



Virtualize More, Manage Less:

IBM System Storage SAN Volume Controller Update SVC 4.3 + TPC

Dominique SALOMON

European Product Introduction Center (PIC) Leader Storage Specialist Certified – IBM Montpellier

Dominique.Salomon@fr.ibm.com

© 2008 IBM Corporation



IBM System Storage SAN Volume Controller

What is Virtualization?

Logical representation of resources not constrained by physical limitations

- Create many <u>virtual resources</u> within single physical device
- Reach beyond the box see and manage many virtual resources as one
- <u>Dynamically change and adjust</u> across the infrastructure





Why Storage Virtualization?

- Not "just another way of helping manage SANs"
- Storage virtualization complements server virtualization
 - Both technologies help increase flexibility and speed responsiveness
- Storage management used to be <u>manually intensive</u>, timeconsuming and <u>disruptive</u> to the business
- Storage virtualization with SVC can help change that to automatic, time-saving and non-disruptive to the business
- Radically changes the way you think about and work with storage to make it fundamentally more flexible than just disk boxes alone



SAN Volume Controller Facts

- IBM has 40 years experience in virtualization technologies
- IBM has shipped over 10 000 SVC engines (SVC nodes) running in more than 4 000 SVC systems (SVC Clusters)
- There are more than 130 customer references and 24 customer case studies for SAN Volume Controller
- SAN Volume Controller is a proven offering that has been delivering benefits to customers for four years
- SAN Volume Controller demonstrates scalability with the fastest Storage Performance Council benchmark results
- SAN Volume Controller can virtualize IBM <u>and</u> non-IBM storage (over 120 systems from IBM, EMC, HP, HDS, Sun, Dell, NetApp, Fujitsu, NEC, Bull)





SAN Volume Controller : Concepts & Features

© 2008 IBM Corporation



IBM System Storage SAN Volume Controller

TRM

SAN Volume Controller Concepts





TRM

SVC : Disk block virtualization concepts





SAN Volume Controller Delivers Value



Reduces the cost and complexity of managing storage

- Creates tiers of storage
- Enables multivendor strategies

Improves business continuity

- Supports data movement without interrupting applications
- Allocate more storage to applications automatically

- Improves storage utilization
- Combines storage capacity into a single resource – from multiple vendors
- Manage storage as a business resource, not as separate boxes

- Improves personnel productivity
- Manage a single storage resource from a central point



Infrastructure Simplification with SAN Volume Controller

Traditional SAN

- Capacity is isolated in SAN islands
- Multiple management points
- Poor capacity utilization
- Capacity is purchased for, and owned by individual processors



SAN Volume Controller

- Combines capacity into a single pool
- Uses storage assets more efficiently
- Single management point
- Capacity purchases can be deferred until the physical capacity of the SAN reaches a trigger point.



Non-disruptive Data Migration with SAN Volume Controller

Traditional SAN

- 1. Stop applications
- 2. Move data
- 3. Re-establish host connections
- 4. Restart applications

SAN Volume Controller

1. Move data

Host systems and applications are not affected.





Business Continuity with SAN Volume Controller

Traditional SAN

- Replication APIs differ by vendor
- Replication destination must be the same as the source
- Different multipath drivers for each array
- Lower-cost disks offer primitive, or no replication services

SAN Volume Controller

- Common replication API, SAN-wide, that does not change as storage hardware changes
- Common multipath driver for all arrays
- Replication targets can be on lower-cost disks, reducing the overall cost of exploiting replication services





SAN Volume Controller : Hardware

© 2008 IBM Corporation

IBM System Storage SAN Volume Controller

SVC Elementary modules (Hardware nodes)

- SVC is a mix of IBM Storage Software and IBM hardware (SVC nodes)
- SVC nodes are based on IBM System x3550 servers which have following characteristics :
 - Two dual-core Intel Xeon 5160 processors at 2.33GHz (Dramatically improved throughput compared with previous generation of nodes)
 - 8GB of cache
 - Four 4Gbps FC ports
 - Fully mirrored cache between pairs of nodes
- Designed to provide high performance
- Scalable solution :
 - Minimum configuration : 2 nodes
 - Upgradable up to 8 nodes (by adding pairs of nodes)
- New engines may be intermixed in pairs with older engines in SVC clusters (Helps protect investments and offers enhanced growth capability)
- Cluster nondisruptive upgrade capability may be used to replace older nodes with new 8G4 nodes.

