IBM

IBM System Storage Solutions for Smarter Systems

- Get an overview of IBM storage technologies, products, and solutions
- Learn about the different aspects of storage efficiency
- Understand the latest concepts in data protection

Bertrand Dufrasne
Karen Lawrence
Urban Biel
Matthew Lee
Rosemary Mccutchen
Libor Miklas
Aslam Mirza

Redbooks

ibm.com/redbooks

International Technical Support Organization

IBM System Storage Solutions for Smarter Systems

July 2011

SG24-7945-00

Note: Before using this information and the product it supports, read the information in “Notices” on page xiii.

First Edition (July 2011)

This edition applies to IBM System Storage products and offerings as of February 2011.

Contents

Preface

It is tremendously important to your business that your data is available efficiently, consistently, and constantly, and that it is protected from misuse and potential disaster. The storage systems that capture, manipulate, store, and retrieve data are a *key factor* in enabling your business to continue. Without these storage systems, your business activities can stall.

IBM® offers storage solutions that address the key needs of storage efficiency and data protection and retention. This IBM Redbooks® publication shows that IBM has the technology, products, and services to help keep your storage systems enabled and your business flourishing.

The information presented is structured as follows:

- Storage efficiency:
  - Virtualization
  - Tiering
  - Thin provisioning
  - Flexible delivery
  - Compression
  - Deduplication
  - Visibility
- Data protection:
  - Backup and restore
  - Continuous operations
  - Retention and compliance

The IBM vision of the Smarter Planet™ has transformed the technologies we offer (see http://www.ibm.com/smarterplanet/us/en/overview/ideas/index.html?re=sph). As an example, several of the technologies that can enable a planet to be smarter can enable your storage solutions to be more efficient and provide data protection. This book explains how.

This book has been written for those most concerned about efficient, protected, and available data, along with maintaining their IT environment at a competitive price. The audience for this book is that of IT professionals, such as CEOs, CIOs, CFOs, practitioners, information architects, and current and prospective IBM clients and Business Partners.

The team who wrote this book

This book was produced by a team of specialists from around the world working at the International Technical Support Organization, San Jose Center.

Bertrand Dufrasne is an IBM Certified Consulting IT Specialist and Project Leader for System Storage® disk products at the International Technical Support Organization (ITSO), San Jose Center. He has worked at IBM in various IT areas. He has authored many IBM Redbooks publications and has also developed and taught technical workshops. Before joining the ITSO, Bert worked with IBM Global Services as an Application Architect. He holds a Masters degree in Electrical Engineering.

Karen Lawrence is a Technical Writer with the IBM ITSO team in North Carolina. She has 25 years' experience in IT, with expertise in application design, change management, and the software development life cycle (SDLC). She has worked with SMEs on leading technologies in global, regulated enterprises, in the areas of storage systems, security, and disaster recovery, and the applications to support these. Karen holds a Bachelors degree in Business Administration from Centenary College in New Jersey.

Urban Biel is an IT Specialist in IBM Global Technology Services (GTS), Slovakia. His areas of expertise include System p, AIX®, Linux®, PowerHA™, DS6000/DS8000/SVC, Softek, and GPFS™. He has been involved in various projects that typically include the implementation of high availability and disaster recovery solutions, using DS8000® copy services with AIX/PowerHA. He also executed several storage and server migrations. Urban holds a second degree in Electrical Engineering and Informatics from the Technical University of Kosice.

Matthew Lee is an IBM Certified IT Architect, currently performing a Lead IT Architect role within IBM Integrated Technology Delivery (ITD) based in Sydney, Australia. He has worked in various industries for the past 25 years, including manufacturing, airline, banking, insurance, education, and oil, and joined IBM in 1999. Matthew started his career with IBM as a Tivoli® Professional Services and Support Professional and has broad experiences in many areas, including UNIX® operating systems, systems management, security, storage, solution design, implementation, and management. He holds a Bachelor of Engineering degree from the University of Western Australia and a post graduate degree in Business Administration from the Singapore Institute of Management.

Rosemary Mccutchen has over 20 years of IT experience and is currently a Certified Consulting IT Specialist working in Advanced Technology Skills (ATS), Storage, at IBM in Gaithersburg, Virginia. She is responsible for XIV® customer demonstrations, proofs of concept, and workshops, as well as XIV beta testing. Rosemary has extensive hands-on experience with XIV and has authored multiple XIV white papers and XIV training documents.

Libor Miklas is a Team Leader and an experienced IT Specialist working at the IBM Global Services Delivery Center in Czech Republic. He demonstrates ten years of practical experience in the IT industry. During the last six years, his main focus has been on backup and recovery and on storage management. Libor and his team support midrange and enterprise storage environments for various global and local clients, worldwide. He is an IBM Certified Deployment Professional of the Tivoli Storage Manager family of products and holds a Masters Degree in Electrical Engineering and Telecommunications.

Aslam Mirza is an Executive IT Specialist with the Systems Technology Group at IBM. He is currently based in New York, working as a pre-sales consultant for enterprise storage solutions. He has more than 35 years of experience with IBM large systems, storage systems, tape systems, and the System Storage Resiliency Portfolio. For the past several years, Aslam has worked with clients on the strategy and design of solutions that address business requirements. He has co-authored two IBM Redbooks, *IBM System Storage Business Continuity Solutions Overview*, SG24-6684 and *IBM System Storage Business Continuity: Part 1 Planning Guide*, SG24-6547.

Figure 0-1 The Team: Libor, Urban, Bertrand, Matthew, Aslam

Now you can become a published author, too!

Here's an opportunity to spotlight your skills, grow your career, and become a published author—all at the same time! Join an ITSO residency project and help write a book in your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks in length, and you can participate either in person or as a remote resident working from your home base.

Find out more about the residency program, browse the residency index, and apply online at:

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us!

We want our books to be as helpful as possible. Send us your comments about this book or other IBM Redbooks publications in one of the following ways:

- Use the online **Contact us** review Redbooks form found at:

  ibm.com/redbooks

- Send your comments in an email to:

  redbooks@us.ibm.com

- Mail your comments to:

  IBM Corporation, International Technical Support Organization
  Dept. HYTD Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400

Stay connected to IBM Redbooks

- Find us on Facebook:

  http://www.facebook.com/IBMRedbooks

- Follow us on Twitter:

  http://twitter.com/ibmredbooks

- Look for us on LinkedIn:

  http://www.linkedin.com/groups?home=&gid=2130806

- Explore new Redbooks publications, residencies, and workshops with the IBM Redbooks weekly newsletter:

  https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm

- Stay current on recent Redbooks publications with RSS Feeds:

  http://www.redbooks.ibm.com/rss.html

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application program interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. These and other IBM trademarked terms are marked on their first occurrence in this information with the appropriate symbol (® or ™), indicating US registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at http://www.ibm.com/legal/copytrade.shtml

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

AIX®
BladeCenter®
DB2®
Domino®
DS4000®
DS6000™
DS8000®
ECKD™
Electronic Service Agent™
Enterprise Storage Server®
FICON®
FlashCopy®
GDPS®
Geographically Dispersed Parallel Sysplex™
GPFS™
HyperFactor®
HyperSwap®
IBM Systems Director Active Energy Manager™
IBM®
Informix®
Lotus®
NetView®
Parallel Sysplex®
Power Systems™
PowerHA™
PowerVM™
ProtecTIER®
Redbooks®
Redbooks (logo) ®
Service Request Manager®
Smarter Planet™
System p®
System Storage®
System x®
System z®
Systems Director VMControl™
Tivoli Storage Manager Fastback™
Tivoli®
TotalStorage®
WebSphere®
XIV®
z/OS®
z/VM®

The following terms are trademarks of other companies:

Linear Tape-Open, LTO, Ultrium, the LTO Logo and the Ultrium logo are trademarks of HP, IBM Corp. and Quantum in the U.S. and other countries.

Snapshot, WAFL, SnapVault, SnapRestore, SnapMirror, SnapManager, SnapLock, NearStore, FlexVol, FilerView, Data ONTAP, NetApp, and the NetApp logo are trademarks or registered trademarks of NetApp, Inc. in the U.S. and other countries.

Microsoft, Windows, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Intel, Intel logo, Intel Inside logo, and Intel Centrino logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Other company, product, or service names may be trademarks or service marks of others.

1

Executive overview

IBM initiated the concept of the IBM Smarter Planet™ for which technology enables the planet to become more interconnected, instrumented, and intelligent amidst the high-speed information highway. The Smarter Planet is an undertaking that IBM envisioned and is making a reality. Briefly, the Smarter Planet is one in which the *key factors* of life (energy, water, transportation, buildings, public safety, and city operations) are handled in a smarter way, by the use of technology.

For more details about this exciting vision, see the following website:

`http://www.ibm.com/smarterplanet/us/en/?ca=v_smarterplanet`

Many initiatives toward making the planet smarter are already doing so.

Several technologies that are, right now, making the planet smarter, include those that render your storage solutions smarter through efficiency and data protection. We believe that it is only with advanced technologies and a superior understanding of these that we can address the demanding requirements of your business.

We do discuss IBM products in this Redbooks publication. However, it is the technologies that we hope will excite you: technologies that can help to relieve the pressure of providing efficient storage systems and protected data to your organization.

1.1 The pressure is on for data efficiency and data protection

As IT professionals, your ever-growing need is to ensure that storage systems are consistently up and running, have bandwidth and space available, and can store massive amounts of data quickly and retrieve it efficiently. Simultaneously, you need to ensure that your data is protected and makes use of the most efficient technologies. Overall, you need to spend less, store more information, access important data efficiently, improve services, maintain flexibility, and ensure that your data is sufficiently protected.

You might have reached out for assistance by researching new products and technologies that promise to improve your IT infrastructure and processes. When you evaluate product investments, two primary concerns need to be addressed:

- "What is the overall business value of the proposed investment to my business, and how might it support my ability to compete?"
- "What are the benefits of improving my IT infrastructure, compared with the current infrastructure and operations requirements?"

In this book, we guide you in answering these questions with confidence.

The amount of data generated and stored is growing exponentially each day.

Data generated by your business is frequently the foundation for business decisions and daily operations, from product development to sales and customer management. This means that any threat to the ability to access your data can pose a huge risk to your business. This includes data recovery from natural disasters and data corruption by hackers.

Facing the challenges of the information explosion, businesses are requiring systems that are smarter: systems that improve efficiency, performance, and cost savings, while simultaneously, lowering management complexity.

This publication is directed to IT professionals, such as CEOs, CIOs, CFOs, practitioners, information architects, and current and prospective IBM clients and Business Partners. We want this book to instill confidence in you that we have the technology, products, and services to help you succeed.

1.2 The concepts behind IBM storage efficiency technologies

Storage efficiency is a key factor that can help reduce costs, improve productivity, minimize risks, and manage growth.

The concepts of IBM storage efficiency technologies and a short description of our applicable products, follow.

1.2.1 Virtualization

Storage virtualization provides the ability to pool storage systems into a consolidated, shared capacity that can be managed from a central point of control. IBM is the industry leader in virtualization technology. IBM storage virtualization can help you to increase utilization and increase administrator productivity.

About the technology

Storage virtualization refers to the abstraction of storage systems as seen from application servers. It is the foundation for many of the other technologies described in this book, such as thin provisioning, tiering, and data protection, which are transparent to the application server. The advantages of using virtualized storage include improved physical resource utilization, improved responsiveness and flexibility, and, as a result, lower Total Cost of Ownership (TCO).

Several storage virtualization technologies are available:

- *Block level storage virtualization* provides storage to operating systems and applications in the form of virtual disks. There are two types of block level virtualization. One is disk level virtualization, whereby an abstraction process moves data from a physical disk level to a Logical Unit Number (LUN) level and is presented as though it were a physical device. Another method is storage level virtualization, which, unlike disk level virtualization, hides the physical layer of Redundant Array of Independent Disks (RAID) controllers and disks, and hides and virtualizes the entire storage system.
- *File level storage virtualization* provides storage volumes to operating systems and applications in the form of files and directories. Access to storage is through network protocols, such as Common Internet File System (CIFS) and Network File Systems (NFS).
- *Tape virtualization* technology virtualizes tape and tape drives by the use of special hardware and software that enhances backup and restore capabilities. The use of tape virtualization enables the tape media to perform as disk devices.

IBM products that support virtualization

The following products support this technology:

- The *SAN Volume Controller* provides an external storage virtualization function that operates in a consistent manner and provides consistent services for all attached servers. It is designed to take control of existing storage, retaining all your existing information.
- The *IBM Storwize V7000* is an innovative midrange disk system with sophisticated, enterprise-class storage functions for mid-sized businesses.
- The *IBM System Storage DS8000* provides high-performance disk storage for business-critical enterprise workloads. It is designed to address the needs of dynamic environments that require the highest levels of availability.
- The *IBM XIV Storage System* is a revolutionary, high-end, disk storage system with a virtualized grid architecture, designed to make storage management easy by eliminating the complexity of storage administration.
- *IBM Scale Out Network Attached Storage (SONAS)* allows you to manage multiple petabytes of storage and up to one billion files in a single filesystem. With SONAS, you can achieve operational efficiency with automated, policy-driven, tiered storage.
- The *IBM Virtualization Engine TS7700* is a family of mainframe virtual-tape solutions designed to optimize tape processing.
- The *IBM TS7600 ProtecTIER®* uses the Virtual Tape (VT) service to emulate the standard IBM tape library that allows backup/recovery applications to benefit from disk performance in typical tape environments.

  The IBM System Storage TS7650G ProtecTIER Deduplication Gateway is designed to meet the disk-based data protection needs of the enterprise-level data center.
- The *IBM Storage System N series* is a modular disk storage system designed to unify data retention and archival storage in an integrated architecture.

1.2.2 Flexible delivery solutions

Implement the IT environment that best fits your organizational strategies by choosing the right delivery model. IBM offers multiple flexible delivery solutions, including remote managed services, cloud computing, and standard delivery options. We offer enterprise-ready storage cloud solutions for storage capacity on demand and for archive.

About the technology

IT services and infrastructures are moving rapidly towards a flexible utility and consumer model by adopting new technologies. One of these technologies is virtualization, and cloud computing is an example of this virtual, flexible delivery model. Inspired by consumer Internet services, cloud computing allows the user to consume Internet offerings and services by using this self-service, on-demand model.

There are generally three cloud service models. In none of these does the customer manage or control the underlying cloud infrastructure (for example, network, servers, operating systems, or storage), but each model offers certain functionality. The American National Institute of Standards and Technology (NIST)[1] briefly defines these as follows:

- *Software-as-a-Service (SaaS):* The consumer uses applications provided over a network.
- *Platform-as-a-Service (PaaS):* The consumer deploys applications onto the cloud infrastructure, using programming languages and tools supported by the provider.
- *Infrastructure-as-a-Service (IaaS):* The consumer rents processing, storage, network, and other resources, with which to run software.

IBM products that support the technology

The following products support this technology:

- *Smart Business on the IBM Cloud* represents standardized cloud services provided by IBM on a pay-per-use basis:
  - *IBM Information Protection Services:* This service provides cost-effective data backup and recovery and continuous data protection.
  - *IBM Smart Business Development and Test on the IBM Cloud:* This cloud offering helps you to assess, plan, design, and implement flexible development and test environments on a private cloud to save on capital and operating costs and to reduce test cycle times, complexity, and risks. Using a self-service test platform provides on-demand physical and virtualized test resources, including IBM and non-IBM components.
- The *Smart Business Cloud*, an IaaS, designed to provide rapid access to security-rich, enterprise-class, virtual server environments that are well-suited for development, test, and other dynamic environments. The IBM Smart Business Storage Cloud (SBSC) can help to alleviate data storage challenges by enabling rapid implementation of a scalable, global file storage system, with flexibility in deployment and management options. The solution provides virtualized storage that enables storage and server consolidation, a unified management platform to help reduce outages, and advanced data replication for cost-effective business continuity and disaster recovery.

[1] Mell, Peter, and Grance, Tim, NIST, Information Technology Laboratory, October 7, 2009, csrc.nist.gov/groups/SNS/cloud-computing/cloud-computing-v26.ppt.

- *Smart Business Systems* are purpose-built, pre-configured cloud infrastructure appliances:
  - *IBM Information Archive (IA):* Data archiving is often required to mitigate organizational risks related to regulatory compliance. IA is such an appliance. It is designed to provide a scalable, secure, cost-effective, data archiving solution. Data is stored securely through enhanced data authenticity and a patent-pending data tampering feature. An indexing and searching mechanism is provided for fast information retrieval.
  - *IBM Cloudburst:* This is a hardware appliance designed to provide access to IBM WebSphere® virtual images and patterns for quickly and repeatedly creating application environments that can be securely deployed and managed in a private cloud. It helps application developers to easily and quickly develop, test, and deploy business application environments. In turn, this ends the need for manual, complex, and time-intensive processes associated with creating application environments. For more information, see this website:

    `http://www-01.ibm.com/software/tivoli/products/cloudburst/`

1.2.3 Automated tiering

Tiering is an intelligent method of reducing the cost of storing your information. It is a method for making your most critical data available quickly, by storing less critical data to less expensive and therefore slower media. Organizations that continue to spend more on disparate disk storage capacity that is not fully utilized can benefit from automated tiering.

About the technology

By implementing automated tiering, you can realize both performance optimization and space management optimization.

For performance optimization, Easy Tier is a dynamic data relocation feature that optimizes performance at the lowest cost. Easy Tier monitors data access patterns, and then intelligently, automatically, and nondisruptively moves the data to the appropriate disk tier. Easy Tier is a feature of the DS8000, the SAN Volume Controller, and the Storwize V7000 products.

Space management optimization, can also be performed through policy-based automation techniques to migrate less active data to lower cost storage. This type of Information Lifecycle Management is available with the Scale Out Network Attached Storage (SONAS) and other Hierarchical Storage Management (HSM) products.

IBM products that support the technology

The following products support this technology:

- The *DS8000 Easy Tier* function can be used in automatic and manual modes, allowing for the highest level of performance optimization and data control.

  In automatic mode, data relocation is performed automatically and transparently, based on I/O activity. Data relocation occurs both within tiers (inter-tier), based on data access activity, and among tiers (intra-tier), based on rank activity.

  Easy Tier automatic mode operates on *hybrid* extent pools. A hybrid extent pool is an extent that contains two storage tier types, including solid-state drives (SSD), Enterprise level (Fibre Channel), and Nearline Serial Advanced Technology Attachment (SATA) drives.

In manual mode, the DS8000 Easy Tier feature allows the user to request the migration of a volume between extent pools, the merging of one extent pool into another, or rank depopulation (removal of all extents from a rank so that the rank can be removed from an extent pool.) All of these Easy Tier manual mode functions are dynamic, that is nondisruptive to host access.

Easy Tier manual mode can be used simultaneously with Easy Tier automatic mode for hybrid extent pools. Easy Tier manual mode can also be used for extent pools where Easy Tier automatic mode is not available (for example, extent pools containing a single disk tier).

Easy Tier statistics can be offloaded using the IBM Storage Tier Advisor Tool (STAT) to understand the workload characteristics and to estimate the potential performance benefits of using SSDs and Easy Tier automatic data placement.

- The SAN Volume Controller and the Storwize V7000 products include *Easy Tier* functionality in two modes, Evaluation and Auto Data Placement. Because most workloads access certain areas of a volume or LUN much more heavily than others, Easy Tier collects performance information at the sub-LUN level. When Easy Tier is activated for a storage pool, SAN Volume Controller or Storwize V7000 continuously gathers extent-level I/O statistics and creates 24-hour moving averages. These statistics can be used to identify the extents that are heavily accessed and can benefit most from relocation to a high performance disk tier (for example, SSDs) and those that are less active and can be relocated to a lower performance disk tier, for example, Hard Disk Drives (HDDs).

  Easy Tier statistics can be offloaded using the IBM STAT to understand the workload characteristics and to estimate the potential performance benefits of using SSDs and Easy Tier automatic data placement.

  Easy Tier is included in the base software for SAN Volume Controller and the Storwize V7000.

- The *IBM Scale Out Network Attached Storage (SONAS)* offering implements feature-rich Information Lifecycle Management (ILM), driven by customizable automation policies for file placement. With the tiered storage policy actions identified, a parallel data movement engine performs physical movement to other storage pools, or to other sites, or external hierarchical storage management (HSM).

- *IBM Hierarchical Storage Management (HSM)* software migrates rarely used files from costly to less expensive storage, typically physical tapes with sequential data access.

  IBM offers HSM on various platforms, incorporated in the following three products:

  – *IBM Tivoli Storage Manager HSM for Windows®:* A product that controls primary disk storage needs by automatically migrating rarely used files on Microsoft® Windows-based platforms to less expensive storage.

  – *IBM Tivoli Storage Manager for Space Management:* This product moves inactive data to reclaim online disk space for important, active data on UNIX and Linux platforms, including the support of General Parallel File System (IBM GPFS™).

  – *IBM Data Facility Storage Management System (DFSMS) for System z:* includes DFSMS for Hierarchical Storage Management (DFSMShsm). This is an automated hierarchical storage manager and productivity tool, designed to support availability management, space management, and disaster recovery.

1.2.4 Thin provisioning

Traditionally, storage is allocated to applications using a *fat provisioning* method, which pre-allocates and dedicates a user-defined amount of storage space for use by the application or host. However, often, not all space allocated to applications is consumed, resulting in wasted *white spaces*.

About the technology

Thin provisioning allows you to present logical or virtual volume sizes that are larger than the allocated physical disk drive capacity. The volume size presented to a server or internal function does not need to be matched by dedicated physical capacity allocation. Instead, physical disk drive capacity is allocated as needed (on demand) for write activity. Unused physical capacity is no longer dedicated to a single volume, but is available for multiple volumes in a storage pool or across the entire storage system.

As a result of thin provisioning, your organization can realize reductions in the need to over-allocate for growth, increase volume size, and add volumes.

IBM products that support the technology

The following products support this technology:

- The *IBM System Storage DS8000* supports thinly-provisioned volumes. Thin provisioning of volumes optimizes physical space utilization of the ranks within a DS8000 Extent Pool. Any real capacity not utilized by a thinly provisioned volume is available for use by any other thinly provisioned volume within the same DS8000 Extent Pool, known as Extent Space Efficient (ESE). The DS8000 also allows the creation of thinly provisioned IBM FlashCopy® volumes (Point-in-Time Images), known as Track Space Efficient (TSE) volumes. Thinly provisioned FlashCopy volumes are supported for both of these open systems, Fixed Block (FB) volumes and mainframe IBM z/OS® Count Key Data (CKD) volumes.
- The *SAN Volume Controller and Storwize V7000* offerings allow for the creation of thinly provisioned volumes referred to as Space Efficient Volumes. Thin provisioning of volumes optimizes utilization of the physical capacity within a SAN Volume Controller or Storwize V7000 Storage Pool. Any physical capacity not utilized by a thinly provisioned volume is available for use by any other thinly provisioned volume within the same storage pool.
- *The IBM System Storage N series* also provides thin provisioning. Configured with SATA disk drives, it is designed to provide terabytes of near-primary storage performance at a cost close to that for tape. Nearline storage can help you to complement and improve existing tape backup, archiving and data protection schemes by inserting economical, easy to use disk-based storage between application storage and tape libraries in a three-tier storage architecture.

1.2.5 Data compression

For businesses to gain and retain their competitive edge and to maintain compliance, there is a clear requirement for the data to remain online and available for longer periods of time. This growing capacity creates a tremendous strain on storage systems. The IBM compression solutions, especially IBM Real-time Compression Appliances, ease that strain.

About the technology

Real time compression is the newest innovation in storage efficiency. Compression enhances throughput and response time by reducing the amount of data processed. The results are improved online storage optimization, a compression of primary storage of up to 80 percent, dramatically lower costs, and an increase in efficiency of your primary, active data.

Data compression is a core storage efficiency technology in most data centers. IBM data compression utilities and product features are commonly used to reduce the cost of storing inactive files, backups, and archives.

IBM products that support the technology

The following products support this technology:

- *IBM Real-time Compression Appliances* are the only storage compression solutions that can shrink primary, online data in real time, without performance degradation. By significantly reducing storage requirements, you can keep up to five times more information online for analytics, use the improved efficiency to reduce storage costs, or achieve a combination of greater capacity and reduced cost.

  All IBM Real-time Compression Appliances apply IBM patented real time data compression techniques to primary and existing storage, delivering optimization and savings throughout the entire storage life cycle. The result is unprecedented cost savings and return on investment (ROI), along with operational and environmental efficiencies.

  IBM Real-time Compression provides enterprises with better storage utilization, lower capital and operational costs, proven technology, better energy efficiency, high availability, disaster recovery, and full transparency, with no change to performance, applications, networks, storage, or processes, and no loss of data integrity.

- *IBM Tivoli Storage Manager client compression*. Tivoli Storage Manager is the number one product of choice as a data protection solution for backup, archiving, disaster recovery planning, space management, database and application protection, bare machine recovery, and record retention. More than 45 operating platforms are supported, using a consistent graphical user interface, and a simple command line interface. Tivoli Storage Manager provides centralized administration, fully automated data protection, high-speed server recovery, data compression and deduplication, and transparent licensing.

- *IBM Enterprise tape systems with Linear Tape-Open (LTO™) Ultrium5 tape drives* are the first generation of the LTO Ultrium5 technology introduced on-chip hardware data compression using the Streaming Lossless Data Compression (SLDC) algorithm. The most remarkable benefit of using LTO data compression (LTO-DC) is that the SLDC algorithm does not apply to incompressible data.

  Every block of data written to tape has a header bit showing if the block is compressed by SLDC or raw format. For each block of raw data the algorithm works on, it caches a copy. After the compression is successful, the algorithm compares the compressed data block to the raw data block in memory and writes the smaller of the two to tape.

1.2.6 Data deduplication

Data deduplication has emerged as a key technology in an effort to eliminate redundant data, increase agility, and improve responsiveness to changing marketplace conditions. Deduplication is the art of intelligently reducing storage needs an order of magnitude better than common data compression techniques make possible. This is done by eliminating redundant data so that only one instance of a data set is actually stored.

IBM has the broadest portfolio of deduplication solutions in the industry, which allows us the freedom to solve customer issues with the most effective technology.

About the technology

Using data deduplication eliminates redundant data, increases agility, and improves responsiveness to changing marketplace conditions. The benefits of deduplication are seen in improved backup and recovery performance and in dramatically reduced amounts of stored data and, therefore, reduced storage costs.

The effectiveness of data deduplication is dependent upon a number of variables, including the rate of data change, the number of backups, and the data retention period. So, when the variables are known, the benefits of data deduplication are measurable. IT organizations have seen a range of measurable benefits from deduplication that includes the ability to move up to 90 percent less virtual machine (VM) data, reduce backup storage by as much as 95 percent, minimize backup windows, and reduce network utilization by up to 90 percent.

Data deduplication can reduce storage footprint and TCO in the following areas:

- Storage capacity (store more data per physical storage system)
- Energy (less data per each disk means less energy spent)
- Replication technology (data deduplication can potentially reduce the data that must be sent across a network to primary storage, for backup replication, and for disaster recovery)

IBM products that support the technology

The following products support this technology:

- *IBM System Storage ProtecTIER Enterprise Edition* software uses IBM HyperFactor®, a patented technology that reduces the phenomenon of missed factoring opportunities. This provides for a more efficient process, rather than comparing the new data to the similar data to identify and store only the byte-level changes. The capacity expansion that results from data deduplication is often expressed as a ratio, essentially the ratio of nominal data to the physical storage used. HyperFactor is able to surpass the reduction ratios attainable by any other data reduction method. See “The concept of data deduplication” on page 102.
- *ProtecTIER Native Replication* is a logical feature that enables clients to replicate and “move” any or all of their virtual tape cartridges from a main site to a remote disaster recovery (DR) location and vice versa. Two of the ProtecTIER Deduplication models (TS7650 and TS7650-G) employ an IP-based replication design. Data is replicated between sites from one, two-node cluster to another.

Other ProtecTIER products are available that offer disk-based and tape-based data protection, along with infrastructure cost reduction, for small, midsize, and enterprise level environments. All of the available IBM ProtecTIER deduplication products use inline (real-time) deduplication processing.

- *IBM System Storage N series Deduplication Gateway:* The IBM N series Storage System offers a deduplication feature (as a native part of its Data ONTAP® operating system) that can be added at no additional cost and uses several benefits already built into the OS of the storage system. This is post-process deduplication within the storage system, so you can realize immediate benefits in existing storage capacity.
- Data deduplication in the *Tivoli Storage Manager:* This is a process in which you identify and then remove duplicate data using server processes, such as reclamation processing of storage-pool volumes. Because duplication identification requires extra disk, I/O, and processor resources, Tivoli Storage Manager lets you control when identification begins and the number and duration of processes.

  You can deduplicate backup and archive data, Tivoli Data Protection data, whole files and files that are members of an aggregate. You can deduplicate data that has already been stored, or configure client-side data deduplication on slow networks. No additional backup, archive, or migration is required. You can deduplicate any type of data except encrypted data.

1.2.7 Visibility

The concept of visibility allows you to have more knowledge about your storage environment than ever before. This is needed because, as data volumes continue to grow, and as systems and existing data are migrated, we are moving toward an increasingly more complex IT infrastructure.

About the technology

Because of the growth in amounts of data, organizations are increasingly adopting smarter storage technologies, such as virtualization, deduplication, thin provisioning, and tiering to provide for the more efficient use of storage and server resources. This new generation of virtualized infrastructures must be managed by smarter tools.

These tools provide storage infrastructure visibility, including the following benefits:

- Enable Line Of Business (LOB) insight into storage utilization and allocation, and enabling easier charge backs
- Allow for more intelligent business decisions about storage efficiency, enabling faster responses to changing business demands, and simultaneously reducing cost.
- Provide a better understanding of application storage performance and utilization patterns, which, in turn, enables better Information Lifecycle Management (ILM) of application data
- Allow proactive, rather than reactive, storage management
- Lead to cost savings for the organization

IBM products that support the technology

The following products support this technology:

- *IBM Systems Director* is the platform management foundation for achieving Smarter Systems. An integral component of the IBM Smarter Systems portfolio, IBM Systems Director enables integration with IBM Tivoli® and third party management platforms, providing the building block for virtualization and integrated services management. IBM Systems Director provides a single-point-of-control, thereby reducing reduce IT management complexity and cost.

Key benefits of the *IBM Systems Director* include these:

- Simplified usability and reduced complexity, in that it eliminates layers of management duplication through a single, unified view of the total IT environment. This reduces system management complexity, cost, and time to problem resolution, and improves visibility, availability, and service.
- Integration with Tivoli and third party management platforms that provides the foundation for virtualization and integrated services management. With an extensive suite of available platform management tasks, and automated tools, IBM Systems Director assists systems management staff in increasing productivity, resulting in improved responsiveness and service.
- Because of its plug-in architecture, many existing technical functions can be integrated seamlessly with IBM Systems Director to provide a “single pane of glass” interface for systems management.

- The *IBM Tivoli Storage Productivity Center (TPC) Suite* is a comprehensive suite of Storage Resource Management (SRM) products, designed to help reduce management complexity by centralizing, simplifying, and optimizing storage tasks associated with storage systems, storage networks, replication services, and capacity management. This suite of products provides the following major benefits:
  - *Cost reduction:* The TPC storage efficiency functions help clients to significantly reduce future storage purchases and optimize their current storage infrastructure.
  - *Availability management:* TPC provides comprehensive event alerting functions designed to minimize storage outages, thus reducing costs.
  - *Performance management:* TPC provides overall performance management for your storage infrastructure, optimizing existing storage assets and deferring new storage purchases.
  - *Storage resource management:* TPC allows you to implement a highly optimized storage infrastructure with significant benefits.

  TPC also provides incident and problem management and improved reporting for an entire heterogeneous storage infrastructure.

1.3 Data protection and retention technologies

For your data protection and retention key factors of your business, this book explains the advances we have made in this arena, including backup, archiving, disaster recovery planning, space management, database and application protection, bare machine recovery, and records retention. One of our offerings, the Tivoli Storage Manager, can support more than 45 operating platforms, using a consistent graphical user interface and a simple command line interface.

The concepts of our data protection and retention technologies, and a short description of our applicable products, follow.

1.3.1 Backup and recovery

Application servers and data often must remain available 24 hours per day. Your business cannot afford to lose data, but you also cannot afford to stop critical systems for hours so you can adequately protect the data. With the rapid increase in the amount of data, the critical business need of the data, and shrinking backup windows, traditional backup and restore methods might be reaching their limits in meeting these challenging requirements. IBM has hardware and software products and services that address your backup and recovery requirements.

About the technology

IBM has the technology to protect your data from hardware failures, errors, and unforeseen disasters by replicating data and making backup and archive copies in auxiliary and off-site storage.

A point-in-time copy using snapshot or FlashCopy can help minimize the impact caused by backups and provide nearly instant restore capabilities. For a point-in-time copy, the original data remains accessible when the copy is being created.

IBM backup and recovery technologies provide centralized, web-based management, intelligent data move and store techniques, and comprehensive, policy-based automation. All of these functions work together to minimize administration costs and the impact to computers and networks. Backup and recovery systems are also scalable, to protect possibly thousands of computers with the following requirements:

- Run more than a dozen operating systems
- Range from mobile computers to mainframes
- Are connected through the Internet, wide-area networks (WANs), local area networks (LANs), or storage area networks (SANs)

IBM products that support the technology

The following products support this technology:

- *IBM Tivoli Storage Manager*. The key difference between Tivoli Storage Manager and other data protection products is its *progressive incremental backup methodology*. Tivoli Storage Manager only backs up new or changed files. It tracks all backups at the file level. Tivoli Storage Manager has a powerful relational database, and so full backups of the latest active version of every file are *not required*, because the latest *active* version of every file is available in the Tivoli Storage Manager server storage hierarchy.
- *Point-in-time copy*, as available with various disk systems as part of their Copy Services functions.

  Data volume snapshot techniques (also known as point-in-time copy) provide nearly instant backup and restore capabilities. A snapshot operation takes much less time than a direct tape backup. It therefore increases the flexibility of backup scheduling and administration. Our point-in-time products are:

  – *IBM FlashCopy:* The FlashCopy function enables you to make consistent, point-in-time, full volume copies of data, with the copies immediately available for read or write access. IBM FlashCopy is available with the IBM DS8000, SAN Volume Controller, and Storwize offerings.

    The DS8000 also offers the FlashCopy Space Efficient, which copies only changed data (before new data is written), saving space and having less impact on the storage system performance.

  – *IBM XIV Snapshot:* IBM XIV Snapshot is based on several innovative technologies that ensure minimal degradation of storage system performance. IBM XIV Snapshots make use of pointers and do not necessarily copy all the data to the second instance of a volume. The XIV system minimizes the impact to the host for write operations by performing a redirect-on-write operation. As the host writes data to a volume with a snapshot relationship, the incoming information is placed into a newly allocated partition. Then the pointer to the data for the master volume is modified to point at the new partition. The snapshot volume continues to point at the original data partition.

These copies can be used for backup, application, or quality assurance testing or other purposes. The optional IBM Tivoli Storage FlashCopy Manager can help deliver the highest levels of data protection for mission-critical IBM DB2®, SAP, Oracle, Microsoft Exchange, and Microsoft SQL Server applications with integrated application-aware point-in-time backup and restore capabilities.

– *SONAS Snapshot:* SONAS Snapshot is a space-efficient, read-only, point-in-time consistent version of an entire SONAS filesystem, frozen at a point-in-time. It enables online backups to be maintained, providing nearly instant access to previous versions of data without requiring complete, separate copies or resorting to auxiliary backups.

1.3.2 Continuous operations

Your IT systems need to be able to prevent and avoid outages and recover from them quickly. IBM offers a range of reliable, high-performance solutions, ensuring that your business remains operational during component failures or significant disasters.

In addition, IBM Business Continuity and Resiliency Services helps to ensure the continuity of business operations and assists with regulatory compliance, improved systems availability, data protection, and the integration of IT operational risk management strategies.

About the technology

A robust, highly available application infrastructure is fundamental to deliver required service level agreements. The objective behind implementing a high availability solution is to provide near-continuous application availability during both planned and unplanned outages.

One method of maintaining continuous operations is by using remote copy, a technology implemented in several of our storage products. This method can be combined with additional methods to allow automation for to either a backup site or backup storage systems, following a failure, and then switching back after the environment has returned to normal (failover and failback). Failover and failback operations need to be as automated as possible, especially in the event of a disaster. The IBM products and services described here can help you achieve that automation.

IBM products that support the technology

The following products support this technology:

- *IBM Copy Services:* This is available with most of our disk storage products (although with variations in their implementation and use). Copy Services enable business continuity capabilities to give you the peace of mind of knowing your mission-critical applications.

  For example, on the DS8000, Metro Mirror is designed to provide a no-data-loss remote mirroring environment for metropolitan distances up to 300 kilometers. Global Mirror is designed to minimize data loss at the recovery site to as low as five seconds or less relative to the production site at virtually any distance. IBM Metro/Global Mirror combines these two capabilities to support a three-site configuration designed for no data loss, regardless of distance.

- *IBM PowerHA™ System Mirror V6.1 for AIX Enterprise Edition:* This is a high availability, key component of business resiliency. Business-critical applications are configured into a cluster, which typically involves at least two systems (nodes). The cluster monitors the critical resources for changes that can indicate a failure, a pending failure, or a possible configuration change.

The Enterprise Edition offers multiple technologies for achieving long distance data mirroring, failover, and resynchronization:

- *AIX Geographic Logical Volume Manager (GLVM):* This component provides data replication and failover to remote sites which is not dependent on storage system capabilities. GLVM mirroring automatically maintains a data replica across geographically remote sites and completely manages data replication during production, failover, and recovery.
- The *Enterprise Edition also supports IBM System Storage DS8000:* This product provides advanced features, such as *Metro Mirror* and *Global Mirror*[2] for data replication between sites, enabling automatic failover between geographically dispersed data centers.

- *Open IBM HyperSwap® for AIX with IBM Tivoli Storage Productivity Center:* is a solution to automatically failover input and output (I/O) from primary logical devices to secondary logical devices in the event of a primary disk storage system failure.
- *The IBM Tivoli Storage Manager Fastback:* Formerly known as Continuous Data Protection (CDP) for files, this is a block-level data protection solution for workstations, laptops, file servers, and select database systems on supported Windows and Linux platforms.
- *IBM Tivoli Storage FlashCopy Manager:* This provides advanced protection and restoration of business critical files on the Microsoft Windows platform (Microsoft SQL Server and Exchange servers) and on the AIX platform (DB2, Oracle, and SAP databases).It is a fast, application-aware backup and restore solution, using an advanced snapshot technology, called FlashCopy. It enables fast, frequent backup of critical applications to limit data loss without disruption to operations. Restore capabilities are nearly instant.

  Optional integration with Tivoli Storage Manager provides access to a broad range of advanced data protection and data reduction capabilities, such as data deduplication, progressive incremental backup, HSM, and centrally managed policy-based administration.
- The *IBM Tivoli Storage Productivity Center for Replication (TPC-R):* This family of products helps to manage the advanced copy services provided by the IBM Enterprise Storage Server® (ESS) Model 800, IBM System Storage DS8000, IBM DS6000™, and the IBM System Storage SAN Volume Controller.

  The IBM Tivoli Storage Productivity Center for Replication V4.2 introduces a feature called Open HyperSwap to help improve the continuous availability attributes of open system servers (IBM AIX®) by managing a set of planned and unplanned disk system outages for disk systems that are capable of Metro Mirror, peer-to-peer remote copy (PPRC). Open HyperSwap allows I/O to the primary volumes in a Metro Mirror relationship to be swapped to the secondary volumes without your intervention. After the swap has occurred, the application automatically writes to the secondary site without noticing the switch.
- *Geographically Dispersed Open Clusters (GDOC):* Continuous availability solutions are integrations of servers, storage, software, automation, and networking. Most of the solutions are based on a form of operating system server clustering to provide application availability. When an application failure is detected, a continuous availability solution can perform a predefined set of tasks required to restart the application on another server, utilizing the mirrored data residing on the secondary storage system.

  GDOC is a solution designed to protect the availability of critical applications that run on UNIX, Windows, or Linux servers. GDOC is based on an Open Systems Cluster architecture, spread across two or more sites, with data mirrored between sites to provide high availability and disaster recovery.

[2] DS8000 Global Mirror support was introduced with PowerHA System Mirror 6.1 SP3 Enterprise Edition.

- *Geographically Dispersed Parallel Sysplex/Peer-to-Peer Remote Copy HyperSwap Manager (GDPS/PPRC HM):* A single point of management and control that provides planned and unplanned switching capabilities to either a backup site or backup storage systems, nondisruptively. GDPS/PPRC HM is designed to work in IBM System z® environments only.

1.3.3 Retention and compliance

Regulations and other business imperatives stress the need for an Information Lifecycle Management (ILM) process and tools to be in place, for which IBM has a broad range of technologies. IBM has designed data protection strategies for maintaining compliance with state, federal, industry-, and company-specific, long-term retention of business critical data, and the processes that require that data be retained for tens of years or even forever.

As the amount of information stored continues to grow, effective data archiving solutions can help you improve productivity, control both infrastructure and legal costs, and manage compliance and operational risks.

About the technology

The IBM Smart Archive strategy, a broad strategy for data retention and compliance, unifies the power and value of IBM software, systems, and services. Our unified, integrated, and coherent strategy offers specific solutions that maximize benefits to you. Information Archive for email, Files and eDiscovery is a specific solution that simplifies and accelerates the implementation of the solution from weeks to days. The Smart Archive strategy identifies six key industry sectors with specific set of retention and data management requirements:

- Financial Services
- Health Care and Life Sciences
- Insurance
- Manufacturing
- Transportation
- Government

For each of these industry sectors, we developed a concrete archiving solution with data management policies that are part of the strategic IBM Information Archive and which significantly help clients to identify their archiving requirements.

When considering retention managed data, consider the following aspects:

- Variable data retention period: This can range from a few months to forever.
- Variable data volume: Many enterprise clients start with 5 TB to 10 TB of storage.
- Data access frequency: Write Once Read Rarely or Write Once Read Never.
- Data read/write: Sector and application affect the read function; volume affects write.
- Data protection: Policies for non-erasable, non-rewritable data, and deletion at expiration.

IBM products that support the technology

The following products support this technology:

- The *IBM Information Archive appliance:* This is an integrated information retention solution. The appliance includes preinstalled servers, disk storage, and the Information Archive software. The appliance provides policy-managed storage for compliance, archiving, and content management applications. It is a universal storage repository for all types of archived information. The appliance can be used to store and manage multiple billions of documents over its deployment lifetime, by archiving data with scalability up to multiple petabytes.

IBM Information Archive archives and retains records in non-erasable, non-rewritable formats. It supports business and legal requirements for retaining and protecting information from premature deletion or malicious modification. The appliance enables archiving applications to move inactive or unnecessary files from primary storage to lower-cost storage tiers.

- *IBM Enterprise tape libraries with LTO Ultrium5:* The LTO Ultrium5 tape drive is the fifth generation of widely used and industry proved LTO Ultrium5 technology. It provides nearly doubled capacity, increased read/write access time to the tape, and enhanced encryption functionality, together with IBM Tivoli Key Lifecycle Manager for data encryption (TKLM), and continues to support backward, full compatibility with the fourth generation and readability of LTO3 tape cartridges.

  LTO5 technology introduces tape partitioning. At present, IBM offers only one filesystem or product which benefits from this functionality: the IBM Long Term File System (LTFS). LTFS is a low-cost, self-describing, tape filesystem that provides an easy storage and distribution solution without additional database application needs.

- The *IBM System Storage TS3500 tape library:* Formerly known as IBM TotalStorage® 3584, this is the only IBM tape automation product available that matches the strict classification of the enterprise tape library. The TS3500 is a highly scalable, automated tape library for backup and archive of mainframe and open systems from midrange to enterprise environments.

  The IBM high density (HD) function of specific IBM System Storage tape libraries and expansion frames is IBM patented technology that allows multiple cartridges to be stored in a tiered architecture. The depth of a cartridge location in an HD slot is known as a tier.

- *IBM System Storage N series SnapLock:* Today's business relies on frequent use of Write Once Read Many (WORM) data storage devices to meet regulatory compliance or simply to add another layer to the data protection roadmap. Using SnapLock addresses issues faced by growing business requirements for WORM data storage and alleviates issues inherent with traditional WORM storage solutions. The IN series SnapLock offers an easy-to-manage, faster-to-access, low cost, disk-based solution.

  The evolution of WORM data storage has now arrived at a new best-of-breed solution in the IBM System Storage N series products configured with SnapLock Compliance and SnapLock Enterprise.

1.4 Why choose IBM?

IBM is the world's largest storage and data services organization, with experts around the world, a broad ecosystem of skilled IBM Business Partners, and service products to help speed time-to-value. IBM has practical experience in providing storage to a number of the world's largest and most demanding environments, including its own information infrastructure, which supports 400,000 employees globally, large outsourced data centers, and IBM cloud services. Statements such as this are reinforced by studies and reports. Other statements that confirm our experience and our commitment to you are as follows:

- No other single vendor offers the broad range of hardware, software, and services solutions as does IBM.
- IBM supports intelligent data migration between storage tiers, supporting every tier from tape to solid-state.
- IBM offers the most efficient options, including virtualization, data deduplication, thin provisioning, space efficient FlashCopy, inline data compression, optimized solid-state storage solutions, tape, and blended disk and tape solutions.
- IBM storage virtualization supports multiple vendor technologies.
- IBM has the world's largest storage services team.

1.5 Conclusion

This Executive Overview has introduced you to IBM technologies and products in the areas of storage efficiency and data protection. The remainder of this book expands on all of these ideas, giving you the knowledge and confidence to reap these benefits:

- Increase the efficiency of your storage systems, in line with your storage strategy
- Ensure that the storage systems you use provide the functionality that gives you the edge over the competition

IBM has the background, experience, and the technologies to assist you in your storage solution.

For more information about these technologies and products, we invite you to visit us at this website:

http://www-03.ibm.com/systems/storage/?cm_re=masthead-_-products-_-stg-allstorage

Part 1

Storage efficiency

In this part of the book, we present and discuss IBM technologies and products that contribute to enhance various aspect of storage efficiency:

- Chapter 2, “Virtualization” on page 21, explains how virtualization is implemented in these IBM products:
  - IBM System Storage DS8000
  - IBM System Storage IBM SAN Volume Controller and IBM Storwize V7000
  - IBM XIV Storage System
  - IBM Scale Out Network Attached Storage (SONAS)
  - IBM Virtualization Engine TS7700, IBM TS7600 ProtecTIER and the IBM System Storage TS7650G ProtecTIER Deduplication Gateway
  - IBM System Storage N series
- Chapter 4, “Automated tiering” on page 45, discusses these products:
  - Easy Tier for the DS8000 and SAN Volume Controller / Storwize V7000
  - Hierarchical Storage Management software
- Chapter 5, “Thin provisioning” on page 67, explains how that function applies to these products:
  - The IBM System Storage DS8000
  - The IBM XIV Storage System
  - The IBM San Volume Controller and Storwize V7000
  - The IBM System Storage N series
- Chapter 3, “Flexible delivery solutions” on page 37, discusses cloud computing technology, with a focus on the service layer of the network architecture known as Infrastructure-as-a-Service (IaaS) and related IBM offerings.
- Chapter 6, “Compression” on page 89, focuses on a newly introduced solution, the IBM Real-time Compression Appliance (RTCA).

- Chapter 7, “Deduplication” on page 101, reviews how deduplication is addressed by IBM technologies and products, including these:
  - IBM ProtecTIER products
  - IBM System Storage N series
  - IBM Tivoli Storage Manager
- Finally, Chapter 8, “Visibility” on page 123 describes and highlights IBM products designed to provide in depth end-to-end visibility and management of complex physical and virtualized IT infrastructure. An overview and introduction is provided for these products:
  - IBM Systems Director
  - IBM Tivoli Storage Productivity Center.
  - IBM Storage Enterprise Resource Planner (SERP)
  - IBM Storwize Rapid Application Storage

2

Virtualization

This chapter describes the concepts and benefits of storage virtualization and how IBM can meet your storage solution requirements with virtualization technologies.

Storage virtualization refers to the abstraction of storage systems from applications or computers. It is a foundation for the implementation of other technologies, such as thin provisioning, tiering, and data protection, which are transparent to the server.

In this chapter, we discuss how virtualization is implemented in the following IBM products:

- IBM System Storage DS8000
- IBM System Storage IBM SAN Volume Controller and IBM Storwize V7000
- IBM XIV Storage System
- IBM Scale Out Network Attached Storage (SONAS)
- IBM Virtualization Engine TS7700, IBM TS7600 ProtecTIER and the IBM System Storage TS7650G ProtecTIER Deduplication Gateway
- IBM System Storage N series

2.1 Virtualization overview

Storage virtualization refers to the abstraction of storage systems from applications or computers. It is a foundation for the implementation of other technologies, such as thin provisioning, tiering, and data protection, which are transparent to the server.

These are advantages of storage virtualization:

- *Improved physical resource utilization:* By consolidating and virtualizing storage systems, we can make more efficient use of previously wasted white spaces.
- *Improved responsiveness and flexibility:* Decoupling physical storage to virtual storage provides the ability to re-allocate resources dynamically, as required by the applications or storage subsystems.
- *Lower total cost of ownership:* Virtualized storage allows more to be done with the same or less storage.

Several types of storage virtualization are available:

- *Block level storage virtualization:* This refers to provisioning storage to your operating systems or applications in the form of virtual disks. Fibre Channel (FC) and Internet Small Computer System Interface (ISCSI) are examples of protocols used by this type of storage virtualization. There are two types of block level virtualization:
  - Disk level virtualization: This is an abstraction process from a physical disk to a Logical Unit Number (LUN) that is presented as of it were a physical device
  - Storage level virtualization: Unlike disk level virtualization, storage level virtualization hides the physical layer of Redundant Array of Independent Disks (RAID) controllers and disks, and hides and virtualizes the entire storage system.
- *File level storage virtualization:* This refers to provisioning storage volumes to operating systems or applications in the form of files and directories. Access to storage is by network protocols, such as Common Internet File Systems (CIFS) and Network File Systems (NFS). It is a file presentation in a single global namespace, regardless of the physical file location.
- *Tape virtualization:* This refers to virtualization of tapes and tape drives using specialized hardware and software. This type of virtualization can enhance backup and restore flexibility and performance, because disk devices are used in the virtualization process, rather than tape media.

2.2 IBM System Storage DS8000 virtualization

In the DS8000, a logical volume also known as a Logical Unit Number (LUN) is created by preparing multiple physical disk drives, or Disk Drive Modules (DDMs), for becoming an entity that can be used by an operating system. When a server has access to a LUN, it has access to the disk files or data as though the disk were a local disk.

Physical layer

The DS8800, the latest of the DS8000 series, uses a switched point-to-point topology to access Serial-Attached SCSI (SAS) disk drives. SAS disks are mounted in disk enclosures, each enclosure consisting of 24 disks. The disk drives are accessed by a pair of device adapters. Each device adapter has access to any disk drive using two independent, switched fabrics, using a Fibre Channel - Arbitrated Loop (FCAL) topology. The ownership of a disk by a disk adapter is defined during the logical configuration process.

DDMs are grouped into *array sites*. An array site is formed from 8 DDMs of the same type, that is, the same capacity and same speed (rpm). Which DDMs are forming an array site is automatically predetermined by the DS8000, as shown in Figure 2-1.

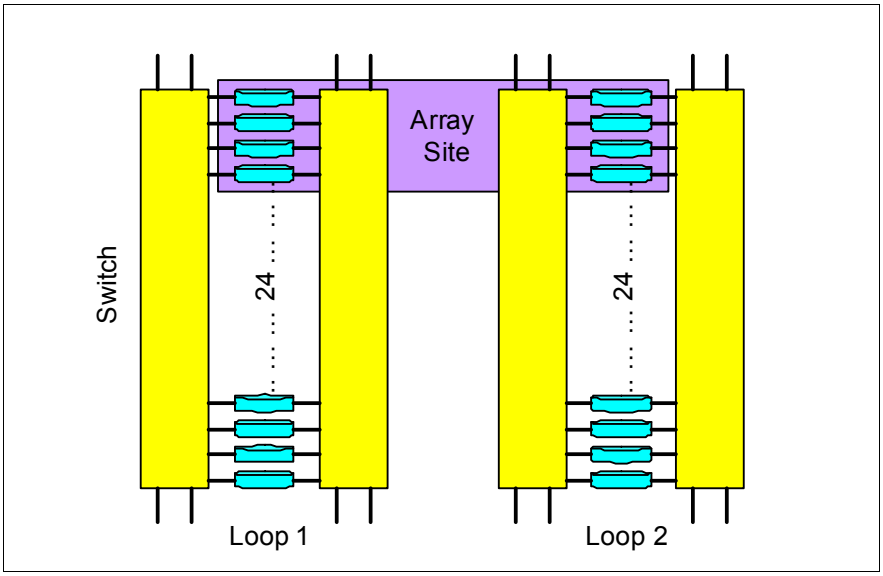

Figure 2-1 Physical disk level - array site

An *array* is created from an array site and is characterized by a specific RAID type by a specific RAID type. Supported RAID types are RAID 5, RAID 6, and RAID 10.

The DS8000 storage family supports hot spare disks. When an array member disk fails, the system automatically replaces the failed member with a hot spare drive and rebuilds the array to restore its redundancy. Spare disks are chosen by the storage system automatically, and more arrays can share one spare disk.

Logical and virtualization layer

In the DS8000 virtualization hierarchy, the first logical construct is called a *rank*, and one array represents one rank. The available space on each rank is divided into *extents*, which are the building blocks of the logical volumes. An extent is striped across all disks of an array. The process of forming a rank has two purposes:

- The array is formatted for either Fixed Block (FB) data for open systems or Count Key Data (CKD) for System z data
- The capacity of the array is subdivided into equally-sized partitions, called extents, with the extent size, depending upon the extent type (FB or CKD)

The next level in the DS8000 virtualization hierarchy is an extent pool that aggregates extents from one or more ranks. A logical volume presented to a server as a LUN is built from these extents, as depicted in Figure 2-2.

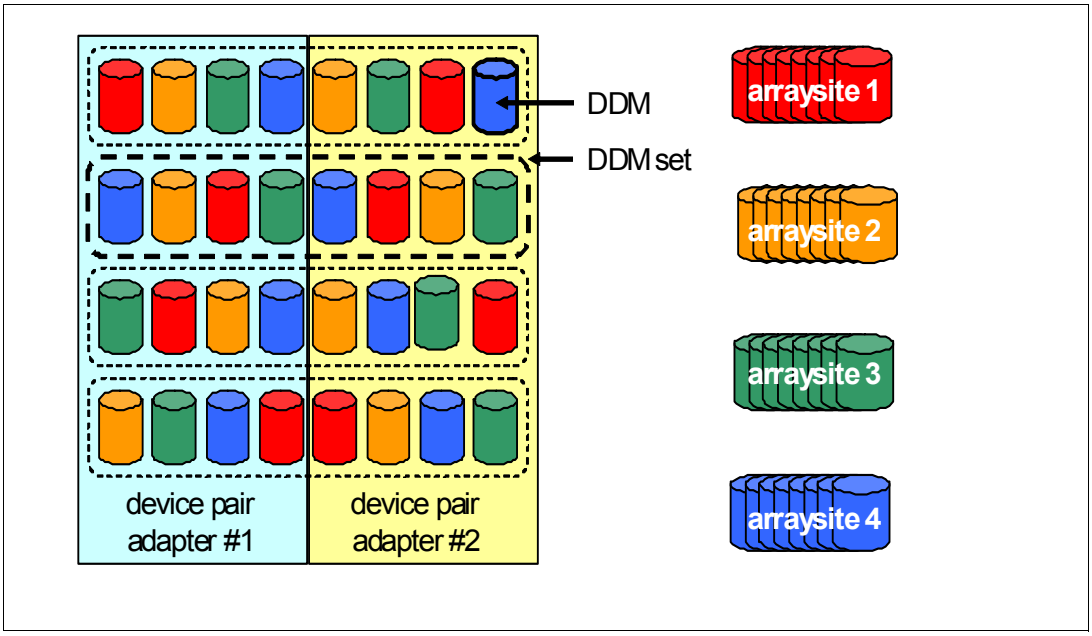

Figure 2-2 The DS8000 virtualization hierarchy, from rank to LUN

For more information, see the Redbooks publication, *IBM System Storage DS8800 Architecture*, SG24-8886, available at this website:

`http://www.redbooks.ibm.com/abstracts/sg248886.html`

2.3 SAN Volume Controller virtualization

The SAN Volume Controller product provides block-level aggregation and volume management for IBM and non-IBM disk storage in a SAN. In other words, the SAN Volume Controller manages a number of LUNs that are presented by back-end storage controllers as one storage pool, and it creates LUNs from that pool that are presented to servers in a SAN. One of the many advantages of this, from the server perspective, is that the storage environment is homogeneous. As shown in Figure 2-3, SAN zoning does not allow servers to access storage directly, because that can interfere with the activities of the SAN Volume Controller.

The SAN Volume Controller is deployed as a pair of nodes that form an *I/O group*. I/O groups form a *cluster* and there can be up to four pair of nodes (four I/O groups) in a cluster. A specific LUN presented to servers is always presented to servers by one I/O group. Within that I/O group, there is one node designated as the preferred node for every virtual disk (VDisk). The server must have access to the LUN using multiple paths. It is the device driver that is managing to which node an I/O request will be issued. Figure 2-3 illustrates this concept.

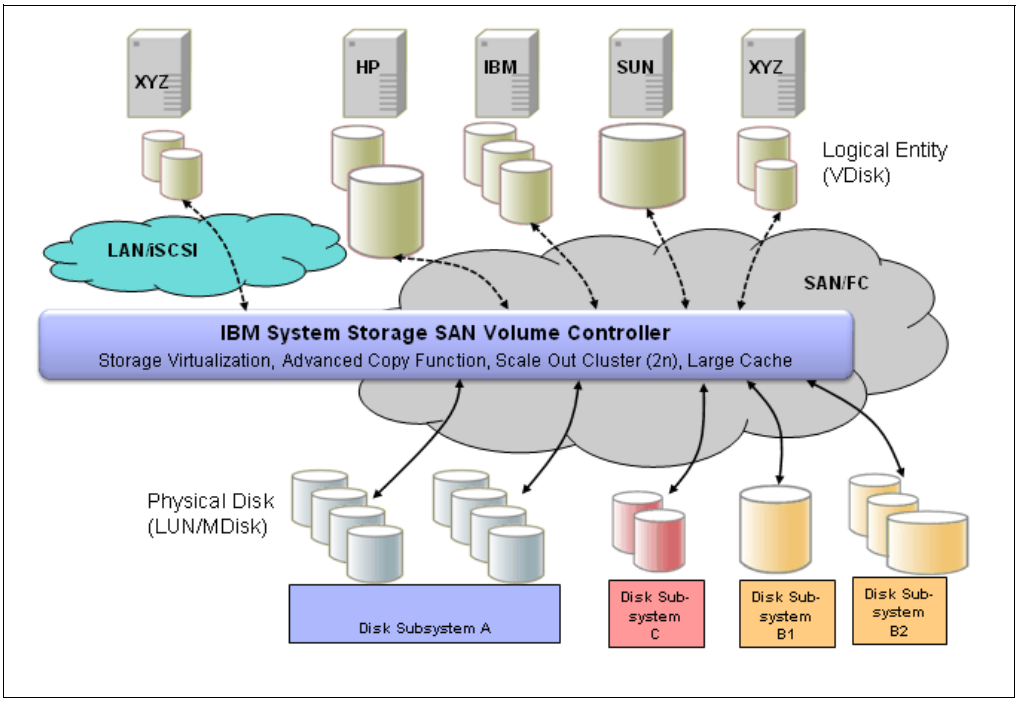

Figure 2-3 Overview of SAN Volume Controller storage level virtualization

The SAN Volume Controller cluster (the SAN Volume Controller I/O groups) sees the storage that is presented to a SAN by the back-end storage controllers as a number of disks called managed disks or *MDisks*. Because the SAN Volume Controller does not attempt to provide recovery from physical disk failures within the back-end controllers, an MDisk is usually, but not necessarily, provisioned from a RAID array. As shown in Figure 2-4, an MDisk can be provided by a SAN disk subsystem or by the solid state drives that are provided by the SAN Volume Controller nodes.

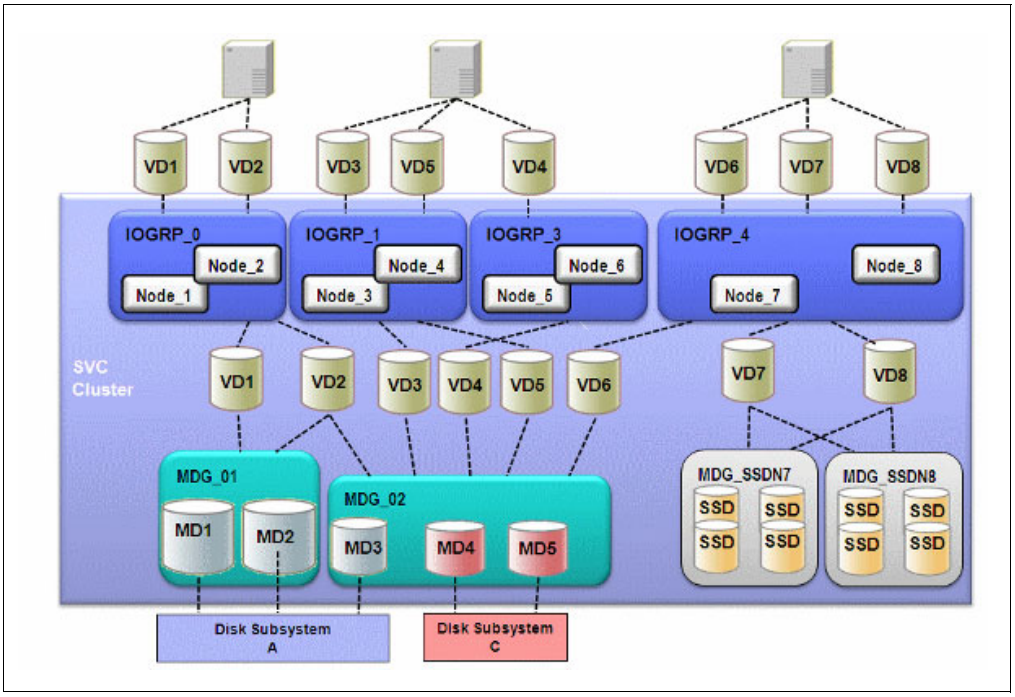

Figure 2-4 SAN Volume Controller storage management overview with MDisk Groups and IO Groups

Each MDisk is assigned to an *MDisk Group* (MDG) and is divided into a number of extents. The size of each extent is selected by the user when the MDG is created and can be from 16 MB (default) up to 2 GB. We suggest using the same extent size for all MDGs in a cluster, which is a prerequisite for supporting VDisk migration between two MDGs. This allows you to, for example, migrate a VDisk between different storage systems. The maximum size of an MDisk is 2 TB, and a SAN Volume Controller cluster supports up to 4,096 MDisks. The maximum size of a VDisk is 256 TB, and a SAN Volume Controller cluster supports up to 4,096 VDisks.

Application servers do not see the MDisks. Rather, they see logical disks called *VDisks*. VDisks are presented by the SAN Volume Controller I/O groups through the SAN (Fibre Channel protocol) or LAN (iSCSI protocol). A VDisk is storage that is provisioned from either one Managed Disk Group (MDG), or, if it is a mirrored VDisk, from two MDGs.

For more information, refer to the Redbooks publication, *Implementing the IBM System Storage SAN Volume Controller V6.1*, SG24-7933, available at this website:

http://www.redbooks.ibm.com/abstracts/sg247933.html

2.4 IBM Storwize V7000 virtualization

The IBM Storwize V7000 leverages disk level virtualization with storage level virtualization. The IBM Storwize V7000 subsystem consists of a set of drive enclosures and control enclosures. Control enclosures contain disk drives and two nodes referred as an I/O group, which are attached to the SAN fabric. Expansion enclosures contain drives and are attached to control enclosures.

The simplest use of the IBM Storwize V7000 is as a traditional RAID subsystem. The internal drives are configured into RAID arrays and virtual disks created from those arrays. The IBM Storwize V7000 can also be used to virtualize other storage controllers.

The Storwize V7000 setup contains a number of internal drive objects, but these drives cannot be directly added to storage pools. They need to be included in a RAID first to provide protection against the failure of individual drives. These drives are referred to as members of the array. RAID levels supported are RAID 0, RAID 1, RAID 5, RAID 6, and RAID 10.

The IBM Storwize V7000 supports hot spare drives. When an array member drive fails, the system automatically replaces the failed member with a hot spare drive and rebuilds the array to restore its redundancy. Candidate and spare drives can be manually exchanged with array members.

Every array has a set of goals that describe the desired location and performance of each array. A sequence of drive failures and hot spare takeovers can leave an array unbalanced, that is, with members that do not match these goals. The system automatically rebalances such arrays when appropriate drives are available.

The next level of IBM Storwize V7000 virtualization uses the same approach as does the SAN Volume Controller (see “SAN Volume Controller virtualization” on page 24).

2.5 IBM XIV Storage System virtualization

The virtualization approach implemented in the IBM XIV Storage System is unique in that it is based on full storage virtualization within the system. This means that all physical storage resources are managed by the storage automatically, and that the virtual logical layer is created without manual intervention.

The traditional approach requires that you carefully plan the relationship between logical structures (such as arrays and volumes) and physical resources (such as disk packs and drives). This will assist you in your efforts to strategically balance workloads, meet capacity demands, eliminate hot spots, and provide adequate performance. The implementation of full storage virtualization employed by the XIV system eliminates many of the potential operational drawbacks that can be present with conventional storage subsystems. At the same time, it maximizes the overall usefulness of the subsystem. The logical structure of the system ensures that there is optimum granularity in the mapping of logical elements to physical disks, thereby guaranteeing an equal distribution of data across all physical resources.

2.5.1 Logical volume

The fundamental building block of a logical volume is a partition. Partitions have the following characteristics on the XIV system:

- All partitions are 1 MB (1024 KB) in size.
- A partition contains either a primary or secondary copy of data.

Each partition is mapped to a single physical disk. This mapping is dynamically managed by the system through innovative data distribution algorithms to preserve data redundancy and load balance. The storage administrator has no control or knowledge of the specific mapping of partitions to drives. Secondary copy partitions are always placed in another module than the one that contains the primary copy partition, as shown in Figure 2-5.

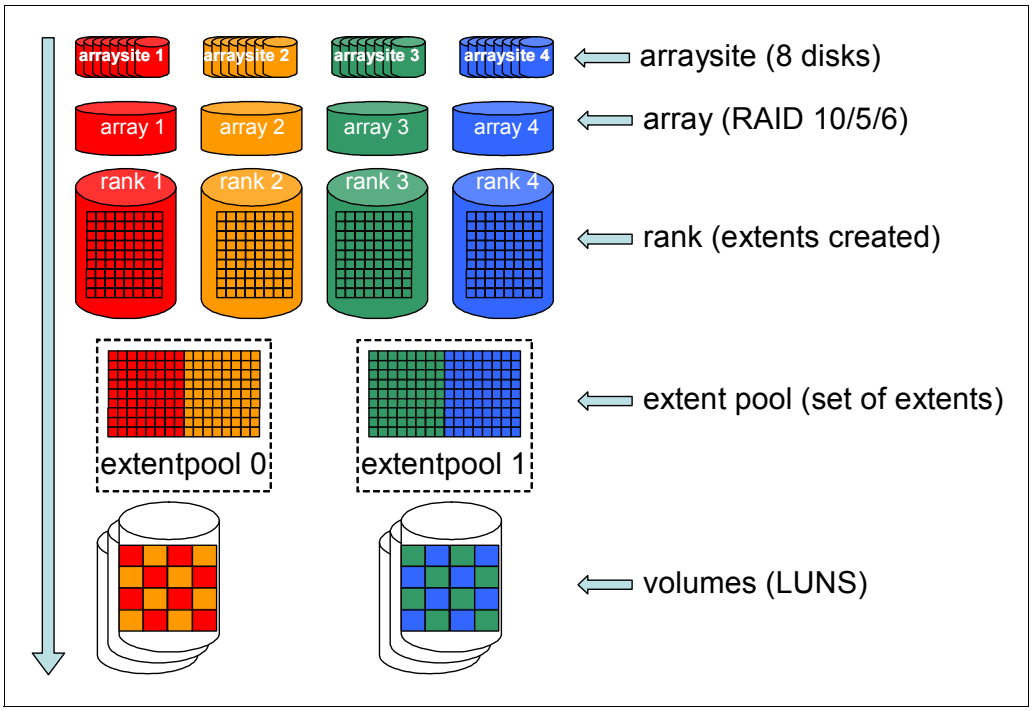

Figure 2-5 Mapping of 1 MB LUN extents to physical disks in data modules

The XIV system manages the distribution of logical volumes over physical disks and modules by means of a dynamic relationship between primary data partitions, secondary data partitions, and physical disks. This virtualization of resources in the XIV system is governed by data distribution algorithms.

