systems can be centralized and functions automated through TPC. Critical information on XIV health, performance, and utilization can be monitored and reported. The IBM SmartCloud family includes SmartCloud Monitoring, SmartCloud Provisioning, and SmartCloud Storage Access.

SmartCloud Monitoring provides visibility into physical and virtual private cloud infrastructure, monitoring health, capacity, and performance of heterogeneous private cloud infrastructure. What-if modeling and analysis of changes made in the environment allow cloud architects to mitigate risk and ensure that SLAs are not compromised. Policy-driven analytics ensures infrastructure is balanced according to performance and security rules. SmartCloud Provisioning enables consistent, dynamic application deployment, automated administration, and virtual machine image management in private and hosted clouds. It can be used to deploy cloud middleware and applications.

SmartCloud Storage Access is a self-service software portal for NAS storage provisioning and monitoring that includes reporting. In a few steps, users can provision storage and share files over the Internet. Administrators can easily monitor and report storage usage and classify storage for different organizational users. SmartCloud Storage Access also works with the IBM SONAS offering and with XIV behind a SONAS gateway, which enables fast time to value for firms deploying private cloud infrastructure-as-a-service capabilities.

WHY XIV FOR PRIVATE AND HOSTED CLOUD DEPLOYMENTS?

The successful deployment of a private or hosted cloud hinges on the right storage architecture. XIV can help firms overcome storage vulnerabilities associated with cloud deployments by offering the following features and benefits. Comments from customers evaluating and using XIV in private and hosted cloud deployments are included.

Agility and Elasticity

The XIV system can handle changing provisioning patterns across mixed, dynamic workloads and provide resource sharing across multiple tenants while maintaining consistently high performance. This is important as tenants enter or exit the cloud. The XIV architecture not only can scale to support peak workload demands that occur at the close of a quarter or as a result of marketing campaigns and seasonal activity but also can optimize for average workloads, thus supporting changing business patterns.

> With XIV, it's easy for me to meet the SLA requirement [for that performance level] because it can absorb load and come back with a relatively fixed latency, and the throughput is relatively consistent. — Storage Architect, Leading eCommerce Company, United States

Starting out with a minimum configuration of 56TB and scaling to 243TB across 15 XIV modules, the system achieves linear growth in *both* capacity and performance. With the XIV management model and user interface, it's quicker to create a virtual volume than a virtual machine, supporting faster time to market. Together, XIV and SONAS support scale-out configurations required by private and hosted clouds today. SONAS, with its global namespace, provides almost infinite scalability, while XIV performance scales with storage capacity. With these features, storage complexity, dynamic workload SLAs, storage provisioning overhead, and performance and capacity obstacles with private and hosted cloud deployments are overcome.

> The challenge was looking for a storage vendor and offering that handled mixed loads, and that had consistent SLAs, for performance, throughput, and latency, as well as tiering them around availability appropriately. We checked, and only three vendors came out that could do what we wanted done. XIV was one of them. — Storage Architect, Leading eCommerce Company, United States

Ease of Management

The XIV architecture provides storage resources that require no tuning. Monitoring and management of the storage array are not required to keep performance consistently high and hotspot free. Storage can be provisioned and used by IT generalists.

> It [storage provisioning] used to take one to two days to initiate. The provisioning operation of running the storage can now be done by email. It takes time to process the email and determine the priority of the request, but now it is down to minutes. The technical time to deliver storage is less because negotiation is no longer necessary. It can be configured in just minutes. — Executive Director, Educational Institution, Germany

Performance is paramount in private cloud deployments, and the XIV architecture ensures that application response times, as measured in milliseconds, can be sustained. This zero tuning approach is essential for cloud providers, where usage is not known, and thus a cloud architect cannot safely tune or optimize the storage. Other storage architectures leverage storage tiers, moving data between tiers based on usage/performance patterns. The XIV approach has no tiers to manage, so intensive storage planning is not required and data does not need to be moved between tiers, introducing performance overhead.