SAN Volume Controller : Copy Services

© 2008 IBM Corporation

IBM System Storage SAN Volume Controller

SVC Advanced Copy Services

FlashCopy :

- Designed to create copies for backup, parallel processing or test purpose which are available almost immediately for use
- FlashCopies across heterogeneous SVC supported disks
- Incremental, Cascaded FlashCopy capabilities
- Up to 256 copies of a single source volume (new with SVC 4.3)
- Space-Efficient Flash Copy (new with SVC 4.3)

Metro Mirror :

- Metropolitan" distance synchronous remote mirroring function
- Up to 300km between sites for business continuity
- Global Mirror :
 - Long distance <u>asynchronous</u> remote mirroring function
 - Up to **8000km** distance between sites for business continuity

Note : SVC is one of the few products to support cascaded copies (copies of copies at later points in time)

SVC Metro Mirror Function

- "Metropolitan" distance synchronous remote mirroring function
- Up to 300km between sites for business continuity
 - As with any synchronous remote replication, performance requirements may limit usable distance
- Host I/O completed only when data stored at both locations
- Designed to maintain fully synchronized copies at both sites
 - Once initial copy has completed
- Metro and Global Mirror delivered as single feature
 - Offers great implementation flexibility
- Operates between SVC clusters at each site
 - Local and remote volumes may be on any SVC supported disk systems

SVC Global Mirror Function

- Long distance asynchronous remote mirroring function
- Up to 8000km distance between sites for business continuity
- Does not wait for secondary I/O before completing host I/O
 - Helps reduce performance impact to applications

- Once initial copy has completed
- Built on Metro Mirror code base
- Metro and Global Mirror delivered as single feature
 - Offers great implementation flexibility
- Operates between SVC clusters at each site
 - Local and remote volumes may be on any SVC supported disk systems

SAN Volume Controller : V4.3 New features

© 2008 IBM Corporation

IBM System Storage SAN Volume Controller

SVC 4.3 Major new features

Space Efficient Vdisks (SEV)

- Can be used for thin provisioning (also referred as storage over-allocation)
- Can be used for Space Efficient Flash Copy (SEFC)
- Can configure Vdisks to hosts in such a way that the sum of the configured vdisk capacity within a cluster is greater than the physical disk capacity in the managed Storage Pools within that cluster.
- Space Efficient FlashCopy
 - SEFC is the capability to have a target of an SVC FlashCopy to only consume the minimum required disk capacity for the changed bits, and not consume capacity equal to the source.

Vdisk mirroring (RAID1)

- A given vdisk has two complete copies of mdisk where the data is saved within the cluster.
- mdisks can come from different mdisk groups and have different mdisk extent sizes
- Improves high availability configurations by allowing vdisks to be mirrored across mdisks, and therefore stay online, when an mdisk group goes offline.
- Mirrored Vdisks can be Space Efficient & used in SVC Copy Services

1) Space Efficient Vdisk SEV – Terminology

- Virtual Capacity
 - The full capacity of the VDisk and what the host sees.
 - For a fully allocated VDisk, this is the only capacity term that applies.

Real Capacity

- The actual physical storage (MDisk extents) pre-assigned to the VDisk. At VDisk creation this value is specified by "-rsize".
 - For a fully allocated VDisk virtual and real capacity are the same.
- Used Capacity
 - The subset of Real Capacity allocated to grains for storing actual customer data

Free Capacity

 The difference between Real Capacity and Used Capacity available to be assigned to grains

Contingency Capacity

 The difference between Real and Used Capacity that SVC tries to maintain as Real Capacity is used. Comes in to play when Auto-Expand is enabled

Space Efficient Vdisk Concept

Users can create VDisks that have different Real and **Virtual Capacities :**

> •Real Capacity defines how much disk space is actually **allocated** to a VDisk

•Virtual Capacity defines how large the VDisk **appears** to the host

Space Efficient Vdisk - details

- If the Used Capacity reaches the Real Capacity then the VDisk will go offline and application I/Os will fail
- Auto-Expand option to have SVC automatically add Real Capacity to the Vdisk when needed
- SVC helps you to avoid exhausting the Real Capacity allowing for alerts to be sent to admins to provide space or by increasing Real Capacity automatically
- User can choose the granularity (Grain) at which space is allocated - 32, 64, 128 or 256K Bytes
- It is never de-allocated even when host deletes files.