The distribution algorithms preserve the load balance of access among all physical disks under all conditions and volume access patterns. The distribution algorithm, although not truly random in nature, in combination with the system architecture, avoids the occurrence of hot spots.

A logical volume is created from partitions. The physical capacity associated with a logical volume is always a multiple of 17 GB (decimal). It is possible to present a logical volume to a host that is not a multiple of 17 GB, but the actual physical space that is allocated for the volume is always the minimum number of 17 GB, the increments needed to meet the block-designated capacity, unless thin provisioning is used. See “XIV Storage System thin provisioning” on page 75 for further details.

2.5.2 Storage pool

Storage pools are administrative boundaries that enable storage administrators to manage separate capacity provisioning for separate applications or departments. Storage pools are not tied in any way to physical resources, nor are they part of the data distribution scheme. The size of a storage pool can vary from 17 GB (one logical volume) to 79 TB, that is, the capacity of the entire system when populated with 1 TB disks. If 2 TB disks are used, you need to have at least two storage pools to cover the physical storage capacity. The storage pool size can by changed dynamically by the storage administrator to meet new needs. For the latest information about storage pool limits, see this website:

http://publib.boulder.ibm.com/infocenter/ibmxiv/r2

Every logical volume is defined within one storage pool and can be moved to any other storage pool, as long as there is sufficient space. This does not trigger any copying of data, and therefore changes are completed instantly and without any system overhead or performance degradation. As a benefit of system virtualization, there are no limitations to the size of storage pools, nor on the association between logical volumes and storage pools.

2.6 IBM SONAS virtualization

IBM Scale Out Network Attached Storage (SONAS) uses the mature technology from IBM high performance computing and is based upon the IBM General Parallel File System (GPFS). Together, we refer to this as the SONAS File System. The file level virtualization and hardware layer abstraction allow the storage administrator to manage storage quite effectively.

2.6.1 Key features of IBM SONAS

These are key features of IBM SONAS virtualization:

- *Global namespace:* File virtualization aids in data sharing by presenting a single namespace (directory and folder structure) for files. This namespace remains constant regardless of where the directory and folder are physically located. This ability not only helps to make it easier to share files among servers. It also facilitates the implementation of tiered storage at the file level.
- *Scalable, enterprise level file server virtualization:* Unified and simplified management of distributed file data through a consolidated logical view. Add and manage data shares, and avoid an impact to users.
- *Grid-based architecture:* Virtualization of hardware failure protection.
- *Cloud storage:* Consolidation of underutilized and fragmented storage into a single, highly utilized storage cloud.

2.6.2 The IBM SONAS architecture

A high level IBM SONAS architecture is shown in Figure 2-6.

From the virtualization perspective, IBM SONAS consists of several storage pods and interface nodes. Each storage pod layer provides highly available, high performance clustered storage, and the interface node layer presets global namespaces to clients using, for example, CIFS, NFS, HTTP, and other protocols. Storage nodes and interface nodes use GPFS high performance cluster technology and low-latency infiniband network to maximize throughput.

Each storage pod consists of high density disk expansion units, connected to two storage controllers that provide a set of RAID 5 or RAID 6 protected LUNs to a storage node. Each storage pod contains two storage nodes that are connected to both storage controllers. There can be from 1 to 30 storage pods and from 2 to 30 interface nodes. There are several types of disks: high capacity SATA disk drives and high performance SAS disk drives.

Storage nodes and interface nodes form a GPFS cluster. Files can be accessed through each of the interface nodes, which providing a highly scalable capability for data access. Additional data access performance can be obtained by adding interface nodes to the limits of SONAS. Collections of files (filesystems) are provided by storage nodes that are gateways to storage controllers and disk drawers. All interface nodes can access all storage on all storage nodes. All storage nodes can send data to any interface node.

Another key aspect of file level virtualization implemented in SONAS is Information Lifecycle Management (ILM) as discussed in “Introduction to SONAS ILM” on page 58. During the life cycle of the file, it can be created on high performance disks, later moved to cost efficient SATA disks, and then, based on policy settings, migrated to tape. From the user’s perspective, the file continues to exist in the original directory and folder.

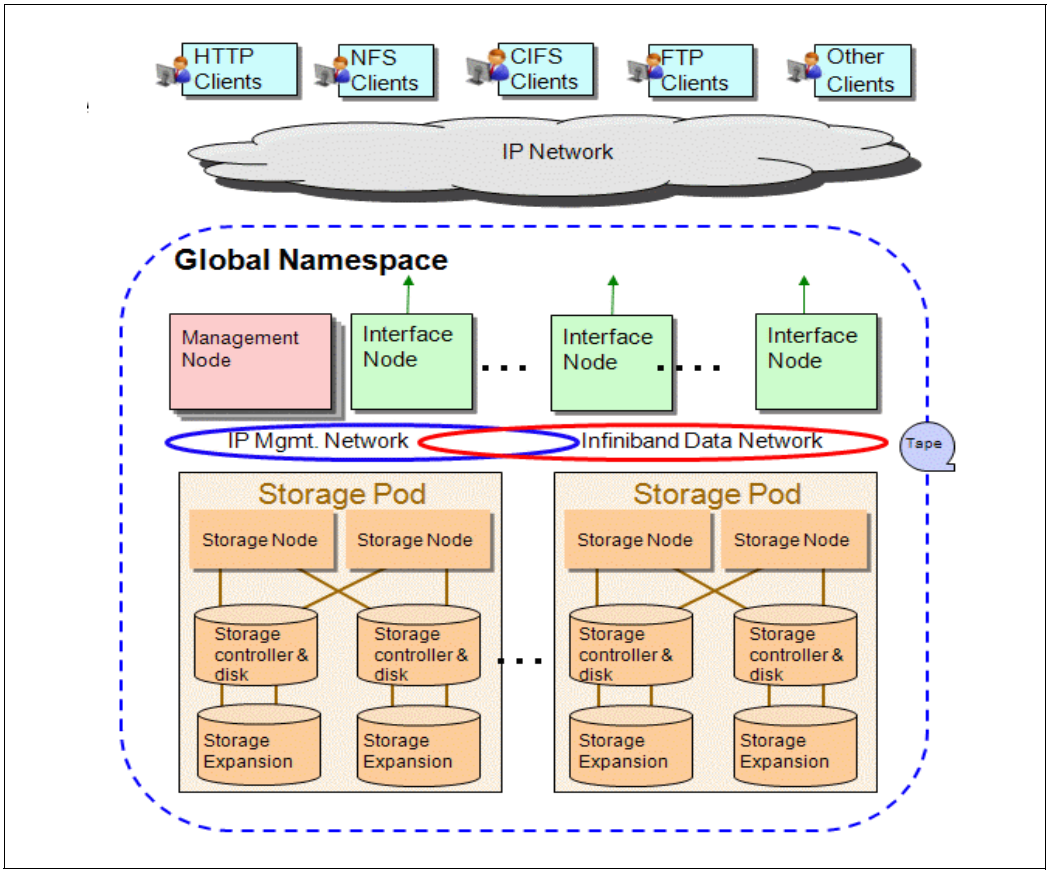

Figure 2-6 SONAS architecture

2.7 IBM Virtualization Engine TS7700

The IBM TS7700 Virtualization Engine is a family of mainframe virtual tape solutions that are designed to reduce or eliminate bottlenecks in a tape environment.

Components and features of the TS7700 are as follows:

- *Virtualization Engine Cluster:* This feature provides the hardware, for example, the controllers, storage system, virtual tapes drives, virtual tapes, optional tape library, and IBM FICON® host attachment required to create the cluster. Up to four ICON channels are available per engine.
- *Virtual drives:* Virtual 3592 tape drives on the TS7700. Up to 256 virtual drives are configurable, per engine.
- *Logical/virtual volumes:* Virtual tape volumes on the TS7700. Virtual drives read and write to virtual volumes. Virtual volumes are stored in cache within the TS7700 for fast retrieval. At a later time, for example, when the physical drives are not busy, data is moved from cache to physical tape drives (for the TS7740) or to a storage expansion frame (for the TS7720). Up to 1,000,000 virtual volumes are configurable, per engine
- *Stacked volume:* Multiple logical/virtual volumes can be written to a physical volume. This volume is called a stacked volume.
- *Tape Volume Cache (TVC):* Disk space provided by the TS7700. Up to 28.8 TB of cache is configurable, per engine, to provide quick access to data.

- *Grid:* This term refers to a network of up to four virtualization engine clusters, available by Internet Protocol (IP), that monitor for and provide disaster recovery and high availability solutions. Data and volume replication between the clusters allows the continuation of production activities, if a cluster becomes unavailable.

Benefits of providing a virtualized tape solution in the mainframe environment include these:

- It provides outstanding backup and recovery capabilities by providing a large number of virtual 3592 tape drives and tapes.
- It allows excellent backup and recovery infrastructure scalability through its cache-centric, modular, and scalable architecture.
- Connecting several TS7700 Virtualization Engines in a *virtual tape grid* architecture virtualizes the location of the virtual volumes, allowing data access and recovery from any node within the grid.
- A tape backup solution provides lower total cost of ownership (TCO).

Figure 2-7 provides an example of a TS7700 grid with four virtualized engine clusters.

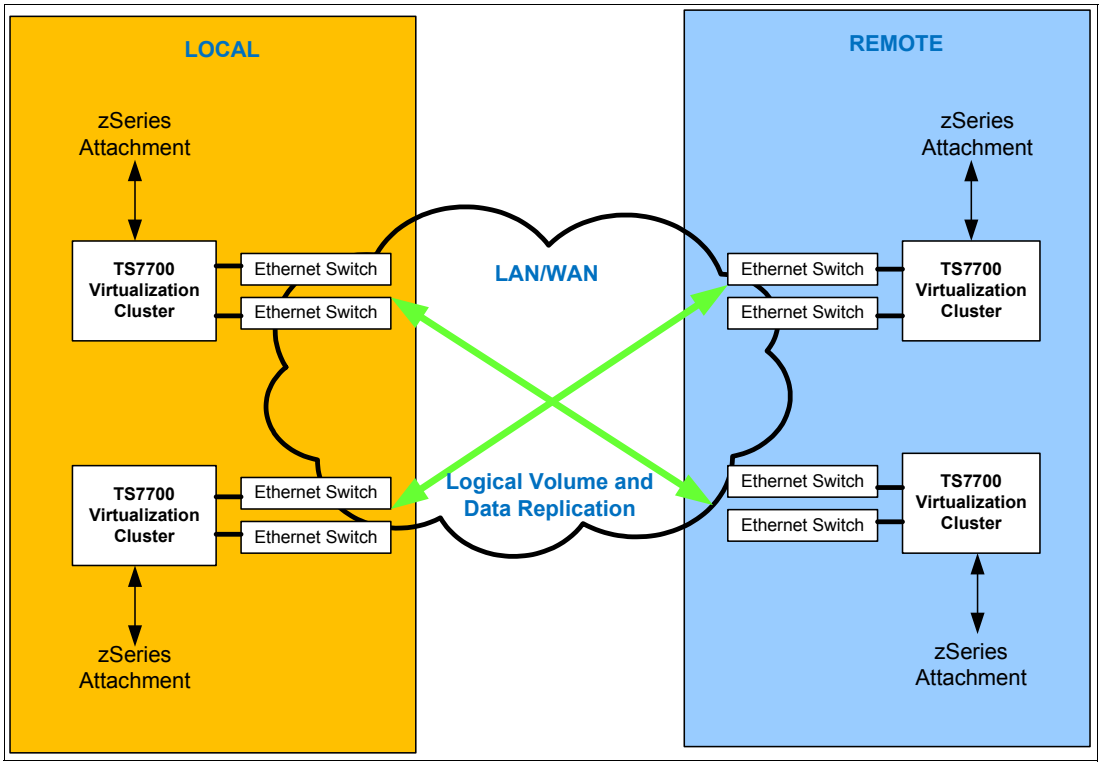

Figure 2-7 TS7700 virtualization engine grid with four clusters

More information about the TS7700 can be obtained from the following websites:

- IBM TS7700 Information Center:

  http://publib.boulder.ibm.com/infocenter/ts7700/cust/topic/com.ibm.storage.ts7740.doc/welcome/ts7740_ic_welcome_8082.html

- IBM TS7700 product website:

  http://www.ibm.com/systems/storage/tape/ts7700/index.html

2.8 IBM TS7600 ProtecTIER virtualization

IBM ProtecTIER is a robust enterprise class data protection system that is capable of tremendously reducing operating expenses and total cost of ownership (TCO) in the area of backup and recovery.

The TS7650 ProtecTIER Deduplication Engine is a Virtual Tape Library (VTL) solution that can backup applications as one or many automated tape libraries. The backup application accesses drives, robotics, and cartridges, just as though it were in a physical tape library environment.

It provides two primary functions:

- A VTL for backup applications. The backup applications access drives, robotics, and cartridges, just as though they were in a physical tape library environment.
- A powerful in-line data deduplication function (using the patented Hyperfactor technology). This is capable of providing up to a 25x reduction in data stored with minimal impact to performance.

2.8.1 The TS7600 ProtecTIER product family

The IBM System Storage TS7600 product family consists of the following products:

For open systems:

- TS7610 Entry Edition: This is an entry level model pre-configured with one VTL and can have up to 5.4TB of disk storage.
- TS7650 Appliance Edition: This is the midrange member of the family with disk storage configurable from 7 to 36TB.
- TS7650G Enterprise Edition or Gateway: This is an enterprise edition available as a single node or a two-node clustered configuration; a maximum of 1 PB of disk storage is configurable in this model.

For mainframe systems:

- TS7680 Gateway Edition for System z: This provides similar functionality to that of the TS7650G in the mainframe environment.

2.8.2 Features of the TS7600 ProtecTIER

The TS7600 provides the following features:

- It provides up to 1000 MB of inline deduplication speed. Using its patented Hyperfactor deduplication technology, only a single, unique instance of redundant data is stored. Hyperfactor provides up to 25X (up to 96 percent) storage capacity reduction, providing direct savings on storage infrastructure costs.
- IBM ProtecTIER can be configured with up to 16 virtual tape libraries, 512 virtual tape drives, and 512,000 virtual tape cartridges.
- It can emulate the Quantum P3000 tape library, the IBM TS3500 tape library, and the DTC VTL.
- It is compatible with many enterprise class backup and recovery solutions such as these:
  - IBM Tivoli Storage Manager
  - IBM Backup Recovery and Media Services

- Hitachi Data Protection Suite - Commvault
- EMC Networker
- HP Data Protector
- Symantec Veritas NetBackup
- Symantec Backup Exec
- CA ArchServe
- Atempo Time Navigator
- Microsoft System Center Data Protection Manager
- Syncsort Backup Express
- BackBone NetVault

- It saves network bandwidth, as only deduplicated data needs to be replicated to a remote location.
- It dramatically improves disaster recovery operations by performing automated electronic replication of data to a remote DR site (and replicated “virtual” tape cartridges can then be cloned).

Figure 2-8 provides a logical overview of several components of the IBM ProtecTIER Gateway.

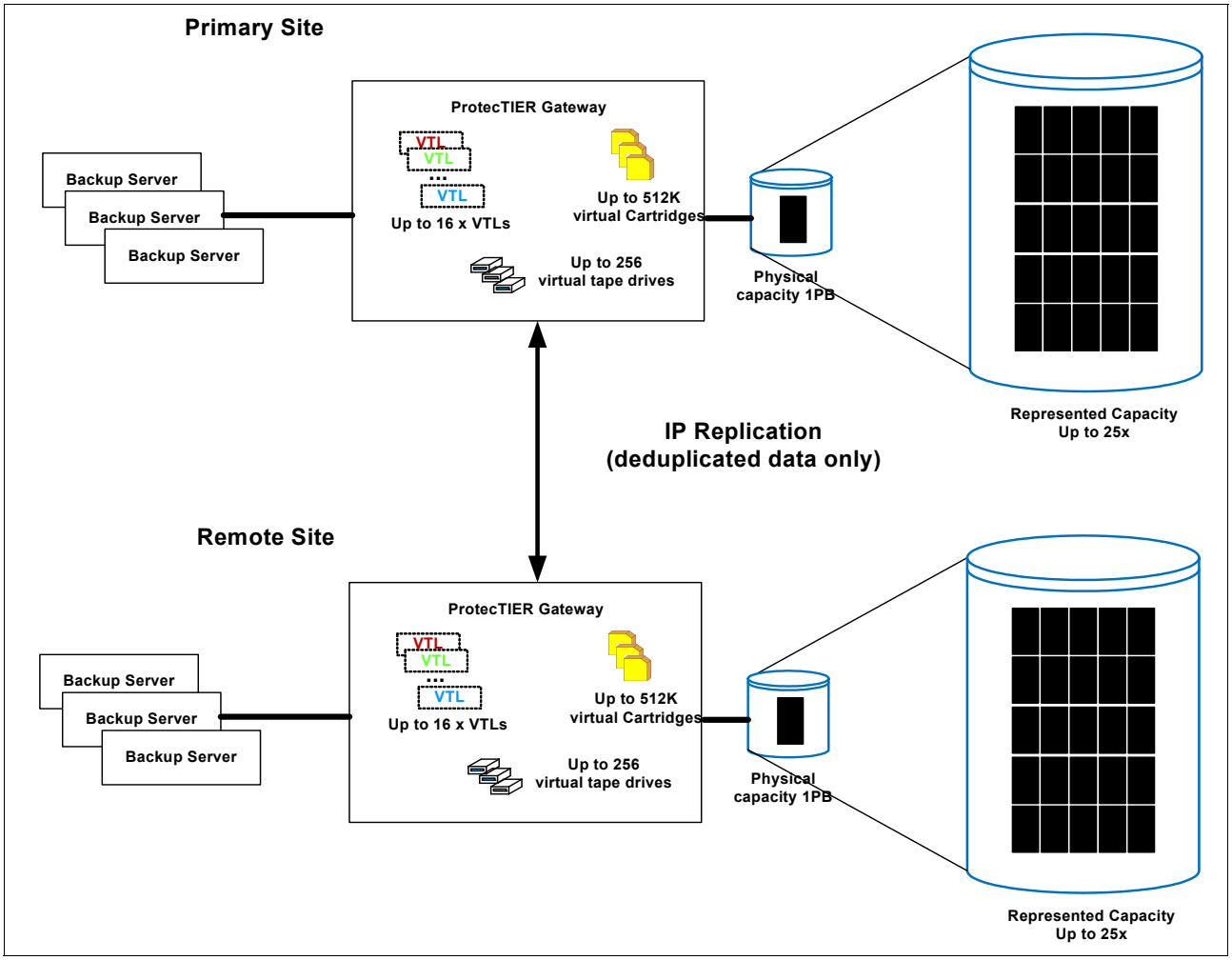

Figure 2-8 Logical representation of the IBM ProtecTIER components

More information about deduplication functions is provided in Chapter 7, “Deduplication” on page 101.

For further details about IBM ProtecTIER for open systems, see the following websites:

- IBM ProtecTIER product website:

  http://www.ibm.com/systems/storage/tape/ts7650g/index.html

- IBM System Storage TS7650, TS7650G, and TS7610 Redbooks:

  http://www.redbooks.ibm.com/redpieces/abstracts/sg247652.html?Open

For further details about IBM System Storage TS7680 ProtecTIER Deduplication Gateway for System z, see the following websites:

- IBM TS7680 product website:

  http://www.ibm.com/systems/storage/tape/ts7680/index.html

- IBM TS7680 Deduplication ProtecTIER Gateway for System z Redbooks:

  http://www.redbooks.ibm.com/abstracts/sg247796.html?Open

2.9 IBM System Storage N series virtualization

The N series portfolio of products is designed to meet the needs of the fast-growing demand for storage in small, medium, and enterprise-level environments. Designed for continuous availability and multifaceted deployments, the N series portfolio is one of the most versatile storage platforms in the industry today. It is based on a single storage architecture and uses a single operating system (N series Data ONTAP), allowing the same management software to manage data center deployments ranging from low-end, back-office to mission-critical.

The N series provides effective storage virtualization, using multiple storage protocols, all within a single, highly available storage system. Protocol support includes Fibre Channel protocol (FCP), Internet Small Computer System Interface (ISCSI), Network File System (NFS), Common Internet File System (CIFS), Hypertext Transfer Protocol (HTTP), and File Transfer Protocol (FTP).

Figure 2-9 illustrates of the use of several protocols in an organization using the N Series.

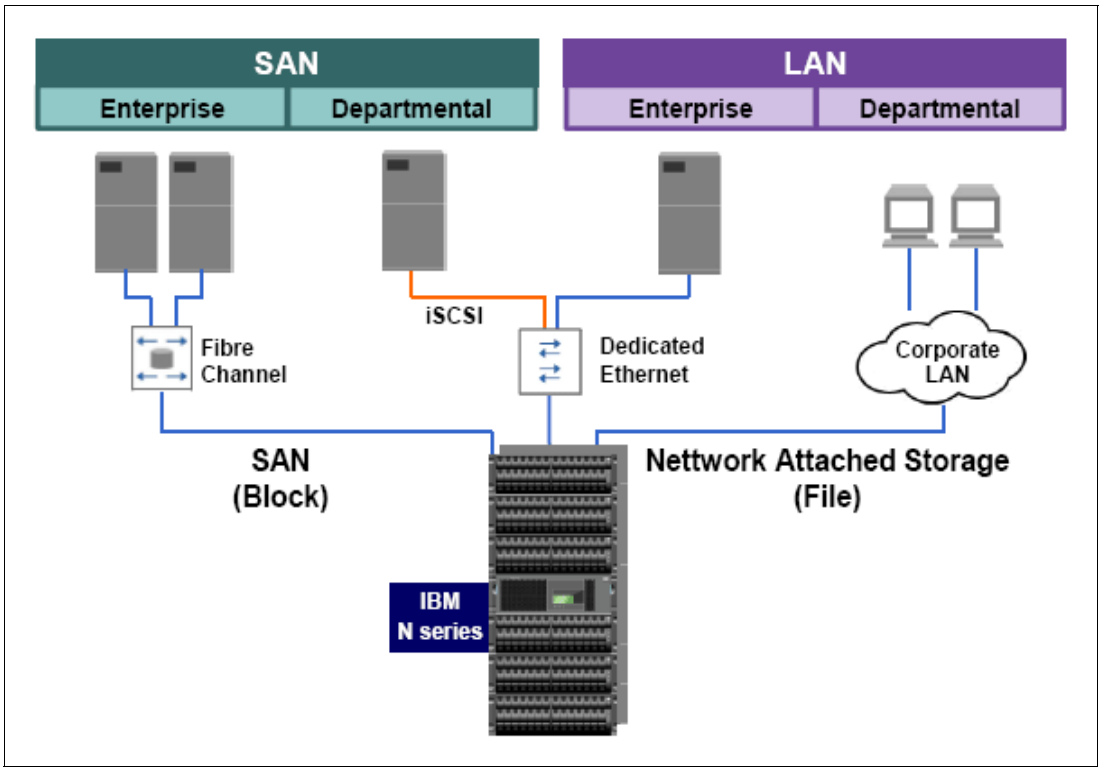

Figure 2-9 Storage virtualization using the IBM System Storage N series

There are multiple levels of virtualization in the IBM N series:

- FlexVol®: These are flexible volumes that contain data that is accessible by the N series, using the multiple network protocols supported by the ONTAP operating system.
- Aggregate: An aggregate is a pool of disks from which space is allocated to FlexVol.
- Vfiler: This is a logical partition of resources within the N series. It allows various departments with separate security domains to share network and storage resources within the same N series hardware.

Figure 2-10 illustrates the virtualization components in the N series.

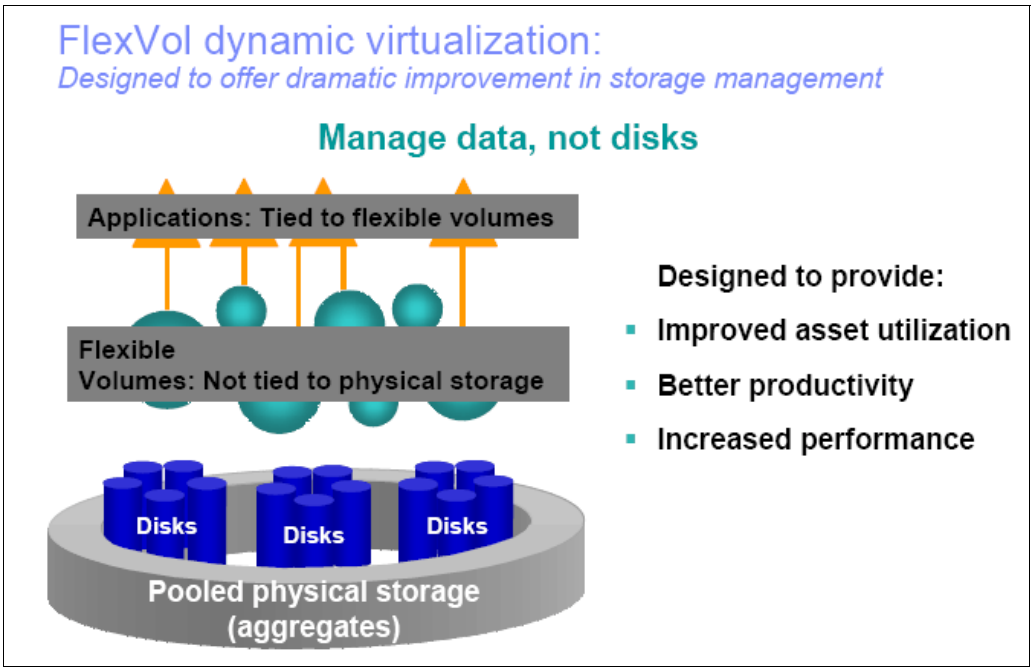

Figure 2-10 N series virtualization components

Complementing the hardware is an extensive list of software designed to provide easy storage management, reporting, data copy services, data deduplication, and disaster recovery functions.

For more details about the N series hardware and software, see the following websites:

- IBM Redbooks website:

  `http://www.redbooks.ibm.com`

- IBM System Storage N series product website:

  `http://www-03.ibm.com/systems/storage/network`

3

Flexible delivery solutions

To cater to the increasing, on-demand needs of business, IT services and infrastructures are moving rapidly towards a flexible utility and consumer model by adopting new technologies. One of these technologies is virtualization. Cloud computing is an example of a virtual, flexible delivery model. Inspired by consumer Internet services, cloud computing puts the user in the “driver’s seat,” that is, users can consume Internet offerings and services by using this self-service, on-demand model.

Cloud computing has the potential to make an enormous impact to your business by providing the following benefits:

- Reducing IT labor costs for configuration, operations, management, and monitoring
- Improving capital utilization and significantly reducing license costs
- Reducing provisioning cycle times from weeks to minutes
- Improving quality and eliminating many software defects
- Reducing end-user IT support costs

From a technical perspective, cloud computing enables these capabilities, among others:

- The abstraction of resources
- Dynamic right-sizing
- Rapid provisioning

This chapter provides an overview of cloud computing technology, with a focus on the service layer of the network architecture known as Infrastructure-as-a-Service (IaaS) and related IBM offerings.

3.1 Cloud computing models

While cloud computing is still a relatively new technology, there are generally three cloud service models, each with a unique focus. The American National Institute of Standards and Technology (NIST) has defined the following cloud service models:

- *Software-as-a-Service (SaaS):* "The capability provided to the consumer is to use the provider's applications running on a cloud infrastructure. The applications are accessible from various client devices through a thin client interface, such as a web browser (for example, web-based email). The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, storage, or even individual application capabilities, with the possible exception of limited user-specific application configuration settings."
- *Platform-as-a-Service (PaaS):* "The capability provided to the consumer is to deploy onto the cloud infrastructure consumer-created or acquired applications created using programming languages and tools supported by the provider. The consumer does not manage or control the underlying cloud infrastructure, including network, servers, operating systems, or storage, but has control over the deployed applications and possibly application hosting environment configurations."
- *Infrastructure-as-a-Service (IaaS):* "The capability provided to the consumer is to provision processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can include operating systems and applications. The consumer does not manage or control the underlying cloud infrastructure but has control over operating systems, storage, deployed applications, and possibly limited control of select networking components (for example, hosts)."

Figure 3-1 depicts some of the cloud computing solutions that are currently available.

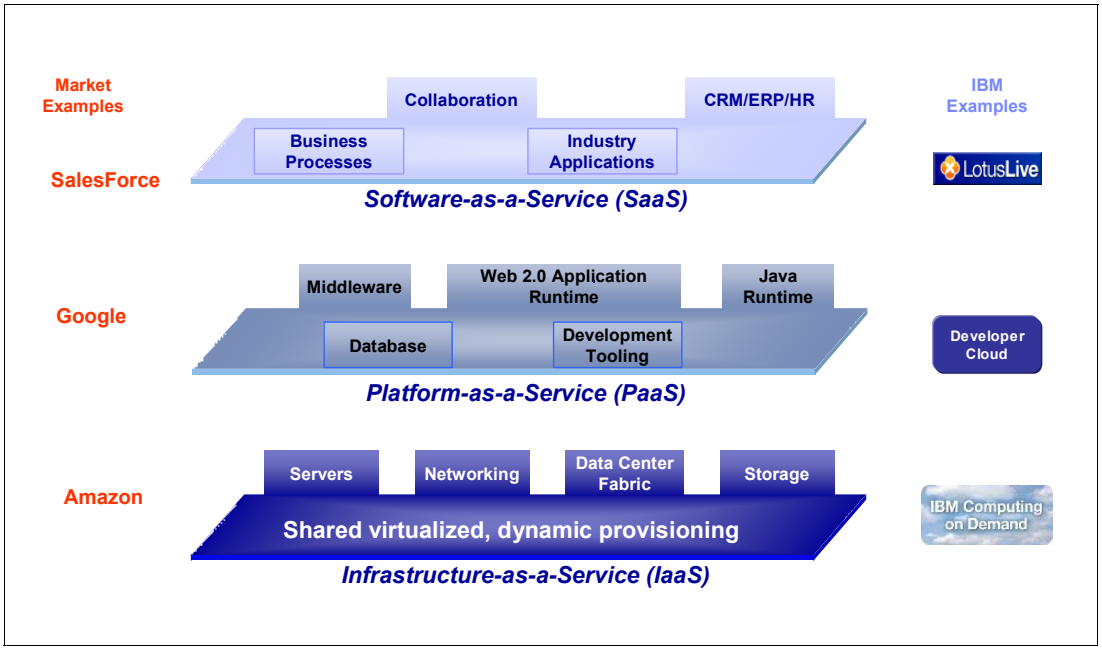

Figure 3-1 Examples of SaaS, PaaS, and IaaS services offered by cloud computing

In addition, NIST has also defined the following models for deploying cloud services:

Private cloud: "The cloud infrastructure is owned or leased by a single organization and is operated solely for that organization."

Community cloud: "The cloud infrastructure is shared by several organizations and supports a specific community that has shared concerns (for example, mission, security requirements, policy, and compliance considerations)."

Public cloud: "The cloud infrastructure is owned by an organization selling cloud services to the general public or to a large industry group."

Hybrid cloud: "The cloud infrastructure is a composition of two or more clouds (internal, community, or public) that remain unique entities but are bound together by standardized or proprietary technology that enables data and application portability (for example, cloud bursting)."

Figure 3-2 summarizes several deployment models used for cloud services.

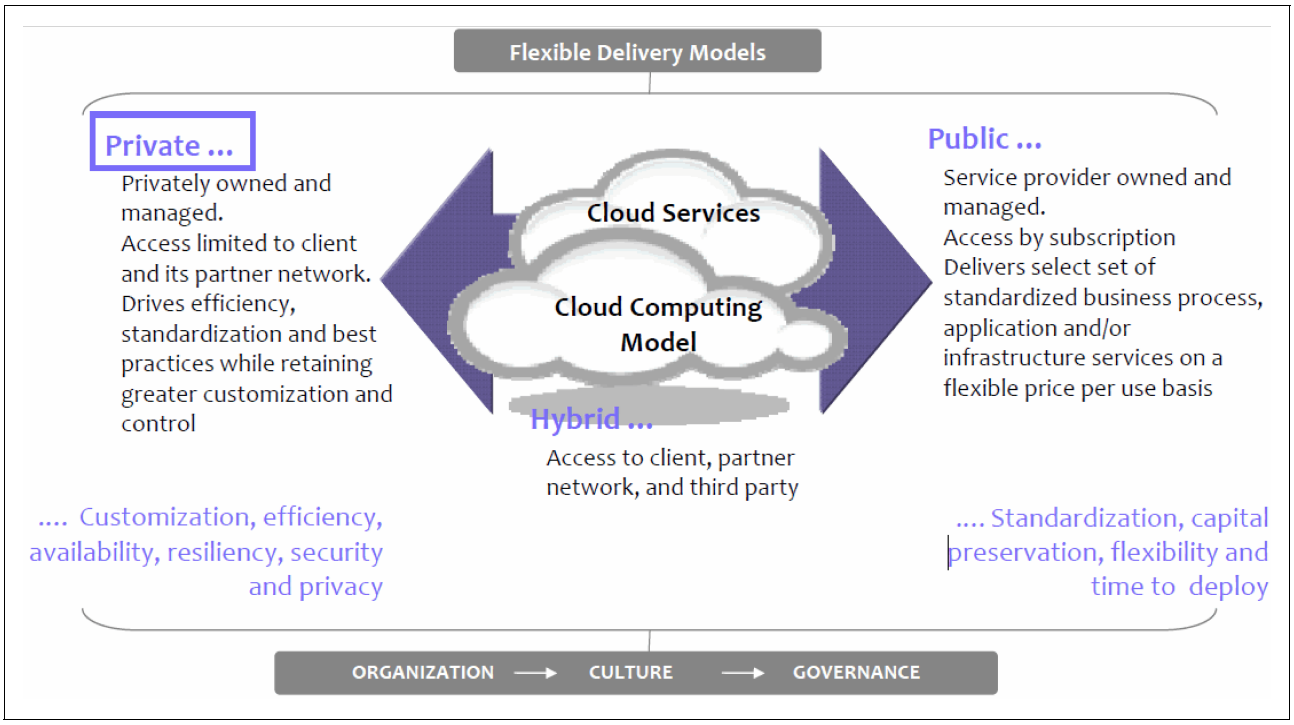

Figure 3-2 Cloud computing deployment models

From a storage perspective, IBM clients, based on their business requirements, can choose to adopt either a pubic or private storage cloud. These storage clouds are defined as follows:

- *Public storage cloud:* This is designed for clients who do not want to own, manage, or maintain the storage environment, thus reducing their capital and operational expenditures for storage. IBM dictates the choice of technology and cloud location, shared infrastructure with variable monthly charges, dynamic physical capacity at the customer level, and security measures to isolate customer data. The public storage cloud allows for variable billing options and shared tenancy of the storage cloud, giving customers the flexibility to manage the use and growth of their storage needs. This is the industry standard view of a storage cloud offering and is comparable to storage cloud offerings by other vendors.

- *Private storage cloud:* With a private storage cloud, customers have the choice of technology and location on a dedicated infrastructure with fixed monthly charges and physical capacity manageable by the customer. Each application can utilize dynamic capacity by sharing the cloud storage among multiple applications.

  Private storage cloud solution technology and services from IBM address multiple areas of functionality, including:

  - Dynamic storage management
  - Scalable capacity and performance
  - Concurrent, multiprotocol data access
  - New levels of manageability

This chapter highlights a few of the technologies and services offered by IBM that make these functionalities available to you.

3.2 IBM cloud computing offerings

IBM offers three types of cloud solutions:

- *Smart Business on the IBM Cloud:* These are standardized cloud services provided by IBM on a pay-per-use basis:
  - *IBM Information Protection Services:* This is an online service that provides cost-effective data backup and recovery and continuous data protection. For more information, see “IBM Information Protection Services - fastprotect online” on page 44.
  - *IBM Smart Business Development and Test on the IBM Cloud:* This cloud offering helps you to assess, plan, design, and implement a flexible development and testing private cloud environment to save on capital and operating costs and to reduce test cycle times, complexity, and risks. Using a self-service test platform, it is designed for the following conveniences:
    - Ease of use
    - Combining service request management, automated provisioning, and configuration management
    - Providing on-demand provisioning of physical and virtualized test resources, including IBM and non-IBM components (for example, operating systems, middleware, storage, network, images, and data)

  This helps you to reduce capital expenses, yet simultaneously gain a dynamically scalable development and test environment to meet changing business needs. For more information, refer to the information available at this website:

  `http://www.ibm.com/services/us/index.wss/offering/midware/a1030965`
- *Smart Business Cloud:* This is an agile cloud IaaS designed to provide rapid access to security-rich, enterprise-class, virtual server environments that are well-suited for development and test activities and other dynamic workloads. The IBM Smart Business Storage Cloud (SBSC) is an example. SBSC offers a storage virtualization solution designed to support your storage optimization efforts. It can help to alleviate data storage challenges by enabling rapid implementation of a scalable, global file storage system, with flexibility in deployment and management options. The solution provides virtualized storage that enables storage and server consolidation, a unified management platform to help reduce outages, and storage management labor demands and costs, and advanced data replication for cost-effective business continuity and disaster recovery. For more information about SBSC, see “IBM Smart Business Storage Cloud (SBSC)” on page 42.

- *Smart Business Systems:* These are purpose-built, pre-configured cloud infrastructure appliances:
  - *IBM Information Archive (IA):* Data archiving is often required to mitigate organizational risks related to regulatory compliance. IA is such an appliance. It is designed to provide a scalable, secure, cost-effective, data archiving solution that uses data deduplication and compression functions to enable efficient data storage. Data are stored securely through enhanced data authenticity and a patent-pending data tampering feature. An indexing and searching mechanism is provided for fast information retrieval. For more information, see “IBM Information Archive” on page 212.
  - *IBM Cloudburst:* This is a hardware appliance designed to provide access to WebSphere virtual images and patterns for easily, quickly, and repeatedly creating application environments that can be securely deployed and managed in a private cloud. It helps application developers to easily and quickly develop, test, and deploy business application environments. In turn, this ends the need for manual, complex, and time-intensive processes associated with creating application environments. For more information, see the information available at this website:

    http://www-01.ibm.com/software/tivoli/products/cloudburst/

Figure 3-3 provides an overview of the IBM cloud computing offerings.

Figure 3-3 IBM cloud computing offerings

For additional details, see these websites:

- http://www.ibm.com/services/us/cloud/leaders/
- http://www.ibm.com/ibm/cloud/services/
- http://www.ibm.com/systems/storage/solutions/cloud/offer.html

3.3 IBM Smart Business Storage Cloud (SBSC)

Storage clouds are the perfect solution for those who need to store, share, process, or serve massive amounts of information. This includes electronic documents (for example, contracts), emails and attachments, presentations, CAD/CAM designs, source code, and web content from images to videos, historical documents, medical images, and photographs. Storage clouds provide massive performance and capacity scalability, along with a rich set of functions to manage, retrieve, replicate, tier, comply, and secure this tidal wave of information.

3.3.1 Storage virtualization solutions and appliances

The IBM Smart Business Storage Cloud (SBSC) is a range of flexible storage virtualization solutions and appliances that helps to alleviate data storage challenges. It does so by enabling quick implementation of highly scalable, global, clustered network attached storage (NAS) systems. It is a pre-packaged service that includes IBM hardware, software, and service components that can be deployed either to your site or the IBM site.

An SBSC is a combination of hardware and software components that are combined to form a system, or a solution, that provides three key features:

- Global namespace
- Clustered filesystem
- Information Lifecycle Management (ILM)

Virtualizing the file sharing system (CIFS, NFS, and so on), using a single, global namespace, means that the user sees one device, rather than the individual servers and storage devices that comprise this virtual NAS device. This is shown in Figure 3-4, as is a summary of the components of SBSC.

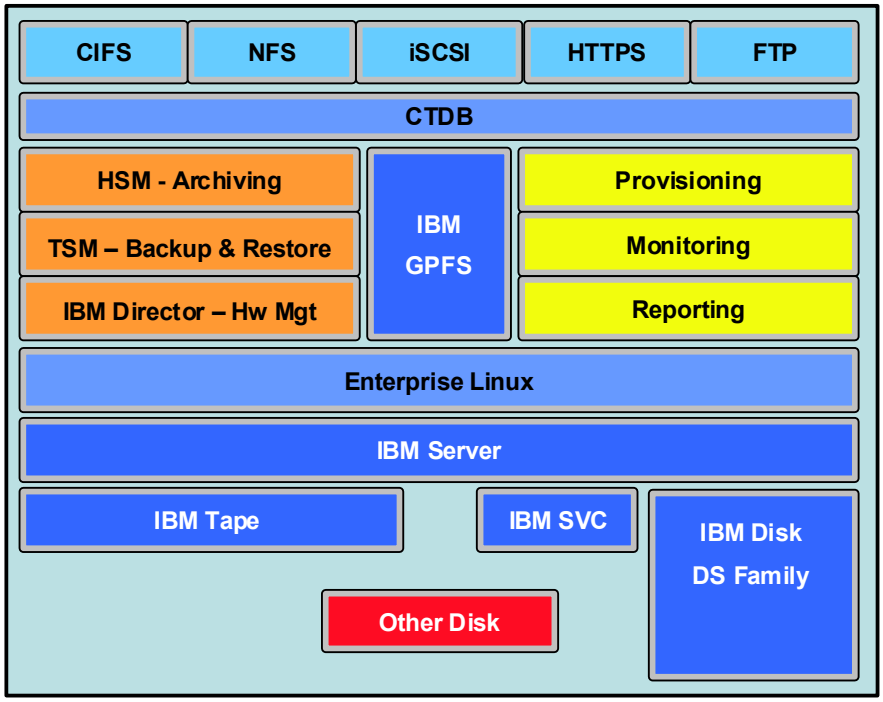

Figure 3-4 Summary of SBSC components

3.3.2 The IBM General Parallel File System (GPFS)

At the heart of the SBSC system is the IBM General Parallel File System (GPFS). GPFS is one of the most scalable commercial filesystems available. It is installed in thousands of nodes and stores petabytes of data in discrete installations. SBSC enables the virtualization of the NFS service by transparently distributing service requests across multiple servers. In turn, this enables the creation of a scale out NAS farm installation.

The product is designed to be highly robust with no single point of failure in the system. It also allows for disaster recovery configurations that use the global file system. These options include synchronous and asynchronous mirroring of data, either on-site or between sites, and cross cluster filesystem mounts.

Application servers connect to SBSC using Common Internet File System (CIFS) and Network File System (NFS) services. Using the cluster trivial database (CTDB), it is possible to provide transparent, nondisruptive failover of CIFS and NFS services between the SBSC nodes. This also provides unparalleled aggregate performance scaling, intelligent load balancing, and simultaneous access to a single file by heterogeneous clients.

From a hardware perspective, the GPFS is made up of servers, SAN switches, and disk storage. Tape storage for archiving data is an optional part of any SBSC solution.

Table 3-1 lists the software used in the SBSC solution.

Table 3-1 SBSC software components

Component	SBSC function
IBM Director	Hardware Management
Tivoli Storage Manager	Backup and Restore
Tivoli Hierarchical Storage Manager	Archiving
DS Storage Manager SMClient	Storage device management
SBSC GUI	User Interface
SBSC Toolset	Management of SBSC system

The IBM SBSC is based on IBM Scale Out Network Attached Storage (SONAS), which offers extreme scale out capability for considerably large storage infrastructures that require high availability.

For additional details about SONAS, see the following websites:

- http://publib.boulder.ibm.com/infocenter/sonasic/sonas1ic/index.jsp
- http://www.ibm.com/systems/storage/network/sonas/index.html
- http://www.redbooks.ibm.com/abstracts/sg247874.html?Open

For additional details about SBSC, see the following website:

http://www-935.ibm.com/services/us/index.wss/offering/its/a1031610

3.4 IBM Information Protection Services - fastprotect online

IBM Information Protection Services - fastprotect online is our managed backup cloud offering. It is designed to provide security-rich, cost-effective, cloud-based data backup for your organization. This is accomplished by providing virtually continuous, scalable protection, along with simplified restore capabilities. It helps to ensure the integrity of critical data by its automatic data backup, its simplified, reliable recovery, and by supporting optimized data protection with scalable pricing.

Your organization faces multiple challenges with regard to safeguarding your data. Consider the expense of maintaining a comprehensive backup data solution for desktops and mobile devices, which might not be protected at all. Add to this the complexities of managing data in diverse locations and with mobile employees, while also maintaining data integrity and availability.

IBM Information Protection Services - fastprotect online helps to provide these benefits:

- Centralized data backup management and the protection of distributed data without monopolizing IT staff
- Speedy recovery in the event of a disaster or data loss
- Simplified execution of backups for desktop and mobile computer users
- A cost-effective pricing structure for a comprehensive backup plan

For additional details about SBSC, refer to the information available at this site:

http://www.ibm.com/services/us/en/it-services/fastprotect-online.html

4

Automated tiering

In modern and complex application environments, the increasing and often unpredictable demands for storage capacity and performance leads to relevant issues in terms of planning and optimization of storage resources.

Most of these issues can be managed by having spare resources available and by moving data, using data mobility tools, or using operating systems features (such as host level mirroring). However, all these corrective actions are expensive in terms of hardware resources, labor, and service availability. Relocating data among the physical storage resources *dynamically*, that is, transparently to hosts, is becoming increasingly important.

IBM Storage Solutions offer two types of automated tiering:

- Automated tiering to optimize performance:

  The Easy Tier feature, available with the DS8000, the SAN Volume Controller, and the Storwize V7000, provides performance optimization. Easy Tier is a built-in, dynamic data relocation feature that provides optimal performance at the lowest cost. Easy Tier is designed to determine the appropriate tier of storage to use, based on data access patterns. Then, it automatically and nondisruptively moves data, at the sub-LUN or sub-volume level, to the appropriate disk tier.

  Easy Tier has been developed by IBM Storage Development in partnership with IBM Research. For more information about the IBM Almaden Research Easy Tier project, see this website:

  `http://www.almaden.ibm.com/storagesystems/projects/easytier/`

- Automated tiering to optimize space management:

  The ability to optimize space management is an Information Lifecycle Management (ILM) function that is available, for instance, with Scale Out Network Attached Storage (SONAS) and with Hierarchical Storage Management, such as that provided by Tivoli Storage Manager and the IBM Data Facility Storage Management System (DFSMS) Hierarchical Storage Management (DFSMShsm).

  Policy-based automation is used to migrate less active data to lower cost storage.

4.1 DS8000 Easy Tier

The DS8000 Easy Tier function is a licensed feature, available at no charge for the DS87000 and the DS8800.

4.1.1 DS8000 Easy Tier overview

The Easy Tier built-in, dynamic data relocation feature, available with the DS8000, operates in the following two modes:

- Automatic mode: DS8000 Easy Tier automatic mode is designed to determine the appropriate tier of storage, based on data access patterns. Then, it automatically and nondisruptively moves data, at the sub-LUN or sub-volume level, to the appropriate disk tier. DS8000 Easy Tier automatic mode also includes automatic intra-tier rebalancing to meet workload requirements.
- Manual mode: DS8000 Easy Tier manual mode allows the user to request migration of volumes between extent pools or the merging of two extent pools. (An extent pool is a logical construct to aggregate the extents from a set of ranks to form a domain for extent allocation to a logical volume.) Manual mode also includes automatic migration of data to other ranks in the extent pool when a user requests that the rank be unassigned from the extent pool.

For more information, refer to the Redbooks publication, *IBM System Storage DS8000 Easy Tier*, REDP-4667, at this website:

http://www.redbooks.ibm.com/abstracts/redp4667.html

4.1.2 DS8000 Easy Tier automatic mode

DS8000 Easy Tier automatic mode includes two types of automatic, transparent data relocation, based on I/O activity:

- Inter-tier relocation: Easy Tier monitors data access activity and then automatically moves heavily accessed data to the higher performing disk tier and lightly accessed data to the lower performing disk tier within an extent pool.
- Intra-tier rebalancing (or Automatic Performance Rebalance- APR): Easy Tier monitors rank activity within an extent pool and automatically moves data between ranks within the same storage tier to balance workload across all ranks in the tier.

With Easy Tier automatic mode, both types of automatic relocation and rebalancing are active simultaneously and work together to optimize performance, by moving data automatically and nondisruptively between two storage tiers or within a storage tier. Both inter-tier relocation and intra-tier rebalancing operate at the sub-LUN or sub-volume level, moving individual extents.

Easy Tier automatic mode operates in *hybrid* extent pools that contain two different tiers from the following three possible storage tiers:

- Solid State Disk (SSD)
- Enterprise (Fibre Channel)
- Nearline (SATA)

Tip: A hybrid extent pool where Easy Tier automatic mode is active is often referred to as a “managed” extent pool.

A hybrid extent pool can be created by initially assigning ranks from two different storage tiers to a newly created pool, by adding ranks of a different storage tier to an existing pool, or by merging extent pools containing ranks from two different storage tiers.

Easy Tier automatic migration activity is nondisruptive and is automatically managed to prevent any performance degradation due to the migration activity. This significantly improves the overall storage-cost performance ratio and simplifies performance tuning and management.

Easy Tier automatic mode requires an Easy Tier license feature, available without charge. When the license feature is activated, Easy Tier automatic mode is enabled by default for hybrid extent pools (pools containing two different disk tiers). Easy Tier automatic mode can be disabled for the system if desired.

DS8000 Easy Tier automatic mode is not supported on encryption-capable systems. (Easy Tier manual mode, described in “DS8000 Easy Tier manual mode” on page 49, is supported for encryption-capable systems.)

DS8000 Easy Tier automatic mode support includes Fixed Block (FB), Count Key Data (CKD) and IBM i volumes.

DS8000 standard (fully-allocated) volumes are supported. Thin provisioned volumes, including Extent Space Efficient (ESE) volumes and Track Space Efficient (TSE) volumes (or the TSE repository) can coexist in an extent pool in which Easy Tier is active, but they will not be managed by Easy Tier.

Easy Tier supports volumes that are the source or target of a DS8000 Copy Services relationship (for example, a FlashCopy source or target volume, or a Metro Mirror source or target volume). Source and target volumes are managed independently by Easy Tier, based on their respective workload characteristics.

Inter-Tier relocation

Easy Tier automatic mode allows you to gain the benefits of using a higher performance disk tier with a much lower cost and without manual intervention. Easy Tier automatic mode monitors data access patterns to determine the appropriate tier of storage and then automatically and nondisruptively moves data, at the sub-LUN or sub-volume level, to the appropriate disk tier.

Sub-volume or sub-LUN granularity is important, because workloads typically concentrate I/Os on only a subset of the data blocks within a volume or LUN (this is often referred to as ‘skew.’) Moving only the high activity (or *hot*) extents to a higher performing disk provides the biggest performance benefit for the least cost. Moving only hot extents rather than the entire LUN or volume reduces the amount of higher cost disk capacity required to improve performance. It also reduces the time and system resources required for migration.

Easy Tier monitors the I/O activity for volume extents in hybrid extent pools over a 24 hour period. Based on these results, Easy Tier creates an extent migration plan and dynamically moves *hot* (high activity) extents from the lower disk tier to the higher disk tier within the hybrid extent pool. Easy Tier automatic mode also moves extents for which activity has dropped off, or *cooled* (low activity), from a higher disk tier to a lower disk tier. Easy Tier automatic mode monitoring and migration continues on an ongoing basis.

One example of Easy Tier automatic mode is shown in Figure 4-1. In this example, the hybrid extent pool contains one SSD rank and three enterprise ranks (for example, Fibre Channel).

Easy Tier has monitored I/O activity for the volume extents for 24 hours, automatically relocated extents as appropriate, including these:

- Migrating hot extents to the higher performance disk tier in the pool (SSD)
- Migrating *warm* (previously hot but no longer hot) or *cold* (inactive) extents to the lower performance disk tier in the pool (Enterprise)

Figure 4-1 shows that four extents are hot and have been migrated from the enterprise storage tier to the SSD storage tier.

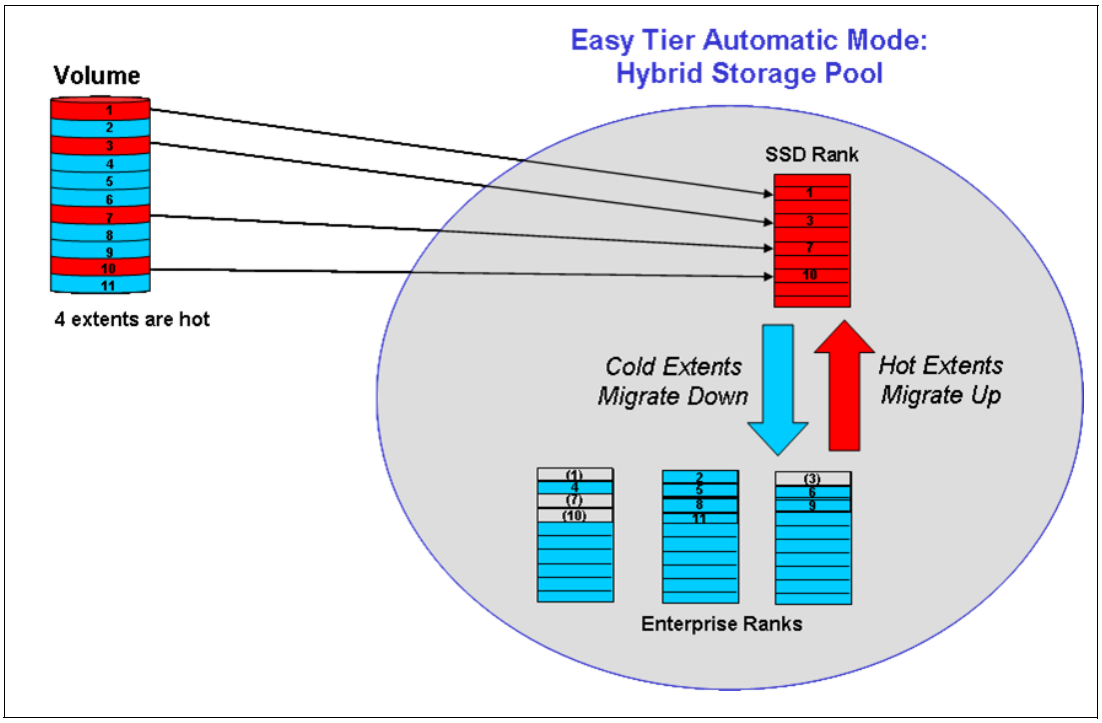

Figure 4-1 Example of Easy Tier automatic mode inter-tier relocation

Easy Tier also rebalances extents within a single disk tier whenever workload is unbalanced across the ranks within a tier, as discussed in “Intra-tier rebalancing” on page 48.

Intra-tier rebalancing

Easy Tier intra-tier rebalancing (or Automatic Performance Rebalance, APR) works in conjunction with Easy Tier inter-tier relocation in optimizing workload performance. Easy Tier inter-tier relocation is discussed in “Inter-Tier relocation” on page 47.

Easy Tier automatic mode intra-tier rebalancing is designed to maximize performance in terms of I/Os and bandwidth *within* a storage tier in an extent pool. Easy Tier automatically moves extents between ranks within the same storage tier to balance workload across all ranks in the tier. Intra-tier rebalancing can eliminate hot spots and improve workload performance without requiring manual intervention.

Easy Tier intra-tier relocation is available for hybrid extent pools managed by Easy Tier.

In Figure 4-2, enterprise Rank 1 on the lower left contains three extents (numbered 1, 2, and 3). These extents are not hot enough to be migrated to the SSD rank, but are producing more workload for Rank 1 than the extents on Ranks 2 and 3. Easy Tier automatic intra-tier relocation will move Extent 2 to Rank 2 and Extent 3 to Rank 3 to re-balance the workload across the three enterprise ranks in the extent pool.

Intra-tier relocation will not only take place in response to naturally occurring workload imbalance, but also in response to changes in the hybrid extent pool. Workload imbalance caused by adding empty ranks, or merging ranks from another extent pool, will also automatically be corrected by intra-tier extent relocation. If deletion or manual migration of volumes results in unbalanced workload across ranks, it will be rebalanced by automatic intra-tier relocation of extents.

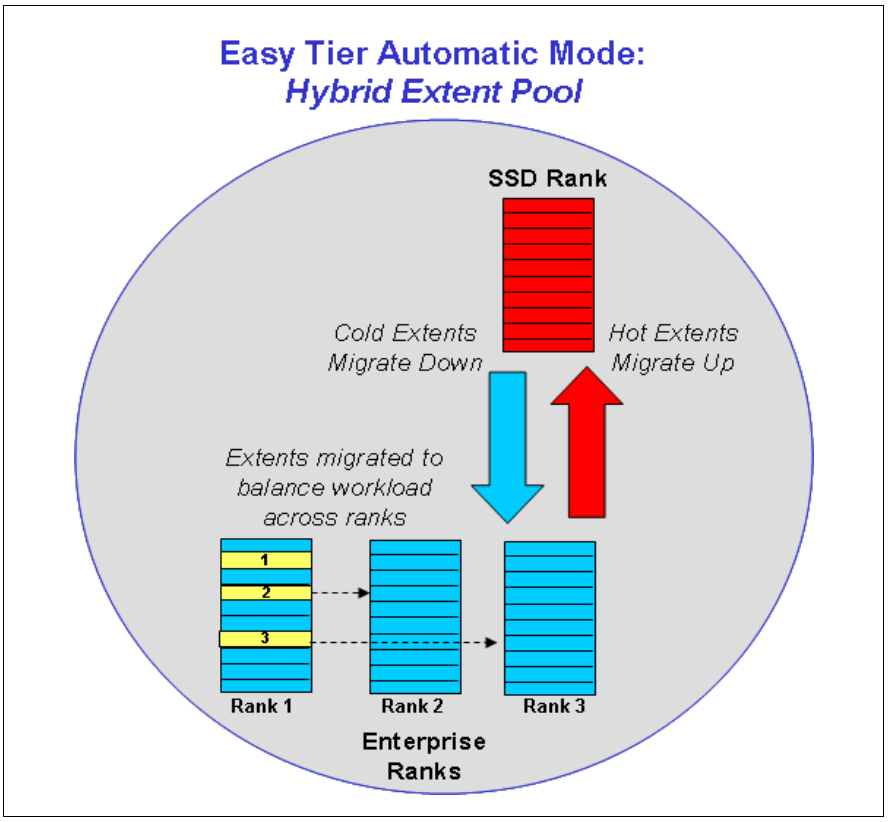

Figure 4-2 Example of Easy Tier automatic mode intra-tier relocation

4.1.3 DS8000 Easy Tier manual mode

Easy Tier manual mode provides the capability for the user to request migration of a volume between extent pools, merging of one extent pool into another, or rank depopulation (removal of all extents from a rank so that the rank can be removed from an extent pool.) All of these Easy Tier manual mode functions are dynamic and nondisruptive to host access.

Easy Tier manual mode can be used simultaneously with Easy Tier automatic mode for hybrid extent pools. Easy Tier manual mode can also be used for extent pools where Easy Tier automatic mode is not available (for example, in extent pools containing a single disk tier.)

Easy Tier manual mode functions require the no-charge Easy Tier licensed feature.

Easy Tier manual mode is supported on encryption-capable systems and non-encryption-capable systems.

Volume migration

Volume migration allows you to migrate a volume to another extent pool or to the same extent pool. Volume migration is dynamic, nondisruptive, and transparent to host access.

A volume can be migrated to another extent pool to redistribute workload across pools (for example, one heavily accessed volume can be moved to an extent pool that does not have other busy volumes.) Manual volume migration can also be necessary, because the performance characteristics available in the target pool (for example, disk drive type or RAID configuration) are more appropriate for the volume. A volume can also be migrated to another pool for managing by Easy Tier automatic mode. For example, a volume that is shown to have hot extents (as discussed in “DS8000 Monitoring and the Storage Tier Advisor Tool (STAT)” on page 54) can be manually moved from a single-tier extent pool to a hybrid (two-tier) extent pool to enable Easy Tier automatic data relocation.

A volume can be migrated to another extent pool to redistribute available disk capacity between extent pools (for example to free up space in the source extent pool, or to take advantage of additional space for expansion in the target extent pool). Volume migration can also be used to change a volume extent allocation method (for example, from rotate volumes to rotate extents.) The extent allocation method can be changed when migrating a volume between pools or by requesting migration within the same pool.

Figure 4-3 shows volume migration to another extent pool. Here, extent pool 0 contains a single rank with three extents for the volume. Using Easy Tier manual mode, the volume is migrated from extent pool 0 to extent pool 2. As part of the migration, the extent allocation method is changed to rotate extents, and the three extents for the volume are spread evenly across the three ranks in extent pool 2.

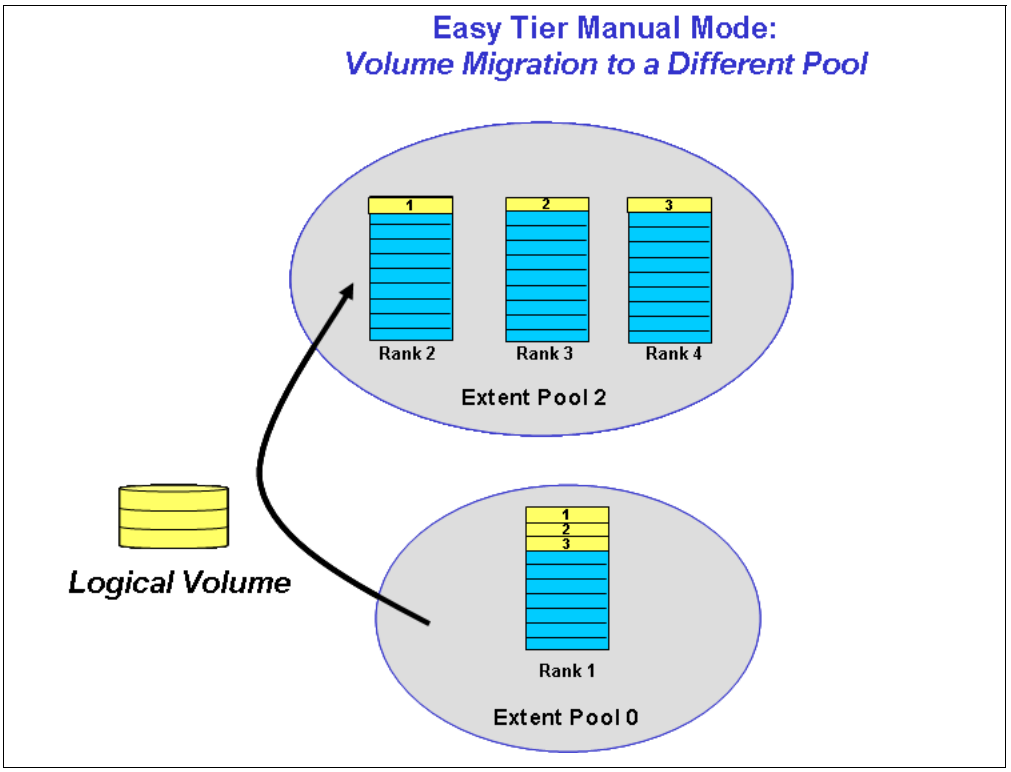

Figure 4-3 Volume migration between two extent pools

Volume migration within a single pool (for example, migration from and to the same pool) can be used to spread extents evenly across all ranks after additional ranks have been added to an extent pool.

Figure 4-4 shows Easy Tier manual mode volume migration within a single pool. At the bottom of the figure, we see extent pool 0 before volume migration. The volume has all three of its extents on a single rank (rank 1), which was the only rank in the pool when the volume was created. Ranks 2 and 3 are new ranks which have been added to pool 0. At the top of the figure, we see extent pool 0 after manual volume migration to and from extent pool 0 is used to spread the three extents for the volume across all three ranks now existing in the pool.

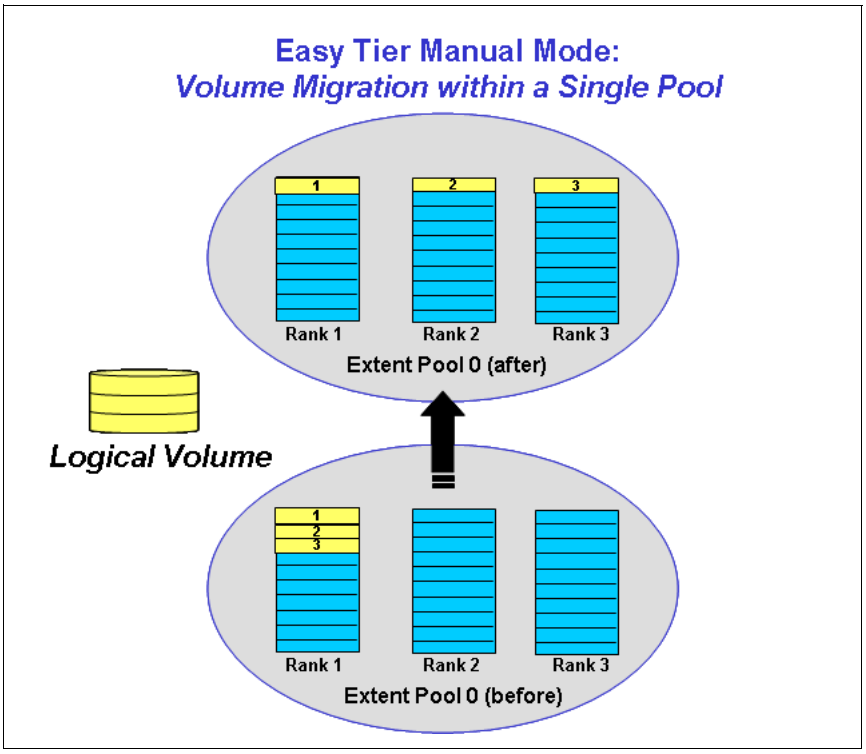

Figure 4-4 Volume migration within a single extent pool

Volume migration (either between pools or within a single pool) is supported for single-tier extent pools that are not managed by Easy Tier automatic mode. Volume migration between extent pools is supported for hybrid extent pools that are managed by Easy Tier automatic mode. Volume migration within a single hybrid extent pool that is managed by Easy Tier automatic mode is not supported. Easy Tier automatic mode will automatically rebalance extents as necessary for performance, both within and across tiers for the extent pools it is managing.

Volume migration is currently supported for standard (fully-allocated) volumes only.

Volume migration is supported between extent pools managed by the same DS8000 server (even-numbered pools managed by Server0 or odd-numbered pools managed by Server1.)

Extent pool merge

Extent pool merge allows you to request that two existing extent pools be combined (one pool is merged into the other). Extent pool merge is dynamic, nondisruptive and transparent to host access.

Figure 4-5 shows extent pool EPFC02 (on the right) being merged into extent pool EPFC00 (on the left.) The ranks and LUNs from extent pool EPFC02 are simply associated with extent pool EPFC00, and then extent pool EPFC02 is deleted.

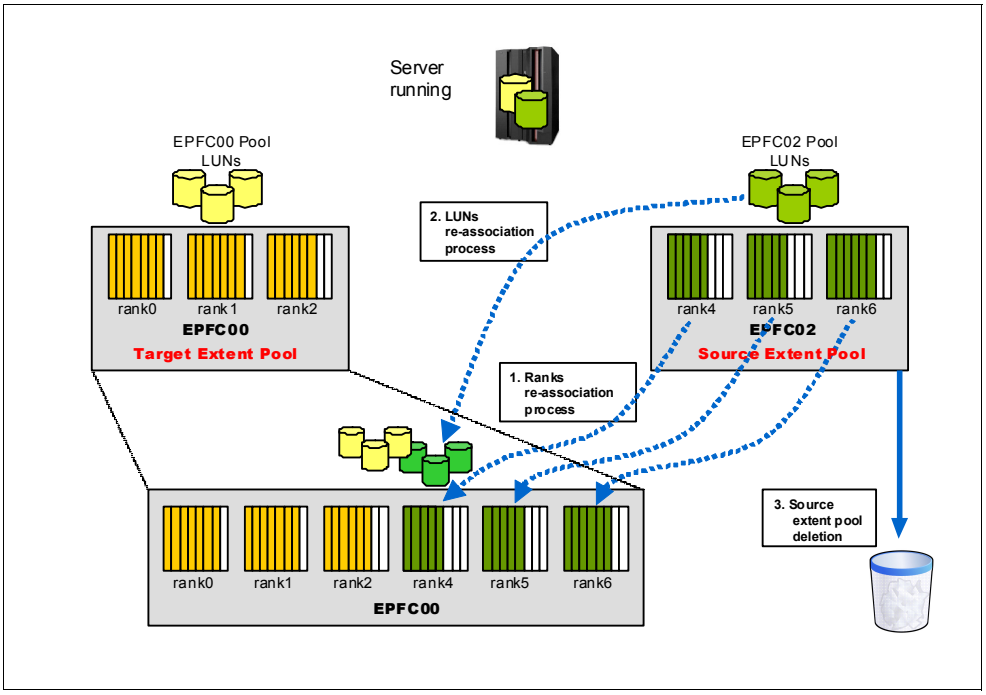

Figure 4-5 Easy Tier manual mode extent pool merge

Two single-tier extent pools with different disk tiers can be merged to create a hybrid extent pool for management by Easy Tier automatic mode. Easy Tier automatic mode inter-tier and intra-tier workload balancing are automatically enabled after the extent pool merge. A single-tier extent pool can be merged into a hybrid pool to enable management by Easy Tier automatic mode. Again, both inter-tier and intra-tier workload balancing is automatically enabled after the extent pool merge. Two hybrid extent pools managed by Easy Tier automatic mode can also be merged, and inter-tier and intra-tier balancing will then take into account the combined workloads and resources.