> XIV's main advantage was it was one big storage [asset] and we could create slices from it and build internal drives. It solved a lot of complexity problems for us. — Executive Director, Educational Institution, Germany

Multiple XIV systems can be managed centrally, via the XIV UI or through IBM TPC. XIV supports integration with cloud management systems, including OpenStack via the Cinder API to provide storage automation, minimizing the effort and overhead associated with storage provisioning and deprovisioning.

> We looked at management. Because of the bar we are setting on APIs, a lot of storage vendors fell off. IBM was willing to make an effort to integrate with Cinder and give us the open stack and onboarding capability. APIs are one of the things that we're standardizing on. — Storage Architect, Leading eCommerce Company, United States

> So in terms of cloud services, we run predominantly private cloud services for our clients. We don't necessarily know what the application stack is that those clients are going to be accessing. We do know that it's XIV in the back. So when a client requires another four or five or six or ten virtual machines, they simply go into a portal, spin up those machines, and they're done. There was no consideration whatsoever about what the storage impact of that was going to be. It's just there, and it just runs. That's probably the best thing I can say about the product; it just works. So, from our point of view, the management overhead of actual sign-in storage on those cloud-based systems is minimal. We provision space once, and then that's it. The rest is automatic." Stephen Bedford, Technical Director, niu Solutions, United Kingdom

The XIV system captures detailed used-capacity information, which can be integrated with external cloud billing systems for showback or chargeback use cases. Although the XIV UI is advanced, cloud architects and providers want to automate storage management functions via CLI commands (or APIs). XIV management tasks can be automated through scripts using these commands. Further, by leveraging SVC, private cloud applications and data can be nondisruptively migrated from legacy storage architectures to XIV.

Availability, Reliability, and Performance

XIV architecture affords extremely high performance. The self-healing architecture and internal data redundancy prevent and minimize downtime, which is a critical SLA factor in cloud deployments. The XIV system can be upgraded in terms of both hardware components and microcode/software without any downtime, thus minimizing performance degradation. The system rebuilds in the event of disk or module failure, and the rebuild times can be extremely fast compared with RAID alternatives. There is almost no performance degradation during rebuilds.

> When there is a drive failure, we rebuild within a short period of time … If we can minimize the exposure time, then the risk to the application is minimized. It's about an hour of rebuild, which is very good. — Storage Architect, Leading eCommerce Company, United States

> We found we would be able to shut down one half of the entire storage system, and the services would still be up and running. We also had high-performance, high-availability storage. — Executive Director, Educational Institution, Germany

Leading private cloud deployments demand millisecond response time for top-tier applications. The overall XIV architecture and the system's even and distributed use of resources, dispersing data across the entire tier, afford impressive price/performance with no hotspots and without the need for tuning.

Industry performance benchmarks, including those from the Storage Performance Council (SPC), found that IBM XIV systems rank high in price/performance, while the International Technology Group (ITG) found that "XIV performance was, by wide margins, reported to be superior to that of existing disk systems."

> XIV added SSD modules into the arrays. One of the SLAs asked for sub 6 milliseconds to get us to this gold standard. XIV proved it can achieve the 6 milliseconds in our large MySQL deployment. IBM XIV became a very good storage offering that allows us to do what most MySQL companies can't do. — Storage Architect, Leading eCommerce Company, United States

> We have seen big advantages in the performance; half the time we found that things ran 50% faster. — Executive Director, Educational Institution, Germany

The XIV system performance optimization enables superior utilization of system resources and automatic workload distribution across all system hard drives with no tuning. The XIV system's ability to maximize load distribution across all disks for all workloads, coupled with a powerful distributed cache implementation, facilitates high performance that scales linearly with added storage enclosures. In addition, because this high performance is consistent, cloud environments get the same performance during the typical peaks and troughs.

> One really big advantage is distribution of hotspots. We don't always know how heavy the concurrent usage might be. Dynamic allocation and monitoring for hotspots on the hard drives, and redistribution to support our workloads, has been a very unique feature. It makes it much easier to plan the information on an SAP system because we do not have to check before how the workload will be for this special SAP system. — Executive Director, Educational Institution, Germany

Low Total Cost of Ownership

As cited previously, cost savings is a large motivator for firms to consider private and hosted cloud deployments. The XIV system offers low cost per TB enhanced by its use of high-density drives. With its standard snapshot services, the XIV system affords optimal capacity utilization through space-efficient differential snapshots. Because XIV delivers consistently high performance independent of capacity, cloud administrators do not need to hoard capacity reserves for the sake of performance. With XIV, 90% capacity utilization can be achieved without noticing performance degradation, whereas traditional storage architectures typically require a 60% utilization safe zone to accommodate space reshuffling and avoid tiering policy violations.