Space Efficient Vdisk – Out Of Space

If a Space-Efficient VDisk runs out of space it goes offline, not read-only

- An error is logged indicating the VDisk has been taken offline
 - Can be configured to generate an email or SNMP trap
- For <u>non Auto-Expand</u> VDisks the user must manually increase the real capacity
- For <u>Auto-Expand</u> VDisks the user must add capacity to the Managed Disk Group or migrate the VDisk

Use the warning thresholds to give advance notice of out of space conditions

Space Efficient Vdisk – When To Use

- Many open system applications typically run at <u>no more than 50% storage capacity</u> <u>utilization :</u>
- Traditional techniques are also to provision as large a disk as maybe needed next year, or the year after.

> In all cases, this generates a lot of <u>wasted disk space</u>.

- With <u>Space Efficient Vdisk</u>, you can start by allocating a small base size and let SVC adding more space but <u>only when really required</u> by the application.
- Warning : Space-Efficient VDisks only save capacity if the host server or application does not write to the whole VDisk
 - Some filesystems (e.g. NTFS) will write to the whole VDisk before overwriting deleted files, others will reuse space in preference to allocating new space
 - Filesystem problems may be mitigated by tools such as "defrag" or by managing storage using host LVMs
 - Some applications only delete log files when the filesystem is nearly full
- So, it is recommended to test this new feature with the applications before using it :
 - IBM Montpellier PSSC can help you to do such validation tests.

Space Efficient FlashCopy

- SVC 4.3 Space Efficient FlashCopy is based on Space Efficient Vdisk feature.
- In many customer environments, there can be multiple copies ("clones") of production systems being used for various purposes such as development, test, QA and training.
 - > With traditional replication techniques, each copy occupies <u>as much disk</u> <u>space as the original.</u>
- Using Space Efficient FlashCopy could significantly reduce the amount of space required because capacity would be used only to store <u>differences between the copies and the original data</u>.

Example

- A customer have a production database of 1TB and needs to create 2 FlashCopy images of this DB
 - With traditional FlashCopy, he would need to allocate 2x 1TB of physical space to create its 2 FlashCopy images
 - With Space Efficient FlashCopy, only physical space required to store modifications between the production DB and the 2 FlashCopy images will be really allocated, which is just a percentage of the production DB space (and not twice its size).

Space Efficient FlashCopy : technical details

Two distinct meanings

- 1 : Space-Efficient Source to Space-Efficient Target
 - Background copy process does not copy unallocated regions
 - Can use incremental feature to refresh mapping (once a full copy of the actually allocated regions is complete)

<u>2 : Fully-Allocated Source to Space-Efficient Target</u>

- Must have zero copy rate to avoid fully allocating the Space-Efficient target
- Provides CDP like function, but no reverse FlashCopy yet
- Try to use same grain size for FC mappings and SEV if possible
 - Beware that defaults are different, 32KB for Space-Efficient VDisks and 256KB for FC mappings

VDisk Mirroring

- VDisk Mirroring allows creation of VDisk with two copies
- The 2 copies can be in different Storage Pools (MDisk groups)
- Anything that can be done today with a VDisk can be done with a mirrored VDisk
- You can combine Space Efficiency and Mirrored Vdisk features
- Other available options of Vdisk Mirroring :
 - Add a copy to an existing VDisk
 - Copy added can be of any structure
 - No provision for combining two VDisks into a single mirrored VDisk.
 - **Remove a copy** from mirrored VDisk
 - Add and remove together can be used for migrations not currently allowed
 - Split a copy from a mirrored VDisk and create a new VDisk
 - Only allowed when copies are synchronized unless use –force