Two single-tier extent pools containing the same disk tier can be merged to increase capacity within the target extent pool. Extent pool merge is supported both for single-tier extent pools that are not managed by Easy Tier automatic mode and for hybrid extent pools that are managed by Easy Tier automatic mode.

Extent pool merge is not currently supported when both extent pools contain thinly provisioned volumes (ESE or TSE) or when thinly provisioned volume extents, virtual space, or a space efficient repository exist on SSD ranks.

Extent pool merge is supported for pools managed by the same DS8000 server (for example, even-numbered extent pools managed by Server0 or odd-numbered extent pools managed by Server1).

Rank depopulation

When you request that a rank be unassigned from an extent pool, rank depopulation provides automatic, nondisruptive migration of extents used for data to other ranks in the extent pool.

A rank can be depopulated with the purpose of changing its RAID type or storage type (FB or CKD), or assigning it as free capacity in the same or another extent pool. Rank depopulation can also be used to upgrade disk drive technology. This can occur, for example, when moving extents off of ranks with old disk drive technology and onto ranks with new disk drive technology or to reconfigure extent pools (for example, to change a pool with multiple disk drive types to a single disk drive type.)

Figure 4-6 shows an example of rank depopulation. At the bottom of the figure we see extent pool 0 before rank depopulation. There are three ranks in extent pool 0 and each rank contains 2 extents. When the user requests that rank 1 be unassigned, Easy Tier manual mode rank depopulation begin, causing extents to be moved off rank 1 and onto the other ranks in the extent pool (ranks 2 and 3.) At the top of the figure we see extent pool 0 after all extents have been removed from rank 1 and rank 1 has been unassigned from the extent pool. Now there are two ranks in the extent pool and each contains three extents.

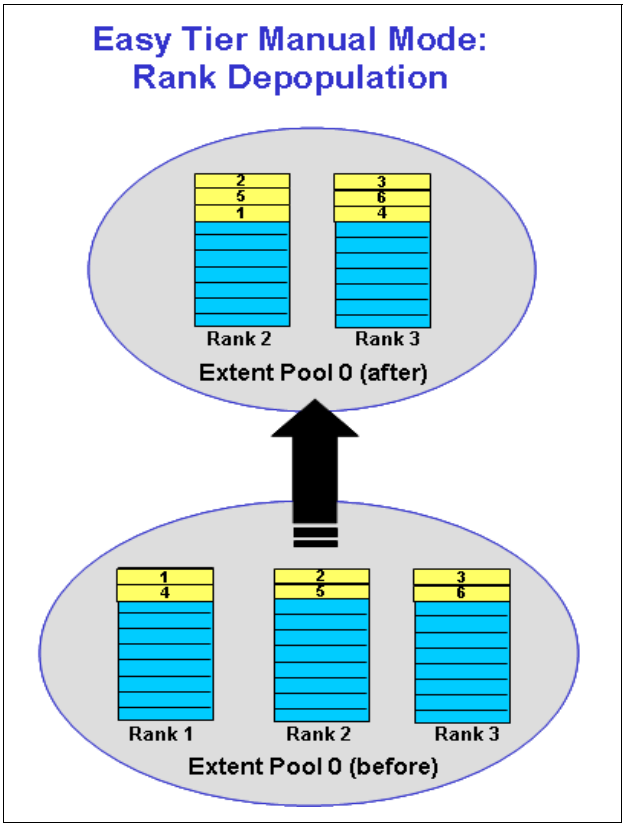

Figure 4-6 Example of Easy Tier manual mode rank depopulation

Rank depopulation is supported both for single-tier extent pools that are not managed by Easy Tier automatic mode and for hybrid extent pools that are managed by Easy Tier automatic mode.

Rank depopulation is not currently supported for ranks with thinly provisioned volumes (ESE or TSE) or related resources (virtual space or repository.)

4.1.4 DS8000 Monitoring and the Storage Tier Advisor Tool (STAT)

Most workloads access certain areas of a volume or LUN much more heavily than others, so the DS8000 provides the capability to collect performance information at the sub-LUN or sub-volume level. These statistics identify the extents that are heavily used and can benefit most from relocation to a high performance disk tier. These statistics also identify extents that are less active or no longer heavily used and need to remain on, or be relocated to, a lower performance disk tier.

DS8000 extent level I/O monitoring does not require the no-charge Easy Tier licensed feature.

DS8000 extent level I/O monitoring is supported for FB, IBM i, and CKD volumes.

DS8000 extent level I/O statistics are gathered continuously and analyzed every 24 hours. These statistics are used as input to Easy Tier automatic mode algorithms for automatic data relocation, but they are collected even if Easy Tier is not active in the system. Statistics are collected for both single-tier extent pools and multi-tier extent pools.

Statistics can be offloaded using the DS8000 GUI or the DS8000 Command Line Interface (DSCLI) and used with the STAT tool to create reports. The tool is a Windows application that creates a set of reports, based on extent performance data collected over a 24-hour cycle. The reports include these:

- *System summary:* This shows the amount (for example, GB) of heavily accessed (“hot”) data in the system, the percentage of data that is hot. It also provides a list of the extent pools with the storage tier or tiers they contain and an indication of whether a storage tier in the extent pool is overloaded in terms of IOPS or bandwidth.
- *System-wide recommendation:* This gives a suggestion for each extent pool on the amount of higher tier disk that can be added to service the hot data with estimated performance improvement and migration time and the amount of lower tier disk that can be added to hold the cold data.
- *Extent pool statistics and improvement recommendation:* This gives a suggested amount of higher tier disk that can be added to the extent pool to service the hot data with estimated performance improvement and migration time, in addition to a suggested amount of lower tier disk that can be added to hold the cold data. It also shows the utilization status in terms of IOPs for each rank in the extent pool and a list of all the volumes in the extent pool with amounts of hot, warm, and cold extents for each volume and the amount of extents on SSD.

The STAT report can be used with Easy Tier automatic mode to view the results and benefits of data relocation. The STAT report can also be used without Easy Tier to estimate the potential benefits of using different storage tiers with Easy Tier automatic mode smart data relocation. This aids in understanding workload characteristics, or can be used as input for other storage administration decisions.

The DS8000 performance monitoring capabilities do not require the no-charge Easy Tier license.

The STAT tool can be downloaded from:

http://www.ibm.com/support/docview.wss?uid=ssg1S4000876&myns=s028&mynp=familyind5385287&mync

4.2 SAN Volume Controller and Storwize V7000 Easy Tier

Easy Tier is a built-in, dynamic data relocation feature available on the SAN Volume Controller and the Storwize V7000. Easy Tier determines the appropriate tier of storage, based on data access patterns and then automatically and nondisruptively moves data, at the sub-LUN level, to the appropriate disk tier in a storage pool containing two tiers of disk.

Easy Tier automatic data placement will allow you to gain the benefits of using a higher performance disk tier at a much lower cost and without manual management. Easy Tier can be enabled for any storage managed by the SAN Volume Controller and Storwize V7000 (for example, internal, external, or a combination of internal and external disk storage). This means that a small amount of more expensive, high performance disk can be leveraged efficiently across a larger amount of less expensive, internal or external disk to optimize performance.

SAN Volume Controller and Storwize V7000 Easy Tier operates in the following two modes:

- Evaluation mode
- Auto data placement mode

Easy Tier is included in the base software for the SAN Volume Controller and Storwize V7000. For IBM Storwize V7000, the Easy Tier feature is available at no charge.

4.2.1 SVC and Storwize V7000 Easy Tier evaluation mode and STAT tool

Most workloads access certain areas of a volume or LUN much more heavily than others, and the SAN Volume Controller or Storwize V7000 Easy Tier provides the capability to collect performance information at the sub-LUN level. When Easy Tier has been activated for a storage pool, the SAN Volume Controller or Storwize V7000 continuously gathers extent-level I/O statistics and creates 24-hour moving averages.

Tip: Extent size for all volumes in a storage pool is determined by the administrator at the time of storage pool creation and can range from 16 MB to 8 GB.

These statistics can be used to identify the extents that are heavily accessed and can benefit most from relocation to a high performance disk tier (for example, an SSD), and those extents that are less active or no longer heavily used and might remain on, or be relocated to, a lower performance disk tier (for example, a Hard Disk Drive (HDD)).

If the storage pool has Easy Tier Auto Data Placement activated, the statistics will be used as the basis for automated extent relocation between disk tiers which is discussed in “SVC and Storwize V7000 Easy Tier auto data placement mode” on page 56. If Auto Data Placement has not been activated for a storage pool (or turned off for an individual volume), or if the storage pool does not contain multiple disk tiers, Easy Tier can be run in evaluation mode to collect the performance statistics. Easy Tier evaluation mode makes the extent-level performance statistics available even before SSDs have been purchased and installed.

The statistics can be offloaded and viewed using the IBM STAT tool to understand your workload characteristics and to estimate the potential performance benefits of using SSDs and Easy Tier automatic data placement. If Easy Tier auto data placement mode is already operational, the statistics can be reviewed periodically to view the results of inter-tier data relocation and to determine whether workload changes require additional high performance disks.

The STAT tool is a Windows application that analyzes statistics collected over a 24-hour cycle and creates set of reports that include these:

- *System summary report:* This shows the total amount of heavily accessed (“hot”) data in the system. It also provides a suggested amount of SSD capacity to be added to the system with an estimate of the resulting system performance improvement.
- *Storage pool recommendation report:* This indicates SSD managed disk(s) suggested to be added to each pool with estimated performance improvements.
- *Volume heat distribution report:* This identifies the amount of hot and cold capacity for each volume and the portion of volume capacity already on SSD (if any).

STAT: The STAT can be downloaded from the following website:

`http://www-01.ibm.com/support/docview.wss?uid=ssg1S4000935`

4.2.2 SVC and Storwize V7000 Easy Tier auto data placement mode

SAN Volume Controller Storwize V7000 Easy Tier Auto Data Placement mode currently operates for multi-tiered storage pools that contain both SSDs and HDDs. The SSDs might be internal disks installed in the SAN Volume Controller Storwize V7000, or external disks contained in another storage system virtualized by SAN Volume Controller Storwize V7000.

Easy Tier automatically and nondisruptively moves frequently accessed (*hot*) data from HDDs to SSDs to improve performance. Easy Tier moves only the hot data to SSDs, providing the biggest performance benefit for the lowest cost. Easy Tier can optimize storage performance for any storage array managed by the SAN Volume Controller Storwize V7000, and can take advantage of a small amount of SSD across a large amount of virtualized backend storage.

Easy Tier Auto Data Placement mode is enabled by default for a multi-tiered storage pool containing SSDs and HDDs and for all striped volumes in the pool. Easy Tier monitors the I/O activity of all volumes in the pool for a 24 hour period and then begins to migrate hot extents to SSDs to optimize performance. The movement of extents is nondisruptive and transparent to the host server and applications, and is automatically managed to prevent any performance impact due to the migration activity.

As shown in Figure 4-7, Easy Tier identifies the most frequently accessed extents (hot extents) and moves them up to the higher-performance SSD storage tier. Over time, any extents in the SSD storage tier that have become less frequently accessed (cold extents) are moved down to the lower-performance HDD storage tier.

If desired, SAN Volume Controller Storwize V7000 Easy Tier can be turned off for an individual volume. This might be done to prevent extents from being moved to a higher or lower tier. Extents for the volume will remain on the disk tier they occupy when Easy Tier is turned off for the volume. I/O activity will continue to be monitored for the volume after Easy Tier is turned off. (The volume’s Easy Tier status will change from Active to Measured.)

SAN Volume Controller Storwize V7000 Easy Tier can also be turned off for a pool (single-tier or multi-tier.) In this case, I/O activity statistics will not be collected for volumes in the pool.

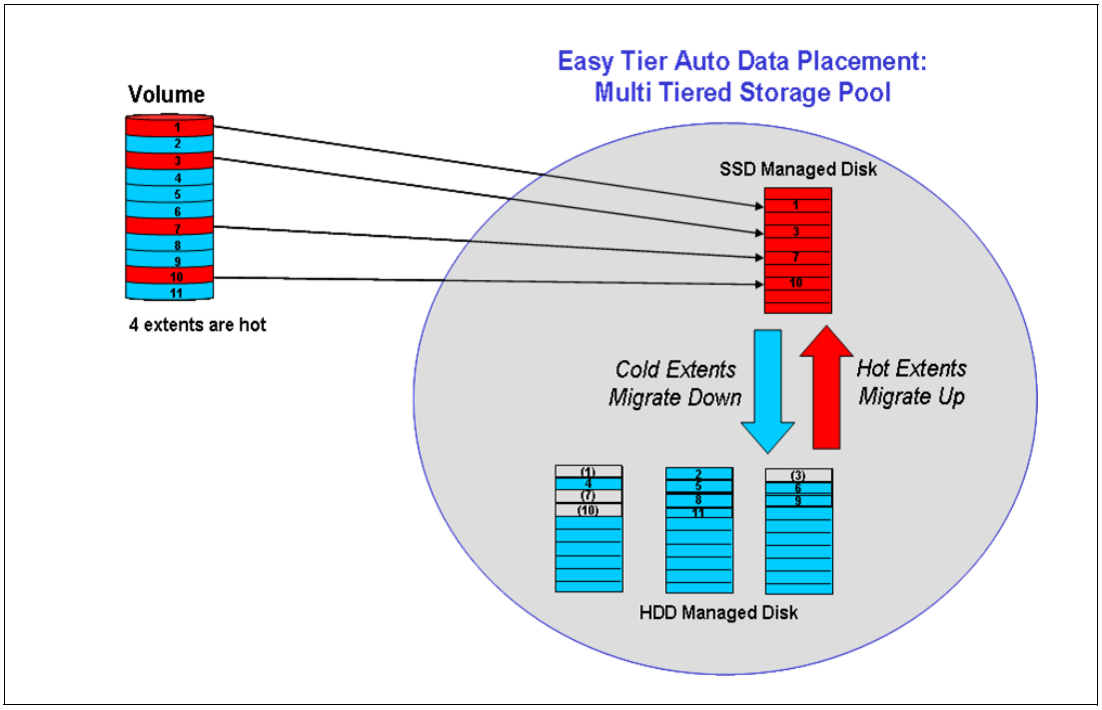

Figure 4-7 Easy Tier auto data placement

The SAN Volume Controller Storwize V7000 support the following volume types for Easy Tier automated data placement:

- Fully provisioned volumes
- Thin provisioned volumes
- Striped volumes
- Mirrored volumes (each part of the mirror is managed independently by Easy Tier)
- Thin Mirrored volumes (each part of the mirror is managed independently by Easy Tier)
- Copy Services volumes (source and target are managed independently by Easy Tier)

Image mode and sequential volumes are not supported for Easy Tier automated data placement. I/O monitoring and collection of performance statistics for image mode and sequential volumes is supported.

For SAN Volume Controller and Storwize V7000 internal SSDs, the minimum number of SSDs required is one per node, with a maximum of four per node.

Both SAN Volume Controller and Storwize V7000 Easy Tier Auto Data Placement support:

- A storage pool with internal SSDs and external Fibre Channel attached HDDs
- A storage pool with external Fibre Channel attached SSDs and external Fibre Channel attached HDDs

Storwize V7000 Easy Tier Auto Data Placement also supports:

- A storage pool with external Fibre Channel attached SSDs and internal HDDs
- A storage pool with internal SSDs and internal HDDs)

4.3 IBM SONAS

IBM SONAS implements a feature-rich Information Lifecycle Management (ILM) function, driven by customizable automation policies for file placement. After the tiered storage policy actions have been identified, the parallel data movement engine can perform physical movement to:

- Other storage pools
- Other sites
- External hierarchical storage management (HSM) storage

The file is always available within its namespace, even during physical data movement. No change to the logical file path is made in the global namespace. Each node participates in parallel for data movement. ILM policies include, but are not limited to:

- File access patterns (data last accessed, date last modified, and so on)
- File characteristics (file name, file size, file type)
- User ID, Group ID
- Any combination of these or other ILM policies

4.3.1 Introduction to SONAS ILM

The top half of Figure 4-8 shows a *logical* view of a global namespace (file and directory structure) as seen by clients. There are four files shown in colors that correspond to their *physical* location which is shown in the bottom half of the figure. There are four storage pools, three internal and one external, each consisting of the following pools:

- Gold storage pool: High performance 15 K RPM serial-attached SCSI (SAS) drives
- Silver storage pool: 1 T Serial Advanced Technology Attachment (SATA) drives
- Bronze storage pool: 2 TB SATA drives
- External storage pool: A tape or tape library

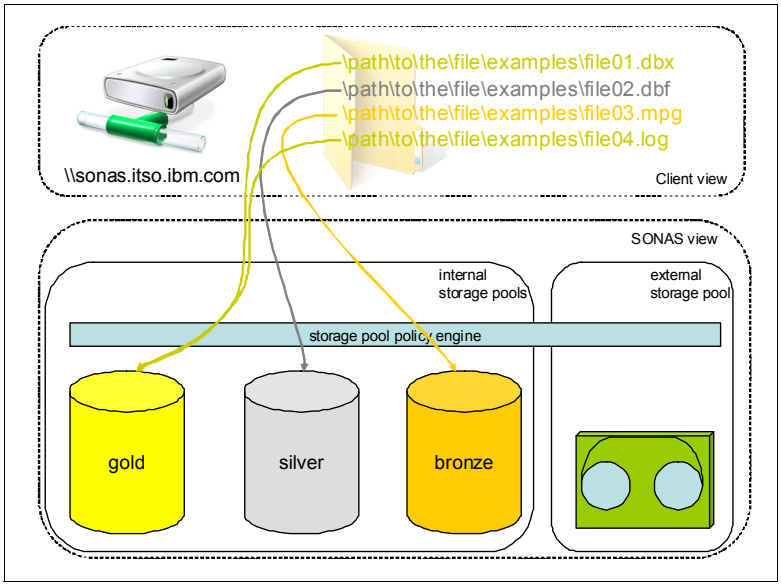

Figure 4-8 File placement and migration policy: logical and physical views

File placement policies are evaluated and applied at the time of file creation. If placement policies are not defined, all new files are placed in the *system* storage pool.

For example, assume that the client is going to create four files within one directory, and the files are to be accessible within this namespace, regardless of physical location. The initial policy creates .dbx and .log files in the gold storage pool, and all other files of a size less then 10 GB are placed in the silver storage pool. The client's file requirements are as follows:

- file01.dbx is a database index file that needs to sustain a high I/O rate
- file02.dbf is a database file with an average I/O rate
- file03 is a large file that is generally read sequentially
- file04.log is a log file that, when active, needs to sustain a relatively high I/O rate

This policy can be implemented as shown in Example 4-1, which defines a default rule and an additional three file placement rules. Any name can be assigned to a storage pool. For illustration purposes, in Example 4-1, the default storage pool is named *silver*; however, in a real-world scenario, this might typically be named, *system*.

Example 4-1 Initial placement policy

```
RULE 'default' set pool 'silver'
RULE 'dbxfiles' set pool 'gold' where upper(name) like '%DBX'
RULE 'logfiles' set pool 'gold' where upper(name) like '%LOG'
RULE 'bigfiles' set pool 'bronze' where KB_ALLOCATED>10000000
```

4.3.2 SONAS automated tiering

Files, such as log files, can become obsolete or not accessed for a period of time. Hence, they do not need to be placed in the highest performing storage pool. Example 4-2 shows the rules for migrating file04.log through the storage pools. Policy rules are Structured Query Language- (SQL)-like statements that specify the conditions for which the rule is to be applied.

Example 4-2 Migration policy

```
RULE 'cleangold'   migrate from pool 'gold'   to pool 'silver' where access_age > 30 days
RULE 'cleansilver' migrate from pool 'silver' to pool 'bronze' where access_age > 90 days
RULE 'cleanbronze' migrate from pool 'bronze' to pool 'tape'   where access_age > 365 days
```

Migration and deletion rules can be scheduled using the chronological scheduler. File migration between pools can also be controlled by specifying thresholds, such as the percent of storage pool used. A migration example is shown in Figure 4-9.

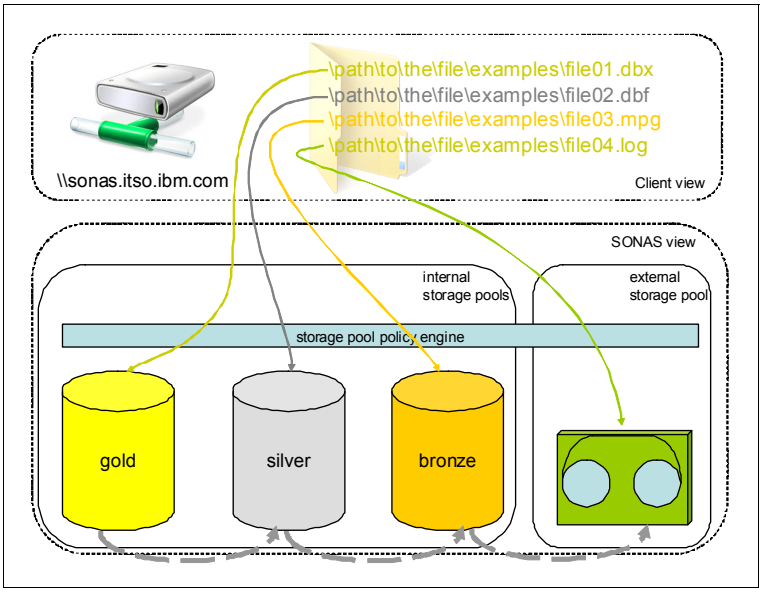

Figure 4-9 Contents of the file04.log file migrated to a tape

File file04.log, after not being updated and accessed for a period of time, is moved transparently between storage pools. Finally, if the file has not been accessed for 365 days, it is moved to tape, but is still visible within the namespace (directory) and can be accessed at anytime.

When using HSM space management on a filesystem, each file in the filesystem can be in one of three states:

- *Resident:* The file resides on disk in the SONAS appliance.
- *Premigrated:* The file resides both on the disk in the SONAS and in Tivoli Storage Manager HSM.
- *Migrated*: The file resides only in Tivoli Storage Manager.

As data is accessed by Common Internet File System (CIFS), Network File System (NFS), or other protocol, when a migrated file is opened and a byte of data that is not in the SONAS cache is accessed, that access triggers a Data Management Application Program Interface (DMAPI) event in the SONAS. Because a recall from physical tape requires waiting for cartridge fetching, tape drive loading, and tape movement to the desired file, physical tape recalls can take a significant number of seconds to start, so the application needs to plan for this delay.

For additional details, see the information available at these websites:

- http://www.redbooks.ibm.com/cgi-bin/searchsite.cgi?query=sonas+AND+architecture
- http://www.redbooks.ibm.com/cgi-bin/searchsite.cgi?query=hsm+AND+tsm

4.4 IBM Hierarchical Storage Management (HSM)

IBM Hierarchical Storage Management (HSM) software migrates rarely used files from costlier to less expensive storage, typically physical tapes with sequential data access. HSM tools can identify and automatically offload rarely used data from primary storage to less expensive storage, yet leaving the reference to the file in the filesystem for eventual transparent access by user. This small reference file left in the client system is also known as *stub file*.

Because the offloaded files are accessed rarely, or possibly never, after their creation or last modification, the delay caused by retrieving the file from tape back to the original location is not the constraint. Furthermore, it can be eliminated by the deployment of cheap random access storage solution, or virtual tape libraries (IBM System Storage TS7720 or TS7740).

IBM offers HSM on various platforms, incorporated in the following three products:

- *IBM Tivoli Storage Manager HSM for Windows:* This product controls primary disk storage needs by automatically migrating rarely used files on Microsoft Windows-based platforms to less expensive storage.
- *IBM Tivoli Storage Manager for Space Management:* This product moves inactive data to reclaim online disk space for important, active data on UNIX and Linux platforms, including the support of General Parallel File System (GPFS).
- *IBM Data Facility Storage Management System (DFSMS) for System z:* includes DFSMS for Hierarchical Storage Management (DFSMShsm). This is an automated hierarchical storage manager and productivity tool, designed to support availability management, space management, and disaster recovery.

The concept of migrating and managing inactive or rarely used data is slightly different among these three products. However, the conceptual model is shared by these products and is depicted in Figure 4-10.

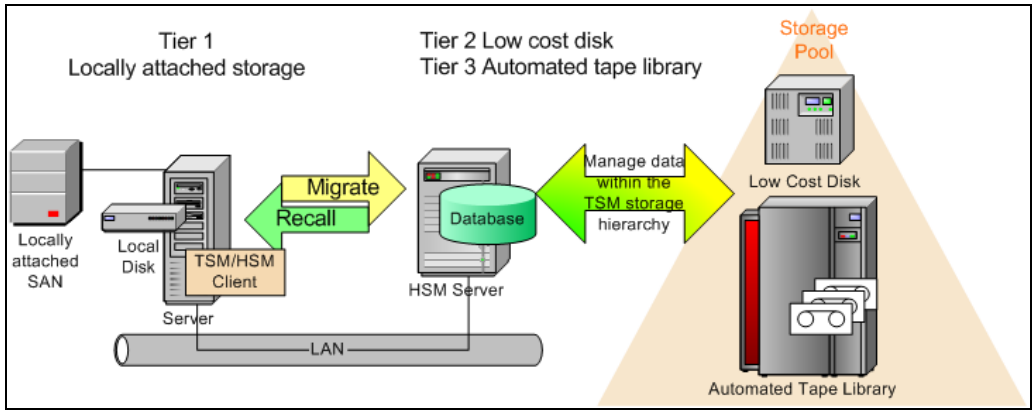

Figure 4-10 Conceptual model of hierarchical space management using various storage tiers

4.4.1 IBM Tivoli Storage Manager HSM for Windows

Tivoli Storage Manager HSM for Windows client provides HSM for Windows New Technology File Systems (NTFS). HSM automatically moves data between high-cost and low-cost storage media.

In effect, HSM turns the fast disk drives into caches for the slower, mass storage devices. The HSM for Windows client monitors the way files are used and lets users automate policies as to which files can safely be migrated to slower devices, such as tape and which files need to remain on hard disks.

The HSM for Windows client manages the migration of individual files, files from parts of NTFS filesystems, or complete NTFS filesystems, to remote storage in Tivoli Storage Manager. Migrated files can be accessed, opened, and updated by the Windows application corresponding to the file extension. The migration of files is transparent to Windows users and applications, with the following caveats:

- Files that have been migrated are replaced by a stub file that displays as follows:
  - In Windows Explorer, with the file name displaying with an overlay icon
  - In a Command Prompt window, with the file name enclosed in brackets
- Access to migrated files can be slower, if the file operation recalls the migrated file from Tivoli Storage Manager server storage, especially if tape is used as a file repository.

Migration overview

Migration is the core process performed with the HSM for Windows client. You can configure two kinds of migration and configure the behavior of migrated files. Migrating a file does not change the last access time stamp of the file.

You can configure *migration jobs* and *threshold migration*. Migration jobs allow you to specify precisely which files can be migrated, but they do not consider the space capacity of the volume. In contrast, threshold migration allows you to control space use of the volume, but allows less control of which particular files are migrated. Here are more details about this:

- A *migration job* defines a set of files and their migration behavior. When you run the job, the files specified in the job are copied to Tivoli Storage Manager storage. HSM for Windows client can delete the original files, replace the original files with stub files, or do nothing to the original files, depending upon your configuration.
- *Threshold migration* provides migration, based on space use. When the used space on a volume reaches a high threshold, migration begins automatically. Files are migrated to free up space until used space falls to a low threshold. The files that are migrated meet a minimum age and size, and are prioritized for migration so that less dynamic and larger files are migrated before more dynamic and smaller files. With proper configuration, threshold migration can automatically prevent the volume from running out of space.

For migration jobs and threshold migration, you configure whether files will be backed up before migrating. We strongly encourage that this feature be enabled so that at least one copy of the files is available in Tivoli Storage Manager in case a user accidentally deletes one or more stub files, thereby removing the original file from storage.

Figure 4-11 represents the results of successful HSM migration and recall test jobs with typical files from multiple applications. All of them demonstrated the similar migration performance parameters, such as commitment to the Tivoli Storage Manager server database, reconciliation with the backup environment, functionality of the stub file, and so on. There are a lot more types of files that are eligible for HSM migration. Consider the benefit of each file in terms of space savings and frequency of use to avoid periodical migration and reciprocal recall of the same files.

Ext.	Application description	Success?
afw	Industrial Aspect Integrator Platform	Y
avi	Audio Video Interleave Movie	Y
doc	Microsoft Word Document	Y
mp3	Compressed Audio File	Y
mpg	MPEG Video File	Y
pdf	Portable Document Format (Adobe Acrobat File)	Y
pps	Microsoft PowerPoint Slideshow	Y
ppt	Microsoft PowerPoint Presentation	Y
rtf	Microsoft Wordpad Document (Rich Text Format)	Y
Vob	MPEG Video File	Y
vsd	Microsoft Visio Document	Y
xls	Microsoft Excel Spreadsheet	Y
zip	Compressed ZIP Archive	Y
bmp	Bitmap Image File (BMP)	Y
m4a	mpeg-4 audio layer	Y
gho	Norton Ghost image file	Y
ghs	Norton Ghost image file	Y
gif	Graphics Interchange File Format File	Y
jpg	Bitmap Image File (JPEG)	Y
rar	Compresses RAR archive	Y
psd	Adobe Photoshop File	Y
tif	Bitmap Image File (TIFF)	Y
wmv	Windows media videos	Y
iso	Raw Images in ISO format	Y
img	Bitmap graphics or MagicISO images	Y
wav	Uncompressed audio format	Y
tiff	Tagged Image format	Y

Figure 4-11 HSM migration test results of typical file types from multiple applications

Features and advantages of HSM for Windows

In addition to the migration and recall of files and the reconciliation of filesystems, the HSM for Windows client provides additional functions beyond the scope of traditional HSM:

- An administrator can define migration jobs for each volume, including or excluding files, based on file type (extension) and various criteria related to the age of a file (creation, modification, and last access).
- An administrator can define recall quotas to limit the number of file recalls during a specified time period. Quotas can apply to the entire system, to user groups, or to specific users.
- The HSM for Windows client can also be used for archiving purposes. Files can be archived in Tivoli Storage Manager, and the original files are either kept on disk or deleted.
- Search and retrieve options are available to the administrator for migrated and archived files. Select files or complete file spaces can be retrieved, either to their original location in the filesystem or to another location in the filesystem.
- Threshold migration monitors filesystem space use and migrates files when space is needed.
- Threshold migration migrates older and larger files from the filesystem. You configure whether the file age or size is a better qualifier for migration.
- Migration jobs can be organized according to the logical structure of a volume (including different parts of the directory structure) and thus potentially reflect the structure of an organization or user groups.
- Migration jobs can be organized according to file type, such as office documents, images, and text files. This organization provides a more logical view of data than does traditional HSM.
- Individual jobs can be executed at various times.

- You can implement migration jobs and threshold migration on the same volume. You can build a policy that is based on both file values (migration jobs) and space use (threshold migration).

4.4.2 IBM Tivoli Storage Manager for space management

Tivoli Storage Manager for space management client for UNIX and Linux provides HSM for UNIX and Linux filesystems and storage pools. The principal operation is similar to HSM for Windows. The goal of the product is to automatically move data between high-cost and low-cost storage media.

Migration overview

The HSM client automatically migrates files to Tivoli Storage Manager storage when space is needed on the local filesystem. On the local system, stub files replace the migrated files. The system administrator can determine the size of these stubs on a per filesystem basis.

At any time, you can migrate particular files. For example, use selective migration if you run threshold migration and some files do not migrate, because their last access dates make them ineligible for migration. The files you specified then migrate immediately to Tivoli Storage Manager server storage.

A file is eligible for automatic or selective migration when it meets the following criteria:

- It is a regular file that you previously have not migrated; character special files, block special files, first-in-first-out (FIFO) special files (named pipe files), and directories are not migrated.
- It is a resident (or premigrated) file on a filesystem for which space management is active.
- It is not excluded from space management in your include-exclude options file.
- The file size is greater than both the stub file size plus one byte and the filesystem block size.
- It meets management class criteria.

You can migrate any file in a set of hard linked files that has not been excluded from space management and to which you assigned a management class that permits automatic or selective migration.

General Parallel File System (GPFS) policy-driven migration

Tivoli Storage Manager for Space Management supports General Parallel File System (GPFS) running on AIX and Linux platforms. It provides a disk-based ILM implementation with the storage pool concept. The Tivoli Storage Manager for Space Management client integrates tape into the GPFS ILM solution. With Tivoli Storage Manager version 6.1 and GPFS version 3.2, the HSM client can be used with GPFS policy-based storage management features to migrate to tape and external storage pools.

GPFS version 3.2 introduces a new feature called *external storage pools*. You can set up external storage pools and GPFS policies allowing the GPFS policy manager to coordinate file migrations from a native GPFS online pool to external pools on the Tivoli Storage Manager server. The GPFS policy manager invokes the migration through the HSM client Command Line Interface (CLI).

Example 4-3 creates a rule named `HsmData`. Files are migrated from the online storage pool named `StoragePool1` to the external storage pool named `hsmpool`. The files are chosen, based on last access time and for those larger than 1,024 bytes. The migration of files starts when the 90 percent (high) threshold is reached. Migration continues until the 80 percent (low) threshold is reached. The difference between the low and premigration thresholds defines the percent of data which is premigrated. (80 percent - 70 percent = 10 percent).

Example 4-3 GPFS threshold-based migration rule with HSM

```
RULE 'HsmData' MIGRATE FROM POOL 'StoragePool1'
THRESHOLD(90,80,70)
WEIGHT( CURRENT_TIMESTAMP - ACCESS_TIME )
TO POOL 'hsmpool'
WHERE FILE_SIZE > 1024
```

For more details about storage pools, policy settings, and rules, see *GPFS 3.2 Advanced Administration Guide*, SC23-5182 at this website:

http://www.ibm.com/support/docview.wss?uid=pub1sc23518201

4.4.3 IBM DFSMShsm for System z

DFSMShsm is a functional component of the DFSMS family. DFSMShsm is a direct access storage device (DASD) storage management and productivity tool for managing low-activity and inactive data on System z. It relieves you from manual storage management tasks and improves DASD use by automatically managing both space and data availability in a storage hierarchy.

DFSMShsm provides space, availability, and tape mount management functions in a storage device hierarchy for both system managed, and non-system managed storage environments. DFSMShsm allows you to automate your storage management tasks, improving productivity by effectively managing the storage devices.

Space management is DFSMShsm function that allows you to keep DASD space available for users so that their service-level objectives for the system can be met. The purpose of space management is to manage your DASD storage efficiently. To do this, space management automatically and periodically performs functions that move low-activity data sets from user-accessible volumes to DFSMShsm volumes and reduce the space occupied by data on both the user-accessible volumes and the DFSMShsm volumes.

The DFSMShsm space management functions are as follows:

- Automatic primary space management of DFSMShsm-managed volumes, which includes:
  - Deletion of temporary data sets
  - Deletion of expired data sets
  - Release of unused, over-allocated space
  - Migration to DFSMShsm-owned migration volumes
- Automatic secondary space management of DFSMShsm-owned volumes, which includes:
  - Migration level cleanup, including deletion of expired, migrated data sets and some migration control data set (MCDS) records
  - Moving migration copies from migration level 1 (ML1) to migration level 2 (ML2) volumes
- Automatic interval migration, initiated when a DFSMShsm-managed volume exceeds a specified threshold

- Automatic recall of user data sets back to DASD volumes
- Space management by command
- Space-saving functions, which include:
  - Small data set packing
  - Partitioned data set compression
  - Data compaction
  - Data set reblocking

DFSMShsm can be set up to run automatically, manually, or from user programs. It can manage data, based on policies defined in Systems Management Server (SMS) management classes and storage groups.

For additional details about z/OS storage and space management, see *z/OS V1R11.0 DFSMShsm Storage Administration* located at this website:

http://publib.boulder.ibm.com/infocenter/zos/v1r11/index.jsp?topic=/com.ibm.zos.r11.arcf000/smspvsm.htm

5

Thin provisioning

This chapter provides an overview of the concept of thin provisioning and how select storage products from IBM support and enable this technology.

Traditional storage provisioning pre-allocates and dedicates physical storage space for use by applications or hosts. However, often, not all of the space allocated to applications is needed, resulting in wasted space, known as *white space*.

Thin provisioning allows applications and servers to see logical volume sizes that are larger than the physical capacity actually dedicated to the volumes on the storage system. Physical capacity is allocated for the volumes only as needed. Thin provisioning offers the following advantages:

- Allows higher storage system utilization which, in turn, leads to a reduction in the amount of storage needed, lowering direct capital expenses (capex).
- Lowers operating expenses (opex), because storage occupies less data center space and requires less electricity and cooling.
- Postpones the need to invest in more storage. As storage prices continue to drop over time, it will cost less when additional capacity is required.
- Simplifies capacity planning, because you can manage a single pool of free storage. Multiple applications or users can allocate storage from the same free pool, thus avoiding situations in which several volumes are capacity constrained, yet others have capacity to spare.
- Makes your storage environment more agile, enabling you to respond to change easily.

5.1 Thin provisioning overview

Thin provisioning allows a server to see logical volume sizes that are larger than the physical capacity actually dedicated to the volumes on the storage system. From the server or application perspective, thinly provisioned volumes appear and function as fully provisioned volumes. However, physical disk drive capacity is allocated only as needed (on demand) for write activity to the volumes. Unallocated physical capacity is available for use, as needed, by all volumes in a storage pool or even across an entire storage system.

In Figure 5-1, two 1 TB volumes are presented to a server. On the left, the fully provisioned volume has its logical volume size matched by dedicated physical disk drive capacity for writing data. On the right, the thinly provisioned volume has a logical volume size that is larger than the physical disk drive capacity reserved for writing data. The logical volume size seen by the server is 1TB, but the physical capacity of the volume is only 200 GB, the amount of space allocated for that data.

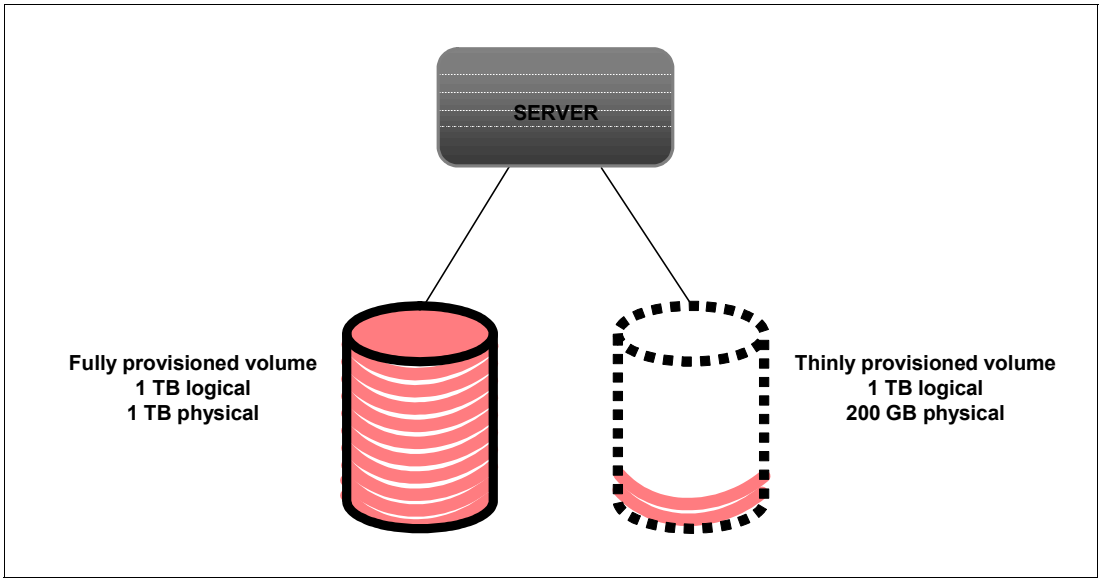

Figure 5-1 A fully provisioned volume (in which the logical volume matches the disk drive capacity) and a thinly provisioned volume (in which the logical volume is larger than the physical disk drive capacity)

5.1.1 Thin provisioning increases storage efficiency

The following aspects of thin provisioning contribute to storage efficiency:

- *Thin provisioning increases utilization ratios:* Thin provisioning increases storage efficiency by increasing storage utilization ratios. Real physical capacity is provided only as it is actually needed for writing data. This results in large potential savings in both storage acquisition and operational costs, including infrastructure costs, such as power, space, and cooling.

  Storage utilization is measured by comparing the amount of physical capacity actually used for data with the total amount of physical capacity allocated to a server. Historically, utilization ratios have been well under 50 percent, indicating a large amount of allocated but unused physical storage capacity. Often, neither users nor storage administrators are certain how much capacity is needed, but they need to ensure they will not run out of space, and also allow for growth. As a result, users request more space than they need and storage administrators might allocate more than is requested, resulting in significant over-allocation of storage capacity.

Thin provisioning increases storage utilization ratios by reducing the need to over-allocate physical storage capacity to prevent out of space conditions. Large logical or virtual volume sizes can be created and presented to applications without dedicating an equivalent amount of physical capacity. Physical capacity can be allocated on demand, as needed, for writing data. Unallocated physical capacity is available for multiple volumes in a storage pool or across the entire storage system.

- *Thin provisioning provides convenience to users:* Thin provisioning also increases storage efficiency by reducing the need to resize volumes or add volumes and restripe data as capacity requirements grow. Without thin provisioning, if an application requires capacity beyond what is provided by its current set of volumes, there are two options:
  - The existing volumes can be increased in size
  - Additional volumes can be provisioned

  In many environments, these options are undesirable because of the steps and the potential disruption required to make the larger or additional volumes visible and optimized for the application.

  With thin provisioning, large virtual or logical volumes can be created and presented to applications. Simultaneously, the associated physical capacity grows only as needed, completely transparent to the application.
- *Thin provisioning makes physical capacity use visible in the storage system:* Without thin provisioning, physical capacity is dedicated at the time of volume creation, and storage systems do not typically display or report how much of the dedicated physical capacity is actually used for data.

  Thin provisioning increases storage efficiency by making it easy to see the amount of physical capacity actually needed and used, because physical space is not allocated until it is needed for data.

5.1.2 Requirements for a successful thin provisioning

For thin provisioning to be successful, there are several requirements:

- A filesystem or application that works well with thin provisioning (that is *thin-friendly*)
- Procedures to monitor physical capacity allocation and to add or free up physical capacity as needed

Thin-friendly filesystems and applications

There are three primary operations in which a filesystem or application affects thin provisioning. Each of these operations can be performed in a way that uses physical storage capacity efficiently *(thin-friendly)* or inefficiently *(thin-unfriendly):*

- *Initialization:* A thin-friendly filesystem or application writes a minimal amount of metadata to a localized area of physical capacity during initialization. A thin-unfriendly filesystem or application writes metadata to many areas of its volumes during initialization, therefore the storage system allocates physical capacity for all of those areas just as though actual user data had been written.
- *Writing:* A filesystem or application that places multiple writes in the same localized area of a volume, is thin-friendly. A filesystem that scatters the writes across large areas of a volume, is thin-unfriendly.
- *Deletion:* A filesystem or application that notifies the storage system that space has been deleted, or uses previously deleted space for new writes, is thin-friendly. A filesystem that deletes data without notifying the storage system of the deletion and writes new data to other unused areas of the volume, is thin-unfriendly.

Figure 5-2 shows how a filesystem view of space use can be the same or different than the storage system view of the same space.

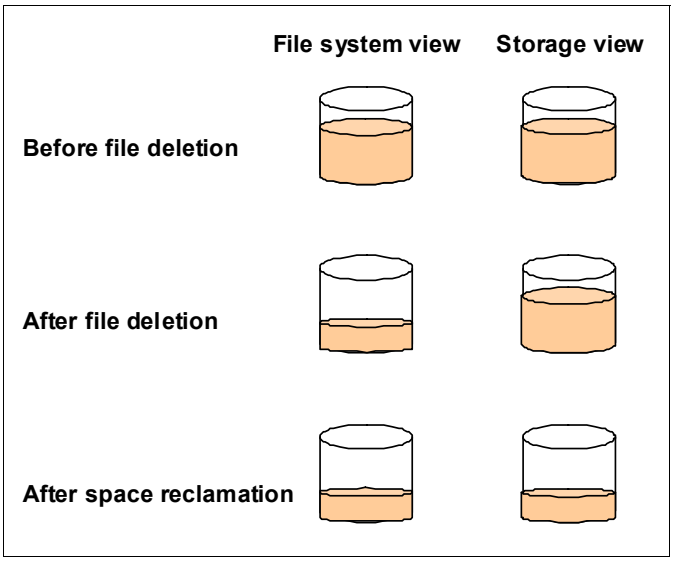

Figure 5-2 filesystem view of data deletion and space reclamation, compared with that of the storage system view

In Figure 5-2, the two views are the same in the beginning (before file deletion.) However, when the filesystem deletes a file, it typically just updates the metadata and does not write zeroes to the empty space. This means that the storage system does not realize that space has been freed. If the filesystem can communicate to the storage system that blocks have been freed, the storage system can reclaim the space.

Monitoring usage and adding physical capacity

Thin provisioning allows allocation of logical or virtual volume sizes that are larger than the physical capacity allocated to the volume. Thin provisioning also allows for an aggregate logical volume capacity that is greater than the physical disk drive capacity available in a storage pool, or even in the entire storage system. This means that it is important to monitor physical capacity allocation and to have procedures in place to provide additional disk drive capacity where it is needed to avoid space exhaustion for a volume, a storage pool, or the entire storage system.

Zero detection

Zero detection means that the storage system knows whether physical disk drive blocks contain data or are empty. If a host requests an empty block that contains zeroes, the storage system does not go to the disk to retrieve it. If a host writes a block of data that is only zeroes, the storage system marks the block empty without writing zeros to the disk. Zero detection improves storage utilization rates by not requiring physical disk drive capacity for empty blocks. Zero detection also enables reporting on physical capacity use for data, in addition to reporting on physical allocation. Zero detection can also enhance performance by reducing disk reads and writes.

5.2 DS8000 thin provisioning

The DS8000 provides the following thin provisioning capabilities:

- DS8000 Extent Space Efficient (ESE) volumes
- DS8000 Track Space Efficient (TSE) FlashCopy targets (Point-in-Time Image)

5.2.1 DS8000 Extent Space Efficient (ESE) volumes

The DS8000 enables creation of thinly provisioned Fixed Block (FB) volumes. DS8000 volumes created with the ESE Storage Allocation Method (SAM) are thinly provisioned and are referred to as ESE volumes. For an ESE volume, a virtual volume of the requested size is created without any physical allocation for data (except for maintaining related metadata). Physical space is allocated only as needed for data written to the ESE volume. If the ESE volume is presented to a host server, the host server will see a volume that is the full size of the virtual volume created, even if physical capacity has not yet been allocated.

DS8000 thin provisioning of volumes optimizes utilization of the physical capacity within a DS8000 extent pool. Any physical capacity not needed by one thinly provisioned volume is available for use by other volumes within the same DS8000 extent pool. Thin provisioning can also reduce physical capacity required in an extent pool, and this physical capacity can then be assigned to another extent pool. DS8000 thin provisioning might even reduce physical capacity required in the DS8000 as a whole, and any physical capacity not required to be purchased in the system results in both acquisition and operational cost savings.

Prerequisites for DS8000 ESE volumes

A DS8000 Licensed Internal Code level providing ESE support is required. A DS8000 thin provisioning license is also required.

Creation of thin provisioned (ESE) volumes

DS8000 thin provisioning is specified at the volume level. A DS8000 extent pool can contain both standard volumes and ESE volumes.

DS8000 volumes created with the ESE SAM are thinly provisioned. A virtual volume will be created at the full size requested and can be presented to a host server. Virtual Space will be allocated in the extent pool for the full virtual volume size.

Physical allocation to an ESE volume

There is minimal initial physical capacity allocation for metadata associated with DS8000 volumes created with the ESE SAM. Physical (real) space is allocated in 1GB increments (extents) as needed for data.

Monitoring and reporting capacity usage

Virtual and real capacity allocation can be displayed in the DS8000 Storage Manager GUI or the DS8000 Command Line Interface (DSCLI).

User-specified thresholds for real capacity use can be set individually for each DS8000 extent pool. When the threshold is reached for the pool, notifications, such as an SNMP trap, are generated. You can take the following actions to make more physical capacity available in a pool:

- Assign more disk drives (ranks) to the extent pool.

- Migrate several of the standard (fully provisioned) volumes to another extent pool using the Easy Tier Manual Mode (Easy Tier manual migration of ESE volumes is not currently supported).
- Delete several of the fully-provisioned FlashCopy target volumes from the extent pool.
- Delete several of the volumes from the extent pool.

If real space in a DS8000 extent pool is completely exhausted, read I/O will be allowed for existing volumes, but any write to an ESE volume that requires allocation of a physical extent will not be allowed.

Considerations for ESE volumes

Considerations for ESE volumes in the DS8000 are as follows:

- There are no performance implications for using DS8000 ESE volumes, compared with standard DS8000 volumes.
- A FlashCopy source volume must be a standard volume (fully allocated.) ESE volumes are not currently supported for use with FlashCopy.
- A FlashCopy target volume must be either a fully allocated standard volume or a thinly provisioned Track Space Efficient (TSE) volume. Thin provisioning of TSE volumes is discussed in “DS8000 Track Space Efficient (TSE) FlashCopy target volumes” on page 72.
- DS8000 does not currently support the use of ESE volumes as source or target volumes for remote mirroring. Thinly provisioned Track Space Efficient (TSE) FlashCopy target volumes are supported with remote mirroring as described in “DS8000 Track Space Efficient (TSE) FlashCopy target volumes” on page 72.
- DS8000 currently supports ESE thin provisioning for FB (open systems) volumes. It does not currently support ESE thin provisioning for Count Key Data (CKD) (mainframe z/OS) volumes.

5.2.2 DS8000 Track Space Efficient (TSE) FlashCopy target volumes

Originally, FlashCopy target volumes were required to be physically allocated at the same size as the source volumes, as shown in Figure 5-4. For a 10 GB source volume, a 10 GB FlashCopy target was defined. This resulted in the allocation of a large amount of physical capacity that was not actually needed.

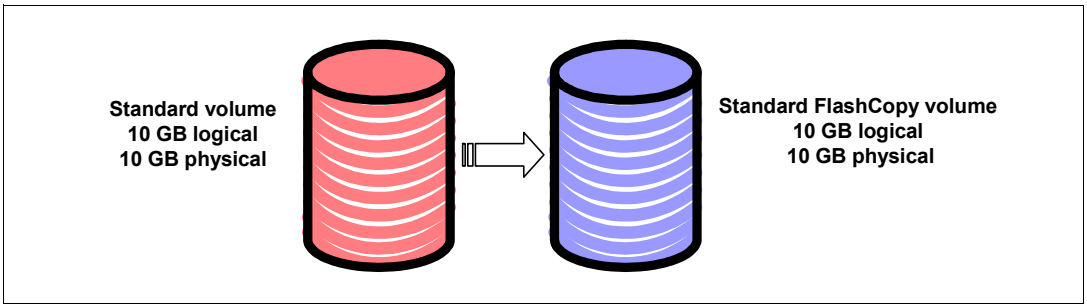

Figure 5-3 Originally, the source volume and FlashCopy target volume were required to be physically allocated at the same size

To alleviate the potential over-allocation associated with fully provisioned FlashCopy target volumes, the DS8000 allows the creation of thinly provisioned FlashCopy target volumes called Track Space Efficient (TSE) targets. TSE FlashCopy target volumes are supported for both open systems FB volumes and mainframe CKD volumes.

TSE FlashCopy target volumes are designed for use with the FlashCopy *nocopy* option. When the nocopy option is specified, the FlashCopy target contains only the differential data required to maintain the Point-in-Time Image.

Figure 5-4 shows a standard source volume on the left, and a TSE FlashCopy target volume on the right.

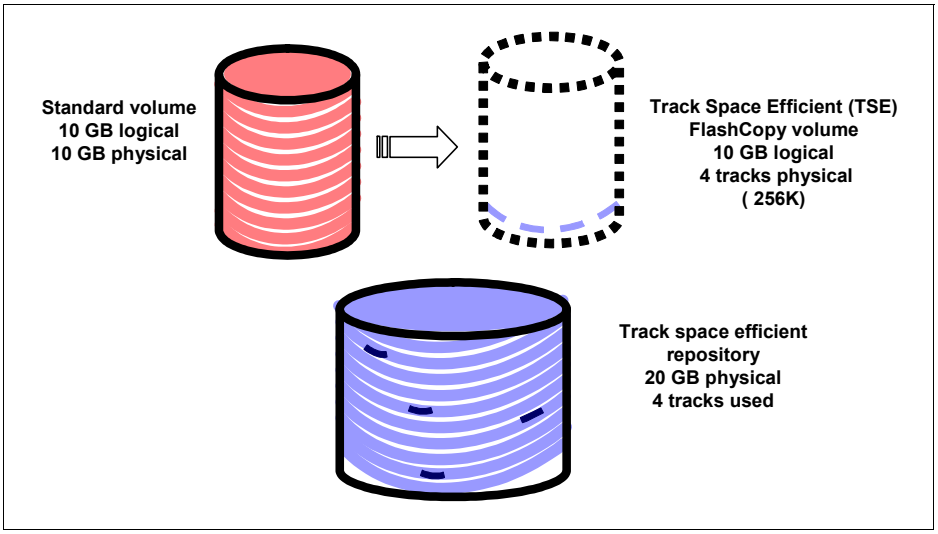

Figure 5-4 Standard volume and Track Space Efficient (TSE) FlashCopy

The standard source volume is 10 GB in size and also has a physical allocation of 10 GB. The TSE FlashCopy target volume is 10 GB in size, which can be seen by a host server, but has a physical allocation of only 256 K, because only 4 tracks have changed after the FlashCopy Point-in-Time Image was created.

Also in Figure 5-4, note that the physical allocation for the TSE FlashCopy target takes place in a fully allocated TSE repository. The TSE repository exists in the same extent pool as the TSE FlashCopy target and can be shared by multiple TSE volumes.

DS8000 TSE FlashCopy target volumes can greatly reduce the space required to store differential Point-in-Time Images, because the granularity of space allocation is a single track (64 K for open systems FB volumes and 56 K for mainframe CKD volumes.)

The space savings are greatest for FlashCopy target volumes that contain the least amount of data (for example, minimal difference between the FlashCopy image and the source volume.) Differences and space utilization will be minimized in any of the following situations:

- The lifespan of the FlashCopy target volume is short (for example, when used only for the duration for a backup to tape or when FlashCopies are done repeatedly on a short interval).
- The FlashCopy target volume will not be changed (for example, when used for a backup).
- The source volume has a low change rate.

Any TSE repository space that is not used by one FlashCopy TSE target is immediately available for use by other FlashCopy TSE target volumes in the same extent pool. Any physical capacity that did not need to be dedicated to the TSE repository can be used for other volumes in the extent pool.

Prerequisites for DS8000 TSE volumes

Prerequisites for DS8000 TSE volumes are as follows:

- A DS8000 Licensed Internal Code level that provides TSE support is required, as is a DS8000 FlashCopy Space Efficient License.
- A TSE repository must exist in the extent pool, as discussed in “Creation of Track Space Efficient FlashCopy target volumes” on page 74.
- The FlashCopy source volume must be a standard volume (fully allocated.) ESE volumes are not currently supported for use with FlashCopy.

Creation of Track Space Efficient FlashCopy target volumes

Before creating TSE volumes, a TSE repository must be created for the DS8000 extent pool in which the TSE volumes will be created. The TSE repository is a specified number of physical extents that will be used to hold any changed tracks associated with the TSE volumes.

DS8000 TSE volumes are created specifying a SAM of TSE. A virtual volume is created at the full size requested and can be presented to a host server. Virtual space will be allocated in the extent pool for the full virtual volume size. Physical space will be allocated in single track increments, as needed, to preserve the Point-in-Time Image.

A single DS8000 extent pool can contain both standard (fully provisioned) FlashCopy volumes and TSE (thinly provisioned) FlashCopy volumes. The extent pool can also contain standard volumes or ESE volumes. TSE FlashCopy volumes can be created in the same extent pool as standard volumes and ESE volumes.

There is no initial physical capacity allocation for DS8000 FlashCopy target volumes created with the TSE SAM.

Physical space available for use by TSE volumes is contained in a TSE repository defined for the extent pool. Individual tracks (64 K for FB TSE volumes, 56 K for CKD TSE volumes) are allocated from the repository for any TSE FlashCopy volume in the extent pool, as needed.

Monitoring and reporting usage

Virtual and real allocation can be displayed in the DS8000 DSGUI or the DSCLI.

A user-specified threshold for available repository capacity can be set for each DS8000 storage pool. When the threshold is reached for the pool, a notification, such as an SNMP trap, is generated. After it is created, the TSE repository cannot be expanded, but repository space can be freed in one of two ways:

- Terminating FlashCopy relationships and deleting FlashCopy volumes from the extent pool
- Re-establishing FlashCopies to release space for differential data

If real space in a TSE repository is equal to or greater than 99 percent full, any write to a TSE volume that requires the allocation of a real logical track is rejected.

5.3 XIV Storage System thin provisioning

The following thin provisioning capabilities are provided by the XIV system:

- Thin provisioning for volumes
- Thin provisioning for snapshots (differential Point-in-Time Images)

Thin provisioning is included with the XIV system. There are no additional Licensed Internal Code level or licensing requirements.

For more information, see the Redbooks publication, *IBM XIV Storage System: Architecture, Implementation, and Usage*, SG24-7659.

5.3.1 XIV Storage System thin provisioning: volumes

XIV thin provisioning is specified at the storage pool level. XIV system volumes inherit the provisioning characteristics of the storage pool in which they are created. All volumes created in thin provisioning storage pools will be thin provisioned volumes. All volumes created in regular (fully provisioned) storage pools will be fully provisioned volumes.

For a thin provisioned XIV system volume, a volume of the requested size is created without any physical allocation reserved. If the thin provisioned volume is presented to a host server, the host server will see the full size of the volume created, even if no physical capacity has been allocated. Physical space is allocated (transparently to the host server) only when data is to be written to the thin provisioned volume.

Figure 5-5 shows an XIV system regular storage pool and an XIV system thin provisioning storage pool.

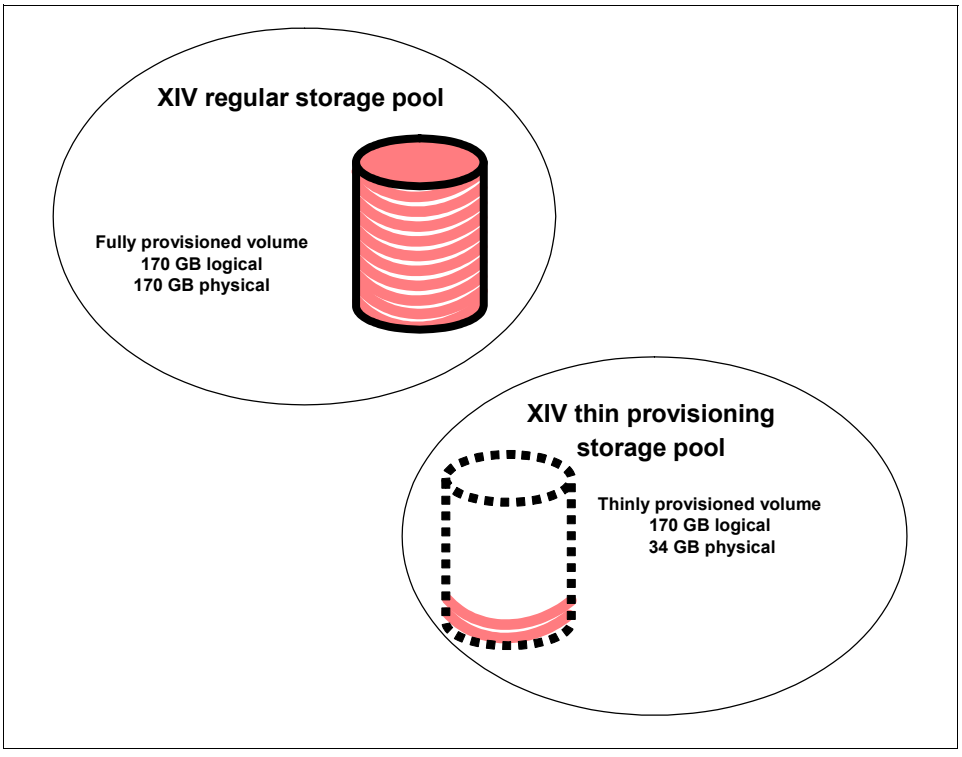

Figure 5-5 XIV system regular storage pool and XIV system thin provisioning storage pool

In Figure 5-5, the volume in the regular storage pool on the upper left has a logical size of 170 GB, and a matching amount of physical capacity is reserved for the volume (170 GB.) The volume in the thin provisioning pool on the lower right has a size of 170 GB, but only 34 GB has been allocated to write the necessary data.

XIV system thin provisioning and storage efficiency

Thin provisioning of volumes efficiently utilizes physical capacity within an XIV storage pool and within the entire XIV system. Physical capacity not needed by any thinly provisioned volume can be used by other thinly provisioned volumes within the same storage pool. Physical capacity that is not needed in an XIV thin provisioning pool can be moved dynamically and transparently to another XIV storage pool.

XIV thin provisioning allows physical capacity to be used only when needed and only where needed, thus reducing capacity requirements across the entire system.

Thin provisioning of volumes is included with the XIV system. There are no additional Licensed Internal Code level or licensing requirements.

Creation of thin provisioned volumes

XIV thinly provisioned volumes must be created in an XIV thin provisioning storage pool. XIV thin provisioning is specified at the storage pool level and inherited by all volumes in the pool. All of the volumes in a thin provisioning pool will be thinly provisioned.

For an XIV system thinly provisioned storage pool, two pool capacities are specified at the time of creation:

- Soft (upper limit of the sum of all volume sizes for the storage pool)
- Hard (actual physical capacity reserved for data for volumes in the storage pool)

For an XIV system thin provisioning pool, the soft capacity is greater than the hard capacity. For an XIV system regular pool (fully provisioned), the soft capacity is equal to the hard capacity.

After an XIV system volume is created in a thinly provisioned storage pool, there is no initial physical capacity allocation. The full logical volume size can be presented to a host server, but physical capacity will only be allocated as needed for writes to the volume.

Physical allocation

There is no initial physical capacity allocation for a thinly provisioned XIV system volume. Physical (hard) capacity is allocated in 1 MB or 17 GB increments, as needed, for data.

The first write to a volume in the pool results in an initial physical allocation of 17 GB. Additional writes to volumes in the pool can result in physical allocations in increments of 1 MB or 17 GB, as determined by the XIV system data distribution algorithm. Multiple thin provisioned volumes might share a 17 GB physical allocation.

Zero detection

The XIV system implements zero detection for 1 MB chunks of data. If a host writes 1 MB of zeroes, The XIV system does not write them to disk, so no physical allocation is triggered. Similarly, a host request for a 1 MB chunk of data that contains only zeroes does not require retrieval from disk. Zero detection not only saves space but improves performance by reducing disk reads and writes, both for host applications and XIV system copy functions. Zero detection also allows the XIV system to report actual use for data for each volume (in units of 1 MB).

The XIV system data migration facility uses zero detection to convert a thick (fully allocated) volume on a traditional storage system to a thin provisioned volume on the XIV system.

The XIV system also executes an ongoing scrubbing process which reclaims 1 MB chunks that contain zeroes as free space.

Monitoring and reporting capacity usage

Soft (logical) and hard (physical) allocations and physical use of data can be displayed or reported on, using either the XIV system GUI or the XIV Command Line Interface (XCLI).

A user-specified threshold for hard capacity use by volumes can be set for all the XIV system storage pools. When this threshold is reached for a pool, notifications, such as an SNMP trap, email, or SMS message, are generated. Actions that you can take to make more physical capacity available in a storage pool include these:

- Adding physical (hard) capacity to the storage pool (from free capacity in the system or free capacity in another storage pool)
- Deleting several snapshots from the pool
- Moving several volumes to another storage pool (the volume snapshots will also be moved)
- Deleting several volumes from the storage pool (the volume snapshots will also be deleted)

If physical capacity in an the XIV system thin provisioning storage pool is completely exhausted:

- The XIV system automatic snapshot deletion function is triggered, deleting the oldest, lowest priority snapshots first, according to user snapshot priority specification
- Access to volumes in the storage pool will be Read Only (default) or No Access, according to the user specification for the pool

Other considerations

Considerations for performance, snapshots, and remote copy in relation to the XIV system thin provisioned volumes are:

- There are no performance considerations for using the XIV system thin provisioned volumes, compared with those of regular the XIV system volumes.
- Thin provisioning of the XIV system snapshots is implemented differently than is thin provisioning of the XIV system volumes. Thin provisioning of the XIV system snapshots is discussed in “XIV thin provisioning: snapshot (Point-in-Time image)” on page 78.
- The XIV system supports thinly provisioned volumes as source and target volumes for synchronous and asynchronous remote mirroring.
- An XIV system volume can be changed from thick (fully provisioned) to thin as follows:
  - Change the storage pool from regular (thick) to thin provisioning
  - Move the volume from a regular storage pool to a thin provisioning storage pool

  These changes can be made dynamically and without impact to host I/O.

Space reclamation

The XIV system supports the Symantec Space Reclamation application program interface (API). This is used in conjunction with the Veritas VxFS filesystem in order for the filesystem to communicate to the XIV system that space has been deleted.

The XIV system also executes an ongoing scrubbing process that will reclaim 1 MB chunks that contain zeroes as free space.

5.3.2 XIV thin provisioning: snapshot (Point-in-Time image)

All XIV system snapshots are differentials. That is, the snapshot contains only the differential data required to maintain the Point-in-Time Image. All XIV system snapshots are also thinly provisioned.

XIV system snapshots do not need to be created in a thin provisioning pool. XIV system snapshots are thin provisioned whether they are in a regular (fully provisioned) pool or a thin provisioning pool.

XIV system thin provisioning of snapshots and storage efficiency

XIV system thinly provisioned snapshots can greatly reduce the space required to store Point-in-Time images.

The space savings are maximized when snapshots store minimal data (in other words, the saved minimal data differs from that on the source volumes). In the following situations. the snapshot space savings are optimized:

- The lifespan of snapshot Point-in-Time Image is short (for example, when used only for the duration for backup to tape, or when snapshots are made repeatedly on a short interval).
- The snapshot will not be changed (for example, when used for backup).
- The master volume (snapshot source) has a low change rate.

Prerequisites for XIV system thin provisioned snapshots

All XIV system snapshots are thinly provisioned. Thin provisioning of snapshots is included with the XIV system. There are no additional Licensed Internal Code level or licensing requirements.

XIV system snapshots do not need to be created in a thin provisioning pool. XIV system snapshots are thin provisioned, whether they are in a regular (fully provisioned) pool or a thin provisioning pool.

Creation of thin provisioned snapshots

A snapshot area must specified for a storage pool (fully provisioned or thin provisioned) before any snapshots can be created. This is the capacity that will be shared by all snapshots in the storage pool.

After an XIV system snapshot is created (in either a thinly provisioned storage pool or a regular storage pool), there is no initial physical capacity allocation. The full logical size can be presented to a host server, but physical capacity will only be allocated as needed to preserve the Point-in-Time Image.

Physical allocation to snapshots

There is no initial physical capacity allocation for an XIV system snapshot. Physical (real) space is allocated in 1 MB or 17 GB increments as needed to preserve the Point-in-Time Image.

The first write within a storage pool to a master volume that has a snapshot (or the first write to a snapshot) results in the allocation of 17 GB of physical capacity in the shared snapshot area for the pool. Any additional writes to the same or other master volumes or snapshots can result in physical allocations in increments of either 1 MB or 17 GB, as determined by the XIV system data distribution algorithm. Multiple snapshots might share a 17 GB physical allocation.

Monitoring and reporting capacity usage

Aggregate hard (physical) capacity allocations for all snapshots within a pool can be displayed using the XIV system GUI or the XCLI.

A user-specified threshold for hard capacity use by snapshots can be set for all XIV system storage pools. When the threshold is reached for a pool, notifications, such as an SNMP trap, email, or SMS message, are generated.

You can take the following actions to make more physical capacity available for snapshots in the storage pool:

- Add hard capacity to the storage pool.
- Delete several snapshots from the pool.
- Move several volumes to another storage pool (volume snapshots will also be moved).
- Delete several volumes from the storage pool (volume snapshots will also be deleted).

If the physical capacity available for the snapshot area in an XIV system storage pool is completely exhausted, the XIV system automatic snapshot deletion function is triggered, deleting the oldest, lowest priority snapshots first.