Other TCO features include XIV's efficient power utilization and effective cooling within a small footprint. In addition, the intuitive XIV management schema minimizes training requirements and allows more administrators to use the XIV system more quickly.

> The usability of the solution … it really has a very good user interface, where you can get monitoring information [and] easily create LUNs that you can mount on your OS, and it's very convenient to handle. I would say this is state of the art in terms of the user interface and how you can manage a storage system. — Executive Director, Educational Institution, Germany

Cloud providers as well as private cloud deployments require secure multitenancy; that is, reliable and performance-sustaining use of shared resources, including storage, among tenants. Tenants can include different organizations or companies. Thus, it is essential that storage architectures enable the use of their resources by different groups and workloads as well as support SLA requirements. The XIV storage architecture ensures that virtual machines and datastores are evenly distributed across every drive in the XIV system, allowing changes in performance or application requirements to be managed without manual tuning. The resource

utilization of the IBM XIV storage system provides ideal load balance — without degrading volume utilization and thus optimizing utilization.

IBM SVC in the cloud eliminates storage islands, improves capacity utilization, and streamlines management, which can limit expenses and conserve time. With SONAS in place, the resulting consolidation, centralization, and automation reduce the administration of petabytes of files for financial and time savings. SONAS floor space conservation helps lower capital and operational expenditures.

CHALLENGES AND FUTURE OUTLOOK

Firms are still analyzing how best to leverage cloud services in a private, public, or hybrid fashion, and they are considering what workloads are best suited to each of these deployment models. Firms that have reached an extensive level of virtualization are looking to take the next step in deploying a private cloud. However, this requires integrating different infrastructure components and streamlining, through integrated workflows, the provisioning of new applications and infrastructure. Today, many of these workflows remain standalone and separate.

However, private clouds offer firms the longer-term opportunity to integrate disparate processes and the ability to manage IT infrastructure as an integrated set of services. IBM can offer workshops and advisory services to help an enterprise formulate a cloud strategy. IBM's range of cloud management and automation offerings can integrate separate workflows around cloud services and provisioning. IBM is taking a leadership position in open source cloud initiatives such as OpenStack.

Driving to cloud is a business decision. The next step firms are taking with their private clouds is adding business logic into the infrastructure and bringing business intelligence into the cloud. This presents a strategic opportunity for IBM to help businesses transition private clouds from an operational advantage to a revenue-generating mechanism. Firms can leverage IBM's core competencies in consulting services, business analytics, and Big Data technologies as they build private and hosted clouds of the future.

CONCLUSION

IT organizations and service providers alike are moving to cloud architectures. Cloud objectives of improved IT efficiency, better business responsiveness, and greater economic advantages can be stymied with the wrong storage architecture and integration. IBM's Smarter Storage strategy eliminates storage vulnerabilities that firms can face in private or hosted cloud architectures. IBM's block and file–based scale-out storage architectures form the underpinning of a highly agile, performance-based, and reliable cloud infrastructure.

IBM XIV storage systems are optimized for cloud deployments because of their ease-of-management, reliability, and performance characteristics. The IBM SmartCloud storage family enables the automation of cloud storage services, including provisioning, monitoring, and access. The TSM family optimizes recovery of data and systems in a cloud deployment. IBM's storage portfolio offers integrated cloud offerings and is well

positioned today and into the future to enable enterprises to deploy cloud architectures. Increasingly, firms are selecting XIV, including solutions with SVC and SONAS, as the underlying storage architecture for private cloud and hosted cloud deployments.

This document was developed with IBM funding. Although the document may utilize publicly available material from various vendors, including IBM, it does not necessarily reflect the positions of such vendors on the issues addressed in this document.