VDisk Mirroring – Using For Migration

- Can be used to migrate VDisk between MDisk Groups with different extent sizes (Vdisk migration restriction)
 - Add a copy in the new MDisk Group
 - Wait for synchronization to complete
 - Remove copy in original MDisk Group
- Migrate from <u>space-efficient</u> to <u>fully-allocated</u> VDisk
- Migrate from image/striped/sequential mode to a sequential VDisk
- Can control copy rate of migration and even suspend/cancel migration
 - Same controls as FlashCopy
 - Default copy rate setting is 50 which is 2MB/s

IPV6 support

- Can have both IPv4 and IPv6 addresses for a cluster and both can be used at once
- 8K Vdisks :
 - Simple doubling of the VDisk maximum
 - 2048 VDisks per IO Group (node pair)
 - 8192 VDisks per 8-node cluster

New Flash Copy Features

- 256 FlashCopy targets
 - In any combination of multiple targets or cascaded mappings from one source
- Auto-delete consistency group
 - Can delete consistency group when last mapping is deleted or moved out
- New maximum of 4096 mappings

SVC 4.3 Licensing changes

- Space-Efficient VDisks and Mirrored VDisks are <u>not licensed</u> <u>features</u>
- FlashCopy licensed now by <u>source</u> VDisk capacity only (and no longer on source + target capacity)

SVC Concurrent Code Upgrade

- Non-disruptive upgrade to new SVC code 4.3 :
 - Download SVC and SVC Console codes from web site (free)
 - Upgrade SVC Console first,
 - Then, upgrade SVC cluster (non-disruptive upgrade)
- Can upgrade directly from any post-3.1.0.5 level to 4.3
- For pre-3.1.0.5 levels, upgrade to 3.1.0.5 first

SVC 2008 & 2009 Roadmap

IBN

Systems

livoli.

SVC Roadmap

2008 :

- Support of DS5000, XiV and other storage subsystems,
- Support of IBM i / VIOS
- New entry-price nodes

2009 (2 releases) :

- iSCSI support
- Copy Services improvements
- Non disruptive Vdisks move between I/O groups
- New High End nodes : 12-48 GB cache, 8x 8GBps FC ports

SAN Volume Controller : References

© 2008 IBM Corporation

IBM System Storage SAN Volume Controller

SVC References

- New SVC case study document
- SVC web site

TRM

SAN Volume Controller demonstration

SVC 4.3 Demo : Physical Configuration

SVC 4.3 Demo : Logical Configuration

SVC 4.3 Demo : <u>Test Scenario</u>

Will be demonstrated during the demonstration :

- Creation of Space Efficient Vdisk without Auto-Expand first :
 - Write on the Vdisk and observe he goes Offline,
 - Change it to Auto-Expand and observe he goes back Online automatically and have Real size automatically allocated,

Create a Mirrored Vdisk

- Show the 2 copies details and how they are automatically synchronized,
- Create a Mirrored and Space Efficient Vdisk
 - Show how to split one copy,
- To finish, show how to create a Space Efficient FlashCopy.

SAN Volume Controller Management with TotalStorage Productivity Center - TPC

© 2008 IBM Corporation

IBM System Storage SAN Volume Controller

Storage Infrastructure Management

The Problem

 Heterogeneous storage infrastructures, driven by growth in file and database data, are consuming increasing amounts of administrative time. IT managers are looking for ways to make their storage administrators more efficient.

The Solution

 Empower administrators with automated tools for managing heterogeneous storage infrastructures

IBM

Four Major TPC Components

TotalStorage Productivity Center for Data

- Data collection and analysis, file systems and databases
- Reporting, chargeback and quotas
- Automated actions
- Support for heterogeneous disk (IBM, EMC, HDS, HP, Engenio)
- IBM 3584 Tape Asset Reporting

TotalStorage Productivity Center for Disk

- Disk subsystem management
- Performance management—IBM and heterogeneous storage
- Storage provisioning—IBM and heterogeneous storage

TotalStorage Productivity Center for Fabric

- SAN topology display and management
- Event reporting, performance reporting
- Security enforcement via zone control
- Heterogeneous fabric support (Brocade, Cisco, McData)

TotalStorage Productivity Center for Replication

- Single point of control for point-in-time and remote volume replication services
- Automated source-target matching
- Cross-device consistency groups, DS8000/ESS and SVC FlashCopy and Metro Mirror