Actions that you can take to make more physical capacity available for the snapshot area in a storage pool include these:

- Increasing the size of the snapshot area (from free capacity in the system or free capacity in another storage pool)
- Deleting snapshots (Point-in-Time Images) from the pool
- Deleting several volumes from the storage pool (volume snapshots will also be deleted)

Other considerations

There is minimal performance difference between accessing the following items:

- A volume without any snapshots
- A volume with one or more snapshots
- A snapshot

XIV system snapshots are supported for use with XIV system remote mirroring. XIV system snapshots can be made for synchronous remote mirroring of source (master) or target (slave) volumes.

XIV system snapshots are created automatically as a fundamental element of XIV system asynchronous replication. Additionally, snapshots can be taken for replication:

- Asynchronous replication of the source (master) volume
- Asynchronous replication of the last replicated snapshot (this creates a duplicate snapshot).

For more information, see the Redbooks publication, *IBM XIV Storage System: Copy Services and Migration*, SG24-7759.

5.4 SVC and Storwize V7000 thin provisioning

SAN Volume Controller and Storwize V7000 are based on the same software and provide the same software capabilities

They provide the following types of thin provisioned volumes (formerly referred to as virtual disks or VDisks):

- Thin provisioned volumes
- Thin provisioned FlashCopy volumes
- Thin provisioned mirror volumes
- Thin provisioned remote mirror volumes

Thin provisioned volumes and thin provisioned FlashCopy volumes are discussed in this section.

5.4.1 SVC and Storwize V7000 thin provisioned volumes

SAN Volume Controller and Storwize V7000 allow the creation of thinly provisioned volumes, as shown in Figure 5-6.

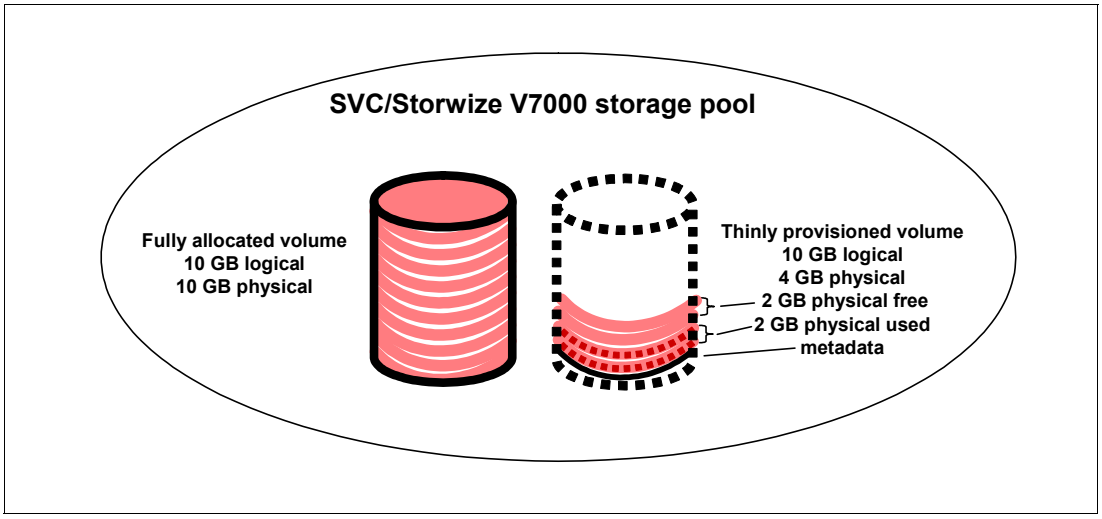

Figure 5-6 A fully allocated volume and a thinly provisioned volume

Thin provisioning of volumes optimizes utilization of the physical capacity within a SAN Volume Controller or Storwize V7000 storage pool. Any physical capacity not utilized by a thinly provisioned volume is immediately available for use by any other thinly provisioned volume within the same storage pool.

Prerequisites for SVC and Storwize V7000 thin provisioned volumes

The SAN Volume Controller Licensed Internal Code level that provides thin provisioned volume support is required. This might be different for other hardware models. No additional license is required for SAN Volume Controller thin provisioned volumes.

The hardware and Licensed Internal Code level providing zero detection (or zero detect) are suggested for optimal thin provisioning efficiency. See the discussion, “Zero detection” on page 70.

All levels of Storwize V7000 Licensed Internal Code provide thin provisioned volume support. No additional license is required for Storwize V7000 thin provisioned volumes.

Creation of thin provisioned volumes

SAN Volume Controller and Storwize V7000 thin provisioning is specified at the volume level. A storage pool (formerly known as a managed disk group) can contain both thin provisioned volumes and fully allocated volumes. When creating a thin-provisioned volume, the user defines the following parameters:

- Real physical capacity allocated to the volume from the storage pool. This is the initial physical capacity reserved for the volume which can be used to store both user data and metadata. It can be specified as an absolute value or a percentage of volume virtual capacity.
- Virtual capacity allocated to the volume from the storage pool. This is the volume size visible to a host server or internal copy functions.
- Grain size to be used to store write data for the volume. The grain size cannot be changed after the thin provisioned volume has been created. The current options are 32 KB, 64 KB, 128 KB, and 256 KB (32 K is the default.) Generally, smaller grain sizes save space but require more metadata access, which can impact performance. If you are going to use the thin-provisioned volume as a FlashCopy source or target volume, specify the same grain size for the volumes and for the FlashCopy mapping function. The grain size also affects the maximum virtual capacity for the thin-provisioned volume.
- Autoexpand (optional). As a thin-provisioned volume uses a portion of its real capacity to write data, the *autoexpand* feature adds free capacity from the pool to maintain a fixed amount of unused real capacity for the volume. The amount of free capacity maintained by *autoexpand* is referred to as the contingency capacity. It will be used when there is no more free capacity in the pool to allocate to the volume.
- The warning level that will trigger notifications when volume usage has reached a certain amount or percent.

  If *autoexpand* is specified, the default is 80 percent of the volume *virtual capacity* (volume size). If *autoexpand* is *not* specified, the default is 80 percent of the volume *real capacity*.

After a thin provisioned volume is created, the following actions occur:

- The specified amount of real capacity is allocated for the volume.
- A small amount of the real capacity (less than 1 percent, independent of volume size) is used to write initial metadata for the volume.

Physical allocation

The initial physical capacity allocation for a thinly provisioned SAN Volume Controller or Storwize V7000 volume is specified by the user as real capacity when the thin provisioned volume is created.

If *autoexpand* was *not* specified for the volume, as data is written to the volume will *use* the real capacity, so the amount of free real capacity will shrink as data is written. If more real capacity is needed, physical capacity can be added manually by an administrator.

If *autoexpand* was specified for the volume, the initial physical capacity is called the contingency capacity. As data is written to the real capacity, additional physical capacity is added to the volume from free capacity within the pool. The amount that is added is what is needed to maintain the contingency capacity (for example, the amount that was *used* to write data.)

For example, if a 100 GB, thin provisioned volume is created with a real capacity of 20 GB and *autoexpand* is specified, the *autoexpand* function will keep adding capacity to maintain 20 GB of free real capacity for the volume. If 5 GB is written to the volume, 5 GB real capacity will be added to the volume in order to keep 20 GB of free capacity. This means that the 100 GB volume now has 25 GB of real capacity allocated. Therefore, 5 GB of the real capacity is used and 20 GB of the real capacity is free.

A user can also manually increase the amount of physical capacity for a thin provisioned volume that has *autoexpand* specified. In our example, if a user manually adds 3GB real capacity to the volume, the volume will have 28 GB real capacity with 5GB used and 23 GB free. The contingency capacity for *autoexpand* becomes the new amount of free real capacity (23 GB), and as real space is used for writes, *autoexpand* will attempt to maintain 23 GB free real capacity.

A thin provisioned volume can also have its real capacity manually reduced by the user. Excess free real capacity can be taken from a volume and either left as free capacity in the pool or given to another volume in the pool.

Increasing or decreasing real capacity of a volume (either manually through the *autoexpand* function) is completely dynamic and transparent to an application accessing the volume.

After a thin provisioned volume is created, a small amount of the real capacity is used for initial metadata. Any host writes to the thin-provisioned volume will cause physical allocation in increments of the grain size (32 KB-256 KB) to store both user data and metadata. Write I/Os to grains that have previously been written will update the same grains and will not result in additional physical allocation.

The metadata storage overhead will never be greater than 0.1 percent of the user data. The overhead is independent of the virtual capacity of the volume.

Zero detection

SAN Volume Controller and Storwize V7000 implements zero detection. If a host writes a grain full of zeroes, SAN Volume Controller and Storwize V7000 will not write them to disk, so no additional physical allocation will be triggered. Likewise, a host request for a grain that contains only zeroes will not require retrieval from disk. Zero detection not only saves space but improves performance by reducing reads and writes. Zero detection also allows the SAN Volume Controller and Storwize V7000 to report actual use for data for each volume (in a granularity of .01 MB by default.)

Zero detection also enables conversion of a fully allocated image mode volume from an external storage system to a thin volume on a SAN Volume Controller or Storwize V7000 using the migration or volume mirroring function.

Monitoring and reporting usage capacity

A user-specified threshold for hard capacity use can be set for each volume.

When the volume usage threshold is reached, a warning event and notifications are generated. Actions that you can take to make more physical capacity available to the volume include these:

- Autoexpansion by the amount of the contingency capacity (as long as there is sufficient free physical capacity in the storage pool)
- Manual addition of physical capacity to the thin provisioned volume by an administrator

A user-specified threshold for hard capacity use can also be set for each SAN Volume Controller or Storwize storage pool.

When the pool usage threshold is reached, a warning event and notifications are generated. Actions that you can take to make more physical capacity available in the storage pool include these:

- Adding physical capacity to the storage pool
- Migrating several volumes to another pool using volume mirroring, then deleting the volumes from the pool
- Reducing real capacity of thinly provisioned volumes in the pool

If the physical capacity of a volume is fully utilized and *autoexpand* has not been specified, the volume will go offline when additional physical capacity is required for data. (For information about the *autoexpand* feature, see “Creation of thin provisioned volumes” on page 76.)

Procedures such as these are available for adding real capacity to the volume, after which the volume can be brought back online:

- Real capacity can be manually increased for the thin provisioned volume from free real capacity in the storage pool.
- Excess real capacity can be taken from other volumes in the pool, then added to the volume that requires additional real capacity.
- Fully allocated volumes can be migrated to another storage pool. The real space associated with the fully allocated volumes is released when migration is complete.
- Volume mirroring can be used to migrate fully allocated volumes in the pool to thin provisioned volumes, after which the fully allocated volumes can be deleted from the pool. The real space associated with the fully allocated volumes is released after the volumes are deleted.

If physical capacity of a volume is fully utilized and *autoexpand* is specified but the free physical capacity in a SAN Volume Controller or Storwize V7000 storage pool is completely exhausted, the volume will go offline as soon as it requires more capacity.

Procedures such as these can be used to increase real capacity in the storage pool. Real capacity can then be added to the volume and the volume brought online:

- Real capacity (unused) can be decreased for other volumes in the pool
- Additional capacity can be added to the storage pool
- Fully allocated volumes can be migrated to another storage pool
- Volume mirroring can be used to migrate fully allocated volumes in the pool to thin provisioned volumes in the pool, and then the fully allocated volumes can be deleted from the pool

Other considerations

Additional considerations in relation to thin provisioned volumes are as follows:

- *Performance:* There is a small amount of performance degradation when using a SAN Volume Controller or Storwize V7000 thin provisioned volume, compared with a fully allocated volume.

  Setting the cache mode to read-write enhances performance. If the cache mode is set to none, the SAN Volume Controller cluster cannot cache the thin-provisioned metadata, which decreases performance.

  Using a larger grain size enhances performance, as discussed in “Creation of thin provisioned volumes” on page 76.

- *Point-in-Time Image (FlashCopy):* Thin provisioning of SAN Volume Controller and Storwize V7000 FlashCopy volumes is implemented in the same way as thin provisioning of other volumes. However, there are several considerations when using thin provisioning for a FlashCopy volume.

  If the FlashCopy source volume is a *fully-allocated* volume, a thinly provisioned volume can be used for the FlashCopy target if a copy rate of zero is specified (the *nocopy* option).

  If a copy rate greater than zero is specified, the background copy will cause the FlashCopy target volume to be considered fully utilized, so there will be no benefit to using a thin provisioned FlashCopy target.

  If the FlashCopy source volume is a *thin provisioned* volume, a thin provisioned FlashCopy target can be used effectively with either a copy rate of zero (the *nocopy* option) or a copy rate greater than zero (the *background copy* option). The currently supported options for FlashCopy grains are:

  – 64 KB
  – 256 KB (default)

  Use the same grain size for the thin provisioned volumes as that of the FlashCopy source and target volumes and for the FlashCopy mapping.

- *Remote Mirroring:* If the remote mirroring source volume is a *fully-allocated* volume, a thinly provisioned volume can be used effectively for the remote mirroring target if *no full copy* is specified. If *with full copy* is specified, the target volume is considered fully utilized after initialization, and there will be no benefit to using a thin provisioned remote mirroring target.

  If the remote mirroring source volume is a *thin provisioned* volume, a thin provisioned remote mirroring target can be used effectively, with either *no full copy* or *with full copy*.

- *Volume Mirror:* If a volume is created with two copies for volume mirroring, both copies must be identical in terms of provisioning, that is: *Both* copies must be either thin provisioned or thick provisioned, as specified at creation.

  If a copy is added to a volume after creation for volume mirroring, the second copy might have the same or another provisioning policy than the original volume. The original volume might be fully allocated and the second copy thin provisioned. Or, the original volume might be thin provisioned and the second copy fully allocated. Or, both copies might be fully allocated or both copies thin provisioned. This means that volume mirroring can be used to convert volumes from thick to thin or thin to thick provisioning by adding a copy after original volume creation. These changes can be made dynamically and without impact to host I/O.

- *Changing from thick to thin or thin to thick provisioning:* A thin provisioned volume can be converted to a fully allocated volume using the volume mirroring function. A fully allocated volume can also be converted to a thin provisioned volume using the volume mirroring function. These changes can be made dynamically and without impact to host I/O.

5.4.2 SVC and Storwize V7000 thin provisioned FlashCopy

Thin provisioning of SAN Volume Controller or Storwize V7000 FlashCopy volumes is implemented in the same way as is thin provisioning of other volumes. However, there are several considerations when using thin provisioning for a FlashCopy volume.

Considerations

If the FlashCopy source volume is a *fully allocated* volume, a thinly provisioned volume can be used for the FlashCopy target if a copy rate of zero is specified (the *nocopy* option.) If a copy rate greater than zero is specified, the background copy will cause the FlashCopy target volume to be considered fully utilized, so there will be no benefit to using a thin provisioned FlashCopy target.

If the Flash Copy source volume is a *thin provisioned* volume, a thin provisioned FlashCopy target can be used effectively with either a copy rate of zero (the *nocopy* option) or a copy rate greater than zero (the *background copy* option).

The currently supported options for FlashCopy grains are as follows:

- 64 KB
- 256 KB (default)

Use the same grain size for the thin provisioned volumes as that of the FlashCopy source and target volumes and for the FlashCopy mapping.

Code level and license considerations related to thin provisioning of FlashCopy are discussed in “Prerequisites” on page 85.

Prerequisites

The SAN Volume Controller Licensed Internal Code level providing both thin provisioning and *zero detect* is required for effective thin provisioning of a FlashCopy target volume. The *zero detect* function might not be available on all hardware models. A SAN Volume Controller FlashCopy license is required. No additional license is required for SAN Volume Controller thin provisioned volumes.

All levels of Storwize V7000 Licensed Internal Code provide thin provisioned volume support, *zero detect* support, and FlashCopy support. No additional license is required for Storwize V7000 thin provisioned volumes, FlashCopy or *zero detect*.

5.5 IBM System Storage N series and thin provisioning

Integral to the N series Write Anywhere File Layout (WAFL®) Data ONTAP operating system is the ability to provide thin provisioning.

Starting with Data ONTAP 7G, N series provides thin provisioning at both the Logical Unit Number (LUN) and volume levels:

- *LUN:* This is a logical entity created within a volume (also known as a Flexible Volume or FlexVol). LUNs are offered to applications or hosts as virtual disks
- *Volume:* This is a logical entity created within an aggregate
- *Aggregate:* This is an abstraction layer between the physical disks and the FlexVol. It is an aggregation of physical disks within an N series system

Figure 5-7 illustrates these logical constructs graphically.

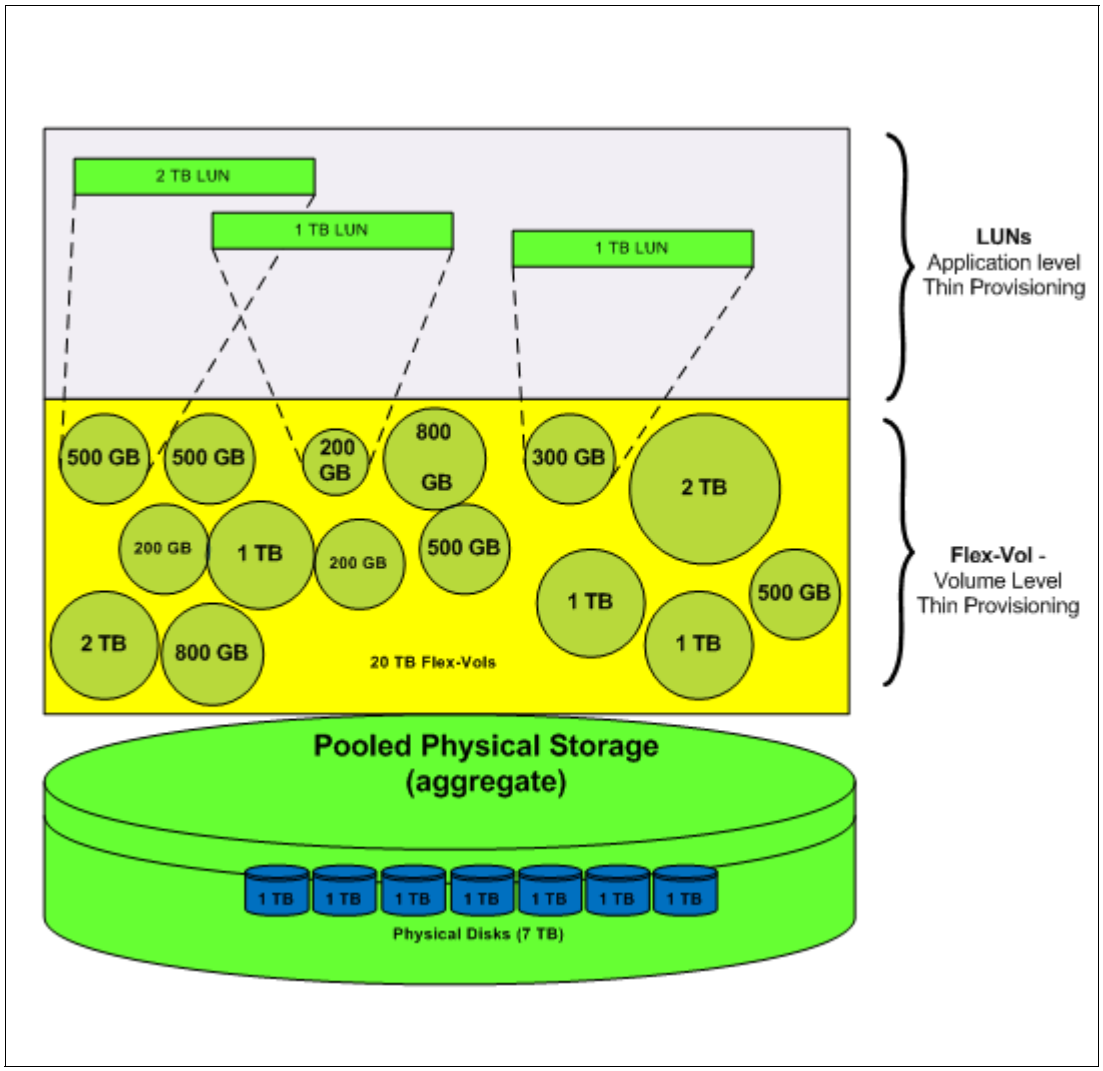

Figure 5-7 Thin provisioning on the N series at the LUN and volume levels

5.5.1 Application or LUN level thin provisioning

By default, LUNs are created with *Space Reserved* setting of 100 percent. That is, the space required by the LUN is pre-allocated from the FlexVol during LUN creation.

Thin provisioning at the LUN level is enabled by disabling space reservation. In this case, physical space is allocated from the FlexVol only when data is written to the LUN.

Using this feature, you can, for example, provision 100, 50 GB LUNs (5 TB) using 1 TB physical storage rather than the full 5 TB.

5.5.2 Volume level thin provisioning

Volume level thin provisioning is configured using the Volume Space Guarantee settings:

- *Volume Space Guarantee - Volume:* The pace for the entire volume is reserved (or guaranteed) by the aggregate
- *Volume Space Guarantee - File:* Space is reserved only for *Space Reserved* LUNs in the volume by the aggregate
- *Volume Space Guarantee - None:* No space (regardless of the LUN *Space Reserved* settings) is reserved or guaranteed by the aggregate

Note that the *Volume Space Guarantee* setting at the FlexVol level interacts with the *Space Reserved* settings at the LUN level.

For more details of thin provisioning on the N series, see these websites:

- IBM System Storage N series product website:

  http://www-03.ibm.com/systems/storage/network/

- Various IBM Redbooks that discuss the N series in greater depth:

  http://www.redbooks.ibm.com/

6

Compression

This chapter provides comprehensive information about IBM strategic solutions for storage efficiency using data compression. The focus is on a newly introduced solution, the IBM Real-time Compression Appliance (RTCA).

The RTCA product resides between Network Attached Storage (NAS) systems and network switches. The RTCA product consists of preconfigured devices that can compress files and databases up to 80 percent and increase application throughput and user response time, in addition to helping improving the efficiency of the application systems and storage systems. These appliances are *not* suitable for environments with Fibre Channel-attached disk subsystems on a Storage Area Network (SAN), including, for example, the IBM System Storage DS8000 family, the XIV system, and the IBM Storwize V7000.

Other products that use data compression technologies are briefly discussed in this chapter:

- IBM Tivoli Storage Manager with client compression enabled
- IBM Enterprise tape automation systems with Linear Tape-Open (LTO™) Ultrium5 tape drive

Widely known software utilities and tools used for data compression are not covered in this chapter, as they do not bring significant value to storage efficiency solutions. They are generally open source tools from multiple vendors and require user intervention. Therefore, they do not contribute to the concept of automating data compression or decompression.

6.1 Compression in smarter data centers

With the continuously growing need for businesses to gain the competitive edge and simultaneously maintain compliance with regulatory and corporate retention obligations, there is a clear requirement for the data to remain online and available for longer periods of time. The consequence is that the amount of data stored will continue to grow exponentially every year.

This growing capacity creates tremendous strain on the IT infrastructure, specifically on storage systems. The IBM compression solutions, especially the IBM RTCA products, moderate that strain. For example, RTCA solutions can help streamline your business efficiency by increasing your overall storage capacity by up to five times, or 80 percent, enabling you to achieve extraordinary storage efficiency throughout the data life cycle, yet also optimizing your primary storage. It allows you to perform more analytics on more data, giving you the competitive advantage, necessary to stay ahead and on top of the today's market, enabling their business to become smarter, faster, and more efficient.

IBM compression solutions fit in a seamless manner into your storage environment and increase your storage efficiency. Simultaneously, these compression solutions reduce operating expenses and power, cooling, and floor space costs. These products also operate transparently with existing IT processes, such as snapshots, replicas, and backups.

Data compression is one of the core storage efficiency technologies in many data centers. IBM data compression utilities and product features are commonly used to reduce the cost of storing inactive files, backups, and archives. In this chapter, we present how IBM achieves the great compression ratio for given data types and how clients can benefit from the competitive and easy-to-implement, preconfigured solutions.

6.2 IBM Real-time Compression Appliances

IBM Real-time Compression Appliances (RTCAs) are strategic IBM storage compression devices that can shrink primary, online data in real time, without performance degradation. By significantly reducing storage requirements, you can keep up to five times more information online for analytics, use the improved efficiency to reduce storage costs, or achieve a combination of greater capacity and reduced cost. The IBM RTCA products can help deliver improved user response time and overall throughput, because applications spend less time waiting for disk requests.

The IBM Real-time Compression Appliance STN6500 and IBM Real-time Compression Appliance STN6800 models increase the capacity of the existing network attached storage infrastructure. All of the RTCA products apply IBM patented, real time data compression techniques to primary and existing TCP/IP attached storage (for example, the N series), delivering optimization and savings throughout the entire storage life cycle. The result is unprecedented cost savings and return on investment, along with operational and environmental efficiencies.

Storage solution: IBM Real-time Compression Appliances (RTCAs) provide a storage efficiency solution for primary, active data that includes files and database systems (DB2, IBM Informix®) residing on an IBM Network Attached Storage (NAS) N series.

The RTCA product software processes incoming data streams (Network File System (NFS) and Common Internet File System (CIFS) protocol data requests) and compresses these data requests as they are on their way to the storage system. The data is stored in the storage system, not in the appliance, and the acknowledgement is sent directly back to the user or application after the write has been committed and acknowledged by the storage system. From a data integrity perspective, it is imperative that it is the storage system that acknowledges that the write has been committed. It is for this reason that the IBM RTCA product software does not utilize a write cache mechanism. All storage commits come from the system, which preserves the integrity of the data between the storage and the application.

IBM Real-time Compression provides enterprises with the following benefits:

- Better storage utilization:
  - Significant reduction (up to 80 percent) in existing storage utilization
  - Delivers real-time compression without performance degradation
- Lower capital and operational costs (CAPEX and OPEX):
  - Slows the growth of storage acquisition
  - Storage agnostic
  - Reduces the amount of storage to be managed throughout the life cycle
  - Easy, transparent deployment
- Proven technology:
  - Based on Lempel-Ziv (LZ) data compression
  - Thirty-five patents pending or approved
  - Hundreds of customers worldwide
- Better energy efficiency:
  - Less to store, power, and cool
- High availability and disaster recovery:
  - Fully redundant components deployed seamlessly into high-availability environments
  - Significantly less data to be replicated to a disaster site
- Fully transparent:
  - No change to performance, applications, networks, storage, or processes
  - No loss of data integrity

Figure 6-1 shows a front view of the 2-unit (2U) base model of the IBM Real-time Compression Appliance STN6500.

Figure 6-1 IBM Real-time Compression Appliance STN6500 model STN6500

6.2.1 How compression appliances work

IBM Real-time Compression provides online storage optimization through real-time data compression, delivering dramatic cost reduction without performance degradation. The IBM RTCA products for NAS are based on *IBM Random Access Compression Engine (RACE)* technology and transparently compress primary storage without changes in performance, storage, applications, networks, or processes.

RACE technology gives IBM RTCA products the following characteristics:

- Random access and real-time to ensure no performance degradation
- Deterministic for reliable and consistent performance
- Lossless to ensure data integrity

The IBM Real-time Compression RACE is made up of three components:

- Random Access Compression Engine (RACE): Enables random-access data compression without compromising performance.
- Unified Protocol Manager (UPM): Enables transparent support of multiple storage and network protocols, including CIFS and NFS.
- Monitoring and Reporting Manager (MRM): Enables online storage compression trending, analysis, and reporting.

RACE takes incoming data streams and compresses the data within these data requests, leaving the metadata intact, to the storage devices. The data is stored in the array, and the acknowledgement that the write has been committed is sent directly back from the array to the user or application.

Figure 6-2 shows a conceptual model of the implementation of either the IBM Real-time Compression Appliance STN6500 and IBM Real-time Compression Appliance STN6800 models, configured in a high-availability environment.

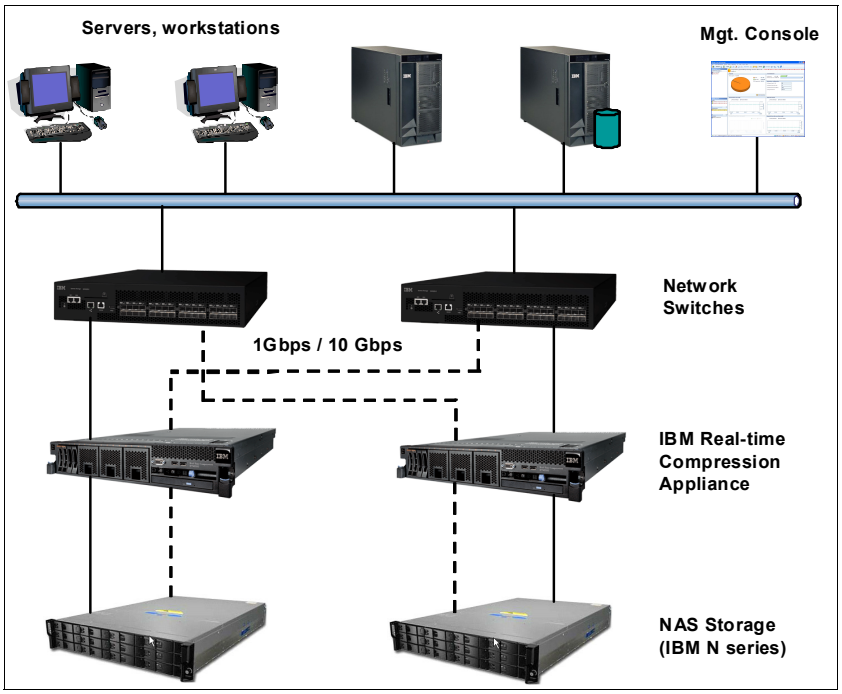

Figure 6-2 IBM RTCA products in a high-availability environment

All requests to read and write data pass through the compression appliance on their way to the storage devices or the user. The appliance is able to operate compression filters. For each file server or NAS storage device, there is a defined list of shares and exports that are included in the compression process. The IBM RTCA product supports both CIFS and NFS filters. Compression filters for CIFS are also referred to as *share filters*, those for the NFS protocol are known as *export filters*. Shares that are listed in a compression filter are called *compressed shares*. Shares that are not listed in a compression filter, and therefore are not compressed by the compression appliance, are called *transparent shares*. By default, all files are transparent unless they belong to shares listed on the Compression Filters page and are not listed in the *File Extensions to Exclude* list.

The compression appliance can connect to a number of file server and storage device interfaces, which are specified by their IP addresses. The compression filter is per file server or storage device. Both CIFS shares and NFS exports can be configured in the same step.

Data flow: Write operation

The typical data write compression process is depicted in the Figure 6-3.

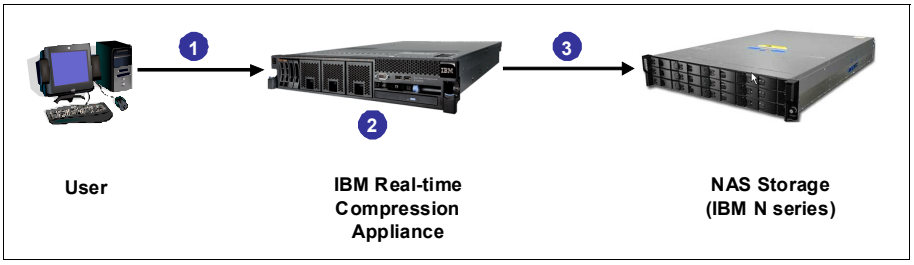

Figure 6-3 Data flow: Write operation with compression

The process is as follows:

1. The user sends a request to write a data block to the storage system. The request is sent over the network from the client through the compression appliance to the storage system.
2. The appliance compresses the data block as the request passes through it.
3. The data block is sent to the storage device, where it is written in its compressed form.

Data flow: Read operation

Similarly, as in the case of the write operation, when the user requests a data read operation, the data flow is as shown in Figure 6-4.

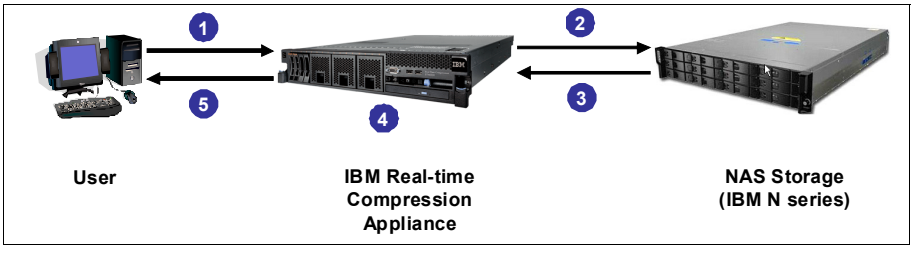

Figure 6-4 Data flow: Read operation with decompression

The process is as follows:

1. The user sends a request to read a data block from the storage system, and the request is sent over the network from the client to the storage system through the compression appliance.
2. The compression appliance sends the request for the data block to the storage system.
3. The storage system returns the relevant compressed data block.
4. The compression appliance decompresses the data block.
5. The decompressed data block is cached in the appliance and returned to the user over the network.

6.2.2 Compression ratio

When an enterprise compresses its files using the IBM compression appliance, it can achieve considerably high compression ratios. This means that available storage space can possibly be doubled, tripled, or even more.

The exact compression ratio depends on the nature of the data. IBM has seen compression ratios as high as 90 percent in certain Oracle database configurations and in the neighborhood of 50 percent with .pdf files. As always, compression ratios vary by data type and how the data is used.

The compression appliance accomplishes this compression in an efficient manner that does not add any latency to the storage server or require allocation of any storage space.

The average IBM Real-time Compression ratios for various types of data are shown in Figure 6-5.

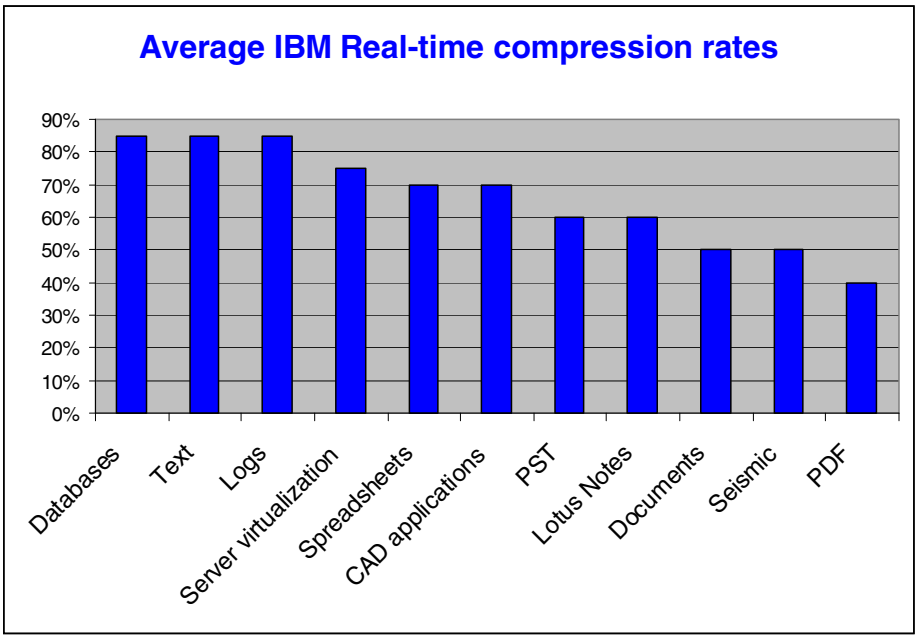

Figure 6-5 IBM Real-time compression ratios

The best way to know what compression ratio you will see for a specific file type is to use any available compression tool that supports the LZ mechanism. This will provide a good approximation of the compressed file size.

6.2.3 Interoperability

This section lists the storage servers and clients supported by the IBM RTCA products.

CIFS shares

Table 6-1 describes the operating systems of clients and NAS storage supported by IBM RTCA products in CIFS environments.

Support: Simple Message Block protocol version 2 (SMBv2) is not supported yet by the compression appliance. This format is available on Windows 2008 servers, Windows 7, and Windows Vista Enterprise Edition with Service Pack 1.

Table 6-1 Supported clients and NAS servers (using CIFS)

Client/storage server	N series Data ONTAP				
	7.0	7.1	7.2	7.3	8.0
Windows 2008 Server	Yes	Yes	Yes	Yes	Yes
Windows 2003 Server SP2	Yes	Yes	Yes	Yes	Yes
Windows 2000 Server SP4	Yes	Yes	Yes	Yes	Yes
Windows 7	Yes	Yes	Yes	Yes	Yes
Windows Vista Enterprise SP1	Yes	Yes	Yes	Yes	Yes
Windows XP SP2 and SP3	Yes	Yes	Yes	Yes	Yes
Samba on Red Hat EL5	Yes	Yes	Yes	Yes	Yes
Samba on SuSE SLED10	Yes	Yes	Yes	Yes	Yes

NFS exports

Table 6-2 describes the operating systems of clients and NAS systems supported for use with the IBM RTCA products in NFS environments.

The compression software supports version 3 of the NFS protocol using TCP or User Datagram Protocol (UDP) running over IP.

Table 6-2 Supported clients and NAS servers (NFS)

Client/storage server	N series Data ONTAP				
	7.0	7.1	7.2	7.3	8.0
Red Hat EL3, EL4, EL5	Yes	Yes	Yes	Yes	Yes
SuSE SLES9, SLES10	Yes	Yes	Yes	Yes	Yes
SuSE SLED10	Yes	Yes	Yes	Yes	Yes
Solaris 8, 9, 10	Yes	Yes	Yes	Yes	Yes
HP-UX 11i v1/v2	Yes	Yes	Yes	Yes	Yes
IBM AIX 5.3, 6.1, 7.1	Yes	Yes	Yes	Yes	Yes
Mac OS X 10.4	Yes	Yes	Yes	Yes	Yes
VMWare ESX 4.0	Yes	Yes	Yes	Yes	Yes

6.2.4 IBM Real-time Compression Appliance offerings

There are two models of IBM Real-time Compression Appliance (RTCA) products. This section outlines the base model configurations and general technical specifications of the products, such as number of ports, memory capacity, and physical attributes. It does not provide deep technical details, such as performance parameters and benchmarking results, because these are available from other resources (see “More information” on page 96).

IBM Real-time Compression Appliance, model STN6500

The STN6500 is a 2U appliance. It comes equipped with 16 one-gigabit copper Ethernet ports and one management port. This is factory-set when the appliance is ordered. At least two cables are required for each NAS. One cable connects the NAS to the STN6500, and another cable connects the STN6500 to the LAN switch for that NAS.

Base components for the STN6500 include these:

- IBM System x3650
- Dual Xeon 4-core 2.40 GHz processors
- 72 GB 1333 MHz memory (RDIMM)
- Dual 146 GB SAS disk drives (6 Gbps)
- Optical drive
- Redundant power supply
- Four Intel® PRO/1000 Quad network interface cards (NICs) (16 network ports)

IBM Real-time Compression Appliance, model STN6800

The STN6800 is a 2U appliance. It comes equipped with 1-gigabit copper Ethernet, 10-gigabit optical Ethernet (10 GBASE-SR), or both ports in the following configurations:

- Four, 10 GbE optical ports and eight, 1 GbE copper ports
- Eight 10 GbE optical ports

In addition, the STN6800 is equipped with a management port. This port configuration is factory-set when the appliance is ordered. At least two cables are required for each NAS. One cable connects the NAS to the STN6800, and another cable connects the STN6800 to the LAN switch for that NAS.

Base components for the STN6800 include these:

- IBM System x3650
- Dual Xeon 4-core 2.40 GHz processors
- 72 GB 1333 MHz memory (RDIMM)
- Dual 146 GB SAS disk drives (6 Gbps)
- Optical drive
- Redundant power supply
- Four Intel PRO/1000 Quad NICs (16 network ports)

6.2.5 More information

For further details about the IBM Real-time compression solutions and appliances, see the following websites:

- Product brochures, marketing materials, and general information:
  - http://www.realtimecompression.com
  - http://www.ibm.com/systems/storage/solutions/rtc/index.html
  - http://www.ibm.com/systems/storage/network/rtc/stn6500/index.html
  - http://www.ibm.com/systems/storage/network/rtc/stn6800/index.html

- Technical details and manuals:
  - STN6500/STN6800 Installation and planning guide:

    http://www.ibm.com/support/docview.wss?uid=ssg1S7003369
  - STN6500/STN6800 Administration Guide:

    http://www.ibm.com/support/docview.wss?uid=ssg1S7003368
- Redbooks publications:
  - Introduction to IBM Real-time Compression Appliances, SG24-7953
  - NAS / N series Solution Design for Real-time Compression Appliances, REDP-4770-00

6.3 Tivoli Storage Manager compression

Tivoli Storage Manager is the number one product of choice for an efficient and effective, enterprise-wide, backup/recovery solution. It provides a data protection solution for backup, archiving, disaster recovery planning, space management, database and application protection, bare machine recovery, and record retention.

Tivoli Storage Manager is the premier storage management solution for mixed platform environments. In this discussion, we focus on data compression using the Tivoli Storage Manager backup/archive client and its application program interface (API) as a common interface for various data protection tools (for SQL, Oracle, DB2, and Mail, for example).

The general concept of Tivoli Storage Manager, along with its features and benefits, are discussed in the Chapter 9, “Backup and recovery” on page 157.

6.3.1 Tivoli Storage Manager basic components

Tivoli Storage Manager, as an enterprise-wide storage management application, provides automated storage management services to workstations, personal computers, and file servers from a variety of vendors, with a variety of operating systems.

The software product consists of two basic functional components:

- Tivoli Storage Manager server with DB2 database engine
- Tivoli Storage Manager clients with common API

Tivoli Storage Manager server

The server program provides backup, archive, and space management services to the clients and manages the storage repository. The storage repository can be implemented in a hierarchy using any combination of supported media or magnetic or optical disk, tape, and robotic storage devices that are locally connected to the server system or are accessible through a SAN. To take advantage of SAN technology, the Tivoli Storage Manager server has features implemented that dynamically share SAN-connected, automated tape library systems among multiple Tivoli Storage Manager servers, and provide (as an option) local area network, or LAN-free and server-free backup.

Tivoli Storage Manager client

Data is sent to the Tivoli Storage Manager server using the Tivoli Storage Manager backup/archive client and complementary products. These products work together with the Tivoli Storage Manager server base product to ensure that any data stored is managed as defined. The Tivoli Storage Manager backup/archive client, included with the server, provides the operational backup and archive functions. The client implements the patented progressive backup methodology, adaptive sub-file backup technology, and unique record retention methods for backup and archive functions.

6.3.2 Tivoli Storage Manager client compression

Tivoli Storage Manager clients benefit from the data compression configured on the source server by reducing the amount of data sent to the Tivoli Storage Manager server over the network and stored within the storage repository. Data compression is enabled by specific settings of Tivoli Storage Manager client parameters that are independent on the Tivoli Storage Manager server operation and data storage hierarchy.

Example 6-1 shows the typical settings in the configuration file (dsm.opt) of a Tivoli Storage Manager client.

Example 6-1 Compression settings of Tivoli Storage Manager client for Windows

```
COMPRESSION YES
COMPRESSALWAYS NO

Exclude.Compression *:\...\*.jpg
Exclude.Compression *:\...\*.zip
Exclude.Compression *:\...\*.tar
Exclude.Compression *:\...\*.rar
```

The Tivoli Storage Manager client contains an option to exclude files from compression processing if the file size is greater after compression. In this case, the file is sent to the server storage repository uncompressed. Typical examples of such files are those that are already compressed archives (for example, .zip, .rar, or .tar files) or compressed pictures (for example, .jpg, or .png files).

Use the *compression* option (as in Table 6-1 on page 95) with the **archive**, **incremental**, **selective**, or **backup image** commands. This option is valid on the initial command line and in interactive mode. The Tivoli Storage Manager server can also define this option which overrides the client value.

If you set the *compressalways* option to *yes*, compression continues even if the file size increases. To stop compression if the file size grows, and resend the file uncompressed, set the *compressalways* option to *no* as shown in the Example 6-1.

When you enable compression, it reduces network utilization and saves server storage, but causes additional processor overhead to the node. Data compression is advised only when there is insufficient network capacity.

Tip: Use either the Tivoli Storage Manager client compression or tape drive compression, but not both. For details, check the following link and navigate to the compression section:

http://publib.boulder.ibm.com/infocenter/tsminfo/v6/index.jsp

Compressing your files reduces data storage for backup versions and archive copies of your files. It can, however, affect Tivoli Storage Manager client/server throughput. A fast processor on a slow network connection benefits from compression, but a slow processor on a fast network connection does not.

6.3.3 More information

For comprehensive information about Tivoli Storage Manager client compression, see the following references:

- http://publib.boulder.ibm.com/infocenter/tsminfo/v6r2/index.jsp?topic=/com.ibm.itsm.srv.doc/c_volumes_data_compression_ulw.html
- *IBM System Storage Solutions Handbook*, SG24-5250
- *Tivoli Storage Manager V6.1 Technical Guide*, SG24-7718

6.4 Enterprise tape drive compression

The first generation of the LTO Ultrium5 technology introduced on-chip hardware data compression using the Streaming Lossless Data Compression (SLDC) algorithm. This data compression basically comes from industry proven Lempel-Ziv (LZ) compression, with several specific enhancements adapted for on-chip tape drive functionality.

The most remarkable benefit of using LTO data compression (LTO-DC) is that the SLDC algorithm does not apply to incompressible data. The typical scenario is that the files already compressed using other software or tools, or the structure of the incoming data, is so random that it can defeat the on-chip compression algorithm.

Every block of data written to tape has a header bit showing if the block is compressed by SLDC or is in raw format. For each block of data that the algorithm works on, it caches a copy of the raw data. After the successful compression of the data, the algorithm compares the compressed data block to the raw data block in memory and writes the smaller of the two to tape.

LTO-DC achieves an approximately 2:1 compression ratio on all available LTO Ultrium generations. For example, the latest technology level (LTO generation 5) indicates native data capacity of 1.5 TB per cartridge. Compressed capacity is reached up to 3 TB per cartridge for specific data formats.

As already mentioned in “Compression ratio” on page 94, the best compression rates are achieved for plain text files, databases files, and raw images. Data that is already compressed does not benefit from the hardware compression algorithm provided by enterprise tape drives (for example, .zip, .rar, .tar, .jpg, ,mpeg, ,avi, and mp3).

For further information, see these websites:

- http://www.ibm.com/systems/storage/tape/
- http://www.ultrium.com/index.html

7

Deduplication

Data deduplication has emerged as a key technology in an effort to dramatically reduce the amount of, and the cost associated with storing, large amounts of data. Deduplication is the art of intelligently reducing storage needs an order of magnitude better than common data compression techniques.

In this chapter, we discuss how deduplication is addressed by IBM technologies and products. including these:

- IBM ProtecTIER products
- IBM System Storage N series
- IBM Tivoli Storage Manager

7.1 The concept of data deduplication

Data deduplication has emerged as a key technology in an effort to dramatically reduce the amount of, and the cost associated with storing, large amounts of data. Deduplication is the art of intelligently reducing storage needs an order of magnitude better than common data compression techniques. This is accomplished by eliminating redundant data so that only one instance of a data set is actually stored. IBM has a broad portfolio of deduplication solutions, and the ability to solve client issues with the most effective technology. Whether it is source or target data, inline or post, hardware or software, disk or tape, IBM has a solution with the technology that best solves the problem.

Data deduplication products read data while looking for duplicate data. Data deduplication products break up the data into elements, using their respective techniques (see, “Types of data deduplication and HyperFactor” on page 103) to create a signature, or identifier, for each data element. Then, data element signatures are compared to identify duplicate data. With duplicate data identified, one copy of each element is retained. Pointers are created for the duplicate items, and then the pointers and duplicate items are discarded.

The effectiveness of data deduplication is dependent upon many variables, including the rate of data change, the number of backups, and the data retention period. For example, if you backup the same uncompressible data one time per week for six months, you save the first copy and do not save the next 24, which results in a 25:1 data deduplication ratio. If you back up an uncompressible file on week one, back up the same file again on week two, but never back it up again, you have a 2:1 deduplication ratio. A more likely scenario is that a portion of your data will change from backup to backup, so that the data deduplication ratio will change over time. With data deduplication, you can minimize your storage requirements.

Data compression is a core storage efficiency technology in most data centers. IBM data compression utilities and product features are commonly used to reduce the cost of storing inactive files, backups, and archives. For data that is backed up repetitively, data deduplication achieves this space reduction by storing a single copy of the data.

The concept of data deduplication is shown in Figure 7-1.

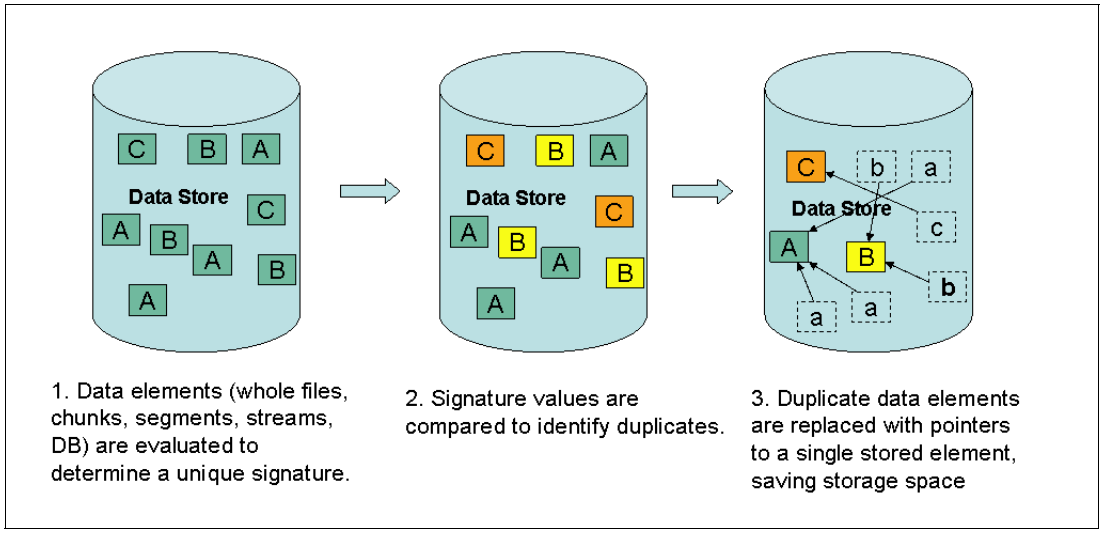

Figure 7-1 The basic concept of data deduplication

Data deduplication can reduce your storage requirements, but the benefits derived are determined by your data and backup policies. Workloads with a high database content generally have the highest deduplication ratios; however, product functions such as Tivoli Storage Manager Progressive Incremental, Oracle RMAN, or Light Speed, can reduce the deduplication ratio. Compressed, encrypted, or otherwise scrambled workloads typically do not benefit from deduplication. Good candidates for deduplication are typically text files, log files, uncompressed, and non-encrypted database files, email files (.pst, .dbx, and IBM Domino®), and snapshots (Filer Snaps, Business Continuance Volumes (BCVs), and VMWare images).

In the following paragraphs, we discuss how IBM technologies and products meet our clients' data deduplication criteria and how our clients can benefit from our competitive, easy-to-implement, preconfigured compression solutions.

Data deduplication can provide greater data reduction and storage space savings than other technologies available today. To further increase disk space, consider combining data deduplication with other technologies, such as Lempel-Ziv (LZ) compression and differencing, a technology used for differential backups. For further information about data compression, see Chapter 6, "Compression" on page 89.

7.1.1 Types of data deduplication and HyperFactor

Many vendors offer products that perform deduplication. Various methods are used for deduplicating data. These are three methods frequently used:

- *Hash-based:* Deduplication uses a hashing algorithm to identify chunks of data. Commonly used processes are Secure Hash Algorithm 1 (SHA-1) and Message-Digest Algorithm 5 (MDA-5). The details of each technique are out of scope for this publication.
- *Content aware:* Deduplication methods are aware of the structure of common patterns of data used by applications. It assumes that the best candidate to deduplicate against is an object with the same properties, such as a file name. When a file match is found, a bit-by-bit comparison is performed to determine if data has changed and the changed data is saved.
- *HyperFactor:* This is a patented technology used in the IBM System Storage ProtecTIER Enterprise Edition software. HyperFactor takes an approach that reduces the phenomenon of missed factoring opportunities, providing a more efficient process, rather than comparing the new data to the similar data to identify and store only the byte-level changes. With this approach, HyperFactor is able to surpass the reduction ratios attainable by any other data reduction method. HyperFactor can reduce any duplicate data, regardless of its location or how recently it was stored. HyperFactor data deduplication uses a 4 GB Memory Resident Index (MRI) to track similarities for up to 1 PB of physical disk in a single repository.

HyperFactor technology

HyperFactor technology, as seen in Figure 7-2, uses a pattern algorithm that can reduce the amount of space required for storage by up to a factor of 25, based on evidence from existing implementations. The capacity expansion that results from data deduplication is often expressed as a ratio, essentially the ratio of nominal data to the physical storage used, as discussed in “The concept of data deduplication” on page 102.

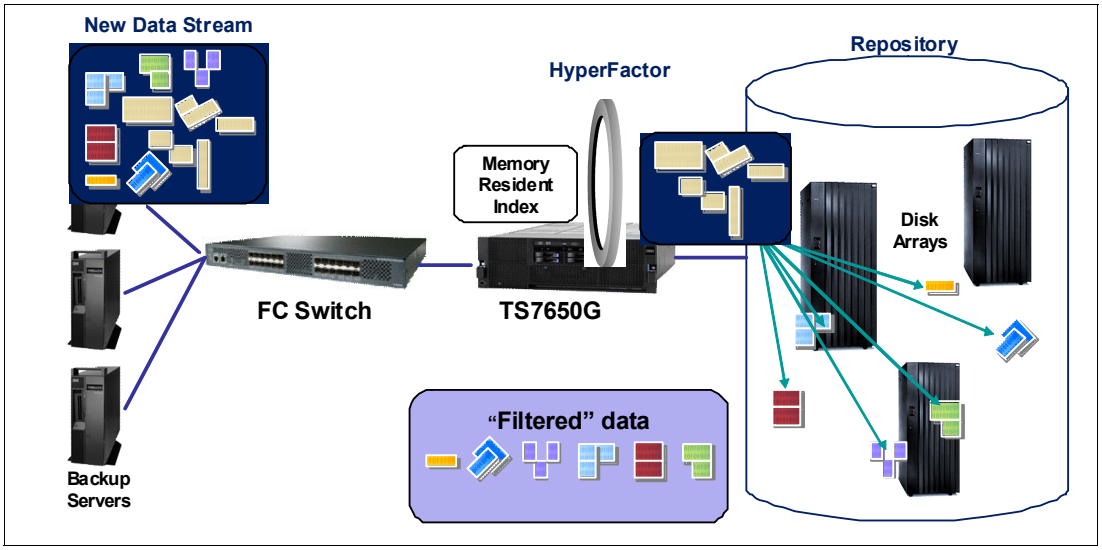

Figure 7-2 IBM HyperFactor technology

Data deduplication processing

Data deduplication can be performed either when the data is being backed up to the storage media (real-time or inline), or after the data has been written to the storage media (post-processing). Each method has positive and negative aspects that must be evaluated by the engineer or technical specialist responsible for the architecture and deployment of the solution. IBM uses inline deduplication processing, as this offers larger target storage space without the need for a disk cache pool for post-processed deduplicated data.

Bit comparison techniques, such as the one used by ProtecTIER, are designed to provide 100 percent data integrity by avoiding the risk of hash collisions.

7.1.2 ProtecTIER native replication

ProtecTIER Native Replication is a logical feature that enables clients to replicate and *move* any or all of their virtual tape cartridges from their main site to a remote disaster recovery (DR) location and vice versa. The TS7650 Appliance and Gateway (TS7650-G) models employ an IP-based replication design. Data is replicated between sites from one, two-node cluster to another.

The user can create policies at the virtual cartridge level to replicate a single barcode or a range of barcodes in a specified time frame and with a set priority. After being replicated (or moved) to the DR site, users can clone these cartridges to real physical tape from within their backup application and outside of their main production stream. In case of a disaster, the TS7650 or TS7650G located at the DR site can become the production (failover) site until the main site comes back online. At that point, the user can replicate or move the newly created tapes back to the main production site (failback). ProtecTIER and Native Replication enables more frequent testing of DR plans.

Figure 7-3 shows the concept of ProtecTIER Native Replication using IBM ProtecTIER Gateway as a hub at the central or disaster site and four spokes at branch offices or remote sites. As mentioned in the figure, each hub supports up to 12 spokes, depending on concrete model and deduplication solution.

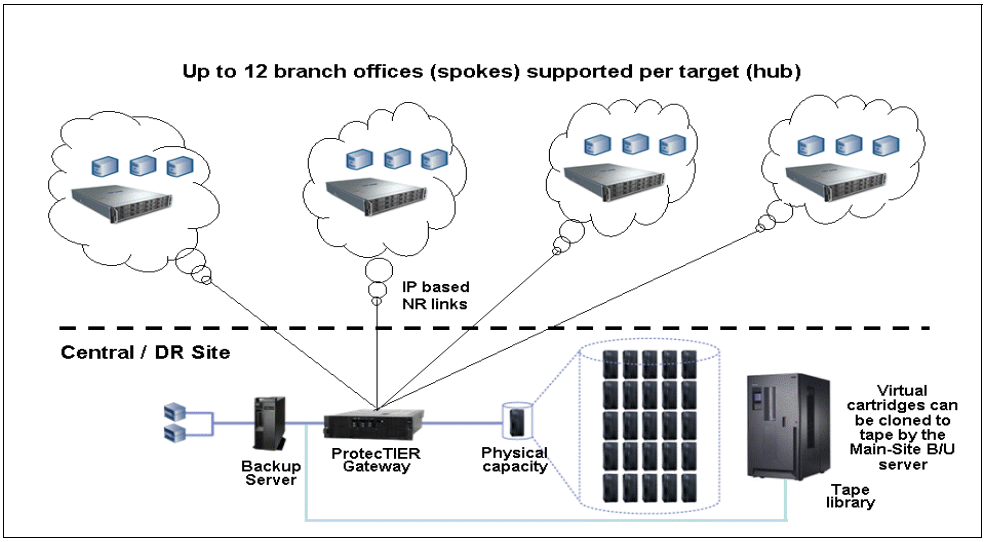

Figure 7-3 Native replication using the IBM ProtecTIER solution

Because ProtecTIER deduplicates data before storing it, only the changes to the data are transferred to the DR site over the replication link. This translates into substantial savings in the bandwidth needed for the replication link.

Tip: All available IBM ProtecTIER deduplication products use inline (real-time) deduplication processing.

Data transfer is started, based on several trigger points, such as policy-based transfer windows, or movement of the virtual tape to a Virtual Tape Library (VTL) export slot. (Note that the ProtecTIER VTL emulates import and export slots, and the opening or closing of the library door to insert or eject a cartridge.) Data verification and validation are done at the DR site to ensure integrity of the transferred data prior to making the virtual cartridge or tape available. ProtecTIER only replicates a partial set of cartridges. ProtecTIER Native Replication is built on top of the ProtecTIER two-node cluster product and is cluster-compliant. Replication can run in parallel to backup and restore activities.

ProtecTIER Native Replication eliminates the risks associated with human intervention and physical tape transportation. Replication occurs automatically (by policy) and simultaneously as data is deduplicated and backed up.

7.2 IBM ProtecTIER products

In this section, we describe the IBM ProtecTIER products in more detail:

- IBM System Storage TS7610 ProtecTIER Express
- IBM System Storage TS7650 ProtecTIER Deduplication Appliance
- IBM System Storage TS7680 ProtecTIER Deduplication Gateway for System z
- IBM System Storage N series

7.2.1 TS7610 ProtecTIER Deduplication Appliance Express

Available in two configuration options, the IBM System Storage TS7610 ProtecTIER Deduplication Appliance Express (type 3959-SM1) is an integrated server and storage hardware platform that ships with the ProtecTIER deduplication software preinstalled. It has a preconfigured repository and VTL interface.

With the HyperFactor 25:1, the IBM System Storage TS7610 ProtecTIER Deduplication Appliance Express is capable of storing up to 135 TiB, or 150 TB, of deduplicated data from a backup server (see Figure 7-4 for details and typical scenario). Both configurations support sustained write performance of more than 80 MB per second (MBps), which usually exceeds the backup data bandwidth in small and mid-range business solutions.

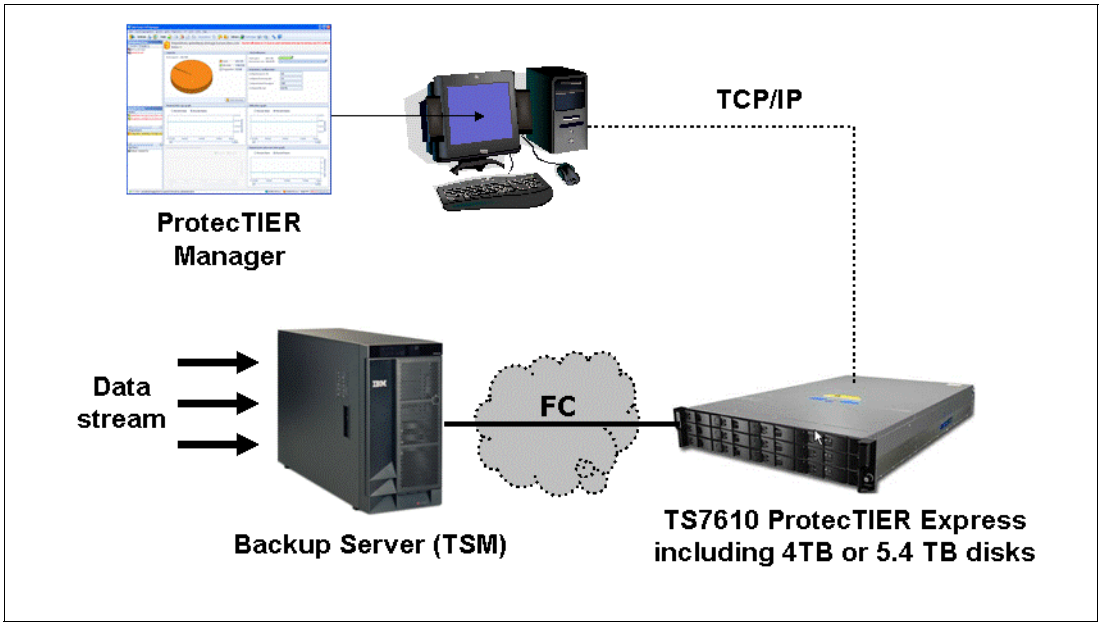

Figure 7-4 Implementation of the IBM TS7610 ProtecTIER Deduplication Appliance Express (conceptual model)

IBM System Storage TS7610 ProtecTIER Deduplication Appliance Express is an ideal solution for clients who have the following needs:

- Experiencing significant data growth in small and mid-range storage solutions
- Targeting to make backup and recovery improvements without making radical changes
- With weekly, full backups 3 TB or less
- With daily incremental backups 1 TB or less

For additional information, see the *IBM System Storage TS7610 ProtecTIER Deduplication Appliance Express User's and Maintenance Guide* located at this website:

`http://www.ibm.com/support/docview.wss?uid=ssg1S7003273`

Interactive documentation for service and support staff is located at this website:

`http://publib.boulder.ibm.com/infocenter/ts7610/serv/index.jsp`

Comprehensive information for our clients and end users is located at this website:

`http://publib.boulder.ibm.com/infocenter/ts7610/cust/index.jsp`

7.2.2 TS7650 ProtecTIER Deduplication Appliance

The TS7650 Appliance (machine type 3958-AP1 on Figure 7-5) is a flexible solution that delivers inline deduplication and scalability that is ideal for IT organizations experiencing rapid data growth and struggling to meet a shrinking backup window.

Figure 7-5 The TS7650 Appliance

The TS7650 Appliance offers the following benefits:

- Significant reduction of backup window
- Restoration time reduction
- Data sharing and resource virtualization optimization
- Operational efficiency
- Emulation of TS3500 tape library with LTO2 and LTO3 virtual tape drives
- Simplified configuration and deployment

TS7650 additional information

For additional information about the TS7650 Appliance, our IBM Client Information Center provides complete and comprehensive information to clients and IT managers at this website:

`http://publib.boulder.ibm.com/infocenter/ts7650/cust/index.jsp`

ProtecTIER product offerings additional information

For more information about the IBM System Storage ProtecTIER products, see ths website:

- *IBM System Storage TS7650 and TS7650G with ProtecTIER*, SG24-7652 located at

  `http://www.redbooks.ibm.com/abstracts/sg247652.html`
- Online documentation, including web-based training sessions, is located at this website:

  `http://publib.boulder.ibm.com/infocenter/ts7650/serv/index.jsp`
- The IBM Client Information Center provides complete and comprehensive information about the TS7650 Appliance to clients and IT managers at this website:

  `http://publib.boulder.ibm.com/infocenter/ts7650/cust/index.jsp`

7.2.3 TS7650G ProtecTIER Deduplication Gateway

The IBM System Storage TS7650G ProtecTIER Deduplication Gateway (Figure 7-6) is designed to meet the disk-based data protection needs of the enterprise data center, and also to enable significant infrastructure cost reductions. The solution offers industry leading inline deduplication performance and scalability up to 1 petabyte (PB) of physical storage capacity per system that can provide up to 25 PB of storage capacity. Combined with IBM storage, the ProtecTIER Gateway solution provides a powerful disk-based repository to improve the retention and availability of backup and archive data.

Figure 7-6 TS7650G ProtecTIER Gateway

The TS7650G is available in stand-alone and clustered configurations. For a stand-alone configuration, one IBM machine type and model 3958-DD4 server is required. For a clustered configuration, two 3958-DD4 or two 3958-DD3 (or any combination) servers are required, along with a Cluster Connection Kit, which includes two required Ethernet switches and one remote network power switch.

Tip: IBM does not support more than one clustered pair of TS7650 Gateway servers in a single frame.

To facilitate backup applications that are designed for use with tape storage, the TS7650G emulates a traditional tape library unit IBM System Storage TS3500 with the support of Linear Tape-Open (LTO) tape drives, LTO2 and LTO3.

The disk storage array attaches to the TS7650G by Fibre Channel connection and holds the repository of factored backup data. The amount of cache available depends on your disk subsystem and configuration. The TS7650G supports the following IBM disk subsystems:

- DS8000 series - all models
- DS5000 series - DS5100 and DS5300
- XIV system 2810
- SAN Volume Controller version 4.3.1 and higher

Distribution across sites

The replication function allows IBM ProtecTIER deployment to be distributed across sites. Network bandwidth remains the most expensive component in most distributed IT environments. The IBM ProtecTIER patented deduplication technology dramatically reduces the amount of bandwidth required by transmitting only new, unique data to the remote location. This feature radically reduces the costs associated with electronically transmitting data and extends the benefits of replication to a larger portion of applications and data.

See Figure 7-7 for typical IBM TS7650G ProtecTIER native replication scenario.

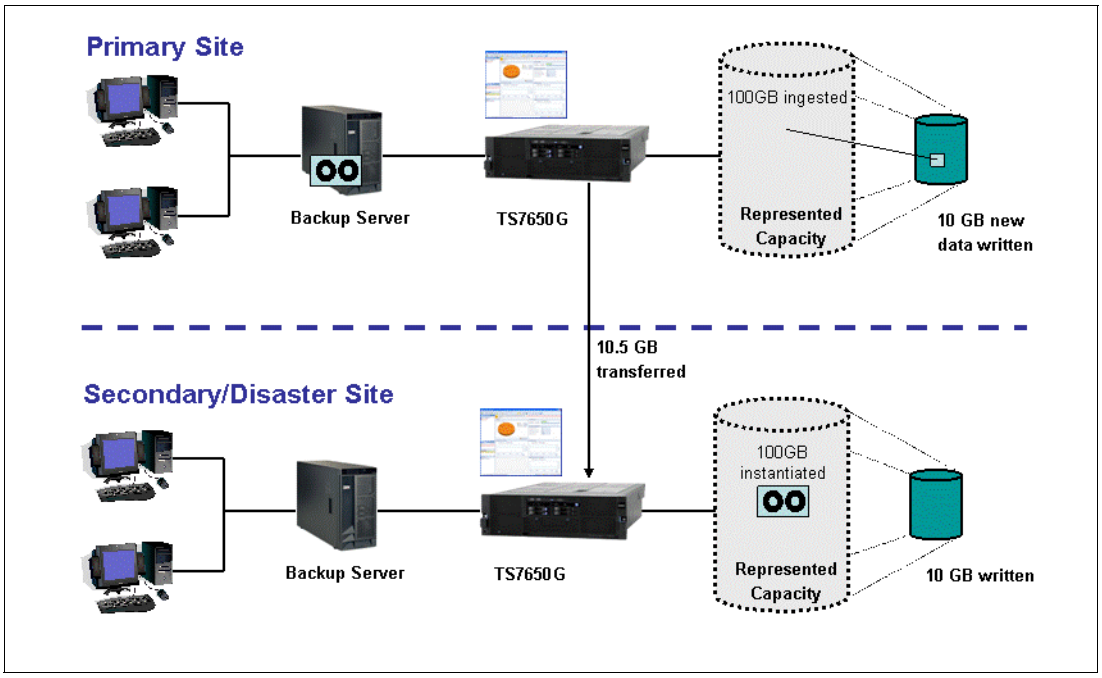

Figure 7-7 ProtecTIER native replication using the TS7650G Gateway

The IBM TS7650G delivers a unique data protection solution that offers these benefits:

- Accelerates backups to meet ever-shrinking backup windows
- Enhances the ability to rapidly restore mission-critical data
- Increases the reliability of backup operations
- Enables proactive verification of backup data integrity
- Increases onsite recovery capacity, thus retaining smaller amounts of storage
- Does not require changes in existing backup processes, procedures, and policies
- Extends disaster recovery capabilities
- Significantly lowers Total Cost of Ownership (TCO)

TS7650G additional information

For additional information, see the following references:

- Online documentation including web-based training sessions is available at this website:

  `http://publib.boulder.ibm.com/infocenter/ts7650/serv/index.jsp`

- The IBM Client Information Center contains comprehensive information about the IBM System Storage TS7650G Deduplication Gateway for clients and end users at this website:

  `http://publib.boulder.ibm.com/infocenter/ts7650/cust/index.jsp`

7.2.4 IBM TS7680 ProtecTIER Gateway for System z

The IBM System Storage TS7680 ProtecTIER Deduplication Gateway for System z (machine type and model 3958-DE2) combines a Virtual Tape Library solution, with IBM unique and patented HyperFactor deduplication technology and integrated native replication technology to provide users an optimal disk-based target for Systems z applications that traditionally use tape.