Standard Edition Bundle

TotalStorage Productivity Center for Disk

Key Capabilities

- Configures multiple storage devices from a single console to improve productivity
- Monitors and tracks the performance of storage devices to optimize the SAN performance

Performance Management

- Collects, stores, alert on performance metrics
- Recommends optimized storage allocation
- Monitors and tunes storage
- Includes Integrated Device Management

TPC Enhances SAN Volume Controller

Asset and Capacity Reporting	 Physical characteristics such as the manufacturer, model, serial number, capacity, etc. Show the allocated and free capacity of every SVC on the network
Configuration Reporting and Management	 Reports on SVC's storage allocated to logical host volumes (which appear to hosts as disk drives) and the managed disks being used on the backside Display the physical managed disks behind what the host sees as a disk drive List all SVC volumes which have been allocated but aren't in use Show which hosts have access to a given SVC volume Show which hosts have access to a given disk drive (within the SVC) Show which SVC volume (and managed disks) a host has access to Discovery, Show the Storage Controllers (ex., ESS/DS4000) that provide volumes to SVC
Performance Management	 Information collection, display, thresholding, alerting
Basic and Automated Provisioning	 Integration with Tivoli Provisioning Manager
44 IBM System Storage	SAN Volume Controller BM System

TPC Disk Manager Functions for SVC

Performance Graphing - SVC

- <u>By subsystem</u>
- By controller
- By I/O Group
- By array
- By Managed Disk Group
- <u>By Volume (aka vdisk)</u>
- By Managed Disk
- <u>By Node</u>
- By Port
- Constraint Violations

Example : SVC Performance Monitoring by Volume

TPC Disk Manager Functions - Thresholds

Threshold types – various subsystems

- Overall Port response time (ms)
- Total port data rate (MB/sec)
- Total port IO rate (ops/sec)
- Cache holding time (sec)
- NVS Full Percentage (%)
- Total Data Rate (MB/sec)
- Total IO Rate (ops/sec)
- Disk Utilization Percentage (%)
- Overall Backend Response time (ms)
- Total Backend Data Rate (MB/sec)
- Total Backend IO Rate (ops/sec)

tem

TPC : SVC Performance monitoring examples

Conclusion

© 2008 IBM Corporation

IBM System Storage SAN Volume Controller

SVC: The Benefits are Real

Key Areas of Cost Saving Observed by Forrester in SVC Customers

Reduction in storage management and administration cost

 Allowing a core group of administrators to control multiple assets across a distributed storage environment (50% efficiency improvement)

Improved storage utilization

- Improve capacity utilization of existing storage assets
- Control the growth of future spending (improved utilization by 30%)

Reduced cost of storage

 Capitalize on being able to purchase the lowest cost storage resources (controlled growth on average by 20%)

Improved customer and end user availability to data-driven applications

 Minimize downtime associated with migrating data between storage assets (\$240,000 in annual savings)

Source: The Total Economic Impact™ Of IBM® System Storage™ SAN Volume Controller

Find Out More

SVC on ibm.com

ibm.com/storage/svc

Virtualization NewsCenter

ibm.com/systems/storage/news/center/virtualization/

SVC Support

ibm.com/servers/storage/support/software/sanvc/index.html

Storage Virtualization Blog

ibm.com/developerworks/blogs/page/storagevirtualization

SVC Sales Kit

IBM System Sales

tinyurl.com/ozhs3

PartnerWorld

tinyurl.com/z5u23

Who to Contact

Worldwide and Geo Sales

Worldwide

- Brian Hamel, VP System Storage Sales
 - bhamel@us.ibm.com
 - Office: + 1 404 238-4850
- Steve Sorce, Director Storage Software Sales
 - ssorce@us,ibm.com
 - Office: + 1 209 736-2773

Americas

- John Miller, Americas Storage Virtualization Sales Mgr
 - millerj@us.ibm.com
 - Office: + 1 610 578-2136
- Ron Broucek, Americas Tivoli Storage Software Sales
 - rjbrouc@us.ibm.com
 - Office: + 1 630 568-7068

Northeast IOT

- John Canny, Storage Virtualization Sales
 - cannyjj@uk.ibm.com
 - Office: + 44 131 558-4268
- David Glover, Tivoli Storage Software Sales
 - gloverw@uk.ibm.com
 - Office: + 44 207 202-3429