The TS7680 provides a connection to System z hosts to transfer your data to and from your repository through Fibre Channel connection. The TS7680 combines two Enterprise controllers and two ProtecTIER servers to provide a high-availability Enterprise deduplication solution with drive and library virtualization. While the Enterprise controllers provide the connectivity to your System z host and control the operation of the system, the ProtecTIER servers provide connectivity to the disk repository and control the identification and the tagging of duplicate data.

IBM System Storage TS7680 ProtecTIER Gateway is available in just one configuration. When using ProtecTIER Native Replication for disaster recovery, the same hardware configuration is required at both sites. See Figure 7-8 for a typical setup with these features:

- Two, ProtecTIER servers
- Two, Enterprise controllers (lower and upper in the same rack)
- One, IBM TS3000 System Console (TSSC) and TSSC Ethernet switch
- One communication module (keyboard, video, mouse

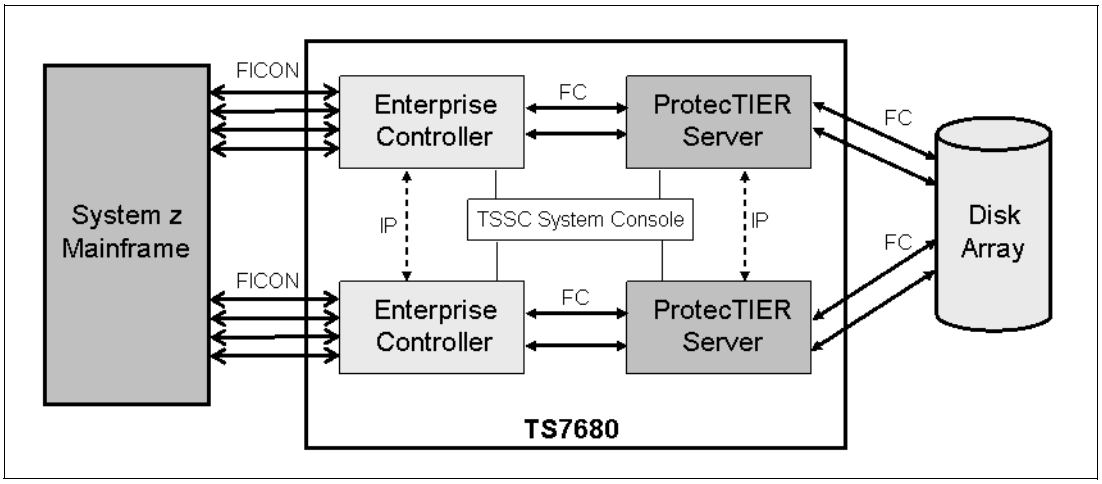

Figure 7-8 Implementation of TS7680 ProtecTIER Gateway

Benefits

A two-node system uses two servers in an active-active cluster and enables a sophisticated system with the following benefits:

- *High availability:* The clustered configuration provides hardware redundancy so that in the event of a node failure, the workload can be distributed to the functioning node, keeping the system available.
- *Increased performance:* Provided that there are sufficient disk resources, the two servers can share the host workload for increased total performance.

Each node connects through front-end Fibre Channel ports to the backup server, and through back-end Fibre Channel ports to the disk repository. The Enterprise controllers enable the nodes to connect to each other and to the ProtecTIER Manager workstation.

Features

The 3958-DE2 (TS7680) provides the following capabilities:

- Compatible with wide area of IBM disk subsystems (IBM DS4000®, IBM DS5000, IBM DS8000, XIV system)
- Up to one million virtual tape volumes
- 128 virtual drive images per Enterprise controller
- Up to 1 PB of raw storage scalability per ProtecTIER server

- Integrated library manager functionality
- Clustering for higher performance and availability
- DFSMS (SMSTAPE) functionality
- Support for z/OS V1R10 and later with program temporary fixes (PTFs)
- Support for IBM z/VM® 5.4 and later with PTFs
- Support for replication functions and disaster recovery (DR) management
- FICON attachment (up to 4 ports per Enterprise controller)
- Scratch processing delete function
- Inline data deduplication performance

The extended information for service support and IT personnel can be found at this website:

http://publib.boulder.ibm.com/infocenter/ts7680/serv/index.jsp

For client-related and general information, but with all technical details, see this website:

http://publib.boulder.ibm.com/infocenter/ts7680/cust/index.jsp

7.3 IBM System Storage N series deduplication

The N series offers deduplication as a native part of its Data ONTAP operating system. This is a feature that can be added for no additional cost and leverages several benefits already built into the storage system OS. Because it is post process deduplication within the storage system, it is possible to realize immediate benefits with existing storage capacity.

The N series deduplication follows a risk-averse post-process methodology and provides owners the benefits of both source and destination deduplication. For environments that rely on multi-tier storage architecture with N series, the deduplication gains will be achieved not only on primary storage, but on backup and archival disk tiers as well.

The N series system uses sub-file, fixed block granularity, which means it is generally file and OS agnostic. Data chunks being compared are at the 4K physical block level. Hashing of incoming data is done by default within the Data ONTAP OS so deduplication is able to take advantage of these existing block fingerprints by simply copying them to a catalog for later comparison. These fingerprints are initially used to detect duplicate blocks after which a byte level comparison between the blocks is done to validate duplication. This two-phase comparison eliminates risk of hash collisions and data corruption.

Because the N series OS is already capable of multiple block referencing, it is simple for the deduplication process to change inode pointers to reference a single physical block rather than multiple, duplicate blocks. After the duplicate blocks are no longer referenced, they are considered to be available again and space savings have been achieved. A maximum of 255 such references can be made to a single physical block. Thus the maximum deduplication ratio that can be achieved is 255:1.

The deduplication license comes at no cost with the N series Filer and Gateway systems. The Gateway systems can be added to existing non-N series disk storage systems to allow these systems to also take advantage of the benefits and features of Data ONTAP.

7.3.1 How deduplication for the IBM System Storage N series works

Regardless of the operating system, application, or filesystem type, *all* data blocks are written to a storage system using a data reference pointer, without which the data cannot be referenced or retrieved. In traditional (non-deduplicated) filesystems, data blocks are stored without regard to any similarity with other blocks in the same filesystem.

In Figure 7-9, five identical data blocks are stored in a filesystem, each with a separate data pointer. Although all five data blocks are identical, each is stored as a separate instance and each consumes physical disk space.

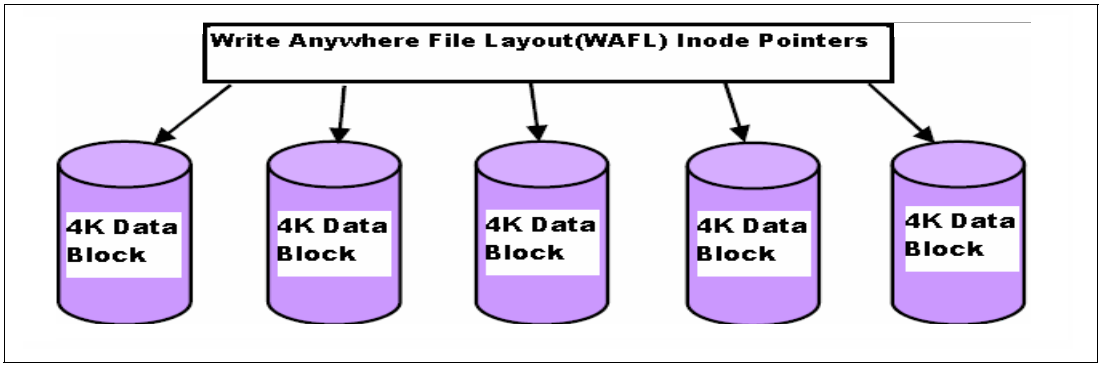

Figure 7-9 Non-deduplicated data

In a deduplicated N series filesystem, two new and important concepts are introduced:

- A catalog of all data block is maintained. This catalog contains a record of all data blocks using a “hash” or fingerprint that identifies the unique contents of each block.
- The filesystem is capable of allowing many data pointers to reference the same physical data block.

Cataloging data objects, comparing the objects, and redirecting reference pointers forms the basis of the deduplication algorithm. As shown in Figure 7-10, referencing several identical blocks with a single master block allows the space that is normally occupied by the duplicate blocks to be made available to the storage system.

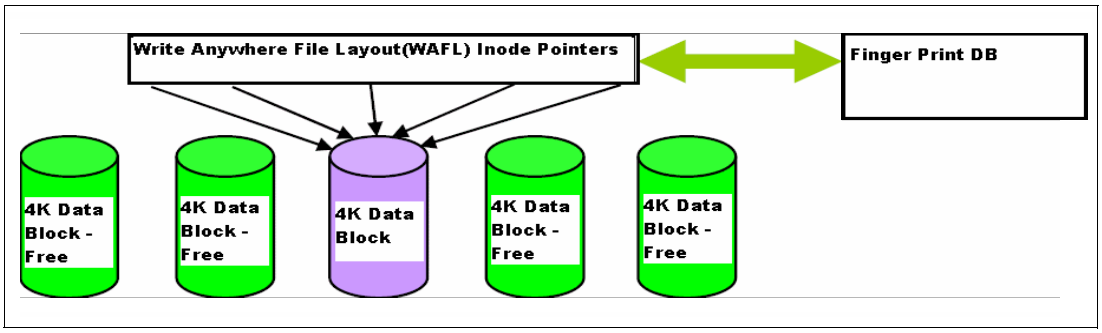

Figure 7-10 Deduplicating identical blocks creates available space

Hashing

Data deduplication begins with a comparison of two data blocks. It is impractical (and arduous) to scan an entire data volume for duplicate blocks each time a new data is written to that volume. For that reason, deduplication creates small hash values for each new block, and stores these values in a catalog.

A hash value, also called a digital fingerprint or digital signature as shown in Figure 2-3, is a small number that is generated from a longer string of data. A hash value is substantially smaller than the data block itself, and is generated by a mathematical formula in such a way that it is unlikely (although not impossible) for two non-identical data blocks to produce the same hash value.

Figure 7-11 shows how hash value is a digital fingerprint that represents a much larger object.

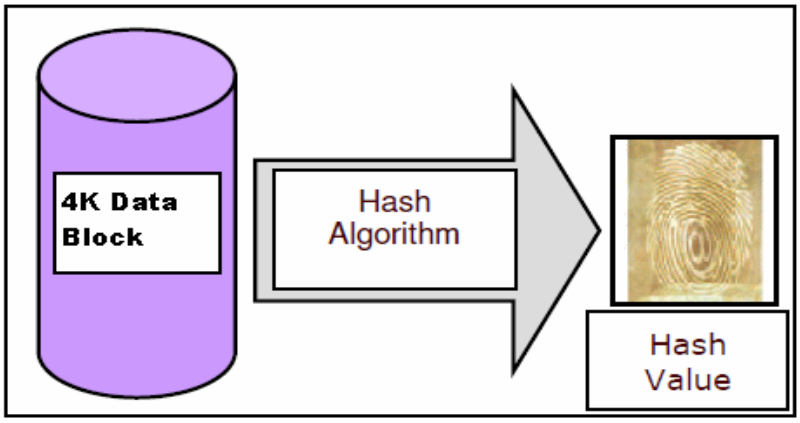

Figure 7-11 In deduplication, a hash value (a digital fingerprint or digital signature) is stored for each new block and represents a much larger object

A hash value is a digital fingerprint that represents a much larger object. This is the deduplication process:

1. The fingerprint catalog is sorted and searched for identical fingerprints.
2. When a fingerprint “match” is made, the associated data blocks are retrieved and scanned, byte-for-byte.
3. Assuming successful validation, the inode pointer metadata of the duplicate block is redirected to the original block.
4. The duplicate block is marked free and returned to the system for re-use.

Hash catalog

A catalog of hash values is used to identify candidates for deduplication. A system process identifies duplicates, and data pointers are modified accordingly. The advantage of catalog deduplication is that the catalog is used only to identify duplicate objects; it is not accessed during the actual reading or writing of the data objects. That task is still handled by the normal filesystem data structure as shown in Figure 7-12.

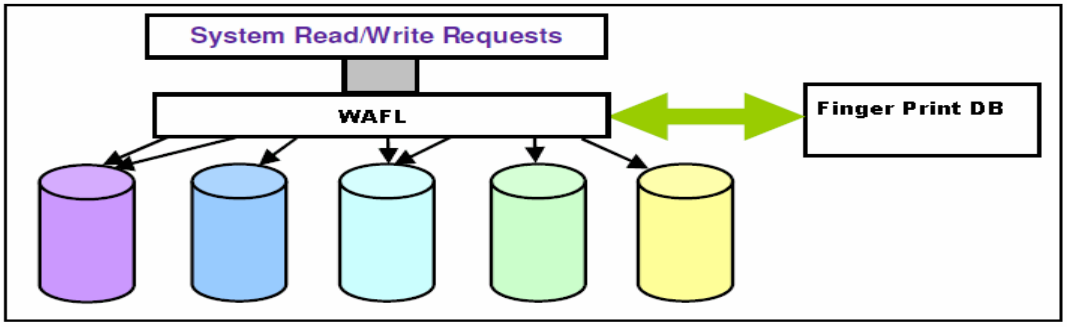

Figure 7-12 Catalog indexing: The filesystem controls block sharing of deduplicated blocks

Deduplication is an N series storage efficiency offering that provides block-level deduplication within the entire flexible volume on N series storage systems. Beginning with Data ONTAP 7.3, N series Gateways also support deduplication. The N series Gateways are designed to be used as a gateway system that sits in front of third-party storage, allowing N series storage efficiency and other features to be used on third-party storage.

Figure 7-13 shows how N series Deduplication works at a high level.

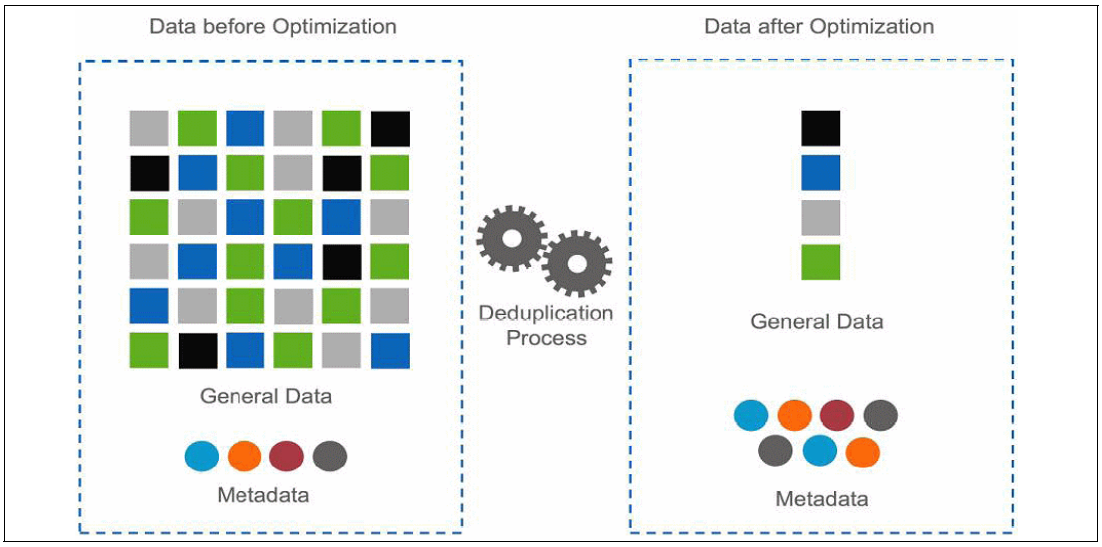

Figure 7-13 High level view of deduplication operations

Essentially, deduplication stores only unique blocks in the flexible volume and creates a small amount of additional metadata during the process. Deduplication has the following features:

- It works with a high degree of granularity (that is, at the 4 KB block level).
- It operates on the active filesystem of the flexible volume. Any block referenced by a snapshot is not made available until the snapshot is deleted.
- It is a background process that can be configured to run automatically, can be scheduled, or can run manually through the command line interface (CLI).
- It is application-transparent, and therefore can be used for data originating from any application using the N series.
- It is enabled and managed through a simple CLI.
- It can be enabled on (and can deduplicate blocks on) flexible volumes with both new and existing data.

In summary, deduplication works as follows:

1. Newly saved data on the N series storage system is stored in 4 KB blocks as usual by Data ONTAP.
2. Each block of data has a digital fingerprint that is compared to all other fingerprints in the flexible volume.
3. If two fingerprints are found to be the same, a byte-for-byte comparison is done of all bytes in the block. If an exact match is found between the new block and the existing block on the flexible volume, the duplicate block is discarded and its disk space is reclaimed.

Deduplicated volumes

Despite the introduction of less expensive ATA disk drives, one of the biggest challenges for disk-based backup today continues to be storage cost. We want to reduce storage consumption (and therefore storage cost per megabyte) by eliminating duplicated data through sharing across files.

The core N series technology used to accomplish this goal is the *dense volume*, a flexible volume that contains shared data blocks. The Data ONTAP filesystem, Write Anywhere File Layout (WAFL), is a filesystem structure that supports shared blocks in order to optimize storage space consumption. Basically, within one filesystem tree, you can have multiple references to the same data block, as shown in Figure 7-14.

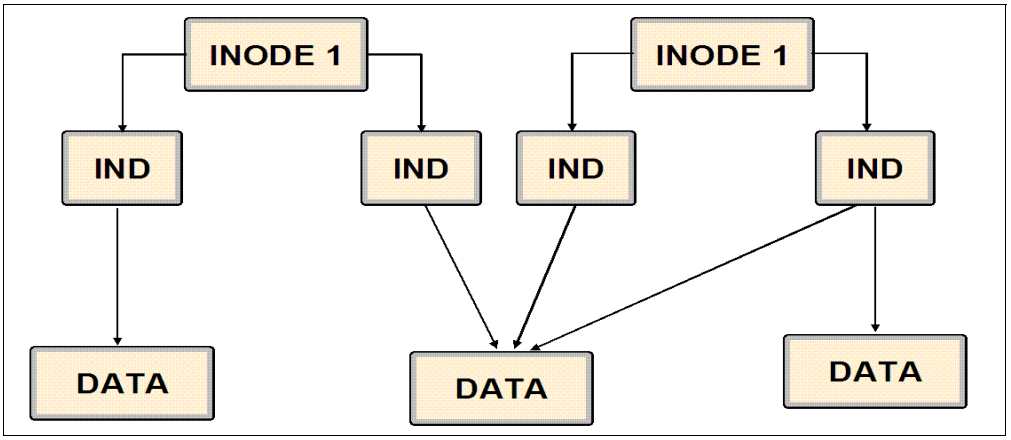

Figure 7-14 Data structure in a deduplicated volume

In Figure 7-14, the number of physical blocks used on the disk is three (rather than five), and the number of blocks saved by deduplication is two (five minus three). In the remainder of this chapter, these blocks are referred to as used blocks and saved blocks, also known as *multiple block referencing*.

Each data block has a block-count reference kept in the volume metadata. As additional indirect blocks (shown as IND in Figure 7-14) point to the data, or as existing blocks stop pointing to the data, this value is incremented or decremented accordingly. When no indirect blocks point to a data block, it is released.

The Data ONTAP deduplication technology therefore allows duplicate 4 KB blocks anywhere in the flexible volume to be deleted.

The maximum sharing for a block is 255. For example, if there are 500 duplicate blocks, deduplication reduces those blocks to only two blocks. Also note that this ability to share blocks is separate from the ability to keep 255 snapshot copies for a volume.

7.3.2 Deduplication metadata

The core enabling technology of deduplication is *fingerprints*. These are unique digital signatures for every 4 KB data block in the flexible volume.

When deduplication runs for the first time on a flexible volume with existing data, it scans the blocks in the flexible volume and creates a fingerprint database that contains a sorted list of all fingerprints for used blocks in the flexible volume. Although deduplication can provide substantial space savings, a percentage of storage overhead is associated with this and needs to be considered when sizing a FlexVol.

After the fingerprint file is created, fingerprints are checked for duplicates and, when found, a byte-by-byte comparison of the blocks is done to make sure that the blocks are indeed identical. If they are found to be identical, the block's pointer is updated to the already existing data block, and the new (duplicate) data block is released.

Releasing a duplicate data block entails updating the indirect inode pointing to it, incrementing the block reference count for the already existing data block, and freeing the duplicate data block.

In real time, as additional data is written to the deduplicated volume, a fingerprint is created for each new block and written to a change log file. When deduplication is run subsequently, the change log is sorted and its sorted fingerprints are merged with those in the fingerprint file, and then the deduplication processing occurs.

Change logs:

- There are two change log files. As deduplication runs and merges the new blocks from one change log file into the fingerprint file, new data that is being written to the flexible volume causes fingerprints for these new blocks to be written to the second change log file. The roles of the two files are then reversed the next time that deduplication is run. If you are familiar with the Data ONTAP use of NVRAM, this process is analogous to when it switches from one half to the other to take a consistency point.
- When deduplication is run for the first time on an empty flexible volume, it creates the fingerprint file from the change log.

7.3.3 Sizing for performance and space efficiency

This section discusses the N series deduplication best practices and behavior.

Deduplication general best practices

The following list contains deduplication best practices and lessons learned, based on internal research and deployments in the field:

- Deduplication consumes system resources and can alter the data layout on a disk. Because of the application's I/O pattern, and the effect of deduplication on the data layout, the read and write I/O performance can vary considerably. The space savings and the performance impact vary significantly, depending on the application and the data contents.
- The performance impact due to deduplication need to be carefully considered and measured in a test setup, and that you consider sizing before deploying deduplication in performance-sensitive solutions. The following are factors that affect the performance of deduplication.
  - Application and the type of data used
  - The data access pattern, such as sequential versus random access, and the size of the pattern of the input and output
  - The amount of deduplicate data: the amount of total and average file size
  - The nature of the data layout in the volume
  - The amount of changed data between deduplication operations
  - The number of concurrent deduplication operations
  - Hardware platform including System Memory and CPU module
  - Load on the system
  - Disk type, such as ATA/FC /SAS and RPM of the disk
- If the amount of new data is small, run deduplication infrequently, because there is no benefit to running it frequently in such a case, and it consumes CPU resources. How often you run it depends on the rate of change of the data in the flexible volume.

  The more concurrent deduplication processes you are running, the more system resources that are consumed.

Given this information, the best option is to perform one of the following actions:

- Use the auto mode so that deduplication runs only when significant additional data has been written to each particular flexible volume (this approach tends to naturally spread out when deduplication runs).
- Stagger the deduplication schedule for the flexible volumes so that it runs on alternative days.
- Run deduplication manually.
- If snapshot copies are required, run deduplication before creating the snapshot to minimize the amount of data before the data gets locked in to the copies. Make sure that deduplication has completed before creating the copy. Creating a snapshot on a flexible volume before deduplication has a chance to run and complete on that flexible volume can result in lower space savings.
- If snapshot copies are to be used, the snapshot reserve is expected to be greater than zero (0). An exception to this might be in an FCP or iSCSI LUN scenario, where it is often set to zero for thin-provisioning reasons.
- For deduplication to run properly, you have to leave at least one percent of free space for the deduplication metadata.

Deduplication is tightly integrated with Data ONTAP and the WAFL file structure. Because of this integration, deduplication is performed with extreme efficiency. Complex hashing algorithms and lookup tables are not required. Instead, Deduplication is able to use N series storage systems with the NearStore® option internal characteristics to create and compare digital fingerprints, redirect data pointers, and free up redundant data areas, as shown in Figure 7-15.

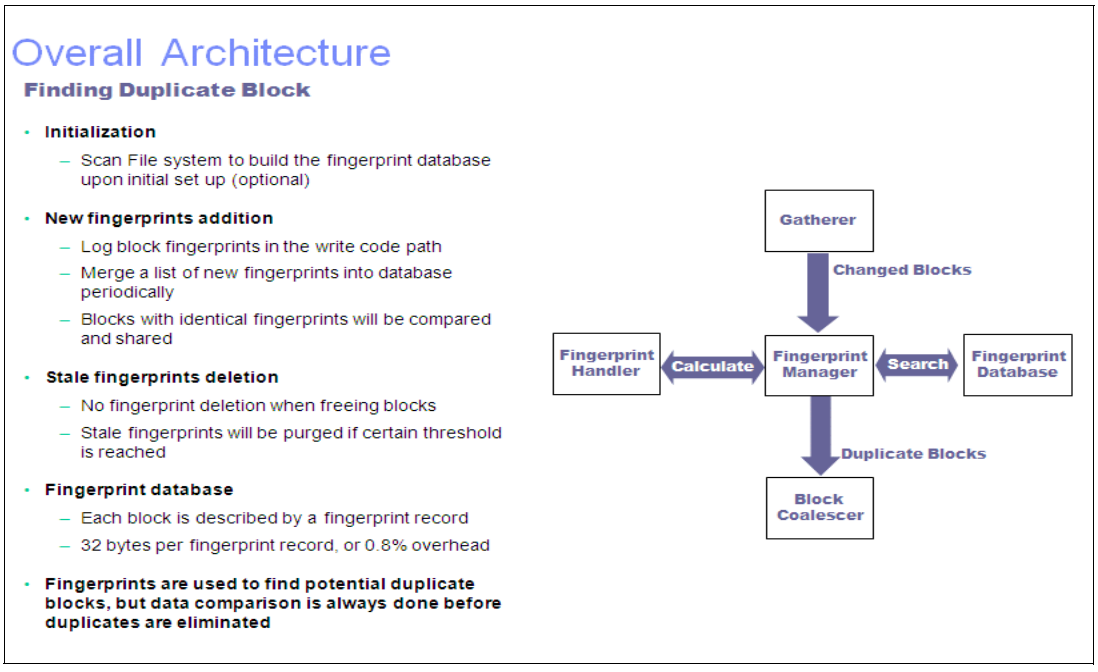

Figure 7-15 Digital fingerprinting

Compressing and deduplicating

It has been shown that combined with other IBM products like ProtecTIER for compressing data, deduplicating offers the greatest space savings. IBM also has a complement to N series Deduplication with IBM Real-time Compression Appliances.

Real-Time Compression is designed to sit transparently in front of your primary storage and reduce the size of every file you create up to 15x depending upon file type. Applications have random, read-write access to compressed data, and the physical capacity required to store a file, or copies and permutations of a file are significantly reduced throughout the entire life cycle including backup. Because less data is written to disk, overall network and storage performance and utilization can also be significantly enhanced.

7.4 IBM Tivoli Storage Manager deduplication

Another option for reducing data storage needs is the native storage pool deduplication available with Tivoli Storage Manager version 6.2 It offers deduplication of backup and archive data.

Native deduplication helps clients store more backup data on the same disk capacity, thereby enabling additional recovery points without incurring additional hardware costs. Tivoli Storage Manager deduplication is especially applicable in smaller environments where scalability is not a primary requirement or where an additional deduplication appliance is not economically affordable. It can be also used in larger environments if appropriate processor, memory, and I/O resources are available on the source servers.

Tivoli Storage Manager version 6.2 and higher offer two types of data deduplication:

- *Server-side deduplication:* Internal Tivoli Storage Manager server processing of data that is already committed by backup clients to a disk storage pool within the Tivoli Storage Manager server storage hierarchy.
- *Client-side deduplication:* The Tivoli Storage Manager server and the backup and archive client work together to identify and store duplicate data in the Tivoli Storage Manager only one time.

Tip: Tivoli Storage Manager version 6.1 does not offer a client-side data deduplication feature. Version 6.2 or higher is required.

7.4.1 Conceptual overview of deduplication in Tivoli Storage Manager

In Tivoli Storage Manager, data deduplication is a method of eliminating redundant data in sequential-access disk primary, copy, and active-data storage pools (FILE type device class). One unique instance of the data is retained on storage media, and redundant data is replaced with a pointer to the unique data copy. With data deduplication, you can save storage space and reduce the overall amount of time required to retrieve data by letting you store more data on disk, rather than on tape.

In order to process data quickly, many storage techniques, including data deduplication, use *hash* functions. A hash function is a process that reads a certain amount of input data (also referred to as a *chunk*), and returns a value that can then be used as a way to refer to that data.

A typical method of deduplication is to logically separate the data in a store into manageable chunks, then produce a hash value for each chunk, and store those hash values in a table. When new data is taken in (ingested) into the store, the table is then compared with the hash value of each new chunk coming in, and where there's a match, only a small pointer to the first copy of the chunk is stored instead of the new data itself.

Typical chunk sizes can be anywhere from 2 KB to 4 MB. The average chunk size is 256 KB. There is a trade-off to be made with chunk size, however, as a smaller chunk size means a larger hash table. If we use a chunk size that is too small, the size of the table of hash pointers will be large, and can outweigh the space saved by deduplication. A larger chunk size means that, to gain savings, the data must have larger sections of repeating patterns, so that, although the hash-pointer table will be small, running deduplication will find fewer matches.

The hashes used in Tivoli Storage Manager deduplication are similar to those used for security products. As examples, Message-Digest algorithm 5 (MD5) and Secure Hash Algorithm-1 (SHA-1) are both commonly used cryptographic hash algorithms, and both are used in deduplication products, along with other more specialist customized algorithms (see "Types of data deduplication and HyperFactor" on page 103).

Before Tivoli Storage Manager chunks the data at a bit file object level, it calculates an MD5 of all the objects in question, which are then sliced into chunks. Each chunk has an SHA-1 hash associated with it, which is used for the deduplication. The MD5s are there to verify that the objects submitted to the deduplication system are reformed correctly, because the MD5 is recalculated and compared with the saved one to ensure the returned data is correct.

7.4.2 Server-side deduplication

Server-side data deduplication in Tivoli Storage Manager is a two-phase process. In the first phase, duplicate data is identified. During the second phase, duplicate data is removed by certain server processes, such as reclamation of storage-pool volumes.

You can deduplicate backup and archive data, Tivoli Data Protection data, whole files and files that are members of an aggregate. You can deduplicate data that has already been stored, or configure client-side data deduplication on slow networks. No additional backup, archive, or migration is required. You can deduplicate any type of data except encrypted data. The process of deduplication is strictly a Tivoli Storage Manager server internal process and does not include any client processing. Therefore, deduplication does not impact system resources of the source devices.

The server-side deduplication can be configured either as:

- *Inline:* Incoming client backup data are immediately processed as they are being stored in the ITSM server storage hierarchy. This option is eligible for smaller and midsize backup environments, where the deduplication process does not significantly impact the backup data commitment to the server storage of Tivoli Storage Manager.
- *Post-processing:* Data is stored in the Tivoli Storage Manager server storage hierarchy as non-deduplicated and in the original structure. Deduplication occurs internally as a scheduled task outside of the backup window. This approach is usually taken in larger backup environments, where a higher number of simultaneous backup sessions can affect the resources of the IBM Tivoli Storage System server. Therefore, the inline deduplication process can significantly impact performance and delay backup jobs.

The conceptual model of server-side data deduplication is explained in Figure 7-16. Data is deduplication internally within the IBM Tivoli Storage System server storage hierarchy without any client involvement or impact. Clients of Tivoli Storage Manager generally do not recognize if the data has been deduplicated or not.

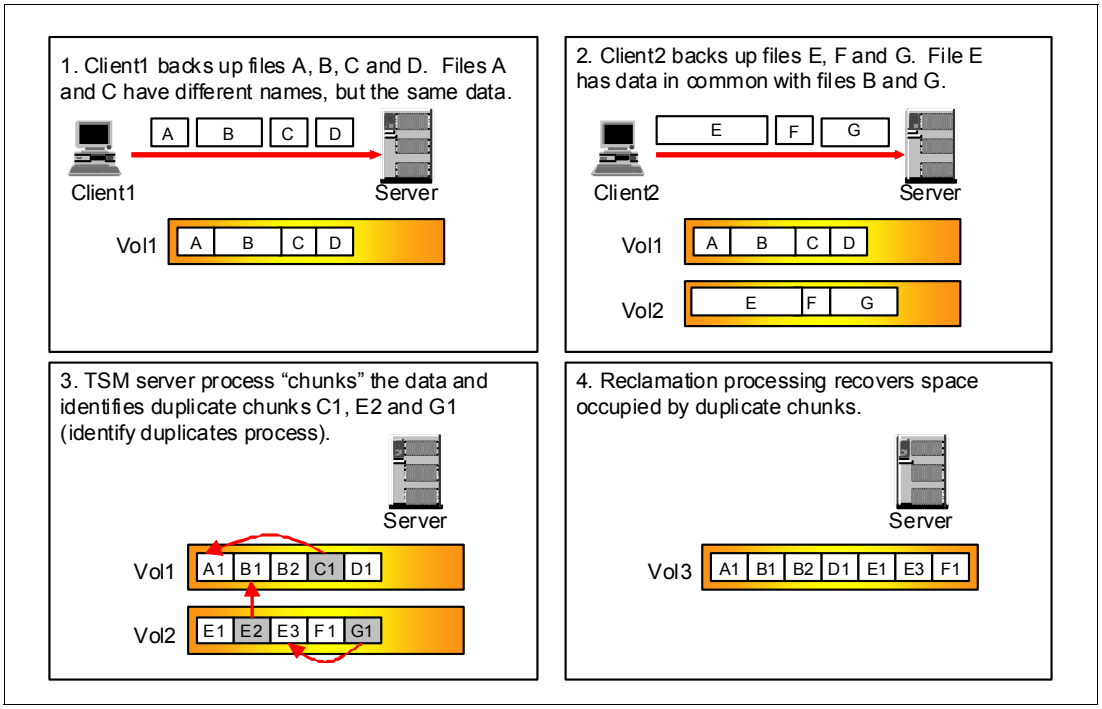

Figure 7-16 Server-side data deduplication with the IBM Tivoli Storage Manager server

7.4.3 Client-side deduplication

In Tivoli Storage Manager version 6.1, only the server can identify and remove redundant data. Starting with version 6.2, you have the option of identifying and removing redundant data during backup and archive processing, before the data is sent to the server over the network. This method of data deduplication is called *client-side data deduplication*. It is available with version 6.2 backup and archive clients and the version 6.2 Tivoli Storage Manager application program interface (API).

Tip: LAN-free backups of large database systems configured on a SAN do not benefit from deduplication, as the target storage repository in Tivoli Storage Manager server is always tape (either physical or virtual).

Tivoli Storage Manager backup and archive client versions earlier than version 6.2 can restore deduplicated, compressed data. That means they can access existing deduplicated data and storage pools that are already set up for data deduplication.

When restoring or retrieving files, the client node queries for and displays files as it normally does. If a user selects a file that exists in a deduplicated storage pool, the server manages the work of reconstructing the file.

Client-side data deduplication uses the following process:

1. The client creates *extents*, which are parts of files that are compared with other file extents to identify duplicates.
2. The client and server work together to identify duplicate extents.

3. The client sends non-duplicate extents to the server.
4. Subsequent client data-deduplication operations create new extents. Some or all of those extents might match the extents that were created in previous data-deduplication operations and sent to the server. Matching extents are not sent to the server again.

Figure 7-17 shows how client-side deduplication works with the Tivoli Storage Manager server.

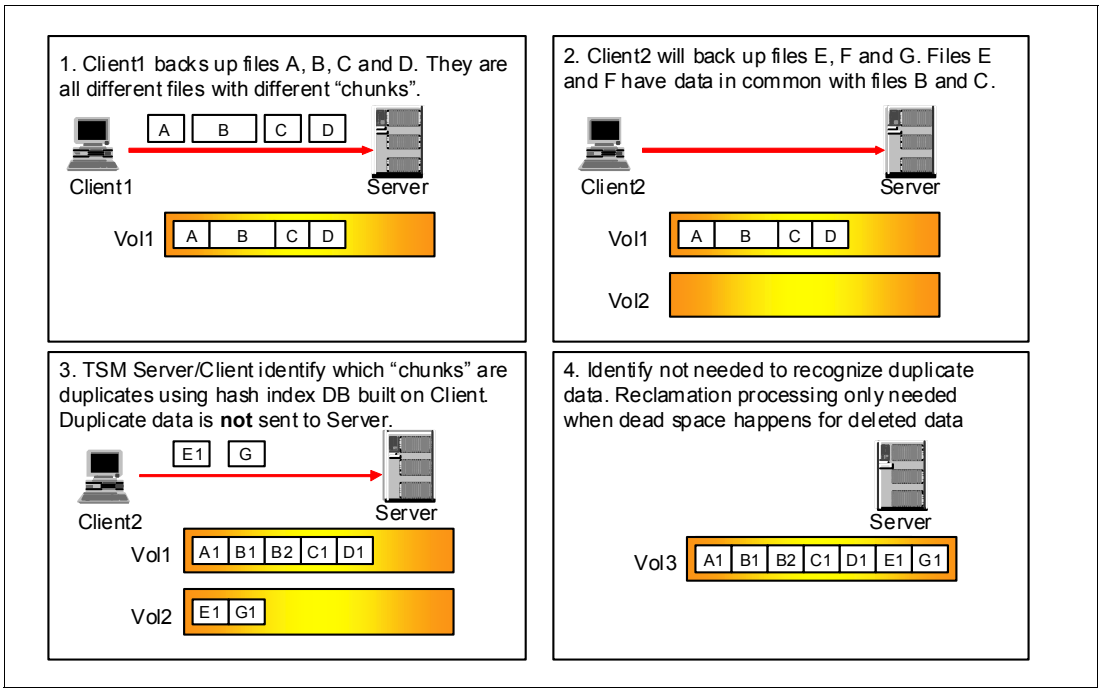

Figure 7-17 Client-side data deduplication with IBM Tivoli Storage Manager server

Client-side data deduplication offers additional features than those of the Tivoli Storage Manager server-side processing. Specifically, you can perform these tasks:

- Exclude specific files on a client from data deduplication.
- Enable a data deduplication cache that reduces network traffic between the client and the server.
- The cache contains extents that were sent to the server in previous incremental backup operations. Instead of querying the server for an extent, the client queries its cache.
- Specify a size and location for a client cache. If an inconsistency between the server and the local cache is detected, the local cache is removed and repopulated.
- Enable both client-side data deduplication and compression to reduce the amount of data that is stored by the server. Each extent is compressed before being sent to the server. Note that these two options, working together, require more system resources (memory and processor) during your backup window.

Attention: Tivoli Storage Manager server and backup-archive client cannot deduplicate encrypted files. Those files are automatically excluded from processing, sent to the Tivoli Storage Manager server in separate transactions, and a warning message logged.

Tivoli Storage Manager does not deduplicate files smaller than 2 KB. They are backed up, but not deduplicated.

7.4.4 Deduplication considerations

This section briefly summarizes scenarios during which backup environments can benefit from IBM Tivoli Storage Manage deduplication, and when they cannot. Both options, server-side and client-site processing, are considered.

When to enable deduplication

These are reasons to consider deploying Tivoli Storage Manager data deduplication:

- You have a large disk storage pool within your Tivoli Storage Manager server backup repository, and you want to deduplicate currently existing data.
- You have enough processor resources and disk I/O available. With the Tivoli Storage Manager new expiration algorithm, your server might have processor or network bandwidth available in contrast with Tivoli Storage Manager version 5.5.
- You cannot consider alternatives for cost, support, power, or other reasons.
- Your backup environment is too small to benefit from an alternative deduplication solution, such as IBM ProtecTIER Gateways.
- The data will reside in a disk storage pool for a long amount of time.

When deduplication is not an option

These are reasons for not considering the deployment of Tivoli Storage Manager for deduplication:

- Your Tivoli Storage Manager server is already processor, memory, or disk I/O constrained.
- You have too many Tivoli Storage Manager servers, or with large amounts of data that needs deduplication (over 6 TB). Here, we suggest another deduplication solution, the IBM ProtecTIER Gateway. Apart from scaling to large sizes (1 PB of non-deduplicated data), IBM ProtecTIER also opens the door for deduplication between single Tivoli Storage Manager servers, perform inline deduplication and IP-based replication, and so on.
- You use the disk storage pool only as temporary storage and will migrate all of the data to tape later.

7.4.5 More information

For additional technical details and descriptions of best practices, see:

- *Implementing IBM Storage Data Deduplication Solutions*, SG24-7888
- The IBM Tivoli Storage Manager 6.2 Information center at this website:

  `http://publib.boulder.ibm.com/infocenter/tsminfo/v6r2/index.jsp`

8

Visibility

Previous chapters of this book presented the benefits for organizations to adopt smarter storage technologies, such as virtualization, deduplication, thin provisioning, and easy tiering, which provide more efficient use of storage and server resources. These new technologies can however lead to a more complex IT infrastructure if they are allowed to grow unchecked.

As such, it is paramount that this new virtualized infrastructure be managed by new generation management tools, because older tools generally lack the required features.

When used properly, these tools can make the adoption of virtualization technology easier and more cost effective. The tools have the following additional benefits:

- Enable line of business insight into storage utilization and allocation, enabling easier departmental charge back.
- Allow for more intelligent business decisions about storage efficiency, enabling you to respond faster to changing business demands, yet reduce costs.
- Provide better understanding of application storage performance and utilization patterns enabling better Information Lifecycle Management (ILM) of application data.
- Allow an organization to perform infrastructure management proactively, through proper capacity management, rather than reactively.
- Improve operational efficiency leading to cost savings for the organization.

This chapter describes and highlight IBM products designed to provide in depth end-to-end visibility and management of complex physical and virtualized IT infrastructure. We provide an overview and introduction for these topics:

- IBM Systems Director
- IBM Tivoli Storage Productivity Center
- IBM Storage Enterprise Resource Planner (SERP)
- IBM Storwize Rapid Application Storage

8.1 IBM Systems Director

IBM Systems Director is the management platform foundation for achieving Smarter Systems. An integral component of the IBM Smarter Systems portfolio, IBM Systems Director enables integration with Tivoli, and third party management platforms, providing the building block for virtualization and integrated services management. IBM Systems Director provides systems management personnel with a single-point-of-control, helping reduce IT management complexity and cost. Key benefits include these:

- *Simplified usability and reduced complexity:* IBM Systems Director eliminates layers of management duplication using a single, unified view of the total IT environment, reducing system management complexity, cost, and time to problem resolution, while improving visibility, availability, and service.
- *Integration:* With Tivoli, and third party management platforms, IBM Systems Director provides the foundation for virtualization and integrated services management. With an extensive suite of available platform management tasks, and automated tools, IBM Systems Director assists systems management personnel in increasing productivity, resulting in improved responsiveness and service.

Instead of using many disparate tools for each platform, IBM Systems Director provides a single user interface (UI), and through its plug-in design provides a seamless end-to-end integrated systems management platform.

IBM Systems Director is available as Express, Standard, and Enterprise Editions to cater to various client requirements and can be installed on multiple operating systems (for example, AIX, IBM i, Windows, Linux on Power, Linux on x86, and Linux on System z).

Table 8-1 provides a high level summary of the functions for each of the editions.

Table 8-1 System Director functions

Functions	Express	Standard	Enterprise
Virtualize physical/virtual system relationships	Yes	Yes	Yes
Monitor system health	Yes	Yes	Yes
Provide threshold and error alerts	Yes	Yes	Yes
Update operating systems and firmware	Yes	Yes	Yes
Simplify deployment with virtual images		Yes	Yes
Control energy use with existing capacity		Yes	Yes
Monitor network system health with servers and storage		Yes	Yes
Automate configuration and placement for new workloads			Yes
Manage workload availability end-to-end			Yes
Understand capacity			Yes
Analyze and report historical performance			Yes

Note that BM Systems Director Enterprise Edition contains the full version and benefits of IBM Tivoli Monitoring (ITM) and IBM Tivoli Application Dependency Discovery Manager (TADDM).

Additional details about IBM Tivoli Monitoring and IBM Tivoli Application Dependency Discovery Manager are available at these websites:

- http://www-01.ibm.com/software/tivoli/products/monitor/
- http://www-01.ibm.com/software/tivoli/products/taddm/

Figure 8-1 illustrates the IBM Systems Director product topology. System Director includes several base managers, implemented as plug-ins, that provide an integrated set of consistent tasks and features to manage discovered resources. Additional plug-ins can be downloaded and installed through purchase of separate license.

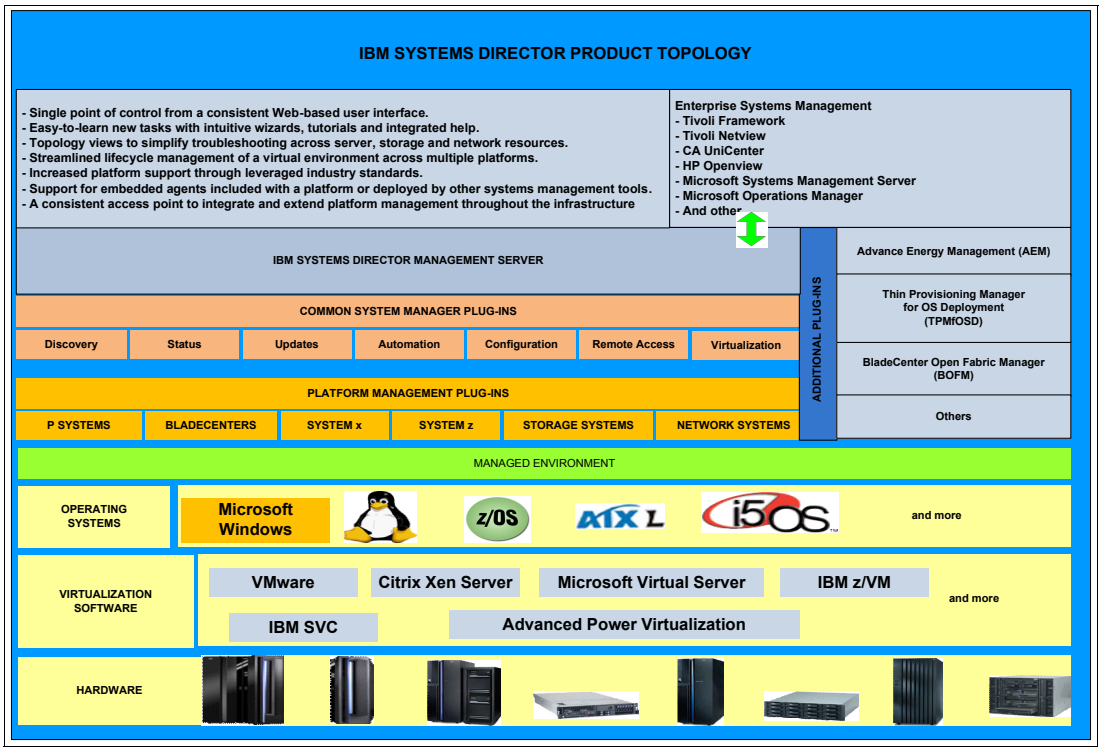

Figure 8-1 Systems Director Product Topology

The next few sections provide an overview, features, and benefits of the plug-in modules.

8.1.1 IBM Systems Director plug-ins

Many plug-ins are available to augment the Systems Director base functionality. Examples of these plug-ins are network-, storage-, and virtualization-specific. Many of these plug-ins complement each other, enabling the integrated management of physical and virtual IT resources with Systems Director.

IBM Systems Director Active Energy Manager (AEM)

The first step in making a data center run more efficiently is to understand the power and cooling characteristics of the individual pieces of equipment using real-time monitoring.

The IBM Systems Director Active Energy Manager™ (AEM) plug-in allows you to perform real-time monitoring of much of your hardware in a data center, allowing you to proactively manage energy costs.

Benefits

IBM Systems Director AEM increases energy efficiency and reduces costs in these ways:

- Providing a single view across multiple platforms including - IBM servers and storage, non-IBM systems, facility providers, facility management applications, PDUs, and equipment supporting the IPv6 protocol – reducing energy management tools and complexity
- Simplifying energy management for IBM servers and storage, providing a single, cross-platform view of energy consumption, to aid power consumption understanding, decision making, and actions
- Controlling power, and limiting power consumption, with tools providing power capping, power savings, power and thermal trending, and group capping
- Providing real-time information about energy use, giving administrators insight needed to reduce power consumption
- Integrating with a wide range of facility provider platforms for monitoring power and cooling infrastructure, to understand impact on servers, storage and network equipment
- Integrating seamlessly with IBM Systems Director management platform, and with higher-level management platforms as Tivoli, to provide centralized, unified management capabilities

Features

The following IBM Systems Director AEM features are provided:

- Monitoring functions include power trending, thermal trending, IBM and non-IBM PDU support, support for facility providers, and energy and environment thresholds
- Management functions include power capping and power savings modes
- Enhanced facility provider support to enable a more complete view of energy consumption in the data center
- Monitoring of energy and environment data from a wide variety of PDUs, meters, and sensors
- Monitoring of power and cooling using integration with a wide variety of infrastructure vendors
- Monitoring of equipment using an IPv6 connection
- Creating relationships from systems to cooling and power units in data centers so power and cooling topologies can now be viewed

Resources managed by AEM are defined as endpoints. AEM communicates with each endpoint to obtain the power consumption or thermal state details and to set power management features. Upon successful discovery, these endpoints are populated into a map showing the relationships between them. Figure 8-2 provides an example of such a map.

For additional details about AEM managed resources, see the information available at this website:

http://publib.boulder.ibm.com/infocenter/director/v6r1x/topic/aem_410/frb0_r_HW_reqs_managed_systems.html

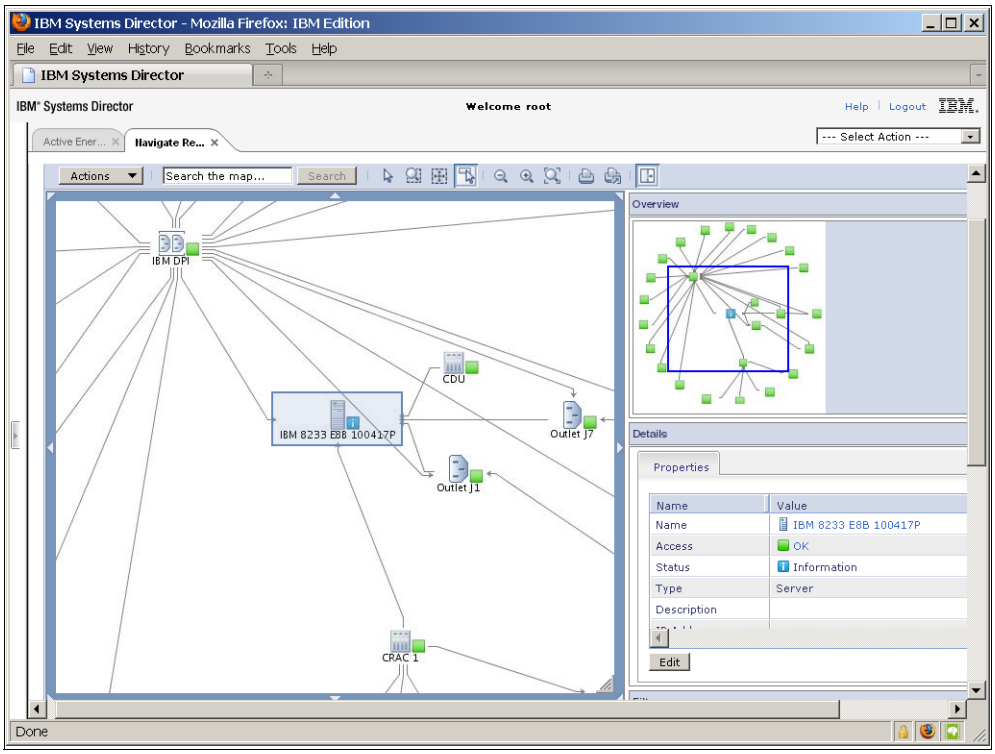

Figure 8-2 Graphical diagram of AEM endpoints within Systems Director

Associated with the discovery process are the following tasks:

- Grouping of AEM resources, such as by endpoints operating system types, functions, locations of endpoints
- Configuring of AEM settings, such as how often endpoints will be polled, how long to keep polled data, information related to calculation of energy costs using the Energy Cost Calculator (part of the plug-in)

Figure 8-3 provides an example of global settings that are required as part of the Energy Cost Calculation. Resource setting can also be performed at an individual endpoint level if required.

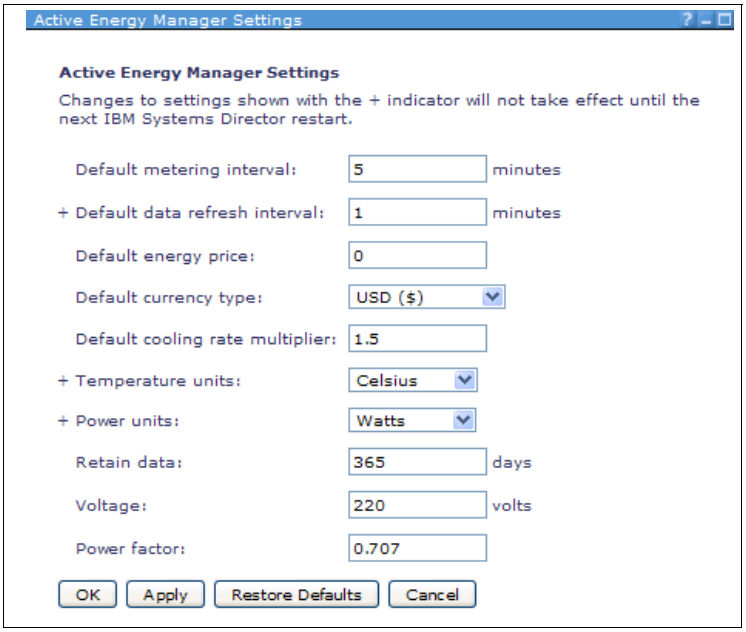

Figure 8-3 AEM global energy settings

- Collecting inventory: After a resource is discovered and access is granted to it, you can collect hardware and software inventory data relating to the resource.
- After data has been collected for a period of time, trending analysis can be performed. AEM collects the following type of trend information:
  - Power consumption
  - Power cap
  - Thermal signature
  - Humidity and dew point
  - CPU speed

Figure 8-4 provides a sample out of power use trend for a specific resource group named called “Power Domain 1”.

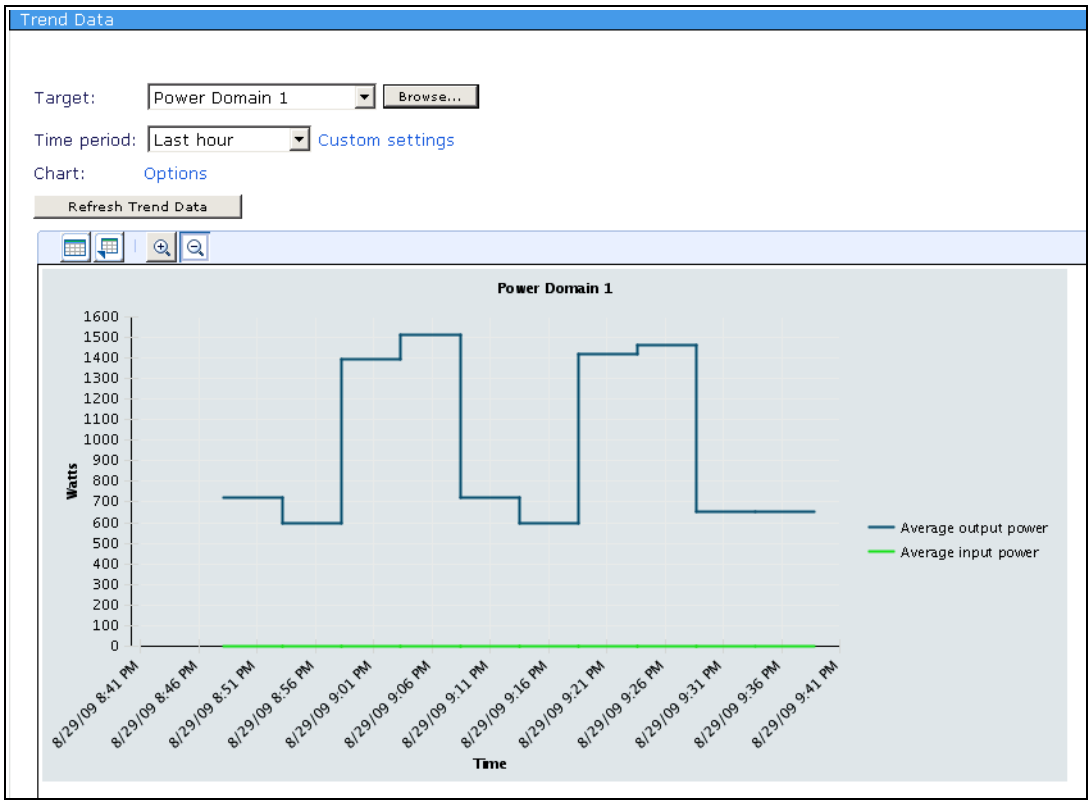

Figure 8-4 Power use output for the “Power Domain 1” resource group

In addition, data can also be saved as bitmap and .CSV files for further analysis if required.

AEM also features an *Energy Cost Calculator* that provides a simple but effective way to calculate the total cost of power consumed by an Active Energy Manager resource or group of resources over a specified time period.

You can use the energy cost calculator on many resource types, including these:

- IBM BladeCenter® chassis
- BladeCenter power domains
- IBM Power Systems™
- IBM System x®
- System z
- Metered devices
  - This is the resource associated with a metering device.
  - A metering device monitors values, such as power use, temperature, humidity, and dew point of other resources).
- Intelligent PDUs (these are also called PDU+)
- Uninterruptible power supplies (UPS)

Figure 8-5 provides an example of the Energy Cost Calculator output.

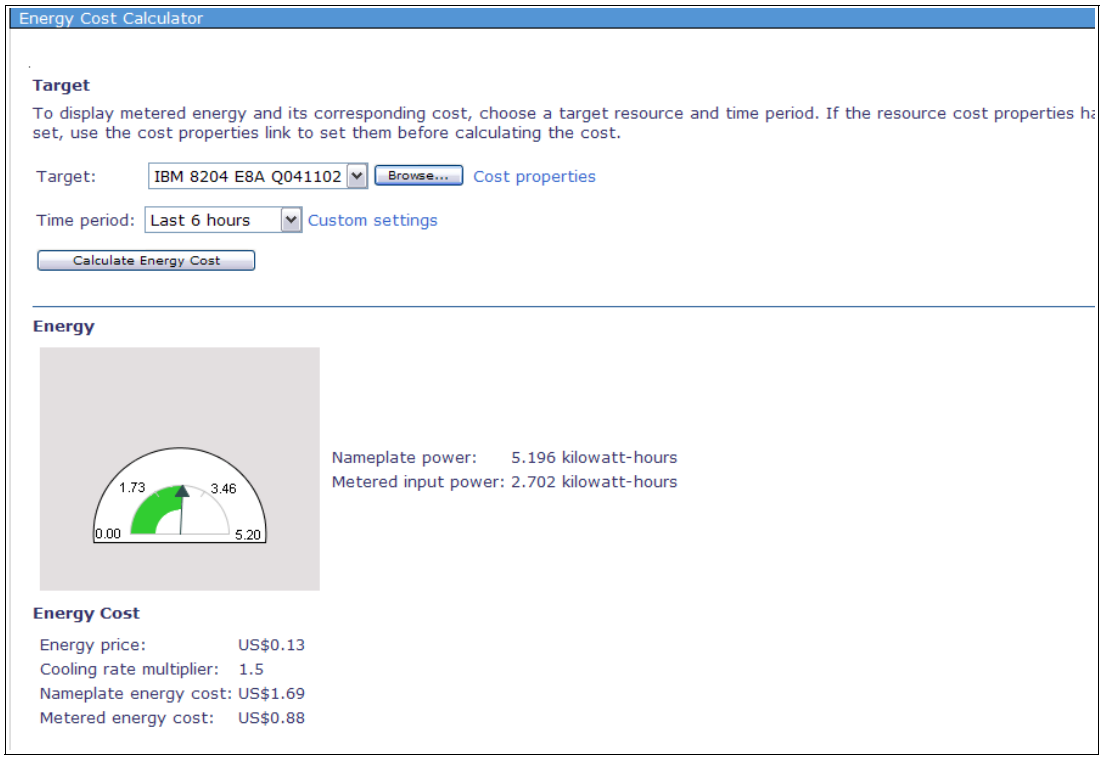

Figure 8-5 Energy Cost Calculator - metered energy cost page

Another feature of AEM is a rich set of energy monitors. Figure 8-6 shows a few of the event monitors that are available.

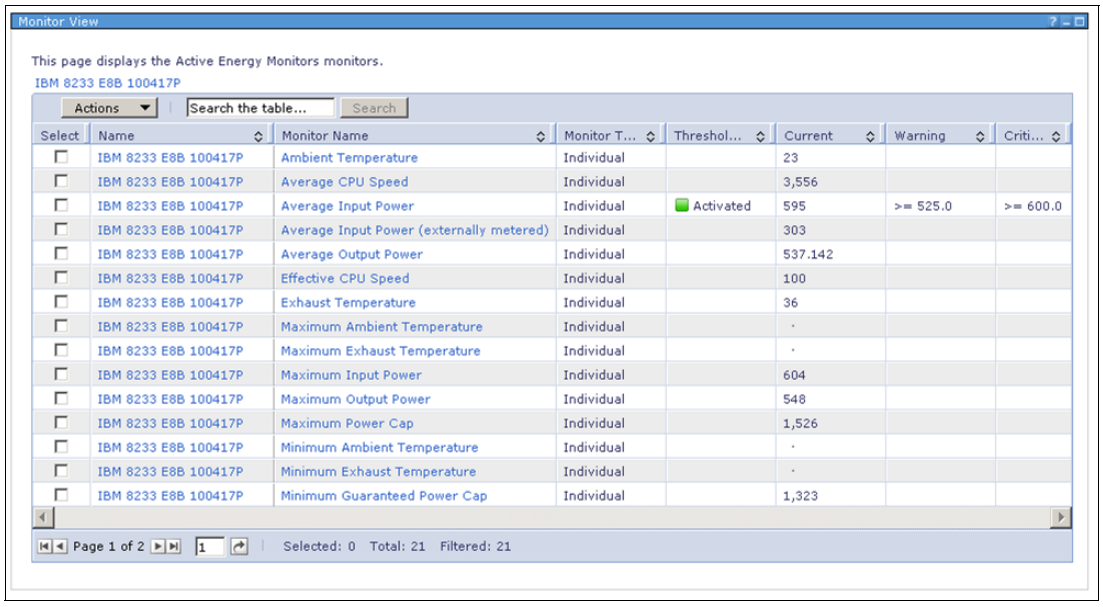

Figure 8-6 Energy Monitors

For each given monitor, a threshold can be set. IBM Systems Director also provides tools for automation plans to be activated when a threshold has been reached. For example, when the exhaust temperature of a server rises above 70 °C, a threshold can be triggered, a critical event created, and an automation plan invoked that causes Active Energy Manager to throttle the CPU back to reduce the exhaust temperature.

Figure 8-7 illustrates an example of events.

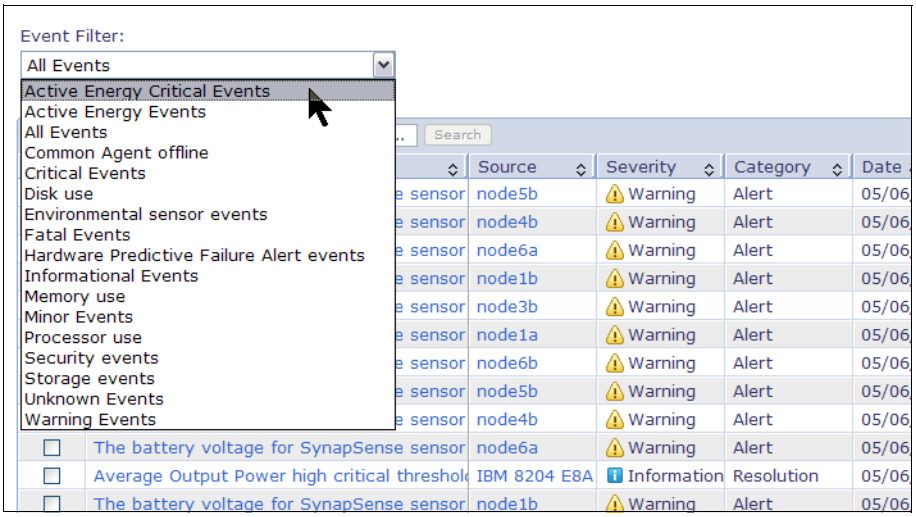

Figure 8-7 AEM events

For additional details about EAM, see the information available at this website:

http://www-03.ibm.com/systems/software/director/aem/index.html

IBM Systems Director Network Control

This plug-in provides integration of server, storage, and network management for virtualization across multiple platforms. Leveraging existing IBM Systems Director functionality and the integration of vendor applications, Network Control can manage, monitor, and configure vendor-based network devices. Here are a few of the technical reasons for implementing IBM Systems Director Network Control:

- Virtualization has resulted in the need to frequently and automatically configure network and storage for virtual systems.
- As environments move to Fibre Channel Over Ethernet (FCoE, a convergence of Ethernet and SAN traffic), storage and network management domain will become blurred. A single tool for both storage and network management can result in easier management.
- Automatic workload redeployment in a virtual pooled environment implies the need to optimize network to meet new workload requirements as well.

Benefits

The following benefits are realized:

- Reduces network management costs by consolidating management functions into a single product
- Improves productivity and service levels by simplifying and streamlining management tasks
- Increases asset utilization through more efficient management
- Provides higher bandwidth and converged fabrics eliminating server, storage and network sprawl

Features

The following features are provided:

- Basic life cycle management of network switches
- Integration of vendor-based device management tools
- Converged Ethernet network device support using native support and vendor tools
- Ability to view systems according to VLAN
- Network device topology collection and visualization (virtual and physical)
- IBM Systems Director Network Control supports these and other network devices:
  - IBM BladeCenter supported switches and Blade NICs
  - Brocade Converged Ethernet network switches, Ethernet switches and routers
  - QLogic converged network adapters
  - Cisco FCoCEE and Ethernet switches
  - Juniper Ethernet switches, routers and appliances
- Virtual switch discovery and inventory for:
  - VMware ESXi and ESX4
  - IBM PowerVM™
  - RHEV-H

As part of the Systems Director base, a network management module is included. This layer focuses on device management by providing component discovery, inventory, monitoring, and alerting, in addition to discovery and inventory of network adapters on all supported servers. Additional functions include these:

- Viewing high-level status of supported devices and lower-level drill-down component status and errors for individual devices
- Defining and viewing monitors and thresholds for all supported network devices
- Discovery and inventory of all IBM-branded and select non-IBM branded switches
- Troubleshooting tools, such as ping and trace route, to assist in diagnosis of network problems
- Providing a set of default groups of network resources to allow users to view the status of a group of related devices or to perform tasks across a set of devices, such as:
  - Ethernet Switches
  - Fibre Channel over Converged Ethernet Switches

The remaining figures in this section illustrate a few of the capabilities within Network Management and Network Control.

Figure 8-8 provides a Network Management summary view containing status of network devices, items managed and configuration/automation section for further configuration of your network devices.

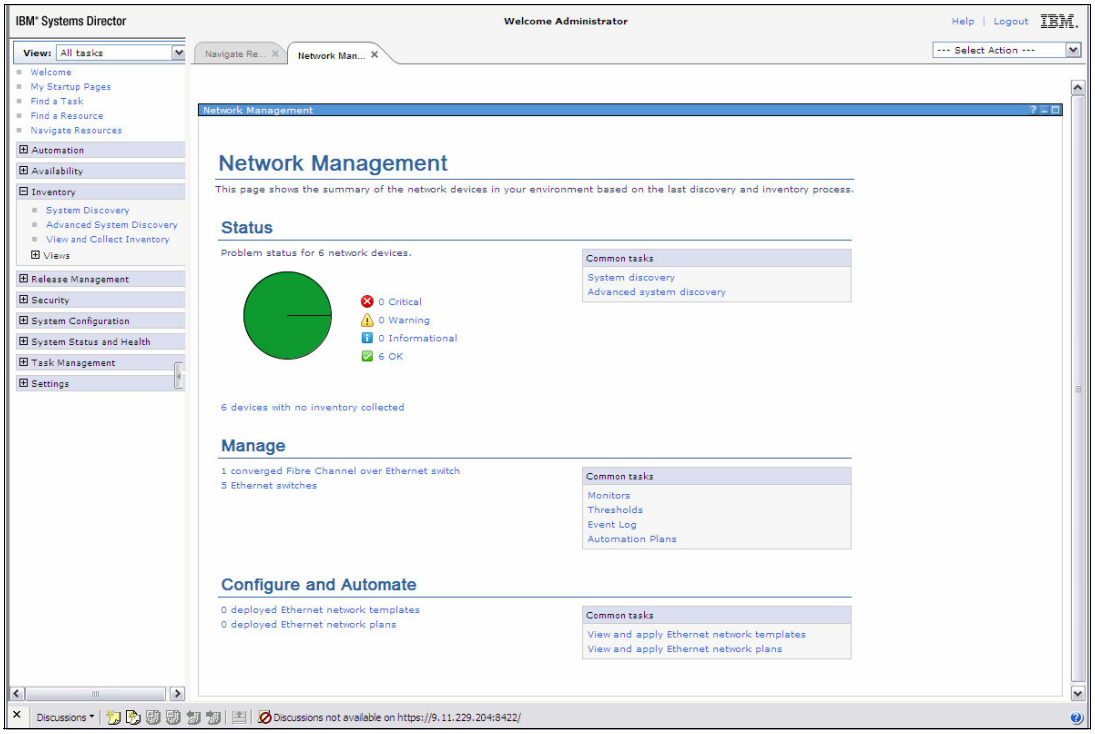

Figure 8-8 Network Management summary page

Network Control provides additional functions to simplify and automate virtualization tasks:

- Integration of Tivoli based technology for advanced discovery and topology
- Graphical view of Layer 2 network connectivity using Director's topology perspective
- Network monitoring at a glance using network topology perspectives with ability to see the components affected by network outages
- Launch-in-Context and Single-Sign-On capabilities
- Logical views of network systems arranged by the subnetworks to which they belong
- Discovery and integrated views of virtual switches and network adapters
- Automated Migration of Virtual Machine Network Parameters with VMControl

In addition to network devices discovered by SNMP, VMControl provides advance network discovery topology through integration with IBM Tivoli Network Management (ITNM). Network objects previously discovered by ITNM are imported into Network Control.

Network topology can be displayed from three perspectives:

- *Port-level:* This shows details of which ports on the servers are connected to which ports on the switches. Also, a network view can be seen from the server or switch perspective.
- *System-level:* This omits port information and provides a simplified view of which servers and switches are connected to one another.
- *Subnet:* This shows all of the devices on an IP subnet.

Figure 8-9 provides a sample topology from a port-level view perspective.

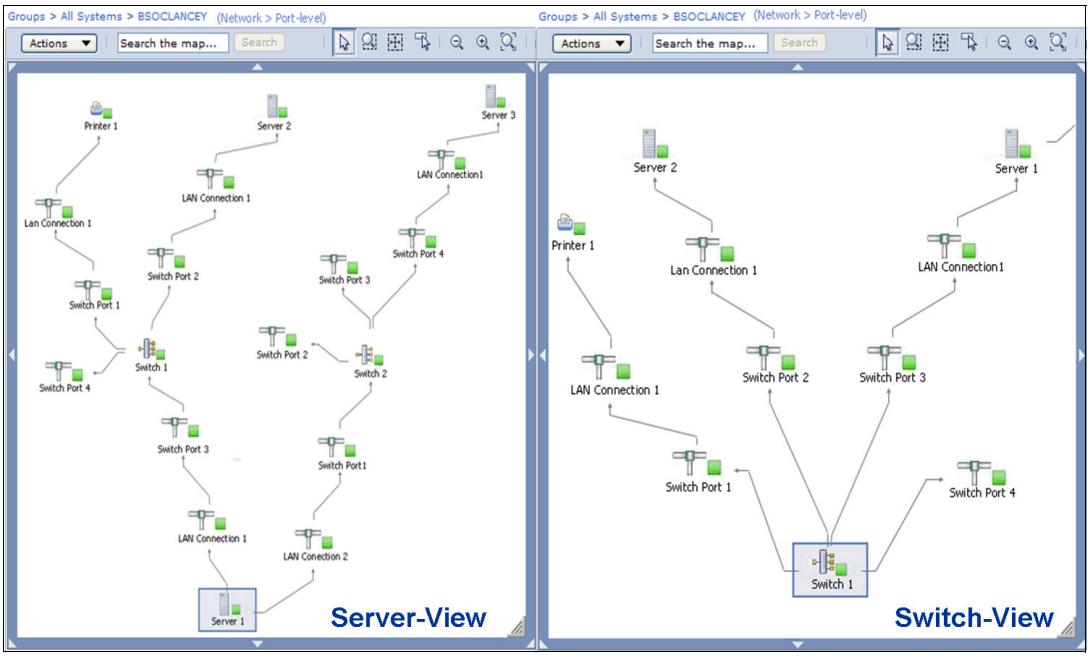

Figure 8-9 Network topology - port -level view

The Network Control plug-in also offers these features:

- *Launch-in-Context:* This is a mechanism by which one application discovers an object which is managed primarily by another application and invokes or launches the application with the context of the object to provide the required management.
- *Single Sign-on capability:* This is a mechanism by which two application share the same authentication so that when one application invokes the other application the user is not required to re-logon.

An example of such an implementation is when a network switch has been discovered by Network Control. In contrast, by selecting that network switch, you can launch the IBM Data Center Manager (DCFM) and automatically log into the software to perform further configuration tasks, such as FCoE configuration.

Figure 8-10 provides an example of Network Control.

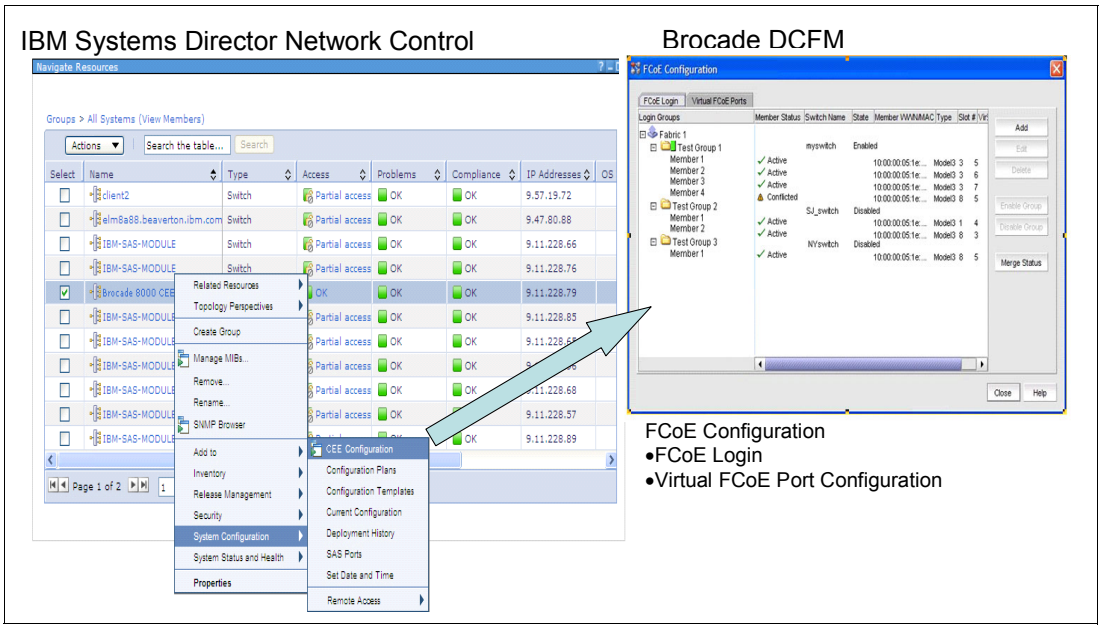

Figure 8-10 Network Control Launch-in-Context and SSO capabilities

Network Control also provides network diagnostics capabilities, such as ping and trace route, through its single interface. For additional details, see the information available at this website:

http://www.ibm.com/systems/software/director/network/

IBM Systems Director VMControl

The IBM Systems Director VMControl™ plug-in helps you gain more from infrastructure wide virtualization, with a deeper level of management. A few of the virtualized environments that are supported include these:

- IBM Power Systems virtualization
- VMware vCenter virtualization
- Windows Server 2008 Hyper-V virtualization
- z/VM virtualization

Traditionally, prior to virtualization, management of resources is done, based on hardware silos with little integration between them.

Adoption of virtualization technologies has introduced new challenges:

- Administrators struggle with several interfaces to manage virtual servers across systems.
- New workloads can be difficult and time-consuming to implement in a virtualized environment.
- Management costs for multiple virtual servers and physical hosts increase, and yet workload resiliency degrades.

The combination of IBM Systems Director and VMControl allows you to reduce the total cost of ownership of your virtualized environment by decreasing management costs, increasing asset utilization, and linking infrastructure performance to business goals.

VMControl is available with IBM Systems Director Editions, or available separately as a plug-in option for IBM Systems Director. There are three editions of VMControl (Express, Standard, and Enterprise) that cater to various requirements. Table 8-2 outlines what is available in each edition.

Benefits

The following benefits are realized:

- Brings together physical and virtual management into a single interface to reduce complexity
- Offers unmatched cross-platform and cross-operating system management, which helps improve service delivery by eliminating isolated silos of virtualization in heterogeneous environments
- Provides faster time-to-value and greater business agility through simplified virtualization management that allows more effective utilization of virtualized resources
- Establishes repeatable accuracy and consistency through automation
- Reduces operational and infrastructure costs through increased efficiency and resource utilization

Features

Table 8-2 summarizes what VMControl features are available with the various IBM Systems Director edition.

Table 8-2 IBM Systems Director VMControl™ features

Features	Express	Standard	Enterprise
Create and manage virtual machines	Yes	Yes	Yes
Virtual machine relocation	Yes	Yes	Yes
Import, edit, create, and delete virtual images		Yes	Yes
Deploy virtual images		Yes	Yes
Maintain virtual images in a repository		Yes	Yes
Manage virtual workloads in system pools			Yes

Figure 8-11 indicates when each VMControl edition can be deployed.

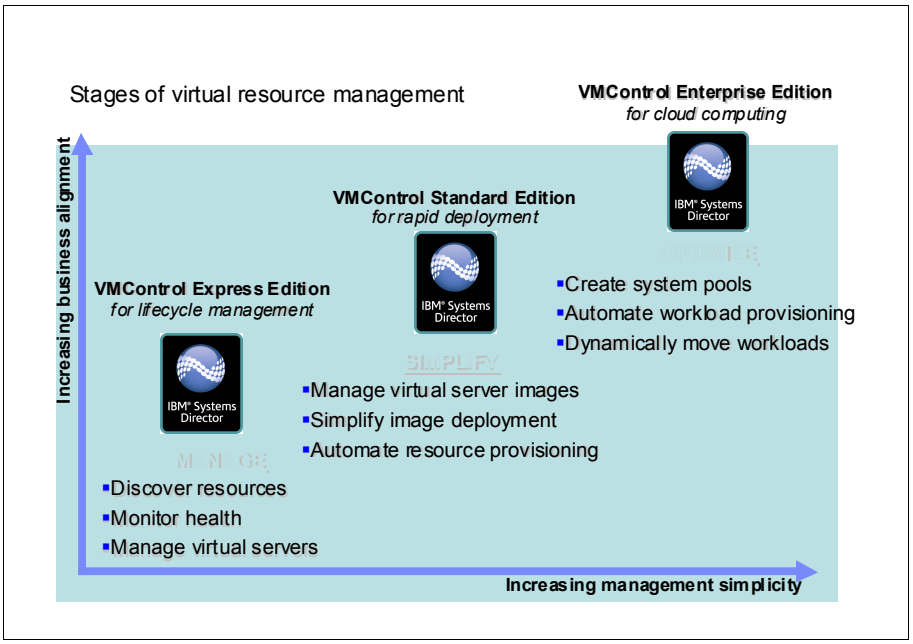

Figure 8-11 VMControl editions

Next, we provide an overview of select VMControl functions on IBM Power Systems.

Virtual server management

After IBM Systems Director has discovered and authenticated the Power Systems server management platform using Hardware Machine Console (HMC) or Integrated Virtualization Manager (IVM), all virtual servers (LPARs) on that system are available for VMControl management.

With the Systems Director interface tasks, such as changing a profile, starting a virtual server, or performing a Dynamic LPAR (DLPAR) operation to change its processor, memory or I/O configuration can easily be performed. If needed, you can interact with the HMC or IVM system entry in Systems Director and launch its native browser interface in a separate window. VMControl can also use Systems Director's topology capability to show the relationships among a physical Power Systems server, its virtual servers, and their assigned resources.

VMControl virtual server management supports all three types of Power Systems virtual servers: IBM i, AIX, and Linux.

For IBM Systems Director VMControl in a Power Systems environment, a virtual server creation is tightly coupled with a virtual appliance creation or use. An empty virtual server can be created without a virtual appliance creation or input, and no workload is generated.

Figure 8-12 shows the attributes required during the creation of a virtual server.

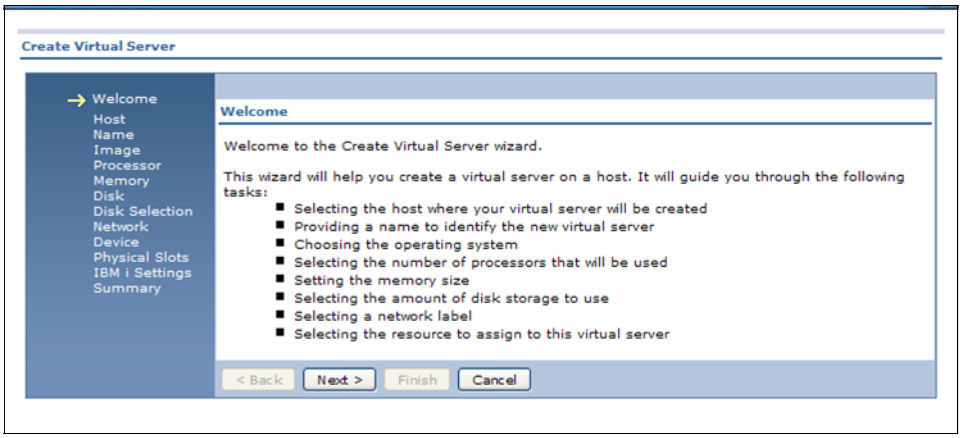

Figure 8-12 Virtual Server attributes

Virtual machine relocation

VMControl systems pools can predict hardware failure problems and relocate virtual servers to maintain resilience. With VMControl, there are three options for relocation:

- Manually relocate virtual servers at any time.
- Activate a resilience policy on the workload to relocate virtual servers automatically to prevent predicted hardware failures from affecting the availability of the workload.
- Create an automation plan to relocate the virtual servers when certain events occur.

Virtual appliance and image management

A virtual appliance is virtual machine image installed within a virtual server.

VMControl supports Open standards virtualization by conforming to Open Virtualization Format (OVF) standard from Distributed Management Task Force (DMTF) organization.

OVF is a packaging standard designed to address the portability and deployment of virtual appliances. OVF enables simplified and error-free deployment of virtual appliances across multiple virtualization platforms.

The VMControl Standard Edition and its Image Manager component address two main challenges when working with virtualization:

- Deploying a new workload in an environment
- Capturing and reusing customized virtual server images

With VMControl Image Manager, you can create a virtual server image repository, capture a running customized virtual server, and store it as a virtual appliance in the repository:

- Capture a virtual server (LPAR or logical partition) on a Power Systems server.
- Capture an existing mksysb image file from the following locations:
  - IBM Systems Director server
  - IBM Systems Director VMControl (Network Install Manager) NIM server
- Capture an existing NIM lpp_source resource (An lpp_source resource represents a directory in which software installation images are stored).

For example, you can capture a virtual server that has the AIX OS and a database application installed and has networking, security, and users configured. When captured, the fully configured virtual server is stored in the image repository and is available for rapid deployment. Figure 8-13 provides a schematic overview.

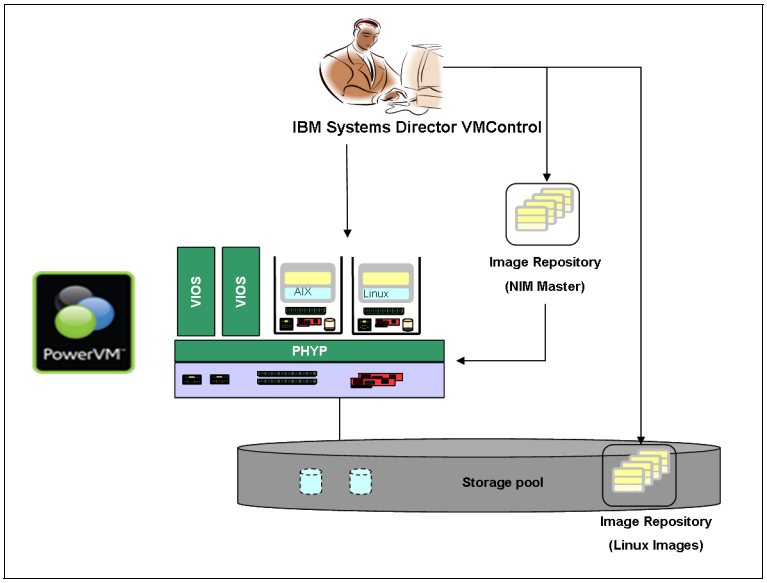

Figure 8-13 Managing PowerVM images within VMControl

When virtual appliances are available in the image repository, deploying a new workload takes minutes instead of days.

For Power Systems servers, VMControl relies on Live Partition Mobility for relocation. VMControl is also capable of managing Linux on z/VM and kernel-based, virtual machine (KVM) environments.

For additional details, see the information available at this website:

`http://www.ibm.com/systems/software/director/vmcontrol/`

IBM Systems Director with Storage Control

This plug-in provides a single management console and system for managing both server and storage systems and is based on IBM Tivoli Storage Productivity Center. It integrates tightly with the IBM Systems Director and VMControl environment providing integrated physical and virtual server and storage management through a single user interface.

The technical reasons for implementing Storage Control are similar to those in Network and Virtual Machine Control. That is, to facilitate the easier adoption of virtualization, it is important that this new breed of management tools provide integration across hardware domains (storage, network, and servers).

Figure 8-15 outline the devices that can be managed by Storage Control plug-in.

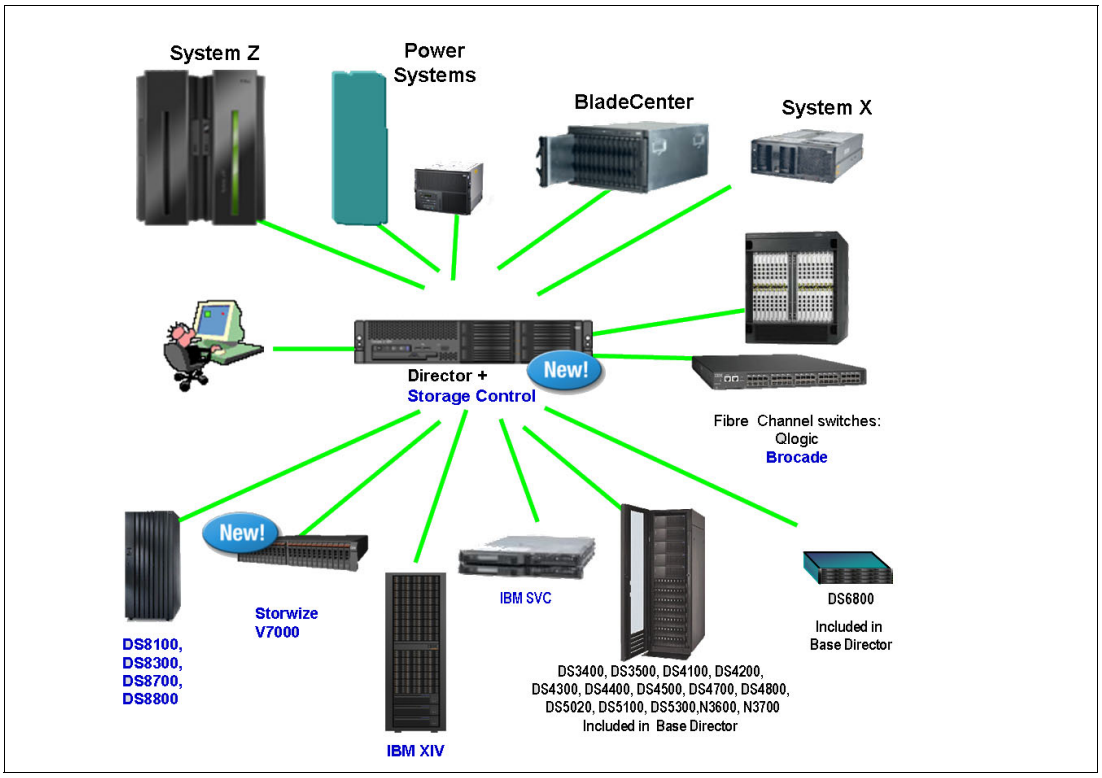

Figure 8-14 IBM Systems Director Storage Control

Benefits

The following benefits are realized:

- Leverage integrated server and storage management to improve service level agreements and responsiveness to business needs
- Reduce complexity and risk of managing storage with the rest of IT operations
- Significantly reduce storage management cost

Features

The following features are provided:

- Extends storage management of IBM Systems Director and VMControl to cover most IBM storage systems
- Storage device discovery and coverage in integrated physical and logical topology views
- Show relationships between storage and server resources
- Ability to configure logical and physical configuration
- Ability to view controller and volume status and to set notification alerts
- Single management server deployment
- Integration with VMControl featuring the following additional functions:
  - Storage provisioning for image creation, deployment, and cloning
  - Ability to manage storage system pool life cycle, take group actions across pool and policy-based storage placement, provisioning, and cloning actions within the pool

Figure 8-16 illustrates a Storage Control summary view that allows the user to view end-to-end capacity management by storage type.

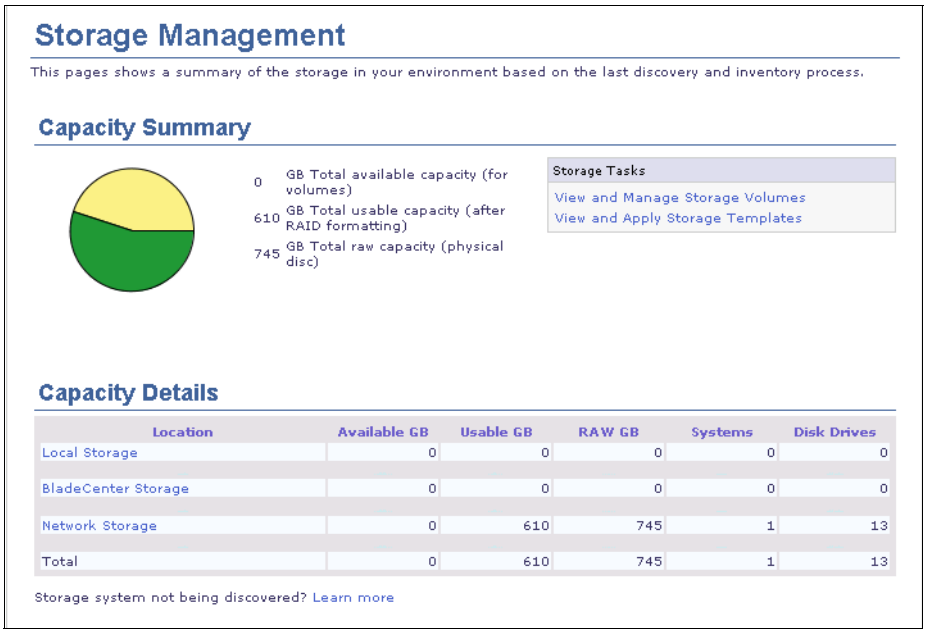

Figure 8-15 Storage summary page view

Through the Systems Director topology views, it allows users to visually identify interconnected virtual machines, servers, network devices, and storage systems. Each device, whether virtual or physical, can be drilled down further to assist users to quickly identify and isolate the root cause of problems at a system component level if required. This drill-down capability can help administrators to spend less time troubleshooting and more time delivering services that add value to the business.

Figure 8-16 provides a drill-down example on the new IBM Storwize V7000 showing details about the storage volumes.

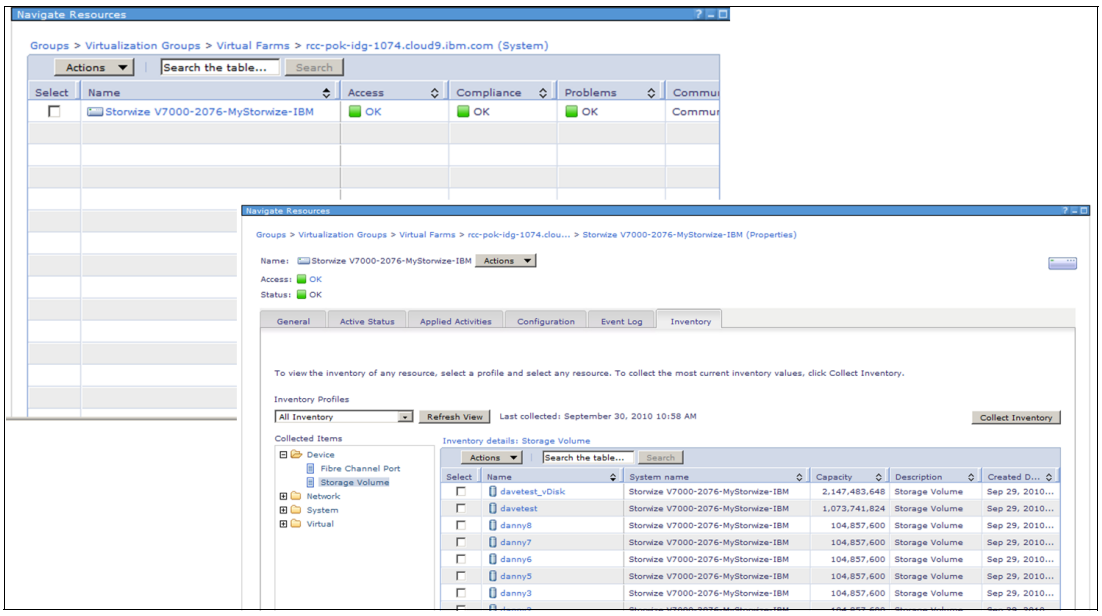

Figure 8-16 Storwize V7000 volume drill-down

For additional details, see the information available at this website:

http://www.ibm.com/systems/software/director/storage/

IBM Tivoli Provisioning Manager for OS Deployment System Director

This Systems Director plug-in edition is designed to provision an operating system through a library of disk images to supported systems in the network. Along with Systems Director, this tool dramatically simplifies the server configuration and operating system installation, BIOS updates and disposal of retired systems for both IBM and non-IBM systems.

For additional details, see the information available at this website:

http://www.ibm.com/software/tivoli/products/prov-mgr-osd-isd/

IBM BladeCenter Open Fabric Manager (BOFM)

This plug-in is designed to help you easily manage I/O and network interconnects on BladeCenters by virtualizing network parameters, such as the World Wide Name (WWN) and Media Access Control (MAC) addresses.

BOFM performs these tasks by controlling the assignment of physical Ethernet MAC and Fibre Channel WWN addresses for the integrated Ethernet interfaces and I/O expansion cards in each blade.

With BOFM, you can configure MAC-based VLANs and security policies, WWN-based zones on switches, and storage partitions on storage systems by using your own pool of MAC addresses and WWNs created in advance. You can then assign these predefined addresses to a blade's I/O ports with BOFM instead of using burned-in hardware addresses that were assigned during manufacture.

For additional details, see the information available at this website:

http://www.ibm.com/systems/bladecenter/hardware/openfabric/openfabricmanager.html

The IBM Workload Partitions (WPAR) Manager for AIX

This plug-in extends IBM Systems Director to manage Workload Partitions. It provides federated WPAR management using a single interface to manage the creation, start, stop, and relocation of WPARs across multiple systems. This can management cost by reducing administrative overheads associated with managing Workload partitions

It enables live application mobility, that is, WPARs can be relocated between systems without restarting the application. This allows for more efficient system load balancing, improving application availability by reducing outages needed for planned system.

Figure 8-17 depicts the use of WPAR relocation for workload balancing where two applications are moved between two servers to balance the load of these servers.

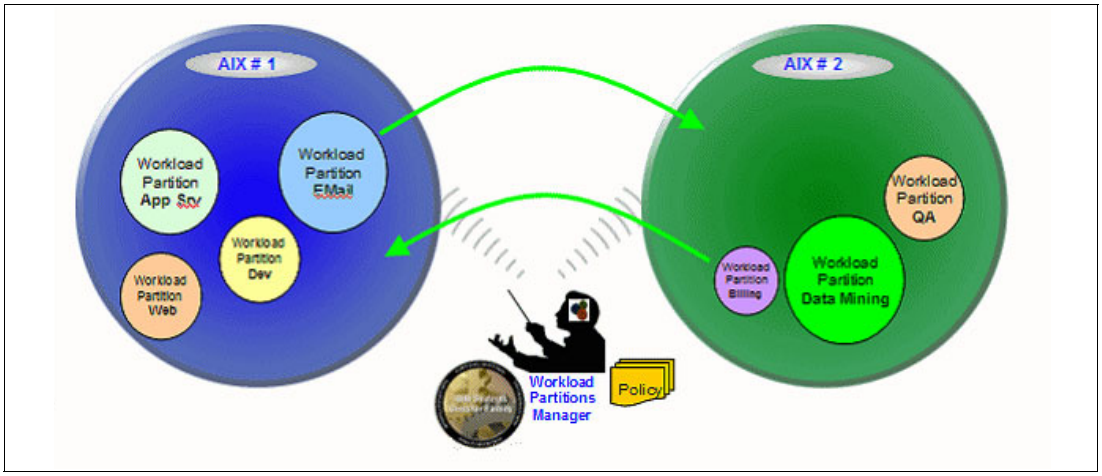

Figure 8-17 WPARs migration

For more information about WPAR, refer to the Redbooks publication, *Introduction to Workload Partition Management in IBM AIX Version 6.1*, SG24-7431

For additional details, see the information available at this website:

http://www-03.ibm.com/systems/power/software/aix/sysmgmt/wpar/index.html

Upward integration modules

This feature lets you make the most of your existing enterprise management structure by upwardly integrating with many workgroup and enterprise-management products.

IBM Systems Director upward integration modules (UIMs) and management packs enable non-IBM workgroup and enterprise-management products to interpret and display data that is provided by Common Agent and Platform Agent. IBM Systems Director UIMs and management packs provide enhancements to the enterprise-management products that collect inventory data, view IBM Systems Director event notifications, and, for some UIMs, distribute IBM Systems Director software packages.

With the IBM Systems Director UIMs and management packs, you can use your enterprise-management software to manage systems that have Platform Agent or Common Agent software installed on them.

You can use Platform Agent software to perform the following tasks:

- Gather detailed inventory information about your systems, including operating system, memory, network adapters, and hardware
- Track your systems with features, such as power management, event log, and system monitor capabilities

Platform Agent uses the latest systems-management standards, including the Common Information Model (CIM), Web-Based Enterprise Management (WEBM) and Extensible Markup Language (XML), to provide compatibility with your existing enterprise-management software. For additional details, see the information available at this website:

http://publib.boulder.ibm.com/infocenter/director/v6r2x/index.jsp?topic=/com.ibm.director.main.helps.doc/fqm0_r_whats_new_in_release_620.html

IBM Systems Director Service and Support Manager

This no-charge *callhome* plug-in, with the integrated IBM Electronic Service Agent™ tool, provides IBM Systems Director the ability to automatically reports hardware problems and collects system service information for monitored systems. The Electronic Service Agent can monitor, track, and capture system hardware errors and service information and report them directly to IBM support for you.

Benefits

The following benefits are realized:

- Automatic problem reporting
- 24/7 direct routing of reported problems to IBM technical support
- Reduced personnel time required for gathering and reporting service information
- Higher availability and shorter downtime

Features

The following features are provided:

- Custom IT management tools enabled
- Secure Internet access
- Accurate solutions with reduced human error in gathering and reporting service information
- Secure web access to your service information
- Consistent IBM worldwide service and support process

For additional details, see the information available at this website:

`http://www-03.ibm.com/systems/software/director/ssm/index.html`

8.2 IBM Tivoli Storage Productivity Center (TPC) suite

The IBM Tivoli Storage Productivity Center (TPC) suite is a comprehensive suite of Storage Resource Management (SRM) products designed to help reduce management complexity by centralizing, simplifying, and optimizing storage tasks associated with storage systems, storage networks, replication services and capacity management.

Specifically, it is designed to assist with the following activities:

- *Cost reduction:* By using the storage efficiency functions of TPC, clients can significantly reduce their future storage purchases and optimize their current Storage Infrastructure.
- *Availability management:* TPC provides comprehensive events alerting functions design to minimize storage outages, thus reducing costs in application outages and personnel costs.
- *Performance management:* TPC provides overall performance management for clients' storage infrastructure, optimizing existing storage assets and deferring new storage purchases.
- *Storage resource management:* TPC allows you to implement a highly optimized Storage Infrastructure with these benefits:
  - Information Life Cycle Management
  - Consistent Storage Policies across the customer's business
  - Significantly improved storage utilization -- deferring storage purchases
  - The ability to manage heterogeneous storage

- *Improved reporting:* TPC provides reporting for an entire heterogeneous Storage Infrastructure - Asset, Performance and Capacity. TPC uses Tivoli Common Reporting (TCR) and additionally has over 400 native reports. TPC is the best performance and reporting tool for SAN Volume Controller and the Storwize V7000.
- *Incident and problem management:* TPC feed Storage Infrastructure data into the IBM Tivoli Integrated Service Management (ISM) family of products. This allows service desks (IBM Tivoli Service Request Manager® (TSRM)) that work off the Tivoli Change and Configuration Management DB (CCMDB) to tightly integrate storage into their knowledge database (using Tivoli Application Dependency Discovery Manager (TADDM).

For more information about IBM Tivoli Integrated Service Management (ISM) products, see the information available at this website:

http://www.ibm.com/ibm/servicemanagement/us/en/

Figure 8-18 provides a high level architecture overview of the TPC product suite.

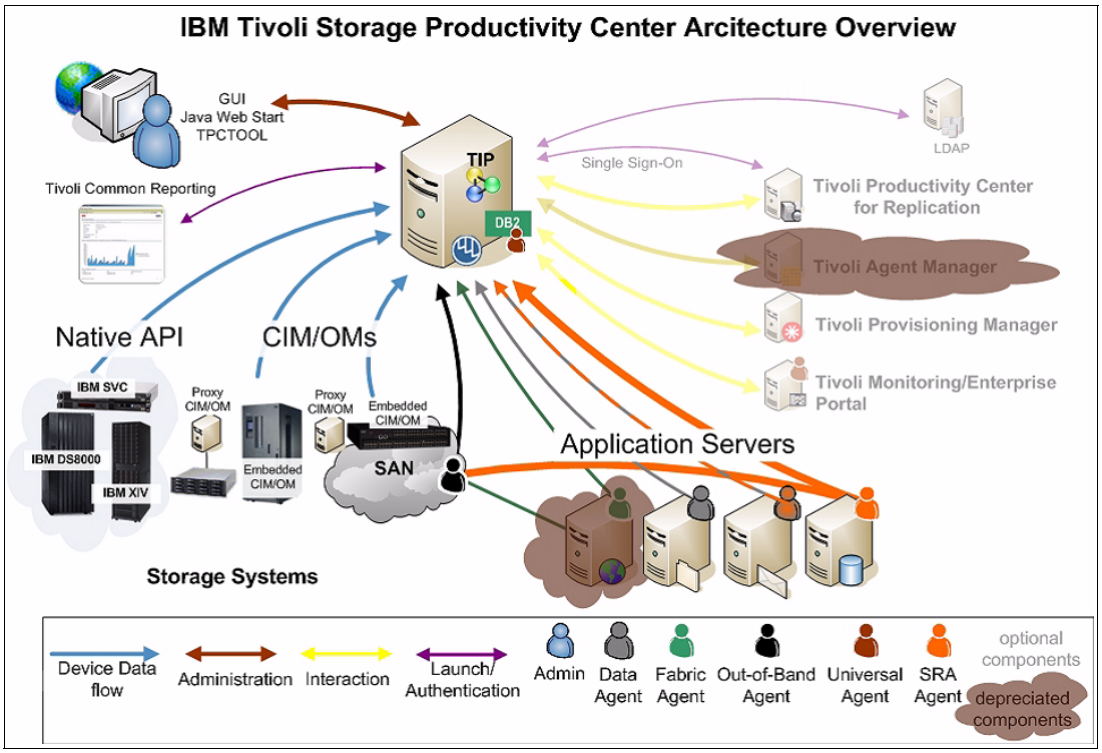

Figure 8-18 IBM Tivoli Storage Productivity Center - architecture overview

The IBM TPC product suite consists of these components:

- IBM System Storage Productivity Center (SSPC)
- IBM Tivoli Storage Productivity Center for Data
- IBM Tivoli Storage Productivity Center for Disk
- IBM Tivoli Storage Productivity Center for Disk Midrange Edition
- BM Tivoli Storage Productivity Center for Replication

8.2.1 IBM System Storage Productivity Center (SSPC)

SSPC is a storage management appliance designed to reduce the number of management servers required for IBM Storage systems management. It is a Windows technology-based component of IBM System x hardware, delivered pre-installed with the TPC Basic Edition and can be upgraded to TPC Standard Edition through purchase of separate license.

SSPC provides the following features:

- DS8000 GUI integration.
- SAN Volume Controller management console and Common CIM Agent.
- Automated device discovery: DS8000 and SAN Volume Controller storage devices can be automatically discovered and configured into TPC environments. These devices are displayed in TPC through a storage topology.
- Asset and capacity reporting: TPC collects asset and capacity information from storage devices on the SAN, which can be kept for historical reporting, forecasting, and used for other tasks, such as analysis and provisioning.
- Advanced Topology Viewer: Provides a linked graphical and detailed view of the overall SAN, including device relationships and visual notifications.

Figure 8-19 depicts a high level overview of the SSPC architecture.

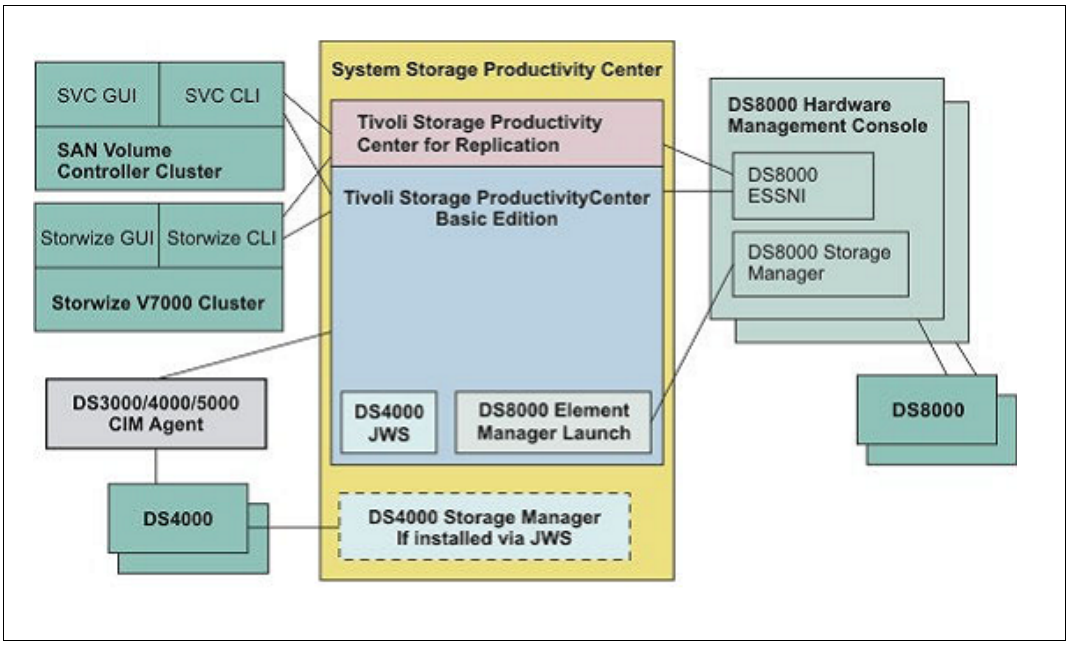

Figure 8-19 SSPC architecture

Licenses: TPC for replication software is bundled with SSPC; however, an additional license is required to activate this software.

TPC Basic Edition has TPC for Disk without performance functions. For more information, refer to “IBM Tivoli Storage Productivity Center for Disk” on page 149.

IBM System Storage Productivity Center (SSPC) uses IBM Tivoli Storage Productivity Center to start graphical user interfaces (GUIs) and external applications, such as element managers, for storage devices. An element manager is vendor-specific software that administers a particular storage device.

For additional details, see the information available at these websites:

http://www.redbooks.ibm.com/abstracts/sg247560.html?
http://www.ibm.com/systems/storage/software/sspc/

Tip: For clients that do not have SSPC, IBM TPC Basic Edition can be procured as a separate product. For additional details, refer to the information available at this website:

http://www.ibm.com/systems/storage/software/center/basic/index.html

8.2.2 IBM Tivoli Storage Productivity Center for Data

IBM Tivoli Storage Productivity Center for Data is designed to help improve capacity utilization of filesystems and databases and add intelligence to data protection and retention practices.

It provides over 400 enterprise-wide reports, monitoring and alerts, policy-based action and file-system capacity automation in a heterogeneous environment.

These reports provide the storage administrator information about the following topics:

- Assets
- Availability
- Capacity
- Usage
- Usage violation
- Backup
- Reports for files, based on criteria, such as data type, size, and last referenced

IBM Tivoli Storage Productivity Center for Data performs the following functions:

- Discovers and monitors disks, partitions, shared directories, and servers
- Monitors and reports on capacity and utilization across platforms to help you to identify trends and prevent problems
- Monitors storage assets that are associated with enterprise-wide databases and issues notifications of potential problems
- Provides a wide variety of standardized reports about filesystems, databases, and storage infrastructure to track use and availability
- Provides file analysis across platforms to help you to characterize your data and, in turn, identify and reclaim space used by non-essential files
- Provides policy-based management and automated capacity provisioning for filesystems when user-defined thresholds are reached or a condition is met
- Generates invoices that charge back for storage use on a departmental, group, or user level

Data Manager, a component of TPC, offers functions designed to help decrease storage costs by:

- Improving storage utilization
- Enabling intelligent capacity planning
- Supporting application availability through computer uptime reporting and application database monitoring
- Provide automation for invoking backups, archiving, scripts, and publishing reports to web servers, etc.

Figure 8-18 on page 145 provides a high level overview of how TPC for Data works.

The process of how TPC and Data works is as follows:

1. It collects data from with storage systems using both Native API (NAPI) and Common Information Model Object Manager (CIMOM).
2. The Storage Resource agent, Data agents, Fabric agents, CIM agents, and Out of Band fabric agents gather host, application, storage system, and SAN fabric information and send that information to the Data Server or Device server.
3. Data agents, such as Storage Resource (SR) and Data and Fabric Agents, are deployed on application servers to collect server-related data. These agents run probes and scans to collect storage related information which is then forwarded to the Data Manager.
4. Data Manager receives information from the agents and then inserts that information in the repository.
5. After integration with Tivoli Provisioning Manager, the automation of resource provisioning, such as filesystem expansion when threshold is exceeded, can now be enabled, yet minimizing human error.

Figure 8-20 displays a view of computer asset details collected using TPC for Data.

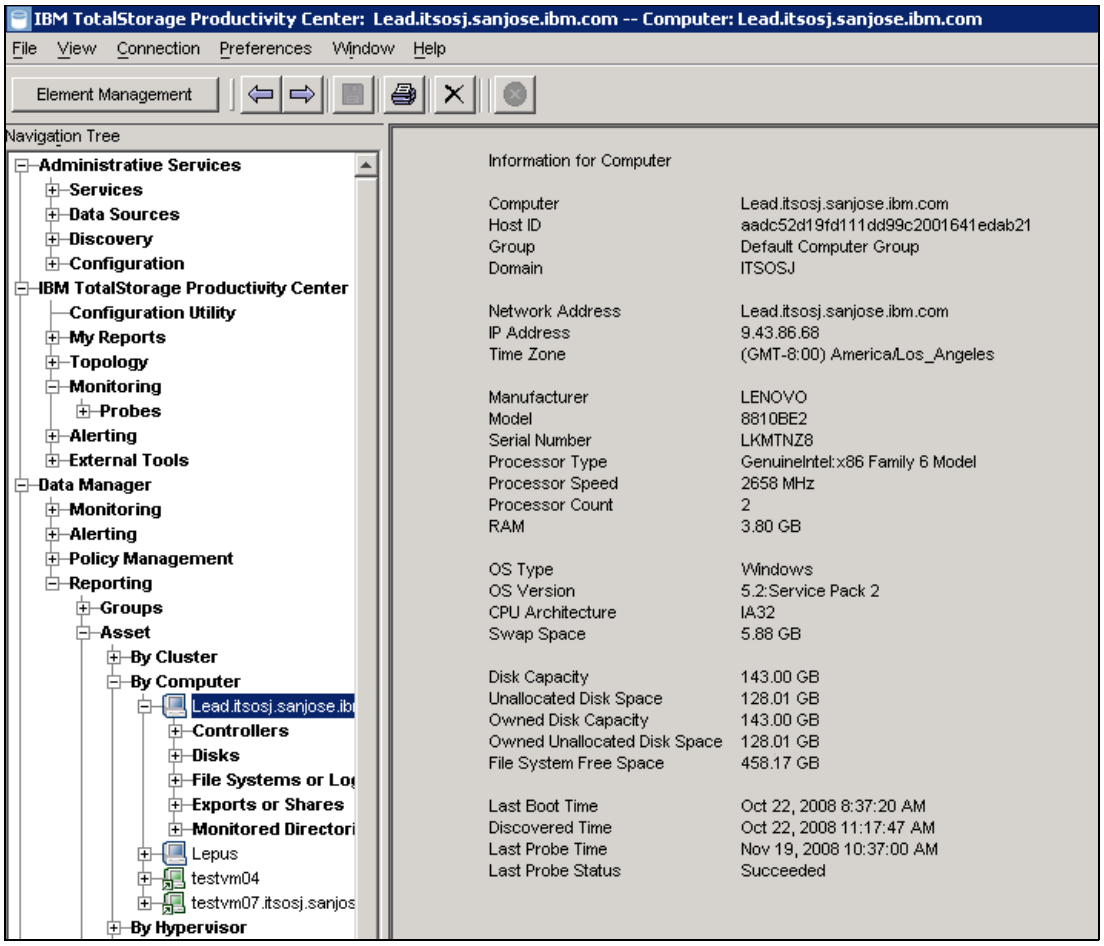

Figure 8-20 Asset details for a computer obtained using TPC for data

For additional details, see the information available at this website:

http://www.ibm.com/systems/storage/software/center/data/index.html

8.2.3 IBM Tivoli Storage Productivity Center for Disk

Tivoli Storage Productivity Center for Disk is designed to provide storage device configuration and management from a single console. It includes performance capabilities to help monitor and manage performance, and measure service levels by storing received performance statistics into database tables for later use. Policy-based automation enables event action, based on business policies. It sets performance thresholds for the devices, based on selected performance metrics, generating alerts when those thresholds are exceeded. Tivoli Storage Productivity Center for Disk helps simplify the complexity of managing multiple SAN attached storage devices.

The ability of the Tivoli Storage Productivity Center to collect asset data and capacity data without deploying resident agents on the servers improves performance and simplifies server data collection. Simultaneously, this reduces the burden on IT administrators to monitor agents. Auto discovery, provisioning, and performance management support for IBM devices, such as the XIV system, brings improved customer time to value.

Figure 8-21 shows a window from TPC for Disk, displaying details of a disk system.

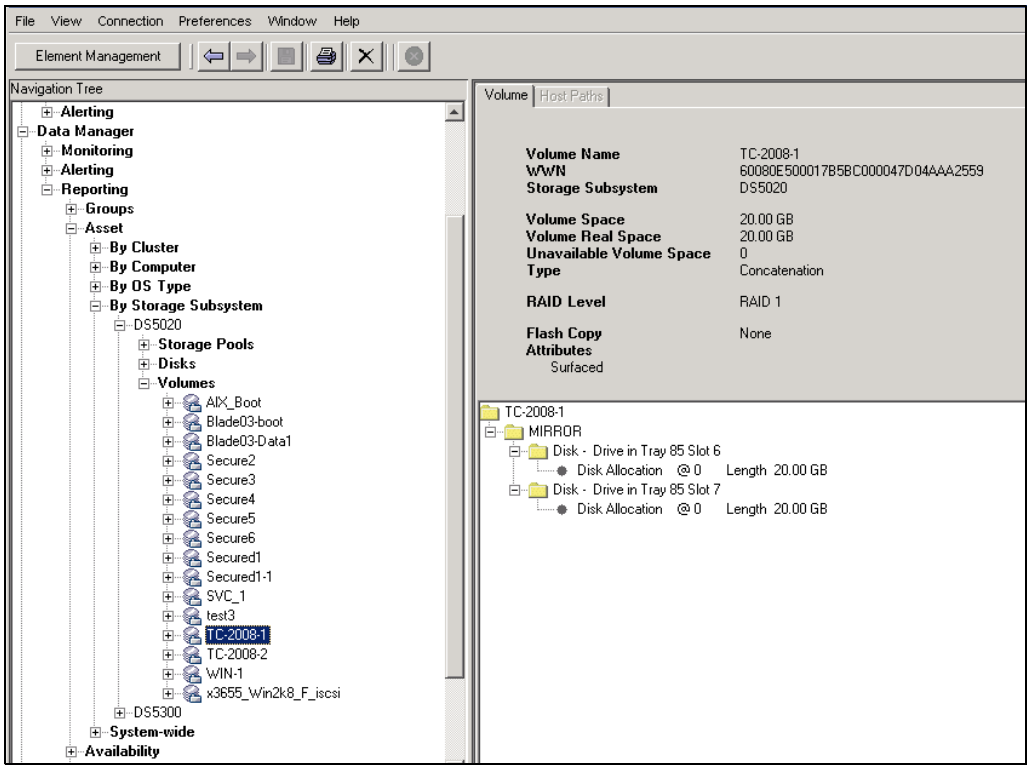

Figure 8-21 Disk system details from TPC for Disk

TPC for Disk provides the following features and benefits:

- Enables proactive performance management with a single, integrated console for performance management of IBM System Storage devices
- Monitors performance metrics across multiple storage subsystems from a single console
- Designed to allow administrators to monitor metrics, such as I/O rates and cache utilization, and support optimization of storage through the identification of the best LUNs across multiple storage subsystems

- Monitors and measures services levels by storing performance statistics for analysis and generation of reports on monitored devices
- Receives timely alerts to enable event action, based on customer policies by setting performance thresholds for the devices, based on performance metrics and the generation of alerts when those thresholds are exceeded
- Provides a Performance Optimization Engine that can help reduce service times of resource-constrained application by an average of 48 percent and up to a maximum of 90 percent
- Designed to help you improve storage return on investment by keeping SANs operational reliably and dependably
- Designed to help reduce storage administration costs by simplifying the management of complex SANs
- Designed to offer comprehensive monitoring and fault identification to improve SAN availability

For additional details, see the information available at this website:

http://www.ibm.com/systems/storage/software/center/disk/index.html

8.2.4 IBM Tivoli Storage Productivity Center for Disk Midrange Edition

IBM Tivoli Storage Productivity Center for Disk Midrange Edition is designed to help reduce the complexity of managing IBM Midrange storage systems (IBM Storwize V7000, DS3000, DS4000 and DS5000 devices) by allowing clients to configure, manage, and monitor performance of their entire storage infrastructure from a single console. IBM Tivoli Storage Productivity Center for Disk Midrange Edition offers equivalent functionality to IBM Tivoli Storage Productivity Center for Disk.

Clients can monitor and analyze performance statistics for these storage systems down to five minute intervals. The performance data can be viewed in real time in the topology viewer, stored for historical reporting, or used to generate timely alerts by monitoring performance thresholds for various device parameters.

It offers the following features and benefits:

- Provides reporting across multiple arrays from a single console
- Receives timely alerts that can enable event action, based on business policies by monitoring performance thresholds for the devices, based on performance metrics and the generation of alerts when those thresholds are exceeded
- Monitors performance metrics, such as input and output (I/O) and data rates, and cache utilization across multiple storage subsystems from a single console
- Helps administrators measure services levels by storing received performance statistics into database tables for later use, and helps analyze and generate reports on monitored devices for display in the central administrative console
- Designed to offer continuous and proactive performance analysis with comprehensive monitoring and fault identification to help improve SAN availability
- Designed to help you improve storage return on investment by keeping SANs operational reliably and dependably
- Designed to help reduce storage administration costs by simplifying the management of complex SANs

- Supports the performance reporting capabilities on the IBM System Storage SAN Volume Controller and SAN Volume Controller Entry Edition with attached DS3000, DS4000 or DS5000 devices
- Supports performance reporting capabilities for any storage virtualized by the IBM Storwize V7000

For additional details, see the information available at this website:

http://www-03.ibm.com/systems/storage/software/center/disk/me/index.html

8.2.5 IBM Tivoli Storage Productivity Center for Replication

The IBM Tivoli Storage Productivity Center for Replication is design to manage multi site data disaster recovery and replication needs efficiently and with less human intervention. It is designed to make management and execution of IBM advanced copy services easy and automatic.

It helps simplify the management of advanced copy services in these ways:

- Automating administration and configuration of these services with wizard-based session and copy set definitions
- Providing simple operational control of copy services tasks, including starting, suspending, and resuming in the same direction or opposite directions
- Offering tools for monitoring and managing copy sessions

TPC for Replication family includes the following products:

- TPC for Replication Two-Site Business Continuity (BC)
- TPC for Replication Three-Site BC
- IBM Tivoli Storage Productivity Center for Replication Basic Edition for System z
- Tivoli Storage Productivity Center for Replication for System z

Each of these replication products is defined as follows:

TPC for Replication Two-Site BC: This supports disaster recovery management through planned and unplanned failover and failback automation for the ESS Model 800, DS8000 and IBM SAN Volume Controller.

TPC for Replication Two Site BC helps you manage replication to a remote backup site through Metro Mirror or Global Mirror. The software is designed to allow you to monitor the progress of the copy services so you can verify the amount of replication that has been completed and the amount of time needed to complete the remaining portion.

Automated failover is designed to help keep your critical data online and available to your users even if your primary site fails. When the primary site comes back on, the software manages failback to the default configuration as well.

TPC for Replication Three-Site BC: This helps provide disaster recovery management through planned and unplanned failover and failback automation for the DS8000 using its Metro Global Mirror (MGM) feature. It features a synchronous copy between two sites combined with an asynchronous copy link to a distant site (perhaps outside the regional power grid). The three-site feature is designed to support fast failover and failback, fast reestablishment of three-site mirroring, data currency at the remote site with minimal lag behind the local site, and quick resynchronization of mirrored sites using incremental changes only.

IBM Tivoli Storage Productivity Center for Replication for System z: This offers the same capabilities and support as the distributed product version, but for the System z environment. It is packaged to run on System z using a mixture of FICON and TCP/IP communications to provide replication management of the DS8000, DS6000, ESS 800, and the SAN Volume Controller, regardless of type of data on them (IBM ECKD™ or FBA).

Tivoli Storage Productivity Center for Replication for System z supports:

- FlashCopy
- Metro Mirror
- Global Mirror
- DS8000 Metro Global Mirror

IBM Tivoli Storage Productivity Center for Replication Basic Edition for System z: This is a product designed to provide a low-cost, single-site, high-availability disk solution that switches to the target volume when there is a fault with the production volume without application impact.

For additional details, see the information available at this website:

http://www.ibm.com/systems/storage/software/center/replication/

8.3 IBM Storage Enterprise Resource Planner (SERP)

SERP extracts data from SRM tools such as TPC and correlates it with the customers' asset management systems, enabling users to maximize utilization of existing systems, plan expansion, reclaim storage, and better manage capacity. SERP automates these features, delivering scores of customizable reports and invaluable insight into a company's storage infrastructure.

It collects data from many instances of SRM and asset tools, and provides a view of assets with respect to geography and business infrastructure that makes the following possible:

- Capacity planning
- Configuration management
- Reclamation and utilization improvement
- ILM and right-tiering
- Deployment planning and reservation
- Topology and connectivity
- Financial transparency
- Chargeback

For additional details, see the information available at this website:

http://www.ibm.com/services/us/index.wss/offering/gts/a1029689

8.4 IBM Storwize Rapid Application Storage (SRAS)

IBM Storwize Rapid Application Storage (SRAS) is an innovative new storage solution that helps improve storage efficiency and application availability, integrates closely with key existing applications, and can be deployed with minimal disruption to applications and to staff.

SRAS provides the following features:

- Improves disk utilization by up to 30 percent through thin provisioning and virtualization
- Improves storage administrator productivity through intuitive and easy to use user interface
- Near instantaneous backup and recovery through the use of Tivoli FlashCopy Manager snapshot technology
- Increases application performance with up to 60 percent better response time through Easy Tiering technology and hot spot analytics
- Easy integration with existing applications: SAP, Microsoft Exchange, IBM DB2, Oracle, and SQL Server
- Data migration utility to import external storage systems into IBM Storwize V7000 system
- Product is designed to be implemented quickly allowing users to realize quicker return on investment through IBM Rapid Deployment Services (RDS)

SRAS is a bundled solution comprising the following components:

- *Solution Implementation Services:* Using the Rapid Deployment Model, SRAS can be deployed in days instead of months as there is no need to integrate the different components. RDS has been designed to standardize and automate product installation, configuration, and integration.
- *Tivoli Productivity Center:* As discussed in "IBM Tivoli Storage Productivity Center for Disk Midrange Edition" on page 150, this is a storage resource management (SRM) tool to configure, manage, and monitor performance disk systems.
- *Tivoli FlashCopy Manager:* This provides near instantaneous backup and restore functions for quick recovery of applications through the use of point in time snapshots.
- *IBM Storwize V7000:* Provides storage with features such as thin provisioning, easy tiering, and storage virtualization.

Figure 8-22 shows the various SRAS components and associated functions.

Figure 8-22 SRAS components

For additional details, see the information available at these websites:

- http://www.ibm.com/systems/storage/news/center/storwize_ras/index.html
- http://www.ibm.com/systems/storage/news/center/storwize_v7000/
- http://www.ibm.com/software/tivoli/products/storage-flashcopy-mgr/

Part 2

Data protection

IBM offers effective, secure solutions that help to address your growing data protection demands and, simultaneously, reduce operational complexity.

In this part of the book, we review IBM storage products and solutions that can contribute to protecting your data in the areas of backup and recovery, continuous operations, and retention and compliance.

- Chapter 9, “Backup and recovery” on page 157 reviews the following products, solutions, and service offerings:
  - IBM Tivoli Storage Manager
  - Copy Services, as available with various disk systems
  - IBM FlashCopy Services
- Chapter 10, “Continuous operations” on page 185 discusses these topics:
  - PowerHA System Mirror for AIX Enterprise Edition (formerly known as High Availability Cluster Multi-Processing Extended Distance (HACMP/XD)
  - Open HyperSwap for AIX with IBM Tivoli Storage Productivity Center for Replication V4.2
  - Geographically Dispersed Open Clusters (GDOC)
  - IBM Tivoli Storage FlashCopy Manager
  - IBM Tivoli Storage Manager FastBack
  - Geographically Dispersed Parallel Sysplex/Peer-to-Peer Remote Copy HyperSwap Manager (GDPS/PPRC HM)
  - Tivoli Storage Productivity Center for Replication (TPC-R) Basic Hyperswap
- Chapter 11, “Retention and compliance” on page 209 describes the following products:
  - IBM Information Archive
  - IBM Enterprise tape libraries with LTO5 tape drives
  - IBM System Storage N series with SnapLock feature

9

Backup and recovery

In today's business world, application servers and data often need to remain available 24 hours per day. You cannot afford to lose data, but you also cannot afford to stop critical systems for hours so you can adequately protect the data. With the rapid increase in the amount of data, the critical business need of the data, and shrinking backup windows, traditional backup and restore methods might be reaching their limits in meeting these challenging requirements.

In this chapter, we discuss the principles and implementation of data backup and recovery techniques in IBM storage systems. We discuss these topics from the hardware and software perspectives.

We review the following products, solutions, and service offerings:

- IBM Tivoli Storage Manager
- Copy Services, as available with various disk systems
- IBM System Storage FlashCopy Manager, which covers these products:
  - IBM Tivoli FlashCopy Manager
  - IBM System z FlashCopy Manager

9.1 IBM Tivoli Storage Manager

Tivoli Storage Manager protects data from hardware failures, errors, and unforeseen disasters by storing backup and archive copies in auxiliary and off-site storage. Tivoli Storage Manager Extended Edition provides centralized web-based management, intelligent data move and store techniques, and comprehensive, policy-based automation. All of these functions work together to minimize administration costs and the impact to computers and networks. The IBM TSM Extended Edition also provides scaling to protect from hundreds to thousands of computers with these characteristics:

- Run more than a dozen operating systems
- Range from mobile computers to mainframes
- Are connected through the Internet, wide-area networks (WANs), local area networks (LANs), or storage area networks (SANs)

Optional software modules allow business-critical applications that are required to run 24x7x365, to utilize Tivoli Storage Manager centralized data protection with no interruption to the services they provide. Optional software extensions also allow SAN-connected systems to use the SAN for data protection, data movements, and to provide Hierarchical Storage Management (HSM) to automatically move unused data files from online disk storage to auxiliary tape storage.

Tivoli Storage Manager is the number one product of choice for an efficient and effective, enterprise-wide storage solution. It provides a data protection solution for backup, archiving, disaster recovery planning, space management, data deduplication, server-forced automated updates of ITSM clients, database and application protection, bare machine recovery, and records retention.

Starting with Tivoli Storage Manager version 6.1, all of the features benefit from the embedded IBM DB2 database engine for enhanced interoperability, performance, and client data throughput capabilities.

Tivoli Storage Manager provides the following capabilities:

- Centralized administration for data and storage management
- Fully automated data protection
- Efficient management of information growth
- High-speed automated server recovery
- Data protection over LAN, WAN, and SAN architectures
- Optionally, specifically designed backup solutions for major groupware, enterprise resource planning (ERP) applications, and database products
- HSM for Windows and UNIX clients
- Disaster recovery planning
- Industry proven storage reduction features, such as data compression and data deduplication
- Full compatibility with multivendor storage devices for data retention, including disk subsystem, tape libraries, and virtual tape products, all from midrange to enterprise-size
- An enhanced reporting and monitoring system, using the Administration Center and working together with the common IBM Tivoli Integrated Portal
- Transparent licensing reflecting the business needs

9.1.1 Tivoli Storage Manager concepts

Tivoli Storage Manager features an embedded DB2 database engine configured automatically as a part of Tivoli Storage Manager server deployment. The database tracks what is backed up, where it is stored, how long it will be stored, how many versions of the data are kept, and the number and period of time a copy of the file is kept after it is deleted from the original filesystem (Tivoli Storage Manager client).

The Tivoli Storage Manager clients send the data to the Tivoli Storage Manager server, either over the LAN or SAN. Most backups occur in accord with their schedule, but clients can perform on-demand backups when required by the administrator. Tivoli Storage Manager clients can also perform their own restores. Tivoli Storage Manager has an improved Administration Center that was introduced in version 5.3 and can be installed either on the same machine as the Tivoli Storage Manager Server, or on a separate machine. Figure 9-1 shows a conceptual diagram of Tivoli Storage Manager.

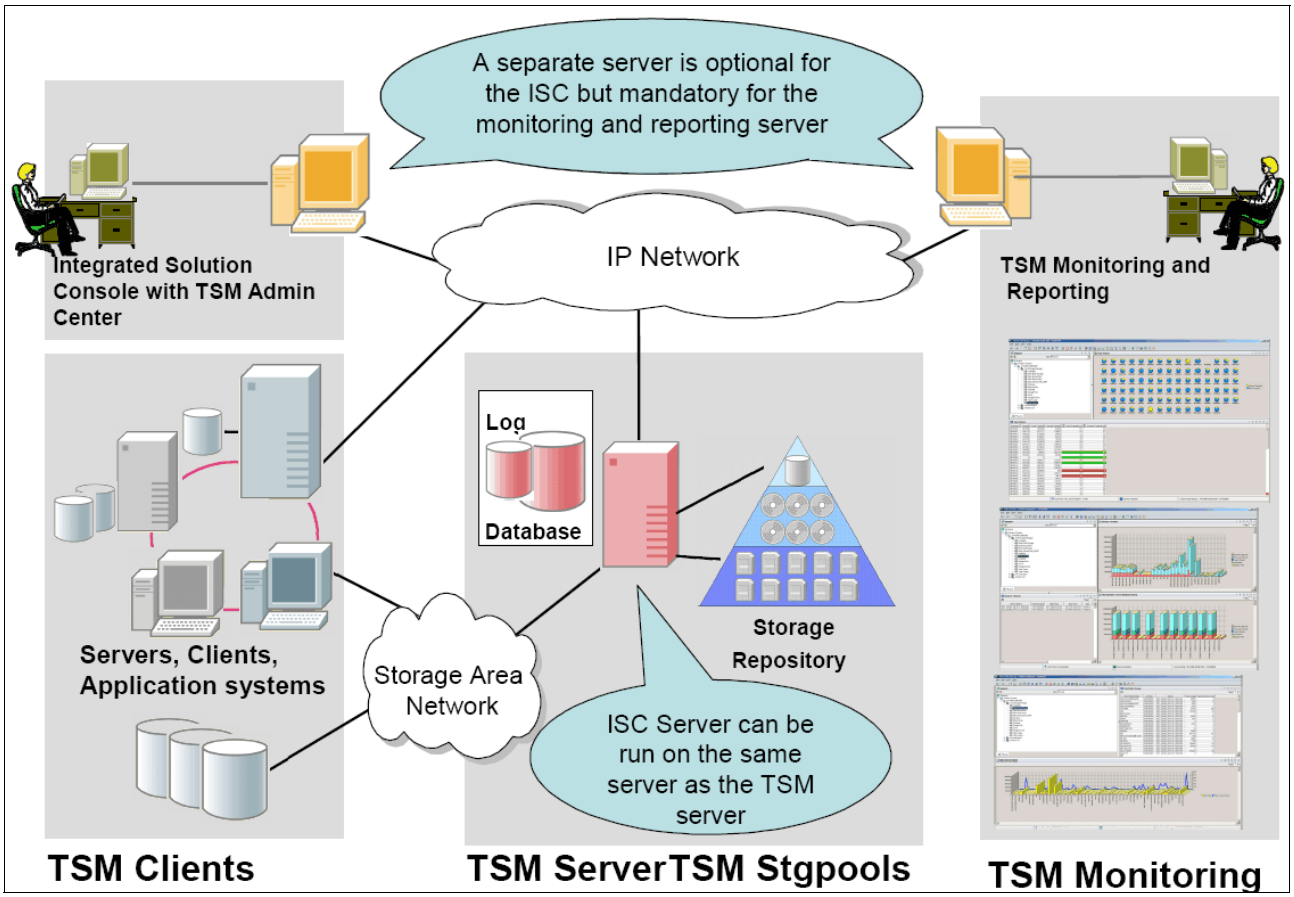

Figure 9-1 The concept of IBM Tivoli Storage Manager architecture

The actual data that the client sends is stored in the Tivoli Storage Manager server *storage pools*. Tivoli Storage Manager is unique in that the storage pools can form a storage hierarchy made up of more than 500 supported devices. This allows for flexibility, longevity, and most importantly, fast backups and fast restores, depending on infrastructure, required Recovery Point Objective (RPO), and Recovery Time Objective (RTO) of lost or corrupted data.

Most businesses back up their data initially to Tivoli Storage Manager server disk storage, which acts as cache, allowing for hundreds of clients to back up at the same time. Then, based on policies, data migrates in a fashion that expedites restores to tape or optical disk.

When the data migrates, all data belonging to one client is moved together to the next pool. By keeping all of that data together, restores are faster because not as much tape positioning is required. This migration process can also accommodate data movement to collocated tapes, which further expedites the restores because each restore contains data from one user only. Collocation can be used to group important functions of your business to ensure rapid recovery in the event of disaster.