Southwest IOT

- Josephine Jagielo, Storage Virtualization Sales Leader
 - josephine@es.ibm.com
 - Mobile: + 34 618 270 942
- Metin Eskicioglu, Tivoli Storage SW Sales Leader
 - metineski@es.ibm.com
 - Office: + 34 91 333 2160

Asia-Pacific

- Adrian Cepak, Tivoli Storage SW Sales Leader
 - acepak@au1.ibm.com
 - Office: + 61 412 821 851

Worldwide Brand Team

- Chris Saul, SAN Volume Controller Product Marketing
 - cbs@us.ibm.com
 - Office: + 1 408 404-6034
- Dorothy Faurot, SAN Volume Controller Product Manager
 - dfaurot@us.ibm.com
 - Office: + 1 919 486-2399

Notice, Disclaimer, and Trademark Information

Copyright © 2007 by International Business Machines Corporation.

No part of this document may be reproduced or transmitted in any form without written permission from IBM Corporation.

Product data has been reviewed for accuracy as of the date of initial publication. Product data is subject to change without notice. This information could include technical inaccuracies or typographical errors. IBM may make improvements and/or changes in the product(s) and/or programs(s) at any time without notice. Any statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

References in this document to IBM products, programs, or services does not imply that IBM intends to make such such products, programs or services available in all countries in which IBM operates or does business. Any reference to an IBM Program Product in this document is not intended to state or imply that only that program product may be used. Any functionally equivalent program, that does not infringe IBM's intellectually property rights, may be used instead. It is the user's responsibility to evaluate and verify the operation of any on-IBM product, program or service.

THE INFORMATION PROVIDED IN THIS DOCUMENT IS DISTRIBUTED "AS IS" WITHOUT ANY WARRANTY, EITHER EXPRESS OR IMPLIED. IBM EXPRESSLY DISCLAIMS ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR NONINFRINGEMENT. IBM shall have no responsibility to update this information. IBM products are warranted according to the terms and conditions of the agreements (e.g., IBM Customer Agreement, Statement of Limited Warranty, International Program License Agreement, etc.) under which they are provided. IBM is not responsible for the performance or interoperability of any non-IBM products discussed herein.

The provision of the information contained herein is not intended to, and does not, grant any right or license under any IBM patents or copyrights. Inquiries regarding patent or copyright licenses should be made, in writing, to:

IBM Director of Licensing IBM Corporation North Castle Drive Armonk, NY 10504-1785 U.S.A.

Trademarks

The following terms are trademarks of International Business Machines Corporation in the United States, other countries, or both. Other company, product, and service names may be trademarks or service marks of others:

IBM, the IBM logo, ON (logo) DEMAND BUSINESS, DB2, Enterprise Storage Server, FlashCopy, POWER5, Tivoli, TotalStorage, TotalStorage Proven, System Storage, System p, System i, System x, System z, AIX, eServer, xSeries, pSeries, iSeries, zSeries, and BladeCenter

Backup slides

IBM Systems

Incremental FlashCopy

- New FlashCopy capability where only changes from either source or target data since last FlashCopy operation are recopied during a target refresh
- Up to 256 incremental and non-incremental targets can exist for same source
- Consistency groups can include both incremental and nonincremental FlashCopy targets
- Helps increase efficiency of FlashCopy operations and can reduce time to refresh copies
- Designed to allow completion of point-in-time online backups much more quickly, thus the impact of using FlashCopy is reduced
 - May enable more frequent backups so enabling faster recovery
 - More frequent backups could be used as a form of "near-CDP"

Start incremental FlashCopy

Start incremental FlashCopy

Only changed data copied by background copy

Cascaded FlashCopy

- New FlashCopy capability to create "copies of copies"
 - Mappings can be incremental or non-incremental
- Allows a vdisk to be both source and target in concurrent FlashCopy mappings
- Maximum number of targets dependent on a single source disk is 16. The example shows 4 targets from source disk 0
- Enables backup of target disks to be made without having to disrupt existing FlashCopy relationships with original source
- Helps reduce time to establish copies of targets, since there is no need to await copy complete of target disk before triggering cascaded copy
- Designed to increase flexibility in use of FlashCopy