Tivoli Storage Manager server

The key component of the Tivoli Storage Manager server is its embedded DB2 database. The database was especially designed to manage the Tivoli Storage Manager client data and does not require any DB2 expertise. All retention policy information, logging, authentication and security, media management, and object inventory are managed through this database. Obviously, this database is fully protected with software mirroring, roll-forward capability, object indexing and expiration, and online backup and restore functions.

For storing client data, the Tivoli Storage Manager server manages a storage repository. The storage repository can be implemented in a hierarchy, using any combination of supported media or magnetic or optical disk, tape, and robotic storage devices, as long as they are locally connected to the server system or are accessible through a SAN. To take advantage of SAN technology, the Tivoli Storage Manager server has features that dynamically share SAN-connected, automated tape library systems among multiple Tivoli Storage Manager servers, and optionally provide LAN, LAN-free and server-free backup services.

An example of the storage hierarchy within the Tivoli Storage Manager infrastructure is shown in Figure 9-2.

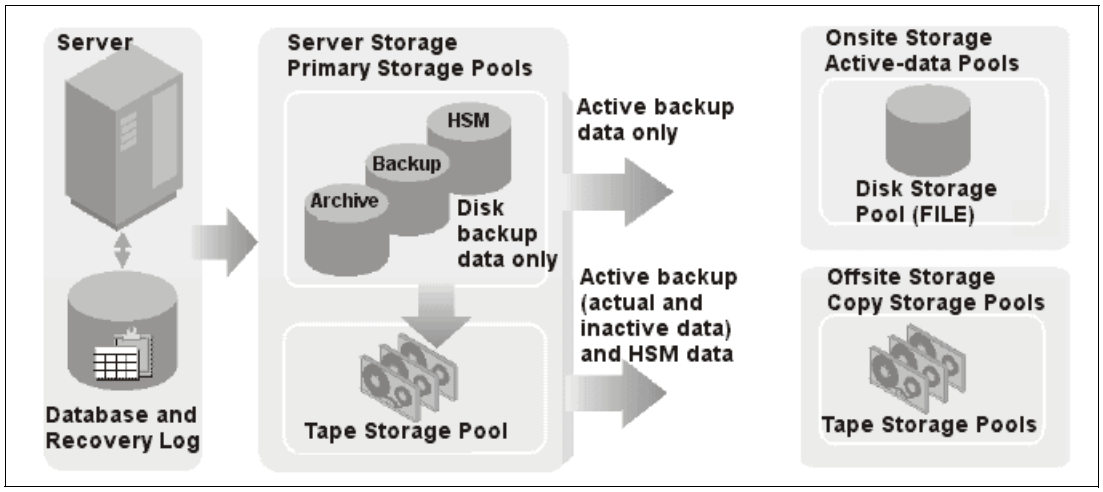

Figure 9-2 IBM Tivoli Storage Manager server storage repository

Tivoli Storage Manager client

Data is sent to the Tivoli Storage Manager server by the Tivoli Storage Manager Backup/Archive client and eventual software enhancement modules. These modules, also known as data protectors, work together with the IBM Tivoli Storage Manager server base product to ensure that any data stored is managed as defined and fully compatible.

The client provides the operational backup and archive functions. The client implements the patented, progressive backup methodology, adaptive sub-file backup technology, and unique record retention methods for backup and archive functions. The Tivoli Storage Manager clients are implemented as multi-session clients, meaning that they are able to take advantage of the multi-threading capabilities of modern operating systems.

To enhance the basic backup and recovery functionality of standard backup and archiving of operating systems, Tivoli Storage Manager introduces additional tools and utilities to support data protection of business critical databases, where an application outage cannot be afforded to perform backup and archive tasks. The following modules are included:

- IBM Tivoli Storage Manager for Mail (either MS Exchange or IBM Lotus® Domino)
- IBM Tivoli Storage Manager for Databases (MSSQL, Oracle, DB2, Mixed)
- IBM Tivoli Storage Manager for Microsoft Sharpened
- IBM Tivoli Storage Manager for Advanced Copy Services
- IBM Tivoli Storage Manager for Copy Services
- IBM Tivoli Storage Manager for Enterprise Resource Planning
- IBM Tivoli Storage Manager for Space Management (supporting AIX and Linux)
- IBM Tivoli Storage Manager Hierarchical Storage Management for Windows (HSM)
- IBM Tivoli Storage Manager for Storage Area Networks (for LAN-free backup)
- IBM Tivoli Storage Manager for Virtual environments (VMware consolidated backup)
- IBM Tivoli Storage Manager for System Backup and Recovery
- IBM Tivoli Storage Manager FastBack (for continuous data protection)
- IBM Tivoli Storage Manager FastBack for Microsoft Exchange
- IBM Tivoli Storage Manager FastBack for Bare Machine Recovery
- IBM Tivoli Storage Manager FastBack for Workstations

Tip: Not all of the products listed have been updated with Tivoli Storage Manager version 6.1 or 6.2. The backward compatibility is guaranteed with the previous version, version 5.5.

9.1.2 Progressive incremental backup

One of the key differences between Tivoli Storage Manager and other data protection products is the *progressive incremental backup methodology*. Tivoli Storage Manager only backs up new or changed files. It tracks all of the backups at the file level. It has no concept of a full backup compared with dependent, incremental, or differential ones.

Because of the powerful relational database in Tivoli Storage Manager, a periodic full backup of the latest active version of every file is not required. This is because the latest *active* version of every file is available in the Tivoli Storage Manager server storage hierarchy. Therefore, it is not necessary to perform a full backup, thus consuming further resources. This progressive, incremental backup using Tivoli Storage Manager is in-line with the IBM strategy for storage efficiency.

Incremental backup by date is also available. Tivoli Storage Manager file level, progressive backup methodology is far superior to other traditional backup methods, such as Full+Incremental or Full+Differential, because progressive incremental backups are never redundant.

9.1.3 More information

For more information, see the following resources:

- Technical details and comprehensive information about all Tivoli Storage Manager features and components, see the Tivoli Information Center at this website:

  `http://publib.boulder.ibm.com/infocenter/tsminfo/v6r2/index.jsp`

- Technical documentation and best practices are located at this website:

  `https://www.ibm.com/developerworks/wikis/display/tivolidoccentral/Home`

9.2 Point-in-time copy

Data volume snapshot techniques (also known as point-in-time copy) provide nearly instant backup and restore capabilities. A snapshot operation takes much less time than a direct tape backup, and so it authorizes more frequent backups and increases the flexibility of backup scheduling and administration.

Many storage systems are now equipped with volume snapshot tools. However, these hardware-based snapshot technologies provide only *crash consistent copies* of data. Many business critical applications, including those that rely on a relational database, need an additional snapshot process to ensure that all parts of a data transaction are flushed from memory and committed, providing consistent copies of data.

9.2.1 Data volume snapshot techniques

Data volume snapshot techniques, such as IBM FlashCopy can help minimize the impact caused by backups and provide near instant restore capabilities. IBM FlashCopy refers to a technology that creates a consistent copy of data at a specific point-in-time, leaving the original data accessible. It is also called Point-in-Time copy.

In this section, we discuss common concepts and features of the IBM FlashCopy function as implemented in the DS8000, the San Volume Controller and the Storwize V7000.

We also discuss the XIV system which uses a somewhat different function, called snapshot. The differences among these offerings are discussed in “XIV system snapshot” on page 164.

Full Volume FlashCopy

The FlashCopy function enables you to make consistent, point-in-time, full volume copies of data, with the copies immediately available for read or write access. A *consistent copy* is one in which the data does not change when the copy is being made. You can use the copy with standard backup tools available in your environment to create backup copies on tape.

FlashCopy creates a copy of a source volume on a target volume. When you initiate a FlashCopy operation, a relationship is created between a source volume and a target volume. A FlashCopy relationship is a mapping of the FlashCopy source volume and a FlashCopy target volume. This mapping allows a point-in-time copy of that source volume to be copied to the associated target volume. The initialization phase is quite fast, in that it only involves setting up new pointers. During the initialization phase, write I/O to source disks is suspended for a very short period of time. Target disks must not be accessed during this phase, as, from the perspective of the server, data on disks will change suddenly. After the FlashCopy is complete, the source and target disks can be accessed independently. The general rules for FlashCopy are as follows:

- The minimum granularity is one logical disk (logical volume, VDisk)
- The FlashCopy target must be the same size as the source disk
- The FlashCopy target must not be accessed during FlashCopy initialization
- The size of the source or target logical disk cannot change after a persistent FlashCopy is established

The FlashCopy is performed at the block level of a storage system, which does not rely on the host operating system, and is therefore transparent to the host. However, to ensure the integrity of the copy, it is necessary to flush the host cache of any outstanding reads or writes prior to performing the FlashCopy operation. Failing to do so produces what is referred to as a *crash consistent copy*, meaning that the resulting copy requires the same type of recovery procedure (such as log replay and filesystem checks) as is required following a host crash.

Several operating systems and applications provide facilities to stop I/O operations and ensure all data is flushed from the host cache. If these facilities are available they can be used to prepare and start a FlashCopy operation. When this type of facility is not available, the host cache must be flushed manually by quiescing the application and unmounting the filesystem or drives. After FlashCopy finishes the copy, the relationship can be terminated or can persist:

- *Non-persistent FlashCopy:* Typical FlashCopy is *non-persistent* meaning that the relationship between a source and a target terminates after the data is fully copied. After FlashCopy terminates, FlashCopy target volumes are stand-alone volumes, with no relationship to the source volumes.
- *Persistent FlashCopy:* After data is fully copied, the relationship persists, and changed blocks on the source volume are being recorded.

It is possible to establish more FlashCopy relationships using the same source. The number varies between storage systems. In other words, a source volume can have more than one target volume. However, a target volume can still have only one source. Furthermore, *cascading FlashCopy* is not allowed (that is, a volume cannot be both a source and a target volume).

FlashCopy Space Efficient (FlashCopy SE)

FlashCopy Space Efficient, or FlashCopy SE, is unlike normal FlashCopy in that it copies only changed data (before new data is written). In this way, it saves space and has less impact on storage system performance.

Typical scenarios in which FlashCopy SE is most useful are as follows:

- Cases where FlashCopy is taken in very short intervals, whereas normal FlashCopy cannot complete the copy in the same amount of time
- Use of FlashCopy for changes over a short period of time, for example, for an incremental backup
- When storage space requirements are not met for a full FlashCopy
- For temporary snapshots during activities such as these:
  - Application development
  - Disaster recovery scenario tests
  - System update testing
- When FlashCopy is used in combination with Metro Mirror of Global Mirror

FlashCopy SE is optimized for use cases where a small percent of the source volume is updated during the life of the relationship. If much more than 20 percent of the source is expected to change, there might be trade-offs in performance as compared with space efficiency.

Because space is allocated as needed on the target volumes, data location on the target volumes is independent of data location on the source volumes. Essentially, all data on FlashCopy SE targets is stored in a random fashion and does not benefit from the locality of data, as can be expected on a standard FlashCopy target. Processes that might access the source data in a sequential manner might not benefit from sequential processing when accessing the target.

With FlashCopy SE, read I/Os against the target volume are always redirected to the source volume. FlashCopy SE relations might fail when no space is left in the storage allocated for the targets or if the storage allocated for the targets becomes unusable. This is to prevent I/O failures on the source volumes. Standard FlashCopy relations require user intervention to remove the relation if the target becomes unusable.

XIV system snapshot

The point-in-time-copy function in the XIV system is called *snapshot*. It is a block-level data copy. XIV system snapshot is based on several innovative technologies that ensure minimal degradation of storage system performance. XIV system snapshots make use of pointers and do not necessarily copy all the data to the second instance of a volume. They efficiently share cache for common data, effectively working as a larger cache than with full data copies. From this perspective, the XIV system snapshot resembles the FlashCopy SE used in, for example, DS8000 storage systems.

With FlashCopy SE, if there is a write I/O attempt to a block of data that has not been copied yet since the creation of a relationship, a first copy of source volume data block is made, and the target volume pointer now points to this new copy. Then the source volume is modified. This is called *copy-on-write*. XIV system uses a *redirect-on-write* mechanism. This means that the pointer of the target volume is not altered, but new data is written to another location, and the source volume pointer now points to this new block.

9.2.2 FlashCopy in combination with other Copy Services

For the DS8000, FlashCopy can work together with Metro Mirror and Global Mirror to provide better data protection. (For an overview of the remote copy function, refer to “Remote Copy Services for continuous operations” on page 188). For example, you can perform a Metro Mirror copy to duplicate data from Site A to Site B, then perform a daily FlashCopy to copy the data elsewhere.

FlashCopy with Metro Mirror or Global Copy

With this option, the FlashCopy target volume can also be a primary volume for a Metro Mirror or Global Copy relationship. You can use this capability to create both a remote copy and a local copy of a production volume.

Figure 9-3 illustrates this capability. The FlashCopy target and the Metro Mirror (or Global Copy) primary are the same: volume B.

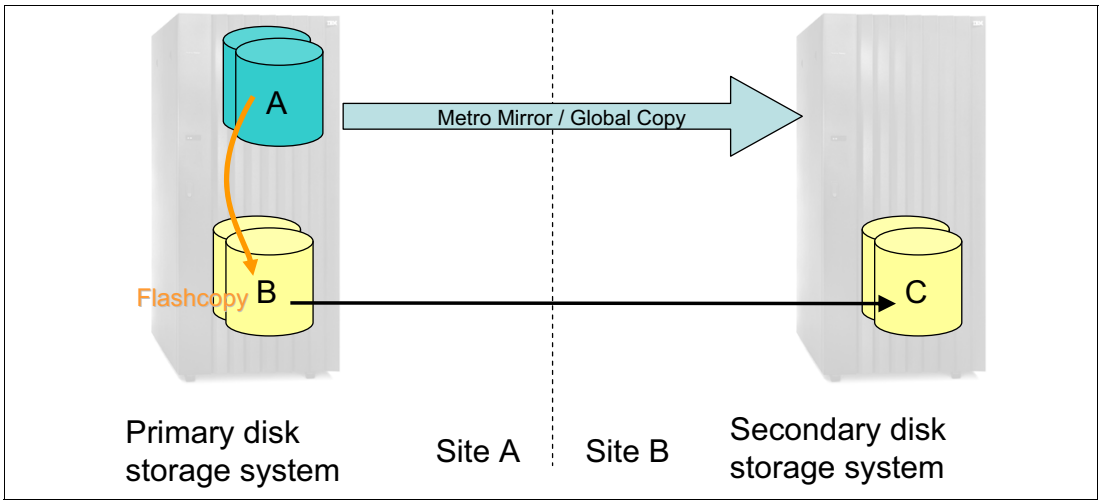

Figure 9-3 The FlashCopy target is the Metro Mirror (or Global Copy) primary

Remote Pair FlashCopy

Remote Pair FlashCopy is also called *preserve mirror*. It has been designed to overcome the shortcomings of the previous solution to FlashCopy onto a Metro Mirror source volume and the loss of the disaster recovery capability during the FlashCopy copy operation. When the Metro Mirror copy is in a synchronous state (usually after initial synchronization or after a switch from asynchronous Global Copy to synchronous Metro Mirror), the state of Metro Mirror pairs is full duplex. During the FlashCopy, the Metro Mirror source volume, for a short period of time, is not in full duplex, and thus disaster recovery is not ensured.

As the name implies, preserve mirror preserves the existing Metro Mirror status of full duplex during the copy operation. Figure 9-4 shows this approach:

1. A FlashCopy command is issued by an application or by the customer to Local A volume with Local B volume as the FlashCopy target.
2. The DS8000 firmware propagates the FlashCopy command through the Metro Mirror links from the local storage server to the remote storage server. This in-band propagation of a Copy Services command is only possible for FlashCopy commands.

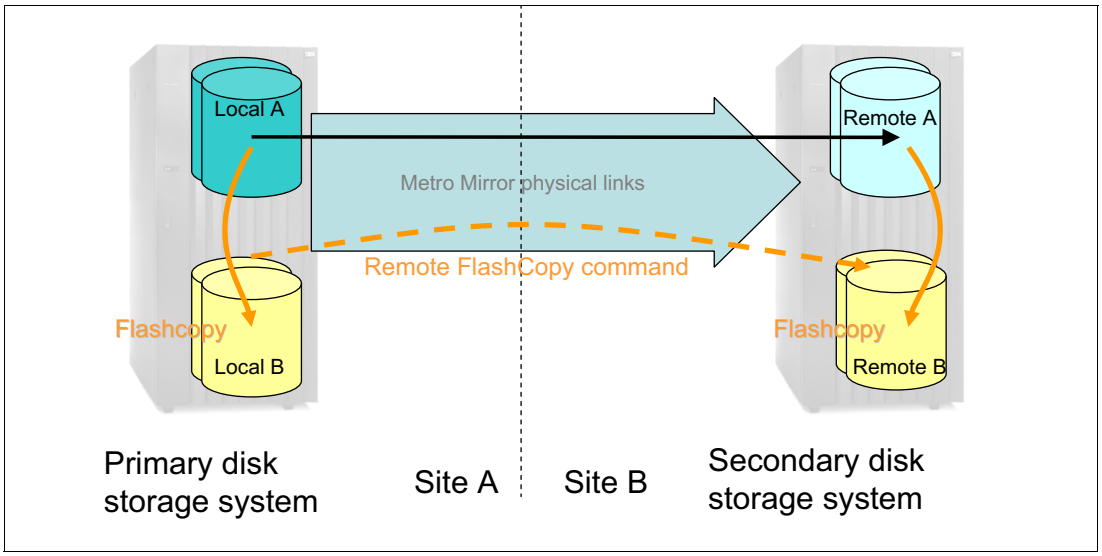

Figure 9-4 Preserve Mirror operation scheme

Independent of each other, the local storage server and the remote storage server then execute the FlashCopy operation. The local storage server coordinates the activities at its end and takes action if the FlashCopy is not successful on both storage servers.

Global Mirror

Global Mirror, as a long-distance remote copy solution, is based on an efficient combination of Global Copy and FlashCopy functions. It is a transparent and autonomic mechanism to intelligently utilize Global Copy in conjunction with certain FlashCopy operations to attain consistent data at the remote site. Details of this copy solution are located in “IBM System Storage DS8000 Global Mirror” on page 191.

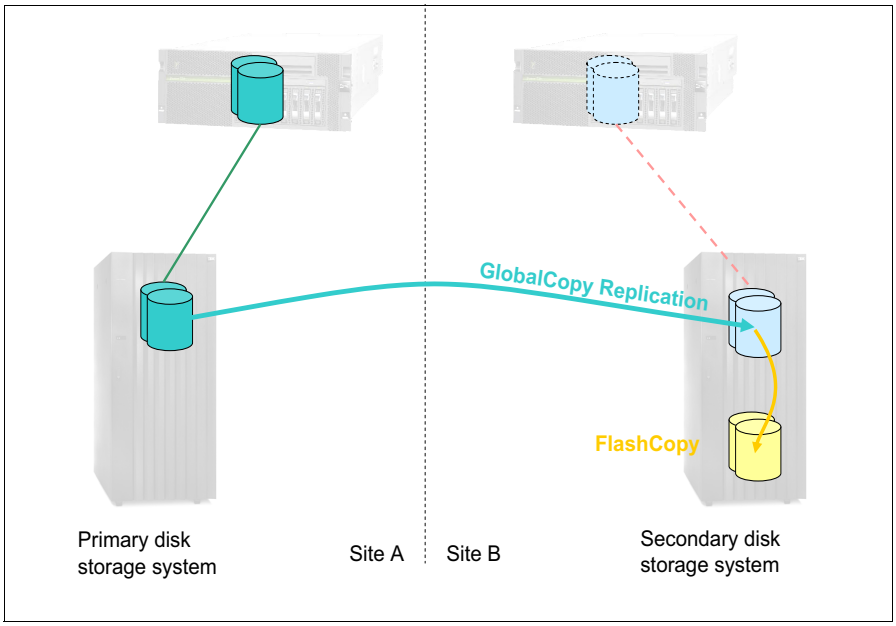

Figure 9-5 Global Mirror operation scheme

Because Global Copy replication target volumes on secondary disk storage system are not in a synchronous relationship with volumes at the primary disk storage system, it needs to pause write IO to the primary volumes and perform synchronization before FlashCopy takes place. After FlashCopy initialization finishes, which is performed virtually immediately, suspended write I/O to the source volumes is released.

Global Mirror takes place in disaster recovery scenarios such as these:

- Distances goes far beyond a campus
- Site interconnection is not sufficient for synchronous remote copy (Metro Mirror)
- Performance at the primary site is important

9.2.3 SONAS snapshot

SONAS snapshot works at the file level, unlike FlashCopy or XIV system snapshots, which work at the block level. SONAS snapshot is a space-efficient, read-only, point-in-time consistent version of an entire SONAS filesystem, frozen at a Point-in-Time. It enables online backups to be maintained, providing near instantaneous access to previous versions of data without requiring complete, separate copies or resorting to auxiliary backups.

Snapshots of a SONAS filesystem are read-only; changes are made only to the active (that is, normal (non-snapshot)) files and directories. Snapshots are only made of active filesystems, and you cannot make a snapshot of an existing snapshot. Individual files, groups of files, or entire directories can be restored or copied back from snapshots.

You can access this directory, but no change is allowed. You can restore files from a specific snapshot by copying them out of the snapshot subdirectory.

The snapshot appears as a special directory. Example 9-1 shows a simple filesystem layout.

Example 9-1 Filesystem before snapshot

```
/fs1/file1
/fs1/file2
/fs1/subdir1/file3
/fs1/subdir1/file4
/fs1/subdir2/file5
```

Example 9-2 shows the same filesystem with one snapshot stored in the /fs1/.snapshots/snap1/ directory.

Example 9-2 Filesystem with one snapshot

```
/fs1/file1
/fs1/file2
/fs1/subdir1/file3
/fs1/subdir1/file4
/fs1/subdir2/file5
/fs1/.snapshots/snap1/file1
/fs1/.snapshots/snap1/file2
/fs1/.snapshots/snap1/subdir1/file3
/fs1/.snapshots/snap1/subdir1/file4
/fs1/.snapshots/snap1/subdir2/file5
```

These are other important SONAS snapshot details:

- Snapshots only consume space when the filesystem changes.
- Snapshots use no additional disk space when first taken.
- Snapshots are enforced to be consistent across the filesystem to a single Point-in-Time.
- For Common Internet File System (CIFS) users, SONAS snapshots are readily accessible using Microsoft Volume Shadow Services (VSS) integration into the Windows Explorer interface, as shown in Figure 9-6.

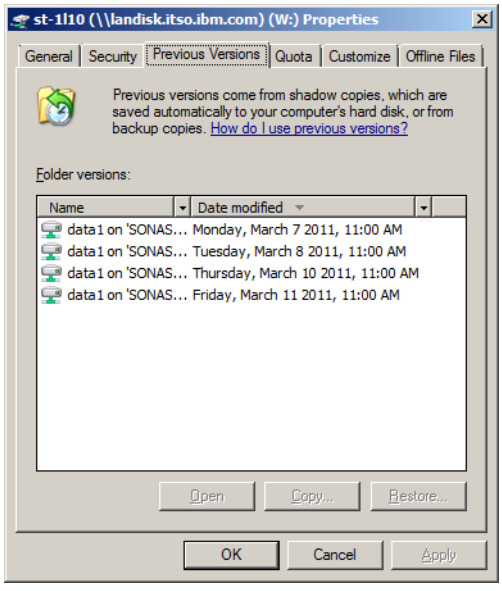

Figure 9-6 SONAS snapshots accessible to Windows CIFS users with a Windows Explorer interface

9.3 IBM System Storage FlashCopy Manager

IBM has two optional offerings that are compatible with the storage system FlashCopy function. Each one helps you to easily produce concurrent and consistent copies of your data:

- IBM Tivoli FlashCopy Manager
- IBM System z FlashCopy Manager

9.3.1 IBM Tivoli Storage FlashCopy Manager

IBM Tivoli Storage FlashCopy Manager is designed to work in a distributed environment. It helps deliver the highest levels of protection for mission critical IBMDB2, SAP, Oracle, Microsoft Exchange, and Microsoft Structured Query Language (SQL) server applications using integrated, application-aware snapshot backup and restore capabilities. This is achieved through the exploitation of advanced IBM storage hardware snapshot technology to create a high performance, low impact application data protection solution.

The snapshots captured by Tivoli Storage FlashCopy Manager can be retained as backups on local disk. Additionally, with the optional integration with Tivoli Storage Manager, you can make use of the full range of advanced data protection and data reduction capabilities. These include data deduplication, progressive incremental backup, HSM, and centrally managed policy-based administration.

Because a snapshot operation typically takes much less time than does a tape backup, the window of time during which the application must be aware of the backup is reduced. This facilitates more frequent backups, which can reduce the time spent performing forward recovery through transaction logs, increases the flexibility of backup scheduling, and eases administration.

Application availability is also significantly improved due to the reduction of the load on the production servers. Tivoli Storage FlashCopy Manager exploits storage snapshot capabilities to provide high speed, low impact, application-integrated backup and restore functionality for the supported applications and storage environments. Automated, policy-based management of multiple snapshot backup versions, together with a simple and guided installation and configuration process, provide an easy to use and quick to deploy data protection solution that enables the most stringent database recovery time requirements to be met.

More details about IBM Tivoli Storage FlashCopy Manager are available at this website:

https://www.ibm.com/software/tivoli/products/storage-flashcopy-mgr/

9.3.2 FlashCopy Manager (FCM) support

FlashCopy Manager (FCM) exploits disk subsystems such as these:

- DS8000
- SAN Volume Controller
- XIV system
- DS 3/4/5k through VSS only

FCM supports applications such as these:

- Microsoft Exchange 2003 and 2007 on Windows 2003 and Windows 2008
- Microsoft SQL Server 2005 and 2008 on Windows 2003 and Windows 2008
- Oracle 10g and11g on AIX 5.3 and 6.1
- DB2 UDB V9.5 or later on AIX 5.3 and 6.1

- SAP releases running on DB2 V9.5 supported by SAP BRTools 7.10 or later on AIX 5.3 and 6.1
- SAP with Oracle

Prior to FCM, Tivoli Storage Manager for Advanced Copy Services was the typical solution used to minimize the back-up window (reducing the time when the systems were offline to guarantee data consistency for the back-up operation).

Figure 9-7 shows the basic operation of the Tivoli Storage FlashCopy Manager.

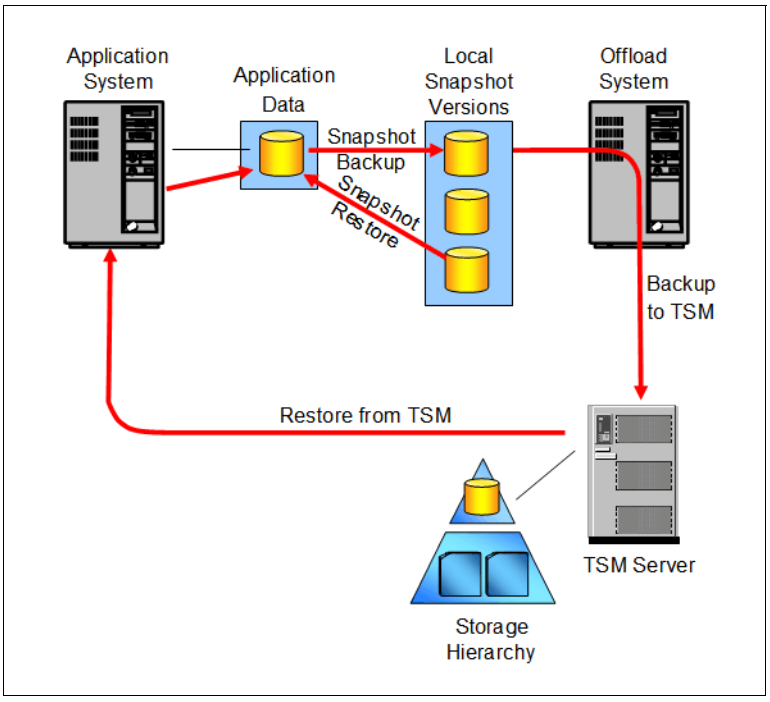

Figure 9-7 Basic operations of IBM Tivoli Storage Manager for Advanced Copy Services

With Tivoli Storage Manager for Advanced Copy Services, you can perform these tasks:

- Trigger a snapshot for backup and map the local snapshot versions to a TSM server
- Transfer outboard of the application server to minimize impact to the application
- Provide long-term retention and disaster recovery with copies on the TSM server
- Provide very fast restore times from the snapshot
- Provide support for multiple, persistent snapshots
- Provide persistent snapshots retained locally
- Provide policy-based management of local, persistent snapshots
- Have differing retention policies for local snapshots and copies on the TSM server
- Automatically reuse local snapshot storage when older snapshot versions expire
- Perform restores:
  - From local snapshot versions
  - From TSM storage hierarchy

In addition to these benefits, for certain environments (such as DB2, SAP, Oracle on AIX, Exchange and SQL on Windows), it is also possible to trigger the snapshot for backup and for restore, and to exploit the IBM storage subsystem, without the intervention of the TSM Server, as shown in Figure 8.

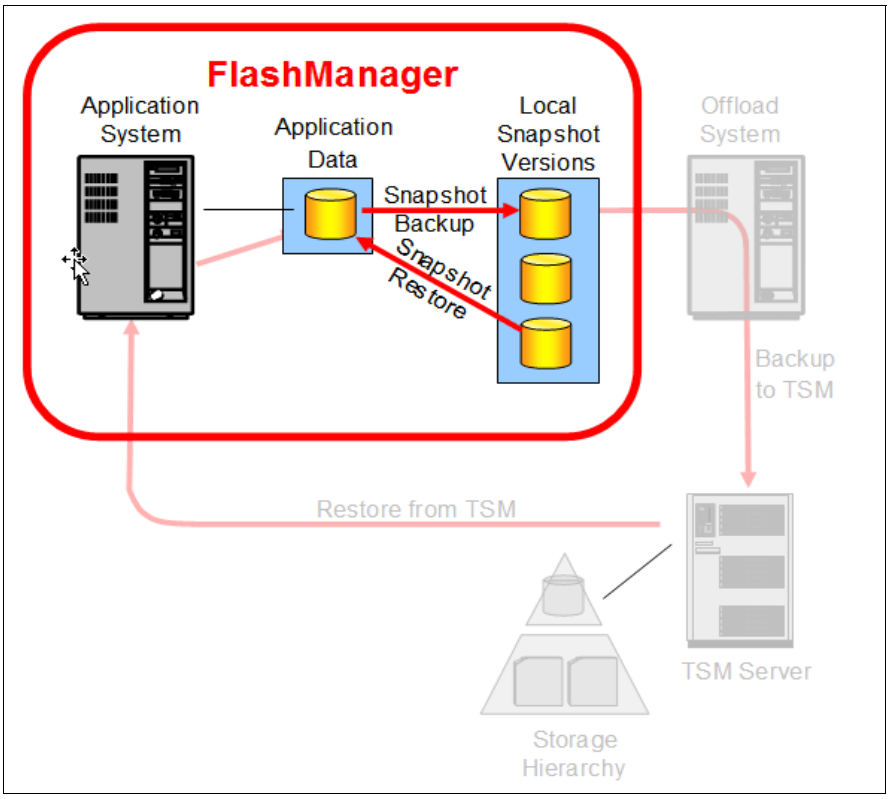

Figure 8 FlashCopy Manager design

9.4 IBM System Storage FlashCopy Manager

IBM System Storage FlashCopy Manager is an IBM program product for IBM z/OS and System z Count Key Data (CKD) IBM storage users. This program assists in the implementation and use of the DS8000, DS6000, and Enterprise Storage Server (ESS) FlashCopy. It is composed of Interactive System Productivity Facility (ISPF), menu-driven application dialogs that generate z/OS batch jobs. In turn, these jobs invoke specialized z/OS programs designed to provide the following functions:

- Simplified data collection, automated configuration, and the creation of z/OS batch jobs that set up and execute a FlashCopy to System z CKD storage command
- Improvement in quality and time to completion for implementing this Advanced Copy Service into any z/OS or System z production and batch CKD storage environment, using any existing z/OS batch job scheduler
- Improvement in System z storage administrator productivity in operating this function
- Ease of use and risk mitigation tool for operationally creating point-in-time copies of CKD volumes concurrently
- Specialized performance enhancements for FlashCopy on System z CKD volumes

In providing these specialized functions, FlashCopy Manager is designed as a supplement to the strategic IBM Business Continuity automation tools of IBM Geographically Dispersed Parallel Sysplex™, FlashCopy Manager can also be used to provide a basic z/OS IBM Copy Services toolkit for environments where the recovery requirement is only for Business Continuity Tier 1 to 4 point-in-time z/OS recovery (that is, less time-critical recoveries, which only require periodic point-in-time copies).

IBM System Storage FlashCopy Manager provides an ISPF, menu-driven application dialog interface, and is designed to provide ease of use. Those who are familiar with ISPF will find the product easy to learn and use.

9.4.1 FlashCopy Manager overview

FlashCopy Manager is a z/OS-based FlashCopy integration, automation, and usage set of tools, invoked through ISPF menu-driven dialogs that are delivered as an IBM software program product. It is designed to provide significant ease of use, automation, performance enhancements, and productivity for z/OS and System z CKD storage users of FlashCopy on the ESS, DS6000, and DS8000.

The FlashCopy Manager tools, Interactive System Productivity Facility (ISPF) menu-driven dialogs, and generated jobs provide the following functions:

- Enable significant FlashCopy performance enhancements in any System z CKD storage environment.
- Provide time and effort savings that lower risk and enable migration from a non-IBM z/OS or System z storage point-in-time copy to FlashCopy.
- Provide a unique set of z/OS-based tools that can accelerate the integration of FlashCopy into any z/OS or System z-based production and batch IBM storage environment.
- Provide a z/OS-based System z enablement toolkit where only lower Business Continuity Tier 1-4 recovery is required (that is, disaster recovery, based on periodic point-in-time copy). See “Typical business continuity tier levels 1 to 4 (for FlashCopy Manager)” on page 171 for further details.

FlashCopy Manager and its companion product, PPRC Manager, provide z/OS batch jobs that are designed to supplement strategic IBM Business Continuity automation tools, such as z/OS TPC for Replication and IBM GDPS®.

Users of z/OS have successfully used the FlashCopy Manager toolkit in large System z IBM storage environments, with thousands of FlashCopy volume pairs.

9.4.2 Typical business continuity tier levels 1 to 4 (for FlashCopy Manager)

FlashCopy Manager provides an interactive set of z/OS-based, ISPF, menu-driven application dialogs, that automatically generate z/OS batch jobs. These z/OS batch jobs invoke specialized z/OS programs for full volume, point-in-time FlashCopy applications. Each batch job:

- Enables significant FlashCopy performance enhancements for z/OS and System z IBM CKD storage
- Enables *operational migration* and provides *risk mitigation* when moving from non-IBM disk point-in-time copy technology to FlashCopy
- Enables simpler integration of full volume FlashCopy operations into z/OS and System z CKD storage customers job streams
- Supports FlashCopy configuration definitions, based on z/OS device numbers, VOLSERs, DFSMS Storage Groups, and VTOC/VTOC indexing matching
- With FlashCopy Manager jobs, provides significant Unit Control Block (UCB) constraint relief by doing a FlashCopy to volumes not in the IOGEN
- Provides the ability to create consistent FlashCopy point-in-time copies from Metro Mirror secondary volumes without impacting Metro Mirror disaster recovery capabilities

9.4.3 FlashCopy Manager solution description

The DS8000, DS6000, and ESS FlashCopy technologies can be exploited to address a wide range of business objectives and requirements. These can include simple tasks, such as creating a point-in-time copy of production z/OS data for transfer from disk to tape. These can also include more complex tasks, such as an on-demand copy of a large online z/OS DB2 database for application development and quality assurance testing, without taking the z/OS DB2 database offline.

FlashCopy Manager is a series of efficient, low overhead REXX, assembler programs and ISPF panels. By using them, the z/OS storage administrator can define, build, and run FlashCopy jobs for any sized z/OS or System z IBM CKD storage FlashCopy environment. In particular, z/OS and System z IBM storage users, who have a large number of FlashCopy pairs, in the hundreds or thousands of FlashCopy pairs, will find FlashCopy Manager of significant value.

FlashCopy Manager provides specialized z/OS FlashCopy invocation programs that provide significantly enhanced FlashCopy initialization times. Particularly, in z/OS or System z IBM storage environments that have a large number of volumes, these performance enhancements can make the use of FlashCopy and FlashCopy Consistency Group (CG) viable; however, it might not have been feasible using a standard Time Sharing Option (TSO), DFSMShsm, or by invoking the ANTRQST macro.

Experience has shown that z/OS FlashCopy environments tend to be highly dynamic. Especially in large z/OS environments, DFSMS allocations are constantly on the move, and database administrators and applications teams are constantly changing and tuning their environments. z/OS and System z IBM storage customers can use FlashCopy Manager:

- To simplify and improve the accuracy of the z/OS or System z IBM storage FlashCopy configuration process, with easy-to-use ISPF dialogs and generated batch jobs.
- To generate z/OS batch jobs to dynamically discover, define, establish, and control the System z IBM storage FlashCopy configuration.
- To support, with FlashCopy Manager jobs, any System z CKD IBM storage volume for which the executing z/OS image can establish shared addressability. FlashCopy Manager has been successfully used in TPF z/VM, and Linux on System z environments.

The FlashCopy Manager toolkit provides significant value to any z/OS or System z IBM storage environment. It provides an automated ability to dynamically discover and reconfigure the System z IBM storage FlashCopy environment on a daily basis, according to the current state of the volumes in the production environment.

You do not need to understand the detailed operation of FlashCopy to use the FlashCopy Manager to configure and manage a FlashCopy environment. You only need to understand how to define and match FlashCopy source and target volumes. The FlashCopy Manager ISPF dialogs provide tools for automated, correct FlashCopy source and target matching. FlashCopy Manager then creates all of the necessary FlashCopy management functions by creating a complete set of z/OS batch jobs.

With this approach, you can submit individual FlashCopy jobs or construct a complex z/OS job stream that includes non-FlashCopy software. You can control the entire process by using standard z/OS job scheduling facilities to accomplish a required set of business objectives. The FlashCopy Manager-created jobs are a key part of that total job stream. In a disk-to-tape process, the FlashCopy Manager-created jobs control the FlashCopy part of the process. The standard disk-to-tape software products complete the disk-to-tape portion of the process. The control data sets that are created by FlashCopy Manager make it easy to dynamically create control commands for the disk-to-tape processes or other processes.

9.4.4 FlashCopy Manager technical highlights

FlashCopy Manager provides a straightforward set of ISPF menu panels, presented in a question and answer format, to obtain the necessary parameters for building FlashCopy jobs.

FlashCopy Manager programs provide the following capabilities:

- Identify FlashCopy source volumes by specification of source or target volumes in the following ways:
  - VOLSER (with wildcard)
  - DFSMS storage groups
  - Device number ranges
- Automatically discover candidate FlashCopy targets by using the same capabilities as described in “FlashCopy Manager solution description” on page 172.
- Correctly match the candidate targets with defined FlashCopy source volumes.
- Build the z/OS jobs that will execute the FlashCopy environment, including (but not limited to) jobs that perform the following tasks:
  - Withdraw FlashCopy relationships
  - Verify the state of active FlashCopy relationships
  - If specified, establish FlashCopy with CG across one or multiple System z IBM disk subsystems
  - If specified, establish Incremental FlashCopy
  - If specified, establish Inband FlashCopy to a remote IBM disk subsystem PPRC secondary
  - Provide significant System z UCB constraint relief by using FlashCopy to copy to volumes not in the IOGEN
  - Query the progress of the FlashCopy and the background copy
  - Run ICKDSF jobs to re-label FlashCopy source volumes
  - Run ICKDSF jobs to re-label FlashCopy target volumes
- Execute FlashCopy Manager jobs from a z/OS image. The FlashCopy source volumes can be any System z CKD format volumes, for which the executing z/OS image can establish shared addressability. FlashCopy Manager has been successfully used on z/TPF, z/VM, and Linux on System z volumes.
- Build jobs for use in a Global Mirror testing and initial implementation in a z/OS or System z IBM storage environment. When used together with PPRC Manager, FlashCopy Manager can also build jobs that perform the following tasks:
  - Perform analysis of the state of the Global Mirror FlashCopy targets, and depending on the results, perform any required processing (REVERT or COMMIT) that is required for individual FlashCopy relationships that need corrective action to create a consistent set of data on all Global Mirror targets (must be CKD volumes).
  - Execute a Fast Reverse Restore FlashCopy process that creates a set of consistent data on suspended Global Mirror PPRC-XD volumes.
- Provide an audit trail of FlashCopy parameter specification.

FlashCopy Manager supports z/OS and System z CKD FlashCopy environments, in which FlashCopy pairs reside in single or multiple IBM storage subsystems. The source and target of any pair must be in the same storage subsystem. Full support of FlashCopy CG functionality provides point-in-time data consistency across multiple storage subsystems.

FlashCopy Manager has special programs that dynamically and quickly obtain device information that is needed to develop the configuration and create the files and jobs necessary for FlashCopy ESTABLISH. FlashCopy Manager is fully aware of, and dynamically discovers and correlates, DS8000, DS6000, and ESS specifics, such as SSID, storage controller serial number, z/OS VTOC structures, and volume size matching. Source and target volumes are matched, based on equal size. FlashCopy Manager discovery programs are also written in assembler language and perform the discovery, matching, and job build process quickly and with low overhead.

FlashCopy Manager jobs use specialized high performance FlashCopy execution programs that initialize FlashCopy with a significantly shorter elapsed time than standard FlashCopy invocation by using DFSMS, TSO, or ANTRQST. This high performance FlashCopy execution capability includes the FlashCopy CG, thus making possible the creation of a point-in-time *consistent copy* of multiple z/OS or System z CKD volumes. FlashCopy Manager's specialized FlashCopy execution programs attain these performance benefits. It does this by using the information gathered during the FlashCopy source-target configuration build process. It then uses the configuration information in specialized invocation programs to achieve a significant reduction in the FlashCopy initialization overhead.

FlashCopy Manager provides an audit trail, so that the definition and running of jobs under its control can be tracked and managed. FlashCopy Manager provides you valuable diagnostic information when each command is executed. It can notify you of any problems involved in issuing FlashCopy commands and report back how long it took to execute the FlashCopy across all volumes or data sets.

9.4.5 FlashCopy Manager components and requirements

FlashCopy Manager has the following software and hardware requirements:

- Hardware:
  - FlashCopy Manager jobs are supported on CKD volumes on DS8000, DS6000, or ESS disk subsystems.

    Support: FlashCopy Manager jobs are not supported on non-IBM storage subsystems.

  - Both source and target FlashCopy volumes must be contained within the same DS8000, DS6000, or ESS. The Point-in-Time or FlashCopy license must be installed.
  - The FlashCopy source volumes must be addressable from the z/OS image running the FlashCopy Manager jobs.
  - FlashCopy Manager jobs support any System CKD IBM storage volume, for which the executing z/OS image can establish shared addressability. FlashCopy Manager has been successfully used in TPF, z/VM, and Linux on System z environments.

    Tip: When using FlashCopy Manager jobs on non-z/OS System z volumes, non-z/OS operating system considerations are outside the scope of FlashCopy Manager.

  - The FlashCopy target devices do not have to be in the IOGEN of the z/OS image that is running the FlashCopy Manager jobs. This can be of significant value for UCB constraint relief in a large System z environment.

- Software:

  At the time of writing, the most current level of FlashCopy Manager is version 5R1.

  - FlashCopy Manager jobs must be run from a z/OS image at z/OS 1.7 or higher, with TSO and ISPF installed. No other pre-requisite software is required.
  - FlashCopy Manager's load library must be Authorized Program Facility (APF) authorized.
  - FlashCopy Manager has no known z/OS system release dependencies.

Tip: FlashCopy Manager only supports volume-level FlashCopy.

9.4.6 Interoperability

The FlashCopy Manager toolkit is designed to be a stand-alone product, but can be used with other products in a z/OS or System z environment, such as Geographically Dispersed Parallel Sysplex (GDPS), and GDPS HyperSwap, z/OS TPC for Replication, or PPRC Manager.

FlashCopy Manager specialized z/OS programs that provide significant FlashCopy initialization performance enhancements can be of high value to any z/OS user. This includes z/OS users of TPC for Replication, GDPS, and GDPS HyperSwap, or Peer to Peer Remote Copy (PPRC, also known as Metro Mirror) Manager.

The following examples indicate the use of FlashCopy Manager with other products in a z/OS environment:

- GDPS, GDPS HyperSwap:
  - FlashCopy of Metro Mirror, Global Mirror, or z/OS Global Mirror primary for local backup purposes
  - Supplemental options for FlashCopy of a Metro Mirror, Global Mirror, or z/OS Global Mirror (also known as eXtended remote Copy (XRC)) secondary
  - Specialized FlashCopy performance enhancements for any z/OS volumes running in the same System z complex
  - Specialized FlashCopy performance enhancements for non-z/OS CKD volumes running in the same System z complex (such as Linux on System z, z/VM, or z/TPF)
- z/OS TPC for Replication:
  - FlashCopy of Metro Mirror or Global Mirror secondary
  - FlashCopy of Metro Mirror or Global Mirror primary
- PPRC Manager:
  - Definition of FlashCopy targets for a Global Mirror configuration
  - FlashCopy of Metro Mirror, Global Copy, or Global Mirror secondary
  - FlashCopy of Metro Mirror, Global Copy, or Global Mirror primary

9.4.7 Summary

FlashCopy Manager provides z/OS and System z users with an improved FlashCopy user interface through easy-to-use ISPF, menu-driven application dialogs, generating batch jobs that invoke specialized z/OS programs for dynamic discovery, build, and execution of System z FlashCopy environments. The goal is to enhance the use of z/OS and System z FlashCopy technology on DS8000, DS6000, and ESS CKD volumes.

FlashCopy Manager is designed to perform the following functions:

- Enable the ability to operationally migrate from non-IBM storage point-in-time copy environments to IBM storage
- Provide an enablement toolkit to accelerate integration and exploitation of FlashCopy into any z/OS production and batch environment
- Enable significant System z FlashCopy performance enhancements

9.4.8 Interoperability with FlashCopy Manager

PPRC Manager is designed to be a stand-alone program product. However, it can be used with FlashCopy Manager to provide a more comprehensive and complete set of tools for defining z/OS or System z CKD DS8000, DS6000, or ESS Advanced Copy Services configurations.

FlashCopy Manager can be used together with PPRC Manager to provide the following tools:

- *Global Mirror:* FlashCopy Manager builds all of the jobs that are required to set up the FlashCopy environment for a Global Mirror configuration. Additionally, FlashCopy Manager builds all the jobs required to perform the Revert/Commit recovery analysis and the Fast Reverse Restore FlashCopy process.
- *InBand FlashCopy:* With the Inband FlashCopy process, FlashCopy commands can be issued to remote secondary devices. This allows for FlashCopy of a PPRC secondary device at a remote site that does not have host connectivity at the remote site.
- *FlashCopy of PPRC primary:* FlashCopy Manager can perform as a FlashCopy of PPRC primary device, where the PPRC primary device is the FlashCopy source. This is useful in cases where a point-in-time copy is desired before data at the primary site is modified.
- *FlashCopy of PPRC secondary:* FlashCopy Manager can perform as FlashCopy of PPRC secondary devices, where the PPRC secondary device is the FlashCopy source. This is useful in cases where a point-in-time copy of the data at the remote site is desired.

9.4.9 Tools and toolkits for FlashCopy Manager

FlashCopy Manager is a *tool*. You might think of it as a *toolkit* that is designed to be an optional z/OS-centric supplement to the previously described strategic business continuity automation offerings. This tool is typically used when specialized System z FlashCopy uses and functions are needed that strategic business continuity automation solutions are not designed to provide.

Purpose of FlashCopy Manager

FlashCopy Manager is designed to supplement the IBM strategic z/OS TPC for Replication and GDPS solutions by addressing these specialized additional requirements. FlashCopy Manager is intended to provide z/OS-based software tools that can perform these functions:

- Enable z/OS and System z storage technology operational migrations from non-IBM copy services to IBM copy services
- Enable z/OS and System z data center migrations
- Enable z/OS IBM disk subsystem reconfiguration, with associated data movement
- Enable System z point-in-time FlashCopy initialization speed enhancements
- Enable integration and exploitation of IBM Advanced Copy Services into the production z/OS batch environment, using any existing z/OS job scheduler

- Address situations where System z client requirement is only for lower Business Continuity Tiers 1 to 4 disaster recovery (that is, only periodic point-in-time copy is required for recovery)

FlashCopy Manager is intended as a supplement, not a replacement, for z/OS TPC for Replication, GDPS, and GDPS HyperSwap Manager.

When FlashCopy Manager is used

FlashCopy Manager is typically used, often as a supplement to z/OS TPC for Replication or GDPS, when a z/OS or System z IBM storage client has the following requirements:

- A z/OS toolkit to enable fast migration and that lowers the risk of the operational migration, from a non-IBM z/OS or System z storage point-in-time copy to FlashCopy
- A set of unique z/OS tools to integrate and dynamically manage FlashCopy into the existing z/OS production and batch environment
- Significant z/OS FlashCopy performance enhancements:
  - Compared to standard IBM TSO, DFSMS, or ANTRQST invocation methods
  - Compared to non-IBM point-in-time copy services
- Significant z/OS or System z UCB constraint relief to enable the use of FlashCopy
- An affordable z/OS-base recovery enablement toolkit for System z, when the requirement is only for lower criticality Business Continuity Tier 1-4 local point-in-time recovery (that is, periodic point-in-time copy for disaster recovery)
- A z/OS-familiar ISPF interface, rather than a GUI, is preferred for the control and management of FlashCopy

9.4.10 Summary

FlashCopy Manager and PPRC Manager are IBM program products that consist of ISPF menu-driven dialogs, automated batch job generator, and specialized programs. You can think of them as toolkits for z/OS and System z CKD IBM storage users to accelerate the implementation and use of DS8000, DS6000, or ESS FlashCopy, Metro Mirror, Global Copy, Metro/Global Copy, and Global Mirror in any z/OS and System z CKD IBM storage environment. These toolkits are designed to provide the following benefits:

- Simplified data collection, automated configuration, and creation of z/OS batch jobs that set up and execute a FlashCopy, Metro Mirror, Global Copy, Global Mirror, or Metro/Global Copy to System z CKD storage
- Improvement in quality and time to completion for implementing these IBM Advanced Copy Services in a z/OS or System z CKD IBM storage environment
- Improvement of z/OS storage administrator productivity to operate these functions
- A z/OS-based set of tools for ease of use and to lower the risk of operationally migrating from non-IBM point-in-time disk copy or a non-IBM disk mirroring to an IBM storage environment
- Specialized performance enhancements for z/OS FlashCopy
- Tools for storage technology migrations
- Tools for data center migrations
- Tools to accelerate integration of IBM Advanced Copy Services into any z/OS production and batch environment and the use any existing z/OS batch job scheduler

9.5 Backup and recovery for the IBM System Storage N series

In this section, we briefly discuss the N series internal backup and recovery features that let users to restore data quickly and transparently, even without any manual intervention by an N series administrator. Using the optional host-based tools, such a SnapManager®, you can easily access your own snapshot backups from the client machine and choose the appropriate point-in-time version to recover.

The following N series backup and recovery tools are introduced in this section:

- Snapshot and SnapManager
- SnapRestore®
- SnapVault®

This section does not discuss features that are included in other chapters of this Redbooks publication, because those functionalities are from other categories of provided services, such as disaster recovery for continuous operation or long-term retention and archiving for business or legal compliance purposes. For this information, see the following links:

- *SnapLock:* Provides file-level data protection for long-term archiving or backup solution for legal or business compliance. This is described in “IBM System Storage N series SnapLock” on page 222.
- *SnapMirror®:* Offers data protection against disaster by replicating the data to another N series, either locally or at a remote site. In this case, continuous operations is the most significant benefit of SnapMirror. Its function is briefly introduced in “Replication to remote site” on page 224.
- *Tape backup:* This is not an integrated part of the N series solution. It utilizes the Network Data Management Protocol (NDMP) to provide external image backups of given filesystem to tape, for example using existing Tivoli Storage Manager backup infrastructure. The solution is presented in section “Tape backup” on page 225.

9.5.1 Snapshot and SnapManager

N series snapshot performs read-only clones of the active filesystem. It enables online backups to be maintained, thus providing nearly-instantaneous access to previous versions of data without requiring complete, separate copies or resorting to auxiliary backups.

Snapshot technology makes extremely efficient use of storage by storing only block-level changes between each successive snapshot. In this way, it can be simply thought of as being similar to any modern source code control system that maintains only the changes made to the original source code to minimize space and maximize recoverability.

Because the snapshot process is automatic and virtually instantaneous, backups are significantly faster and simpler. Snapshots can also be coordinated with outside applications to ensure highly consistent data states (as viewed from the application), prior to performing snapshot and other backup procedures. For example, flushing data from a production database prior to snapshot creation is a generally recognized best practice.

The N series uses the Write Anywhere File Layout (WAFL) feature to implement snapshots as a copy-on-write technique to minimize the disk space that snapshots consume.

The WAFL feature creates and deletes snapshots automatically at scheduled times, and keeps up to 255 snapshots (as illustrated in Figure 9-9) online simultaneously to provide easy access to old versions of files.

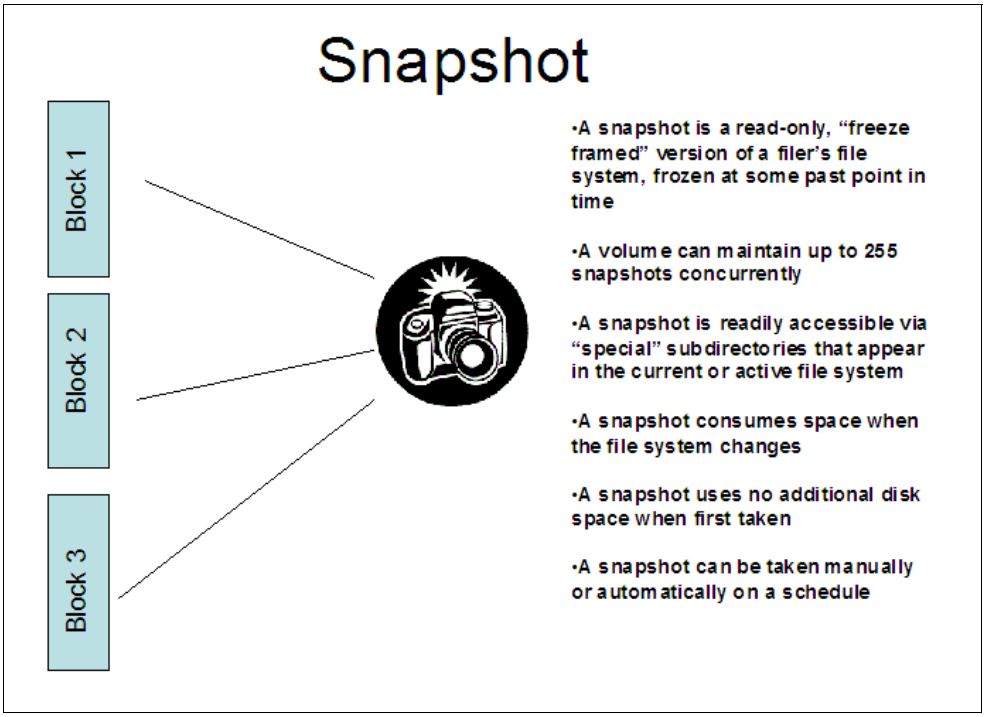

Figure 9-9 Snapshot features

Snapshots

Snapshots use a redirect-on-write technique to avoid duplicating disk blocks that are the same in a snapshot as in the active filesystem. Only when blocks in the active filesystem are modified or removed do snapshots containing those blocks begin to consume disk space.

Users can access snapshots to recover files that they have accidentally changed or removed, and system administrators can use snapshots to create backups safely from a running system. In addition, WAFL uses snapshots internally so that it can restart quickly even after an unclean system shutdown.

Figure 9-10 shows an example of the snapshot configuration window in FilerView® of the N series. You can configure, perform, or delete any of the snapshots in this window, including the job scheduling window.

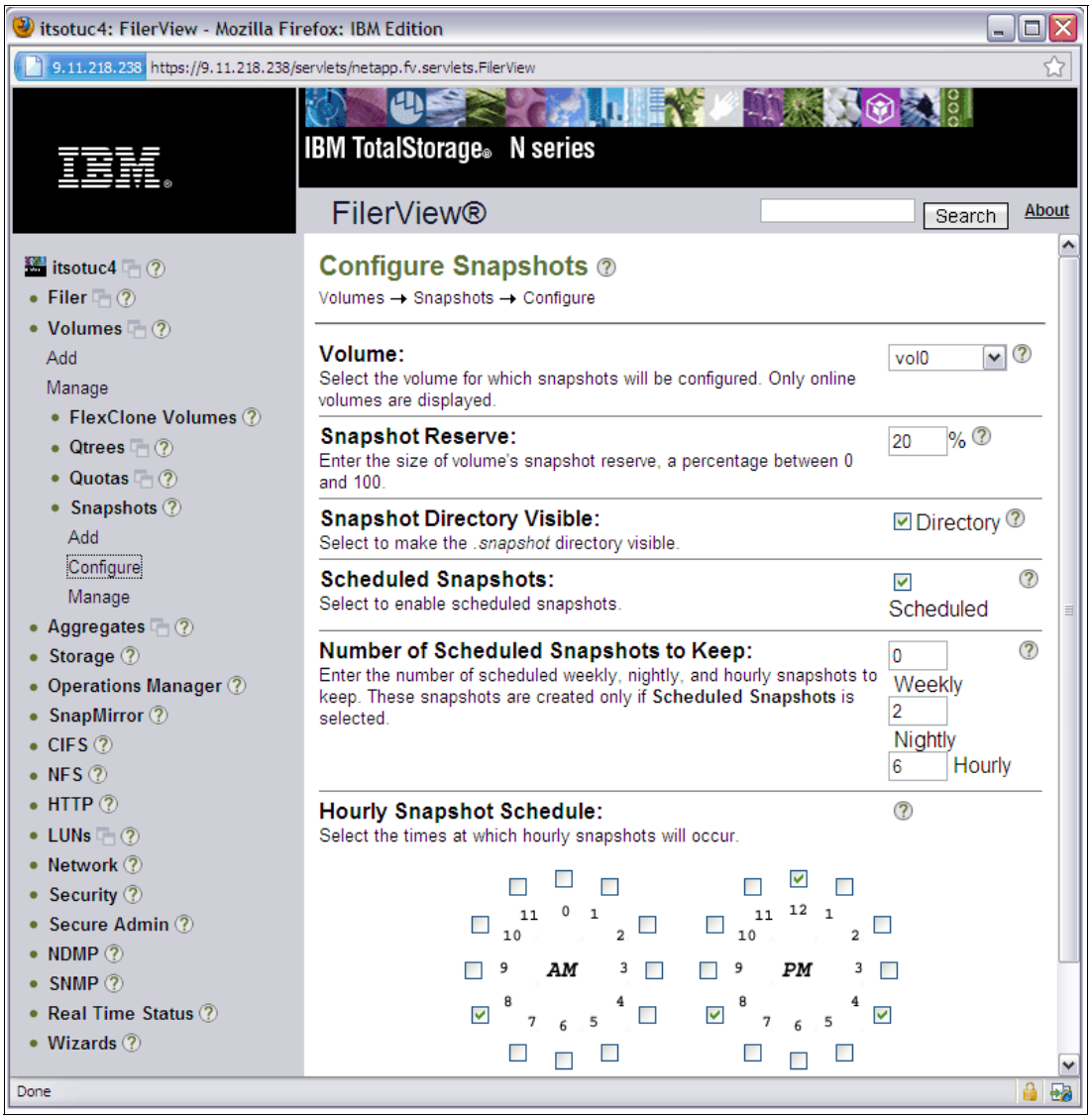

Figure 9-10 Snapshot configuration window, viewed with the N series FilerView

SnapManager

The N series SnapManager is the host software for managing snapshots for backup and restore operations. There are multiple versions of SnapManager that integrate easily with critical applications, such as these:

- SnapManager for MS Exchange
- SnapManager for MS SQL Server
- SnapManager for MS SharePoint
- SnapManager for Oracle
- SnapManager for SAP
- SnapManager for Hyper-V
- SnapManager for Virtual Infrastructure which automates and simplifies backup and recovery of primary storage used by VMWare Virtual Infrastructure

9.5.2 SnapRestore

The N series SnapRestore software is designed to help an enterprise recover data quickly in the case of a disaster. SnapRestore software can help recover data in amounts as small as an individual file, up to a multi-terabyte volume, such that user operations can be quickly resumed.

SnapRestore capabilities

SnapRestore uses the snapshot feature of Data ONTAP software to restore a file, the entire filesystem, or a LUN to an earlier state. It is used to recover a damaged or deleted file, or to recover from a corrupt database, application, or damaged filesystem.

The system administrator can restore a file, or the entire filesystem, LUN, or entire volume, from any existing snapshot without rebooting. The restored file, volume, filesystem, or LUN is available for full production use, having returned to the precise state that existed when the snapshot was created. From a single home directory, to a huge production database, SnapRestore does the job in seconds, regardless of the size of the file or volume.

Using SnapRestore, you can roll back data to the instant a snapshot was taken. That means that, rolling back in time by restoring from a snapshot, has the effect of wiping out any changes made to the volume after the snapshot.

SnapRestore is a data recovery facility available for the N series. It can be used in the following scenarios:

- Disaster recovery
- Database corruption recovery
- Application testing, such as a development environment using large data files

If a client application corrupts data files in a volume, you can revert the volume to a snapshot taken before the data corruption. After you revert a volume to a specific snapshot, you will lose snapshots that are more recent than the snapshot used for the volume reversion.

SnapRestore command

The N series FilerView does not support the restore of snapshots, so we use the Data ONTAP Command Line Interface (CLI) for this.

The **snap restore** command has the following syntax:

```
snap restore [ -f ] [ -t vol | file ] [ -s snap_shot_name ]
[ -r restore_as_new_path ] vol_name restore_from_path
```

This command reverts a volume to a specified snapshot, or reverts a single file to a revision from a specified snapshot.

The **snap restore** command is only available if your storage system has a SnapRestore license. If you do not specify a snapshot, the N series prompts you for the snapshot.

A volume cannot have both a volume SnapRestore and a single file SnapRestore executing simultaneously. Multiple single file SnapRestores can be in progress simultaneously, however.

In a situation where a CIFS-share, NFS-share, file or LUN is recognized as being damaged and a SnapRestore from an existing snapshot is being considered, then we need to know which snapshots exist, so that we can pick the right one to restore from.

Use the **snap list** command to check existing snapshots for a given volume (Example 9-3).

Example 9-3 Available SnapShots for SnapRestore operation

```
itsosjc> snap list cifs_vol7945
Volume cifs_vol7945
working...

  %/used       %/total  date          name
----------  ----------  ------------  --------
  0% ( 0%)    0% ( 0%)  Apr 05 12:00  hourly.0
  0% ( 0%)    0% ( 0%)  Apr 05 08:00  hourly.1
  0% ( 0%)    0% ( 0%)  Apr 05 00:00  nightly.0
  0% ( 0%)    0% ( 0%)  Apr 04 12:00  hourly.2
  0% ( 0%)    0% ( 0%)  Apr 04 08:00  hourly.3
```

The snapshot schedule is determined at the creation of the volume. The default snapshot schedule saves six hourly and two nightly snapshots.

9.5.3 SnapVault

The N series SnapVault software is a reliable and economical way to protect enterprise data, and it offers many significant advantages over traditional backup methods. Although SnapVault can be deployed in configurations designed to emulate the previous backup methods it replaces, the full value of the solution can be realized only by making a significant shift in the way you think about backup and recovery. SnapVault is so useful that it renders many common backup policies and schedules obsolete.

SnapVault is a separately licensed feature in Data ONTAP that provides disk-based data protection for storage systems. The SnapVault server runs on the IBM System Storage N series platform. However, you can use an N series as a SnapVault client as well.

SnapVault replicates selected snapshots from multiple client storage systems to a common snapshot on the SnapVault server, which can store many snapshots. These snapshots on the server have the same function as regular tape backups (see Figure 9-11). Periodically, data from the SnapVault server can be dumped to tape for extra security, using, for example, the existing Tivoli Storage Manager SAN infrastructure.

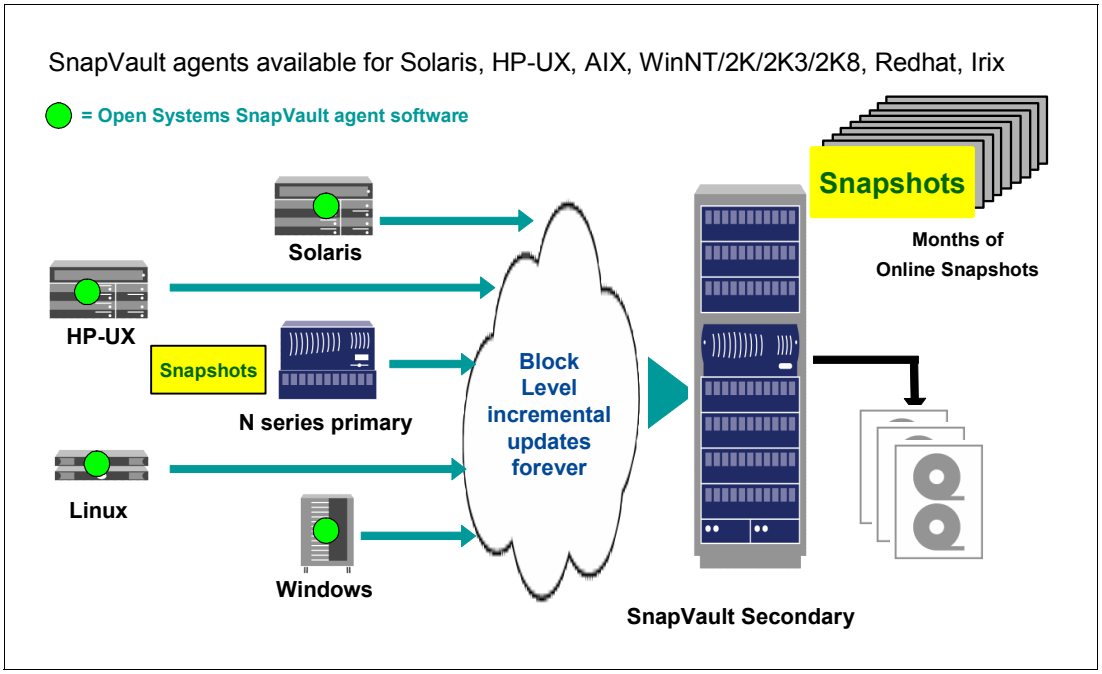

Figure 9-11 SnapValue conceptual model

SnapVault purpose

SnapVault is a heterogeneous, disk-to-disk data protection solution, ideal for use with N series. A SnapVault primary system corresponds to a backup client in the traditional backup architecture. The SnapVault secondary system is always an N series that is running Data ONTAP. SnapVault software protects data that resides on a SnapVault primary system.

All of this heterogeneous data is protected by maintaining online backup copies (or snapshots) on a SnapVault secondary system. The replicated data on the secondary system can be accessed through the NFS or CIFS, just as regular data can be. The primary systems can restore entire directories or single files directly from the secondary system. There is no corresponding equivalent to the SnapVault secondary in the traditional tape-based backup architecture.

SnapVault benefits

In summary, the N series SnapVault protects data on a SnapVault primary system (also known as a SnapVault client) by maintaining a number of read-only versions of data on a SnapVault secondary system (known as SnapVault server). The following benefits are visible:

- Forever incremental backup: Copies only the changes in a data set. The first initial backup is considered a full backup, then SnapVault performs only incremental backups to cover all changes in the filesystems. This approach is similar to the progressive, incremental backup of Tivoli Storage Manager.
- Self-service restore: Users do not require special software or privileges to perform a restore of their own data. Users who want to restore their own data can do so without the intervention of a system administrator, saving time and money. To restore from a SnapVault secondary volume, connectivity to the secondary must be in place. comment: duplicate of text in “Backup and recovery for the IBM System Storage N series” on page 178.
- Consistent security: SnapVault stores backup copies of the data in a WAFL filesystem, which replicates all of the file permissions and access control lists held by the original data. Users who are not authorized to access a file on the original filesystem are not authorized to access backup copies of that file. This allows the self-service restores described earlier to be performed safely.

9.5.4 More information

For more information, see the following documentation:

- Additional details about backup and recovery configuration can be obtained from the comprehensive Redbooks publication, *IBM System Storage N series Software Guide*, SG24-7129
- General information is available at this website:

  `http://www.ibm.com/systems/storage/network/`

10

Continuous operations

In this chapter, we discuss IBM products that can help ensure continuous operations in your environment. This means continuous application and data availability during exceptions, failures, or disasters.

We provide a general approach to the concept of remote copy and how it is implemented in several IBM storage products. This chapter does not include disaster recovery scenarios. However, we describe mirroring technologies and actions that can be taken when switching to either a backup site or backup storage systems, following a failure, and then switching back after the environment has returned to normal (failover and failback).

Rather than using manual processes, it is better for the failback and failover operations to be automated as much as possible, especially in the event of a disaster. The solutions or products described in this chapter can help you achieve that automation include these:

- PowerHA System Mirror for AIX Enterprise Edition (formerly known as High Availability Cluster Multi-Processing Extended Distance (HACMP/XD)): A stretched cluster for AIX, based on the DS8000 Metro Mirror and Global Mirror.
- Open HyperSwap for AIX with IBM Tivoli Storage Productivity Center for Replication V4.2: A TPC-R based solution to automatically failover I/O from primary logical devices to secondary logical devices in the event of a primary disk storage system failure.
- Geographically Dispersed Open Clusters (GDOC): A dispersed cluster solution, based on the Veritas cluster server.
- IBM Tivoli Storage FlashCopy Manager: Fast application-aware backup and restore solution using FlashCopy technology for IBM DB2, Oracle, SAP, Microsoft Structured Query Language (SQL) Server, and Exchange.

- Tivoli Storage Manager FastBack: Provides advanced data protection and near-instant recovery for critical Windows and Linux servers, including selected database systems, and offers bare-machine recovery in the event of hardware failure.

  Tip: Tivoli Continuous Data Protection for Files has been replaced by Tivoli Storage Manager Fastback, which extends the supported platforms to Linux.

- Geographically Dispersed Parallel Sysplex/Peer-to-Peer Remote Copy HyperSwap Manager (GDPS/PPRC HM): A single point of management and control that provides planned and unplanned switching capabilities to either a backup site or backup storage systems, non disruptively. GDPS/PPRC HM is designed to work in System z environments only.
- Tivoli Storage Productivity Center for Replication (TPC-R) with HyperSwap: Combines the TPC-R GUI-based management and control technology with the capability to swap volumes between servers within seconds, providing planned or unplanned switching functionality to the backup site.

10.1 Continuous operations

Today's users expect to be able to tap into information at any time of day, from any location. At the same time, businesses must be increasingly sensitive to issues of client privacy, data security, and regulatory requirements.

To keep your operations running and your business competitive, you need a comprehensive strategy that addresses the three primary aspects of business continuity: high availability, disaster recovery, and continuous operations.

But how do you determine the resiliency and recovery requirements of your business? How do you identify and integrate critical business and IT priorities into a comprehensive continuity and resiliency program? How do you maintain continuous operations?

IBM Business Continuity and Resiliency Services helps to ensure the continuity of business operations and assists with regulatory compliance, improved systems availability, data protection, and the integration of IT operational risk management strategies. The business continuity and resiliency services portfolio includes these services:

- *IBM Resiliency Consulting Services:* Help you to assess, design, plan, implement, and test the resiliency of your business infrastructure
- *IBM Managed Resiliency Services:* Provide you with a wide range of proactive and event-driven managed services to meet your recovery time and Recovery Point Objectives (RPOs) in the event of a disaster
- *IBM Information Protection Services:* Enable you to quickly back up, restore, and archive critical data, information, and systems on demand
- *IBM Infrastructure Recovery Services:* Help you to recover IT and business infrastructures, including work areas, from disruptions, with the goal of reducing costs and timeframes
- *IBM System Storage Resiliency Family:* Solutions that can help you to protect critical business assets, and align recovery costs, based on patented business impact and information value

IBM provides a wide range of proactive and event-driven managed services, so you can select the capabilities you need, based on the criticality of your data and processes, including lower-cost options. By managing and operating these services for you, either fully or partially, IBM can enable you to balance workloads, lower application and system downtime, and reduce data loss. At the same time, IBM can help you to avoid capital expenses, monitor and manage operational expenses and service levels, and reduce the burden on your IT staff.

With businesses becoming more complex and interconnected, the risk and cost of disruption extends well beyond IT, and into every aspect of your business processes.

10.2 Remote Copy Services for continuous operations

Table 10-1 lists the IBM storage products and the name of the remote copy technology that implements the specific type of copy service. Each technology and product is discussed in this chapter.

Table 10-1 Remote Copy Services functions in IBM storage products

	Synchronous with data consistency	**Asynchronous**	**Asynchronous with data consistency**
DS8000	Metro Mirror	Global Copy	Global Mirror
SAN Volume Controller	Metro Mirror	Global Mirror	
V7000	Metro Mirror	Global Mirror	
XIV system	Synchronous Remote Copy	Asynchronous Remote Copy	Asynchronous Remote Copy
SONAS			Asynchronous Replication

Tips:

- For the XIV system to have asynchronous remote copy volumes consistent, the volumes must be defined in the same consistency group (CG).
- Global Mirror is the name for asynchronous replication *without* data consistency for the SAN Volume Controller and V7000, but it is name for asynchronous replication *with* data consistency for the DS8000.

10.2.1 Synchronous remote copy services

Synchronous remote replication ensures consistent data copy from primary storage to secondary storage, in most cases to a remote data center.

Metro Mirror

Metro Mirror, formerly known as Synchronous PPRC, provides real-time mirroring of logical volumes between two storage systems that can be located up to 300 km from each other. It is a synchronous copy solution, where write operations need to be acknowledged and successful on both the local and remote sites before they are considered complete. This mirroring technique is typically used for applications that cannot suffer any data loss in the event of a site or storage failure.

As data is synchronously transferred, the distance between the local and the remote disk has significant impact on application response time.

A write operation has the following phases (here specifically described for the DS8000):

- Write to source volume (storage system cache and Non Volatile Storage, or NVS)
- Write to target volume (storage system cache and NVS)
- The remote target storage system acknowledges that the write is complete
- The primary disk storage system acknowledges that the I/O is complete

The write operation is depicted graphically in Figure 10-1.

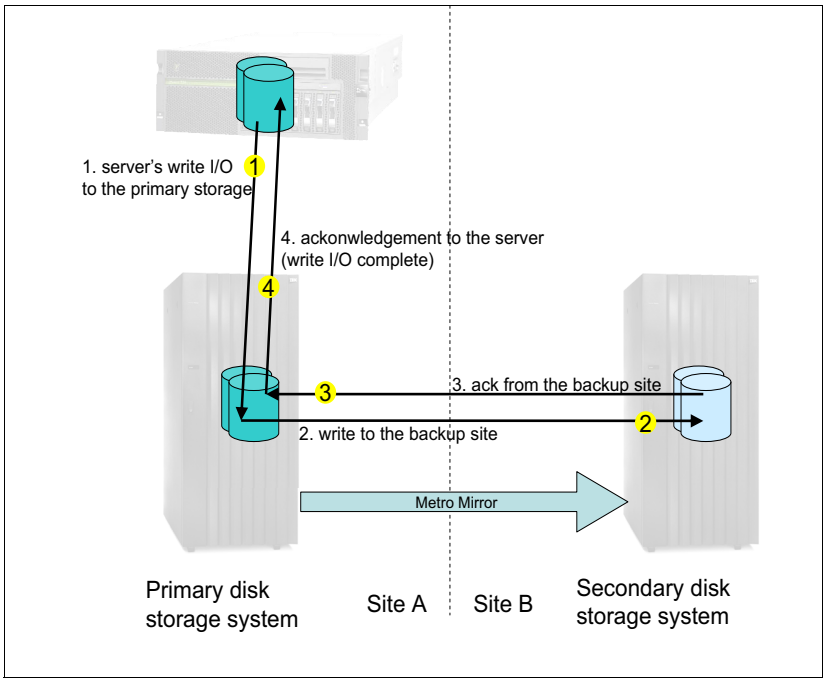

Figure 10-1 Metro Mirror write I/O phases

In the event of a disaster, it is unlikely that the entire complex will fail at the same moment. Failures tend to be intermittent and gradual, and a disaster can occur over many seconds, even minutes. Because a portion of the data might have been processed and other data lost in this transition, data integrity on the target volumes is exposed. This is called a *rolling disaster*. The mirrored data at the recovery site must be managed so that cross-volume consistency is preserved during the intermittent or gradual failure. In order to restart applications at the remote site successfully, the remote site volumes must have consistent data. For Metro Mirror, consistency requirements are managed through use of the *Consistency Group (CG)* option.

Because the disk attached to the server is being mirrored using Metro Mirror, this offers improved opportunities for high availability solutions. For open system environments, IBM offers several solutions in this area, including GDOC for Open System environments, and the PowerHA for AIX, both discussed in "PowerHA System Mirror for AIX Enterprise Edition" on page 195.

XIV system Synchronous Remote Copy

The Remote Mirroring function of the XIV system provides a real-time copy between two or more storage systems supported over Fibre Channel (FC) or iSCSI links. This feature provides a method to protect data from site failures. Remote Mirroring can be a synchronous copy solution where write operations are completed on both copies (local and remote sites) before they are considered to be complete. This type of remote mirroring is normally used for short distances to minimize the effect of I/O delays inherent to the distance between sites.

Principles for synchronous remote copy are the same as described in section "Metro Mirror" on page 188. For data consistency reasons, if there is more than one volume that needs to keep dependent write order, source to target volumes relationships are coupled in a logical object, called a Continuity Group (CG). The CG ensures that the write order to the source volume is preserved in the write order to the target volumes.

Two XIV systems, each at a separate site, might have different volumes mirrored in a bidirectional configuration. This configuration can be used for situations where there are two active production sites, with each site providing a disaster recovery solution for the other. Each XIV system is active as a production system for certain peers and as a mirroring target for other peers. XIV system mirroring provides an option to automatically create slave volumes.

XIV system mirroring is data aware. Only actual data is replicated. If 100 GB volume has only 10 GB of data written, only that 10 GB is replicated.

After a cross-site connection failure, synchronous mirroring automatically resynchronizes volume pairs when the connection recovers. Before a resynchronization process is initiated, the system creates a snapshot of the slave volume/CG. A snapshot is created to ensure the usability of the slave volume CG in the event of a primary site disaster during the resynchronization process. If the master volume CG is destroyed before resynchronization is completed, the slave volume CG might be inconsistent, as it might have been only partially updated. To handle this situation, the secondary XIV system always creates a snapshot of the last consistent slave volume CG after reconnecting to the primary XIV system and before starting the resynchronization process. The snapshot is preserved until all volumes of the CG are synchronized.

10.2.2 Asynchronous Remote Copy Services

This section describes asynchronous remote copy services for theDS8000, the XIV system and SONAS respectively.

IBM System Storage DS8000 Global Copy

Global Copy, formerly know as PPRC Extended Distance (PPRC-XD), is a non-synchronous remote copy function for open systems and System z environments. It is used for longer distances than are possible with Metro Mirror. It is purposed for remote data migration, off-site backups, and transmission of inactive database logs at virtually unlimited distances.

As shown in Table 10-1 on page 188, the naming convention for asynchronous replication within our products differs. For the DS8000, it is called Global Copy. For the SAN Volume Controller and the V7000, it is called Global Mirror. Global Mirror does exist for DS8000 as well, but it stands for asynchronous replication *with consistent data* on a target.

Data is written locally on a storage system, and a write acknowledgment is returned immediately. A local storage system keeps track of data changes until the data is replicated to the secondary storage system.

A storage system algorithm evaluates which data is changed frequently. It is inefficient to send frequently changed data to the second storage system immediately, as this wastes resources in capacity and performance of the primary storage. Although the secondary copy is not synchronous with the primary data, it is a near-synchronous copy.

Figure 10-2 depicts this process.

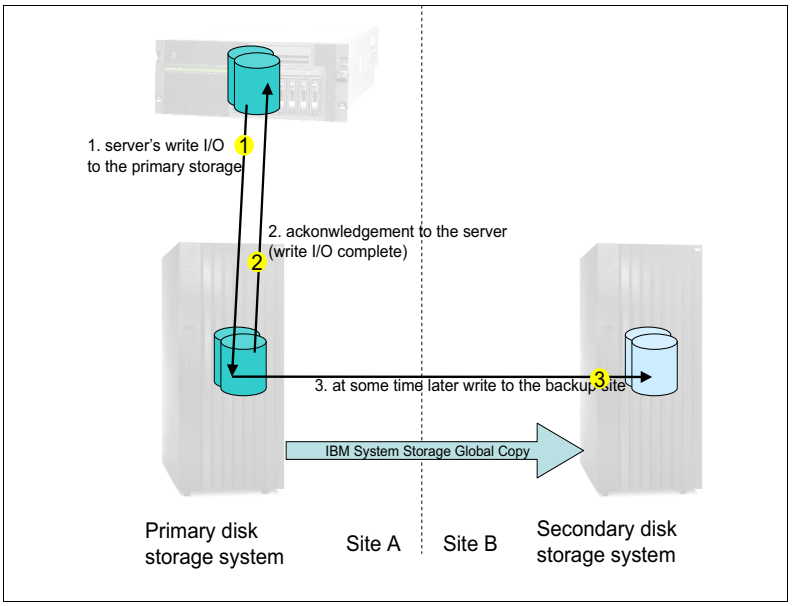

Figure 10-2 Asynchronous remote copy

IBM System Storage DS8000 Global Mirror

As shown in Table 10-1 on page 188, the name *Global Mirror* has different meanings for different products. For the DS8000, it is an asynchronous mirroring with consistency. For the SAN Volume Controller and V7000, the target copy cannot be considered as consistent. See "IBM System Storage DS8000 Global Copy" on page 190.

In the DS8000, Global Mirror uses Global Copy technology to send data to the second storage system and uses FlashCopy or FlashCopy SE to periodically create a copy of data as shown in Figure 10-3.

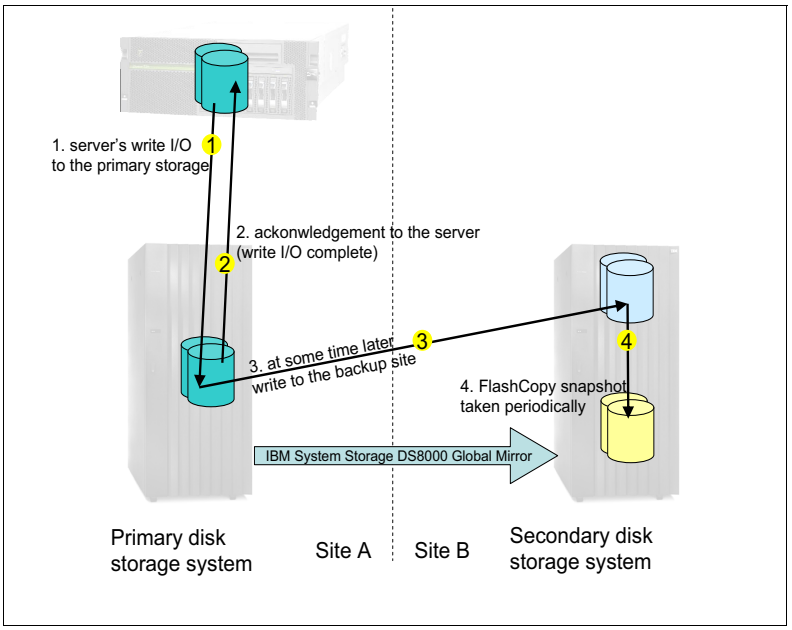

Figure 10-3 IBM System Storage DS8000 Global Mirror

To create the consistent data copy, all writes to the source volumes are paused for a brief period of time. Then, all out-of-sync data is transmitted to the secondary volumes (drained) and FlashCopy automatically initializes. With this approach, a Recovery Point Objective (RPO) of seconds can be achieved, with a very low impact on the primary storage response time.

The interval between consistent data snapshots (FlashCopies) at the secondary storage system is tunable and can be defined from 0 seconds (the default) to 18 hours. Zero seconds means that, immediately after FlashCopy initialization finishes, another consistent data copy is created. You can have an RPO of seconds in this way.

If out-of-sync data cannot be copied within an acceptable time (defined as *drain timeout*, which is 30 seconds by default), then Global Mirror operation will fail. This can happen if, for example, the connection between the sites fails, or the bandwidth is insufficient for the amount of changed data in the snapshot interval. In these cases, primary storage switches from Global Mirror to Global Copy, and the update to secondary storage continues. FlashCopy will not be run; thus, the last consistent data on the secondary storage system is intact. The primary disk storage system will evaluate when it might be possible to continue in Global Mirror and will switch to this mode automatically at that time.

XIV system asynchronous Remote Copy

Asynchronous mirroring enables replication between two XIV system volumes or CGs. This type of mirroring does not suffer from the latency inherent in synchronous mirroring. For this reason, asynchronous mirroring yields better responsive time from the source system.

Asynchronous mirroring is initiated at a defined interval, known as a sync job schedule, or a sync job. A sync job is the synchronization of data updates that have been recorded on the master volume since the last successful synchronization. The sync job schedule is defined for both the primary and secondary system peers in the mirror. This provides a schedule for each peer and is used when the peer takes on the role of the master. The purpose of the schedule specification on the slave volume is to set a default schedule for an automated failover scenario.

The system supports schedule intervals from 20 seconds up to 24 hours. Each sync job schedule needs to be based on your RPO requirements.

A schedule set to *NEVER* means that no sync jobs will be automatically scheduled. In addition to schedule-based snapshots, a dedicated command to run a mirror snapshot can be issued manually. These ad-hoc snapshots are issued from the master volume and initiate a sync job that is queued behind any outstanding sync jobs.

When a link failure occurs, the primary system starts tracking changes of the mirror source volumes, so that these changes can be copied to the secondary system, after being recovered. When recovering from a link failure, asynchronous mirroring sync jobs proceed as scheduled. Sync jobs are restarted and a new, most-recent snapshot is taken. The primary system copies the changed data to the secondary volume.

SONAS asynchronous replication

Asynchronous replication allows filesystems within an SONAS file name space to be defined for replication to another SONAS system usually in another location over the customer network infrastructure.

The asynchronous replication process looks for changed files in a defined filesystem of the source SONAS since the last replication cycle was started against it, and uses the modified rsync tool to efficiently move only the changed portions of a file from one location to the next.

In addition to the file contents, all extended attribute information about the file is also replicated to the remote system.

The file based copy allows the source and destination file trees to be of differing sizes and configurations, as long as the destination filesystem is large enough to hold the contents of the files from the source.

File shares defined at the production site are not automatically carried forward to the secondary site and must be manually redefined by the customer for the secondary location. These shares must be defined as R/O until a disaster occurs on the primary site.

Asynchronous replication is defined in a single direction per file set, such that one site is considered the source of the data, and the other is the target. The replica of the filesystem at the remote location can be used in a Read-only Mode only, until it is to become usable as a main copy in the event of a disaster.

Tivoli Storage Manager HSM (Tivoli Storage Manager HSM) stub files are replicated as regular files, and a Tivoli Storage Manager HSM recall is performed for each file, so that each can be omitted using the command line. For more information, see “SONAS automated tiering” on page 59.

10.2.3 Three-site mirroring

The DS8000 supports three-site mirroring, either through a combination of Metro Mirror and Global Mirror, or Metro Mirror with z/OS Global Mirror.

IBM System Storage DS8000 Metro Mirror and Global Mirror

Metro Mirror and Global Mirror can provide a three-site, multi-purpose, continuous availability replication solution for both System z and Open Systems data. As shown in Figure 10-4, Metro Mirror provides high availability replication from a local site (Site A) to an intermediate site (Site B), whereas Global Mirror provides long distance disaster recovery replication from an intermediate site (Site B) to a remote site (Site C). This is called the *cascading three-site solution*.

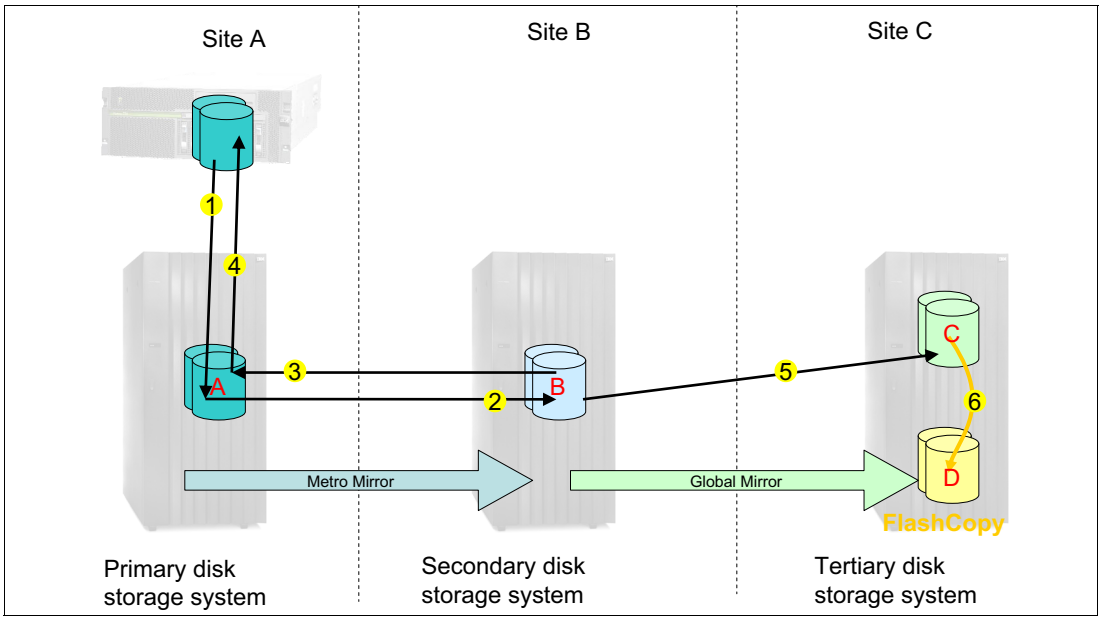

Figure 10-4 DS8000 Metro Mirror and Global Mirror replication solutions

The intermediate site (Site B) to remote site (Site C) component is identical to Global Mirror, described in “IBM System Storage DS8000 Global Copy” on page 190, with the following exceptions:

- Writes to intermediate site volumes are Metro Mirror secondary writes, rather than application primary writes.
- The intermediate site volumes are both source (Global Mirror) and target (Metro Mirror) at the same time.

The intermediate site disks are collectively paused by the Global Mirror master disk subsystem to create the consistent set of updates, which takes 3 milliseconds (ms) every 3 to 5 seconds. After these few milliseconds, the updates continue to be sent to the remote site (Site C) volumes. The intermediate site to remote site drain is expected to take only a few seconds to complete. When all updates are drained to the remote site, all changes since the last FlashCopy from the C volumes to the D volumes are copied again, with FlashCopy, to the D volumes.

If there is a storage failure (or disaster) at the intermediate or local site, or even a loss of connectivity between the local and intermediate sites, data can no longer be cascaded to the remote site. However, if we establish physical connectivity as well, directly between the local and remote sites, we can, in the case of failure at the intermediate site, still copy data from the local to the remote site. The *incremental resync* process provides the capability to use this connection with a Global Mirror relationship between the local and remote sites to copy data after a failure at the remote site,

IBM System Storage DS8000 Metro zGlobal Mirror (MzGM) for System z

Metro zGlobal Mirror (MzGM) is a three-site, continuous availability disaster recovery solution for System z. Data is replicated synchronously from Site A to Site B using Metro Mirror. Data is also replicated asynchronously, using eXtended Remote Copy (XRC, also known as zGlobal Mirror (zGM)). This is called the *multi-target, three-site solution*.

Figure 10-5 provides a pictorial overview of MzGM.

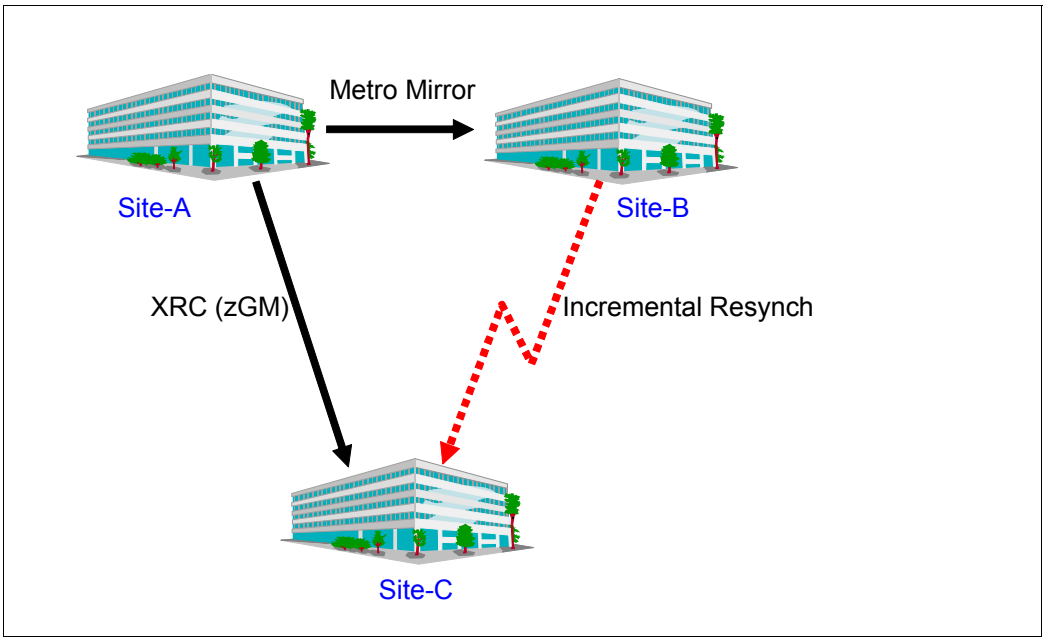

Figure 10-5 Overview of Metro zGlobal Mirror (MzGM) for System z, available on the DS8000

To provide data consistency, zGM requires a processor at Site C.

If there is a failure at Site B, replication continues from Site A to Site C, asynchronously. When Site B is restored, replication can begin again from Site A to Site B.

If Site A fails, the *incremental resync* process ensures that data at Site C is the same as that at Site B, and asynchronous replication continues from Site B to Site C. In this case, with GDPS (see “GDPS/PPRC HM” on page 198), you can dynamically switch access to the disk systems at Site B, non disruptively, using HyperSwap.

This solution offers near-zero data loss and protects against a geographical disaster that might render Sites A and B unusable.

In distributed environments, MGM can be managed by TPC-R. Additionally, in System z environments, MGM and MzGM can also be managed by GDPS.

10.2.4 PowerHA System Mirror for AIX Enterprise Edition

The IBM PowerHA System Mirror V6.1 for AIX Enterprise Edition (PowerHA) protects geographically dispersed sites so that business continuity is ensured, even if an entire site fails during a disaster. The Enterprise Edition is essential for protection from site-wide failures or large-scale disasters. Business-critical data is replicated across multiple sites and the Enterprise Edition enables automated failover to these remote locations. The Enterprise Edition offers multiple technologies for achieving long distance data mirroring, failover, and resynchronization:

- The *AIX Geographic Logical Volume Manager (GLVM)* component provides host-based synchronous and asynchronous data replication and failover to remote sites, a functionality that is not dependent on the storage system capabilities. GLVM IP-based mirroring automatically maintains a data replica across geographically remote sites and completely manages data replication during production, failover, and recovery.
- The Enterprise Edition supports DS8000 and SAN Volume Controller with either Metro Mirror or Global Mirror for data replication between sites, enabling automatic failover between geographically dispersed data centers. It combines storage-based data replication and server-based cluster functionality to protect applications against potential disasters including site failures.

  The Enterprise Edition, in combination with Metro Mirror data replication, manages a clustered environment to allow for *synchronous* mirroring of critical data to be maintained at all times across remote sites up to 300 km apart.

  The *asynchronous* Global Mirror option extends failover to backup resources at geographically dispersed sites at virtually unlimited distances with minimal impact on performance in the production application environment.

Figure 10-6 shows a basic example of four-node PowerHA cluster that is attached to two DS8000 storage systems, using Metro Mirror for data replication between remote sites. For data access, nodes are attached to the local storage system only, although for high availability reasons, a SAN connection between nodes and storages can be also established, cross-site.

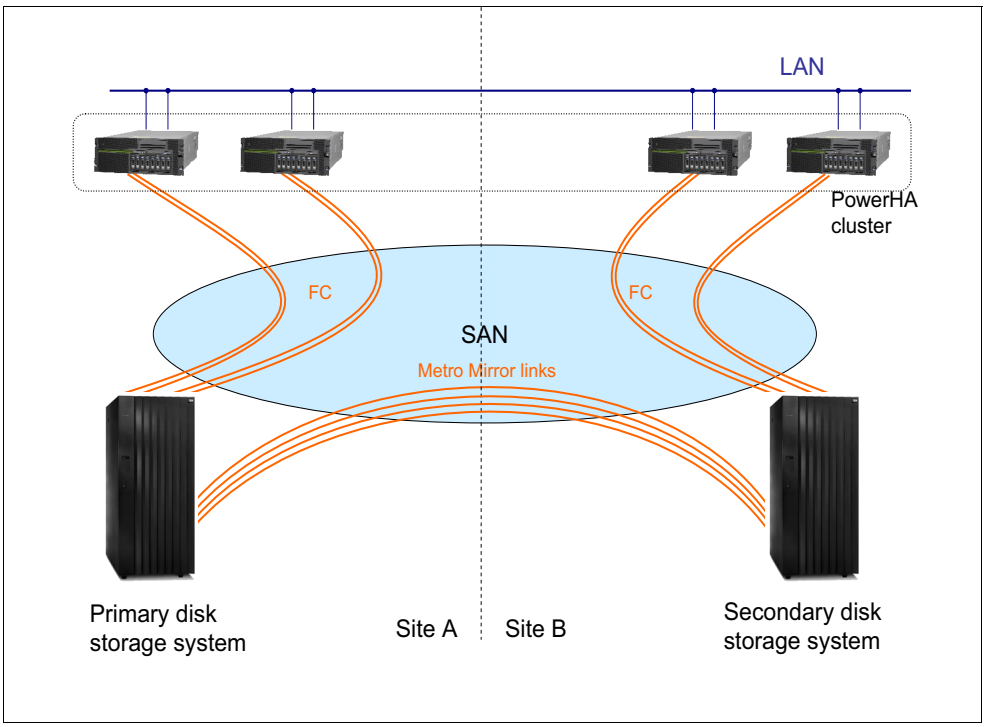

Figure 10-6 IBM PowerHA System Mirror for AIX Enterprise Edition using DS8000 Metro Mirror for data replication

PowerHA System Mirror V6.1 for AIX Enterprise Edition provides integrated support of Metro Mirror data replication for the Enterprise Storage Server, the DS6000, DS8000, and SAN Volume Controller, in addition to supporting Global Mirror for DS8700 (R5.1 or later) and the SAN Volume Controller.

For additional details, see the information available at this website:

`http://www.ibm.com/systems/power/software/availability/aix/index.html`

10.2.5 Open HyperSwap for AIX with TPC-R

The IBM Tivoli Storage Productivity Center (TPC) for Replication helps to manage the advanced copy services provided by the DS8000, and the SAN Volume Controller.

Starting with V4.2, the IBM Tivoli Storage Productivity Center for Replication (TPC-R V4.2) introduced a new feature called Open HyperSwap. Open HyperSwap (comparable to HyperSwap for System z) helps to improve the continuous availability attributes of open system servers, such as IBM AIX, by managing a set of planned and unplanned disk system outages for Metro Mirror PPRC capable disk systems. This is illustrated in Figure 10-7.

This provides a business continuity solution that protects customers from storage failures, with minimal application impact. Open HyperSwap allows I/O to the primary volumes in a Metro Mirror relationship to be swapped to the secondary volumes without manual intervention. After the swap has occurred, the application is automatically writing to the secondary site without noticing the switch.

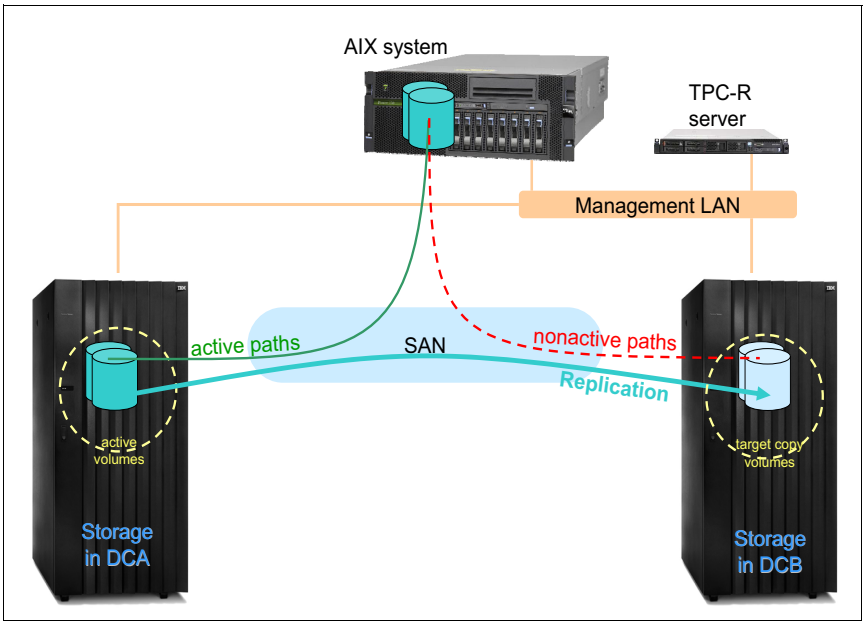

Figure 10-7 Operation scheme of Open Hyperswap

Open HyperSwap replication is a special replication method, based on Metro Mirror which is designed to automatically failover I/O from the primary logical devices to the secondary logical devices in the event of a primary disk storage system failure. Failover is intended to cause minimal disruption to the application. Open HyperSwap replication applies to both planned and unplanned replication swaps.

When a session has Open HyperSwap enabled, an I/O error on the primary site automatically causes the I/O to switch to the secondary site without any user interaction and only minimal application impact. In addition, when Open HyperSwap is enabled, the Metro Mirror session also supports disaster recovery. If a write is successful on the primary site but is unable to get replicated on the secondary site, IBM Tivoli Storage Productivity Center for Replication suspends all replication for the session, thus ensuring that a consistent copy of the data exists on the secondary site. If the system fails, this data might not be the latest data, but the data will be consistent and allow the user to manually switch host servers to the secondary site.

You can control Open HyperSwap from any system running IBM Tivoli Storage Productivity Center for Replication, on the AIX, Windows, Linux, or z/OS operating systems. However, the volumes of the Open HyperSwap session must be attached to an AIX host, and an IP connection between the AIX host system and the Tivoli Storage Productivity Center for Replication server must be available.

For more information, see these references:

- The Redbooks publication, *IBM Tivoli Storage Productivity Center V4.2 -Release Guide*, available at this website:

  `http://www.redbooks.ibm.com/redpieces/pdfs/sg247894.pdf`

- The IBM Tivoli Storage Productivity Center Information Center's, *Managing a session with HyperSwap and Open HyperSwap replication*, available at this website:

  `http://publib.boulder.ibm.com/infocenter/tivihelp/v4r1/index.jsp?topic=/com.ibm.tpc_V42.doc/frg_t_manage_hs.html`

10.3 GDPS/PPRC HM

GDPS and PPRC are two separate products. GDPS is a management tool, designed to provide automation and a single point of control for your replication environment. PPRC is the former name for the DS8000 Metro Mirror synchronous replication technology that propagates data from one site to another.

GDPS/PPRC HM is designed for System z environment only.

GDPS/PPRC HM is designed to extend the availability attributes of an IBM Parallel Sysplex® to disk systems, whether the Parallel Sysplex and disk systems are in a single site, or whether the Parallel Sysplex and the primary and secondary disk systems span across two sites.

GDPS/PPRC HM provides the ability to transparently switch primary disk systems with the secondary disk systems for either a planned or unplanned disk reconfiguration. It also supports disaster recovery capability across two sites by enabling the creation of a consistent set of secondary disks in the case of a disaster or suspected disaster.

However, unlike the full GDPS/PPRC offering, GDPS/PPRC HM does not provide any resource management or recovery management capabilities. GDPS/PPRC HM provides a subset of GDPS/PPRC capability, with the emphasis being on remote copy and disk management capabilities. At its most basic, GDPS/PPRC HM extends Parallel Sysplex availability to disk systems by delivering the HyperSwap capability to mask disk outages that are caused by planned disk maintenance or unplanned disk failures. It also provides monitoring and management of the data replication environment, including the freeze capability.

In the multi-site environment, GDPS/PPRC HM provides an entry-level disaster recovery offering. Because GDPS/PPRC HM does not include the systems management and automation capabilities as does the GDPS/PPRC, it cannot provide, in and of itself, the short Recovery Time Objective (RTO) that is achievable with GDPS/PPRC. However, GDPS/PPRC HM does provide a cost-effective route into full GDPS/PPRC at a later time if your RTOs change.

There are two distinct types of disk problems that trigger a GDPS-automated reaction:

- *Mirroring problems (FREEZE triggers):* There is no problem with writing to the primary disk system, but there is a problem mirroring the data to the secondary disk system.

  GDPS uses automation, keyed off of events or messages, to stop all mirroring when a remote copy failure occurs. In particular, the GDPS automation uses the IBM PPRC `CGROUP FREEZE` and `CGROUP RUN` commands which have been implemented as part of Metro Mirror.

  GDPS/PPRC HM uses a combination of storage system and Sysplex triggers to capture, at the first indication of a potential disaster, a data-consistent secondary site copy of your data using the `CGROUP FREEZE` function. In this way, the consistent image of the data is ensured on the secondary copy at the earliest sign of a disaster, even before production applications are aware of I/O errors.

- *Primary disk problems (HyperSwap triggers):* There is a problem writing to the primary disk, because of either a hard failure, or the disk system is not accessible or not responsive.

  GDPS/PPRC and GDPS/PPRC HM deliver a powerful function known as “HyperSwap.” HyperSwap provides the ability to non disruptively swap from using the primary volume of a mirrored pair to using what is currently the secondary volume. The HyperSwap function is designed to be completely automated, thus allowing all aspects of the site switch to be controlled by GDPS.

There are two ways that HyperSwap can be invoked:

– Planned HyperSwap:

  A planned HyperSwap is invoked manually by operator action using GDPS facilities. One example of a planned HyperSwap is when a HyperSwap is initiated in advance of planned disruptive maintenance to a disk system.

– Unplanned HyperSwap:

  An unplanned HyperSwap is invoked automatically by GDPS, triggered by events that indicate the failure of a primary disk device.

In both cases, the systems that are using the primary volumes will experience a temporary pause in processing. During this pause, the disk mirroring configuration is changed to allow use of the secondary volumes (and mirroring can be established in the opposite direction, depending on the option selected), the Unit Control Blocks (UCBs) for the primary devices are updated to point to the formerly secondary volumes, and then the systems resume operation.

In benchmark measurements at IBM using currently supported releases of GDPS, the I/O hold time for an unplanned HyperSwap is generally less than 30 seconds for even very large configurations (for example, a ten-way Sysplex with approximately 20,000 mirrored volume pairs, or even a 30-way Sysplex with a more moderately sized disk configuration). Most implementations in the world are actually much smaller than this, and typical I/O hold times using the most current storage and server hardware are generally measured in seconds. Although results will obviously depend on your configuration, these numbers give you a high-end figure for what to expect.

A HyperSwap in GDPS/PPRC HM affects *all* mirrored LSSs with devices in the configuration (single CG). For example, if one single mirrored volume were to fail, and HyperSwap is invoked, processing will be swapped to the secondary copy of *all* mirrored volumes in the configuration, including those in other, unaffected, disk systems. The reason is that, to maintain disaster readiness, all primary volumes *must* be in the same site. If HyperSwap were to only swap the failed LSS, you will have certain primaries in one site and the remainder in the other site.

10.3.1 Managing the GDPS environment

The bulk of the functionality delivered with GDPS/PPRC HM relates to maintaining the integrity of the secondary disks, and being able to non disruptively switch to the secondary volume of the Metro Mirror pair.

However, there is an additional aspect of remote copy management that is available with GDPS/PPRC HM, namely, the ability to query and manage the remote copy environment using the GDPS panels.

In this section, we describe this other aspect of GDPS/PPRC HM. Specifically, GDPS/PPRC HM provides facilities that give you these capabilities:

- Be alerted to any changes in the remote copy environment
- Display the remote copy configuration
- Stop, start, and change the direction of remote copy
- Stop and start FlashCopy

Note that GDPS/PPRC HM does not provide script support, so all of these functions are only available using the IBM NetView® interface for GDBS.

GDPS also provides monitors to check the status of disks, Sysplex resources, and so on. Any time there is a configuration change or something in GDPS that requires manual intervention, GDPS raises an alert. GDPS uses the Status Display Facility provided by System Automation as the primary status feedback mechanism for GDPS. It is the only dynamically updated status display available for GDPS.

The Status Display Facility provides a dynamically-updated, color-coded panel that reflects changes and exception conditions in the GDPS replication environment.

10.3.2 Testing

To facilitate testing of site failover and failback processing, consider installing additional disk capacity to support FlashCopy in Site 1 and Site 2. The FlashCopy can be used at both Site 1 and Site 2 to maintain disaster recovery checkpoints during remote copy resynchronization. This ensures that a consistent copy of the data is available if a disaster-type event were to occur during the testing your site failover and failback procedures.

In addition, FlashCopy can be used to provide a consistent point-in-time copy of production data to be used for nondestructive testing of your system and application recovery procedures. FlashCopy can also be used, for example, to back up data without the need for extended outages to production systems, to provide data for data mining applications, and for batch reporting.

The latest GDPS/PPRC HM prerequisite information is located at this website:

http://www.ibm.com/systems/z/advantages/gdps/getstarted/gdpspprc_hsm.html

10.3.3 Summary

GDPS/PPRC HM is a powerful offering that can extend Parallel Sysplex availability to disk systems by delivering the HyperSwap capability to mask planned and unplanned disk outages. It also provides monitoring and management of the data replication environment, including the freeze capability. It can provide these capabilities either in a single site, or when the systems and disks are spread across two data centers within metropolitan distances.

In a multi-site configuration, GDPS/PPRC HM can also be an entry level offering, capable of providing zero data loss. The RTO is typically longer than what can be obtained with a full GDPS/PPRC offering. As time goes by, if your business needs to migrate from GDPS/PPRC HM to the full GDPS/PPRC offering, this can be achieved as well.

In addition to disaster recovery and continuous availability capabilities, GDPS/PPRC HM also provides a much more user-friendly interface for monitoring and managing the remote copy configuration.

For additional information, refer to *GDPS Family - An Introduction to Concepts and Capabilities* at this website:

http://www.redbooks.ibm.com/redpieces/abstracts/sg246374.html

10.4 TPC-R Basic HyperSwap

Basic HyperSwap is a z/OS enhancement that contributes to high availability. Basic HyperSwap is not a disaster recovery function. It is exclusively a high-availability function and cooperates with the Metro Mirror product, which provides synchronous data replication.

Basic HyperSwap is a storage-related activity that executes in the IBM System z server and, more precisely, within the z/OS. The management interface to Basic HyperSwap is Tivoli Storage Productivity Center for Replication (TPC-R).

TPC-R handles HyperSwap through a special session type, called a HyperSwap session. This TPC-R-related HyperSwap session contains an underlying Metro Mirror session and is basically managed as a TPC-R Metro Mirror session.

TPC-R controls all activities during the creation and construction of a HyperSwap session. These activities include defining the HyperSwap session itself, populating the session with the appropriate Metro Mirror volume pairs (that is, TPC-R copy sets), and starting the session to fully initialize all copy pairs. Eventually, all TPC-R copy sets within the session are in FULL DUPLEX state. When in FULL DUPLEX state, the session configuration is transferred to the Input Output Supervisor (IOS) and the IOS takes over the active session under its own control.

IOS handles and manages all further activities in the course of the subsequent swap of all device pairs, which are defined within the TPC-R HyperSwap session. These activities include a FREEZE and FAILOVER in the sense of Metro Mirror management. The FREEZE guarantees an I/O quiescence to all primary devices. After that, the IOS performs the actual swap operation and swaps the content of the UCBs between the primary device and secondary device. The swap occurs in the same manner as an unplanned swap.

10.5 Geographically Dispersed Open Clusters (GDOC)

Continuous availability solutions are integrations of servers, storage, software and automation, and networking. Most of the solutions are based on a form of operating system server clustering to provide application availability. Data replication and failover are automated. When an application failure is detected, a continuous availability solution can perform a predefined set of tasks required to restart the application on another server, utilizing the mirrored data residing on the secondary storage system.

Geographically Dispersed Open Clusters (GDOCs) is a solution designed to protect the availability of critical applications that run on UNIX, Windows, or Linux servers. GDOC is based on an Open Systems Cluster architecture, spread across two or more sites, with data mirrored between sites to provide high availability and disaster recovery.

This discussion provides a concise overview of the GDOC function, its benefits, and applications. This solution is delivered by IBM Global Services.

Figure 10-8 provides a GDOC architecture overview.

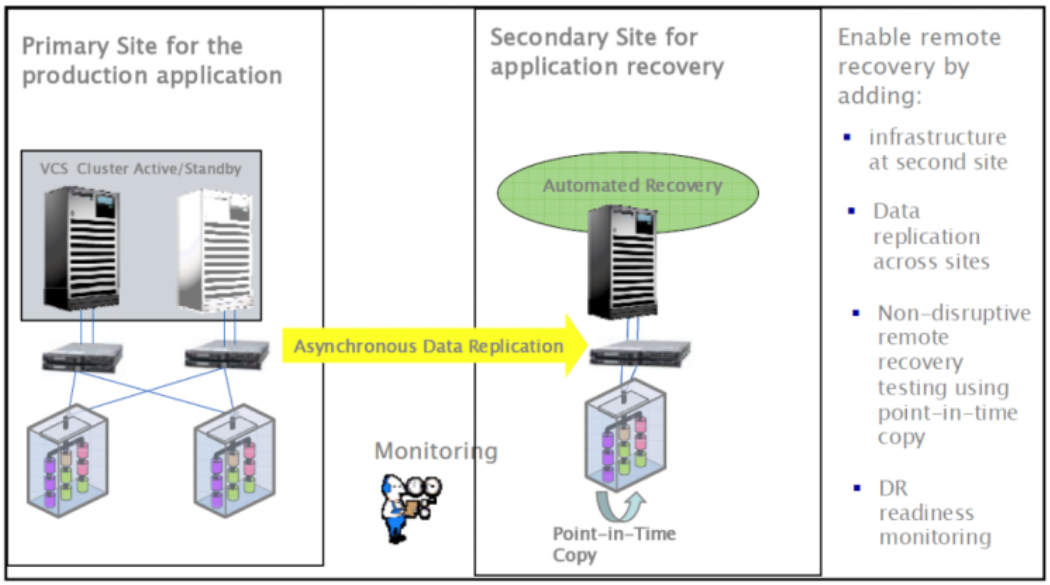

Figure 10-8 GDOC architecture

10.5.1 GDOC services details

Combining IBM services expertise and Symantec software, IBM Implementation Services for GDOC provides automation, testing, and management to support world-class, high availability and near-transparent application recovery for open systems.

Here we list highlights of the IBM Implementation Services for GDOC:

- Automates application recovery across remote mirrored sites
- Accommodates heterogeneous, open system environments
- Provides for regular recovery testing
- Provides guidance and support during the full course of the implementation
- Capitalizes on the extensive knowledge and experience of IBM in designing and implementing multi-data center, high-availability solutions

10.5.2 Automating application recovery across geographically dispersed sites

GDOC typically starts with a cluster of servers managed by Symantec software to provide high availability at a single site. It then adds mirrored copies of the data, along with more servers, at a second, geographically distant site. A data replication solution keeps the data between the two sites synchronized, enabling an organization with GDOC to recover from an outage quickly with minimal data loss. In fact, the recovery time with this solution ranges from minutes to hours—in contrast to the 24- to 48-hour recovery time that is typical with traditional recovery solutions.

Software automates recovery from an outage both within a site and across sites and provides for regular recovery testing. By eliminating manual recovery processes, organizations can greatly increase their ability to survive a local or regional outage. If an outage does occur, automated recovery processes take over to accelerate recovery time, minimize human error and enable the environment until IT staff is available.

Because the recovery time with this type of solution is much shorter than that with a tape backup, the solution is especially appropriate for business-critical applications that require rapid recovery.

10.5.3 Accommodating heterogeneous environments

Designed specifically for today's open systems, this solution is ideal for environments that include one or more servers, based on the following operating systems: IBM AIX, HP-UX, Sun Solaris, Linux, and Microsoft Windows.

A graphical user interface (GUI) enables IT staff to manage the entire cluster configuration from a single point of control. In the event of an unplanned outage, the IT staff can administer and monitor all systems, either individually or through a single point of control.

A GDOC solution also accommodates a wide variety of storage and replication technologies, including products from IBM, Symantec, Hitachi Data Systems, EMC, Oracle, and NetApp®.

10.5.4 Providing full life cycle guidance and support

With its wide array of integration, consulting, and project management skills, IBM is well qualified to assist you in:

- Implementing and managing data replication across sites
- Automating application recovery at a secondary site
- Managing, monitoring, and testing the recovery solution.

Ideally, GDOC begins with a technology consulting workshop in which IBM assesses the solution requirements. IBM consultants make sure they understand your business, your current environment, and RPOs, which are usually just minutes or possibly seconds. They also look at your RTO, which is usually a few hours or less. One common discovery is that organizations that have already deployed an in-house recovery solution often lack the necessary automation, monitoring, and testing capabilities needed to maintain business continuity in the event of an outage.

During the technology consulting workshop, IBM also defines the desired state, along with the steps needed to achieve that state. In addition, IBM provides:

- Project management
- Assessment of client availability requirements
- Application recovery planning, configuration, and implementation
- Automation of application recovery
- Testing of application recovery
- Onsite skills transfer

These efforts provide the full range of services and support organizations needed for a timely, high-quality, multiplatform availability, and recovery solution.

10.5.5 Guarding against costly downtime

Although it is impossible to eliminate all human, software, and technology errors, proactive solutions, based on GDOC can help keep these errors to a minimum, thereby reducing downtime, and speeding recovery from unforeseen catastrophes. These solutions provide the ability to perform a controlled site switch for both planned and unplanned site outages with minimal data loss. The result is full data integrity across multiple volumes, disk systems, and open system platforms.

This timely resumption of critical operations is especially crucial for enterprises that are highly dependent upon key information systems. Such organizations not only require a multi-site data replication solution. They also need the full automation, testing, and monitoring capabilities necessary to provide a high level of confidence that their solution does and will work.

To learn more about IBM Implementation Services for Geographically Dispersed Open Clusters, see this website:

http://www.ibm.com/services

10.6 Tivoli Storage Manager FastBack

IBM Tivoli Storage Manager FastBack™ gives you an option to easily protect and quickly recover data in remote and branch offices, to do so cost-effectively, and to simultaneously prevent the unwanted effects of data recovery. Examples of unwanted effects are lost or deleted files or folders, a virus or hacker attack, a corrupted database, a disk, or server crash.

The Tivoli Storage Manager FastBack product family includes the following components:

- *IBM Tivoli Storage Manager FastBack:* This is a local disk-based, block-level, incremental continuous, data protection solution for Windows and Linux servers. It includes selective replication for off-site disaster recovery and business resilience, plus near-instant recovery of any data asset, from a single file to an entire volume.
- *IBM Tivoli Storage Manager FastBack for Microsoft Exchange:* This provides granular recovery of any Exchange email object, including, messages, attachments, contacts, calendar entries, notes, and tasks.
- *IBM Tivoli Storage Manager FastBack for Bare Machine Recovery:* This delivers restoration of the Windows server operating system volume, even to a completely different hardware platform or to a virtual machine, within an hour. Then, using the near-instant data volume recovery capabilities of Tivoli Storage Manager FastBack, you can get back up and running a few seconds later.
- *IBM Tivoli Storage Manager FastBack for Workstations:* This simplifies the backup and recovery of valuable information about employees' personal computers.

Because Tivoli Storage Manager FastBack performs backup at the block level, it protects the data of any application, including Microsoft Exchange, SQL Server, SharePoint, IBM DB2, SAP and Oracle. The backup agent is so processor efficient that it is an excellent solution for protecting virtual machines directly, without the need for a proxy server.

Application recovery is where Tivoli Storage Manager FastBack provides return on investment, providing near-instant access to any data that you need to restore. You can recover individual files and folders, from any previous point-in-time, with a simple drag-and-drop operation. Or you can restore full access to an entire disk volume – even terabytes of data – within a couple of minutes. Tivoli Storage Manager FastBack also includes built-in disaster recovery, enabling you to restore your operations in another location when a facility becomes unavailable.

Tivoli Storage Manager FastBack can be integrated seamlessly with Tivoli Storage Manager, IBM enterprise-class data management solution. For example, you can launch and manage FastBack operations from the Tivoli Storage Manager Admin Center.

10.6.1 Continuous data protection with FastBack

Tivoli Storage Manager FastBack is a data protection solution for user workstations, laptops, and file servers and selected database systems on supported Windows and Linux platforms. This product offers transparent, real-time replication and traditional backup services, and can operate independently on Tivoli Storage Manager. Files can be replicated to both the local disk and a remote target. When network connectivity is temporarily unavailable, files are queued up and replication is resumed when the remote target is available, thus providing continuous data protection.

The FastBack client provides a flexible, easy to use file protection system. Your most important files can be continuously protected. Your less important files can be protected at scheduled intervals to save time and storage space. Email files can also be protected. And you can prevent any changes (including deletions) to files in folders that you designate as vaults.

The concept of continuous protection

Continuously protected files are backed up to a local drive, so that backup copies are created even when network conditions prevent storing backup copies on remote storage locations. Continuously protected files can also be stored on remote storage locations, when network connections allow. If a remote location is not available when you change a continuously protected file, the Tivoli Storage Manager FastBack client makes a backup copy on that device as soon as the device becomes available.

Scheduled backup copies are created on the interval that you configure (hourly, weekly, daily, or monthly). If the remote device for scheduled backups is not available at the time of the backup, the FastBack client makes backup copies on the remote location as soon as that device becomes available.

Every time you change a file, a backup copy is created by FastBack client. This allows you to choose which version of a protected file you want to restore. You configure how many backup copies to save. The diagram in Figure 10-9 provides an overview of how the Tivoli Storage Manager FastBack client protects your data.

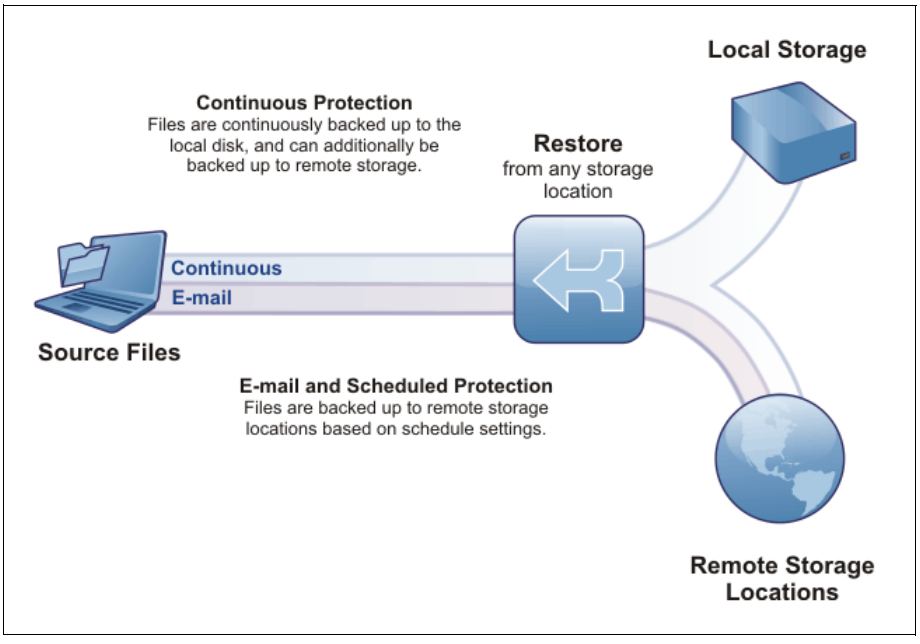

Figure 10-9 Tivoli continuous data protection conceptual diagram

After installation of a Tivoli Storage Manager FastBack client, the client immediately provides continuous protection for a pre-configured list of files. You can see the backup copies in the \RealTimeBackup\ folder in the root of your primary drive, and in the list of files that you can restore with the client restore wizard. The default space allocated for your backup copies is 500 MB.

Tip: Tivoli Storage Manager FastBack can store backup copies on a Tivoli Storage Manager server, but there is no requirement to use Tivoli Storage Manager. It is a stand-alone product and has no dependencies on Tivoli Storage Manager.

Types of protection

The Tivoli Storage Manager FastBack client offers three types of protection for your files:

- *Continuous protection*: This means that every time a file is saved, a backup copy is created. Hence, the backup copy exactly matches the original file as you last saved it. If you choose to save more than one version of a backup copy, the previous backup copies match the previous versions of your file.
- *Scheduled protection:* This occurs on files that are copied to the remote storage area on a regular schedule. They are not backed up every time you save them, as are continuously protected files. Hence, scheduled protection yields fewer backup copies. If a file is lost between the time it is saved and the time it is backed up, you are able to restore only a previous version of the file.
- Vaulting protection: Tivoli Storage Manager FastBack client can interact with the Tivoli Storage Manager backup/archive client to move the data on tape for offsite *vaulting*.

Table 10-2 compares all three options of the data protection with FastBack.

Email files are protected on a schedule only. If the storage area is unavailable when a protected file is saved, the client notes that the file has changed. When the storage area becomes available, the client makes a backup copy of the most recent version of the file.

Table 10-2 Comparison of the three types of protection

Feature	Continuous protection	Scheduled protection	Vaulting
Suggested for what files	Suggested for your most important files; not suggested for large, dynamic files, such as email or databases	Suggested for large, dynamic files, such as email or databases	Suggested for files that you do not want to be changed or deleted
How protected	Backup copies are created on a storage area	Backup copies are created on a storage area	Vaulted files and folders cannot be modified, nor deleted
Backup frequency	Each file is backed up when it is saved	File is backed up only at the scheduled time, and only if it has not been saved since the previous schedule	No backups
Backup copy area	Local or remote	Remote only	Not applicable

10.6.2 Product summary

Tivoli Storage Manager FastBack is continuous backup software that protects your most important files in real-time. FastBack offers the following benefits:

- Captures every save of a file when it occurs
- Minimizes the backup and recovery window
- Optimizes RPOs and RTOs
- Easy to implement and manage, it runs transparently in the background with a light footprint
- Works with multiple backup and replication targets (for example, local disk, web, file server, Network Attached Storage (NAS) device, Tivoli Storage Manager)
- Deletes old versions of files to make room for new versions
- Supports various database systems, such as SQL, DB2, Oracle, and email files
- Bare-machine recovery functionality
- Disaster recovery solution in remote site

10.6.3 More information

More information about Tivoli Storage Manager FastBack is available at this website:

http://www.ibm.com/software/tivoli/products/storage-mgr-fastback/

The IBM strategy of continuous protection of business critical data is available at this website:

http://www.ibm.com/systems/information_infrastructure/leadership/data_protection.html

11

Retention and compliance

In this chapter, we discuss IBM data protection strategies with regard to these needs:

- Data retention and compliance with legal and business requirements, such as the Sarbanes-Oxley act
- Long-term data retention of business-critical data
- Processes that require that data be retained for tens of years or even forever

The IBM solutions for data retention and compliance include the following products described in this chapter:

- IBM Information Archive
- IBM Enterprise tape libraries with LTO5 tape drives
- IBM System Storage N series with SnapLock feature

We do not cover other IBM products related to content collection or content management.

11.1 Retention and compliance management

Retention and compliance management helps to improve your overall compliance posture and enables a faster response time to legal inquiries, therefore reducing risks and associated costs. Retention and compliance management from IBM provides integrated collection and archiving, records management, and electronic discovery (eDiscovery), delivered using flexible, modular, on-ramps that can be deployed in any order or in any combination. Compliance management solutions are the integral parts of the IBM Smart Archive strategy.

Retention managed data is typically kept in an archive or a repository. In the past it has been known as archive data, fixed content data, reference data, unstructured data, or as other terms that imply that it is of a read-only nature. It is often measured in terabytes and kept for long periods of time, sometimes forever.

Not only are there numerous state and federal regulations that must be met for data storage, but there are also industry-specific and company-specific ones as well. With these regulations constantly being updated and amended, organizations need to develop a strategy to ensure that the correct records are maintained for the correct period of time, in line with the most current regulations. These records must be readily retrievable and accessible when they are requested by regulators or auditors. It is easy to envision the exponential growth in data storage that will result from these regulations and the associated requirement for a means of managing this data. Overall, the management and control of retention-managed data is a significant challenge in the IT industry when considering factors, such as cost, latency, bandwidth, integration, security, privacy, and more.

Continuous amendments to these regulations and to other business imperatives stress the need for an Information Lifecycle Management (ILM) process and tools. IBM has unique experiences with a broad range of ILM technologies.

IBM provides a comprehensive, open set of solutions that provide content management, data retention management, and sophisticated storage management, in addition to storage systems to house the data. To specifically help clients with risk and compliance efforts, the IBM Risk, Retention, and Compliance framework is another offering designed to illustrate the infrastructure capabilities needed to help address the myriad of compliance requirements. Using this framework, your organization can standardize the use of common technologies to design and deploy a compliance architecture that can help you to deal more effectively with compliance initiatives.

IBM has identified six key industry sectors with a specific set of retention and data management requirements:

- Financial Services
- Health Care/Life Sciences
- Insurance
- Manufacturing
- Transportation
- Government

Based on the analysis of each of these industry sectors, IBM has developed and prepared concrete industry-adapted archiving solutions with specific sets of data management policies. These policies are part of the strategic IBM Information Archive, and they can significantly help in identifying archiving requirements specific to the retention and data management requirements of your industry.

11.1.1 Characteristics of data that is managed for retention and compliance

Storage characteristics of retention-managed data are as follows:

- *Variable data retention period:* This is usually a minimum of a few months, to forever
- *Variable data volume:* Many clients start with 5 TB to 10 TB of storage in an enterprise, with storage usually consisting of a large number of small files
- *Data access frequency:* Write Once Read Rarely or possibly Write Once Read Never
- *Data read/write performance:* The write characteristic handles volume, and the read characteristic varies by industry and application
- *Data protection:* The requirement is for non-erasability, non-rewritability, and deletion when the retention policy expires

Data characteristics of retention-managed data are as follows:

- *Data life cycle:* Typically, the data is frequently accessed near its creation, and then the access frequency diminishes exponentially, near zero (certain industries have peaks that require access, such as check images during annual tax preparation)
- *Data rendering:* For data that is stored in an old data store, there is the risk that the format used to store the data is no longer supported
- *Data mining:* With the tremendous amounts of data being saved, there is intrinsic value in the content of the archive that can be exploited using data mining (this implies indexing and categorization of the data when it is initially saved)

11.1.2 IBM Smart Archive strategy

The IBM Smart Archive strategy is a broad strategy for data retention and compliance that unifies the power and value of IBM software, systems, and services.

IBM Smart Archive elements

The IBM Smart Archive strategy and its Information Archive for Email, Files and eDiscovery consist of three crucial elements:

- IBM Information Archive
- Content Collector Discovery Analytics Starter pack, which includes Content Collector for Email, Content Collector for File Systems, eDiscovery Manager, and eDiscovery Analyzer
- Content Manager Enterprise Edition

In this chapter we focus on IBM Information Archive as a key storage repository for archived data that is identified and collected by Content Collectors and managed by Content Manager.

The solution consists of pre-configured software, hardware, and implementation services from a single vendor. The total solution configuration uses our experience in client archiving deployments, simplifying implementation, and providing quick time-to-value. Additionally, proven service packages are available to manage ongoing administration and maintenance of the solution.

IBM Information Archive for Email, Files and eDiscovery is ideal for organizations that are:

- Required to capture and retain email and filesystem content
- Demonstrating compliance with retention mandates
- Experiencing soaring storage costs and need to lower these costs
- Required to quickly and efficiently respond to eDiscovery requests

The IBM strategy is built upon a highly scalable, secure, high performance infrastructure with the flexibility to meet multiple requirements in a single solution. As a result, your ability to locate, collect, organize, analyze, trust and use information across the organization is improved.

IBM Smart Archive components

These are the three primary components of the IBM Smart Archive strategy:

- *Optimized and unified ingest capabilities:* This enables a deeper understanding of what information to retain by using discovery- and analytics-based assessment technologies. These technologies reduce solution complexity and costs by unifying data and content archiving through common collection and classification technologies. This phase is also known as *content collection and archiving*.
- *Flexible and secure infrastructure:* The infrastructure allows you to easily implement and enforce policies and to protect the integrity of your business records throughout the information life-cycle. You can speed up the time-to-value by using multiple modular, yet integrated, solutions, including choice of management and delivery models (for example, traditional on-premise software, preconfigured appliance, and cloud-based and hybrid options). This component includes the following items:
  - Content repositories
  - Storage repositories
  - Archive services
- *Integrated software, supporting compliance, analytics, and eDiscovery:* This component reduces risk, responds more quickly to legal inquiries, and establishes trust. It also leverages information using integrated software to support your information life-cycle governance, analytics and eDiscovery requirements. The key feature of this component is called *compliance management.*

11.2 IBM Information Archive

IBM Information Archive is a consolidated, preconfigured hardware and software product, enabling you to archive data with scalability up to multiple petabytes.

This product is designed to archive and retain records in non-erasable, non-rewritable formats. It supports businesses and legal requirements for retaining and protecting information from premature deletion or malicious modification. IBM Information Archive is a universal storage repository for all types of archived information. The focus of IBM Information Archive is to help mid- and enterprise-level clients reduce storage costs and improve operational efficiencies in their data centers.

IBM Information Archive can help reduce the need for primary storage by enabling archiving applications to move inactive or unnecessary information off of the primary storage tier, to lower-cost storage tiers. When your primary storage tier has less information, applications run faster, and backup and recovery operations can be completed more quickly.

The IBM Storage System Archive Manager (SSAM) is the software component at the core of the IBM Information Archive. In addition to data encryption capabilities offered by SSAM, each data collection within the IBM Information Archive maintains a set of tamper-proof audit logs. These provide an immutable and retention-protected record for documents in the collection. Audit logs track document ownership and system life-cycle events, including document creation and deletion, changes to retention policies, and system software upgrades.

IBM Information Archive is depicted in Figure 11-1.

Figure 11-1 IBM Information Archive, front view

IBM Information Archive features the option to use the retention policies from archiving applications (such as the Tivoli Storage Manager client) or to assign your own retention policies. There are three built-in retention policy enforcement options:

- *Basic:* This policy level enables applications and users to delete documents before the retention period or retention hold expires. Users can also increase or decrease a retention policy and modify the protection level of a collection. This option can be used when information is archived for general business reference purposes or to meet broad corporate governance requirements.
- *Intermediate:* This policy level allows users to increase or decrease retention periods, but prevents deletion before the expiration of the retention period or the retention hold requirements. Users can increase the protection level to Maximum, but cannot decrease it to Basic. Clients looking to save important project information or enforce strict corporate governance requirements are likely to use this protection level.
- *Maximum:* This policy level only allows users to increase retention periods and prevents information from being deleted until the retention period or retention hold has expired. This protection level cannot be altered and is ideal for companies that have to archive data to meet strict regulatory compliance requirements. This protection level might be compared to tape storage systems configured in Write Once Read Many (WORM) format.

The IBM Information Archive base appliance frame has a drawer with a 16-drive capacity, and has one disk controller. For more storage, the controller can be attached to a maximum of six expansion drawers. A fully populated base frame of one controller and six expansion drawers yields a capacity of 112 Terabyte (TB) on Serial Advanced Technology Attachment (SATA) disks.

The base rack can be supplemented with an expansion frame that has the capacity for two additional disk controllers, plus a maximum of ten expansion drawers. Each document collection requires another disk controller, so two additional controllers represent two additional document collections. A fully populated expansion rack has a capacity of 192 disks and provides 254 TB of total usable storage space. A primary rack, combined with an expansion rack, provides the total physical storage capacity of 304 TB hard disks.

The hard disk drives (HDDs) operate in a Redundant Array of Independent Disks (RAID) 6 configuration to maintain data integrity. Because of the RAID 6 configuration, two of every eight drives are reserved for parity. This drive configuration yields a maximum of 81 usable 2 TB drives in the primary rack, and 70 usable 2 TB drives for each of the two possible collections in the expansion rack.

11.2.1 IBM Information Archive concepts

The IBM Information Archive appliance is an integrated information retention solution. The appliance includes preinstalled servers, disk storage, and the Information Archive software. The appliance can be used to store and manage multiple billions of documents over its deployment lifetime.

The Information Archive appliance provides policy-managed storage for compliance, archiving, and content management applications. Both time-based and event-based retention options are available. Stored data can be compressed and deduplicated before it is housed within the appliance. Optional features include Enhanced Remote Mirroring (ERM) for disaster recovery, high-availability server configurations, and migration from disk storage to tape.

Depending on the configuration chosen, files can be archived using the Network File System (NFS) protocol or the Common Internet File System (CIFS). In addition, HTTP protocol supports the retrieval of files from the appliance. Another option is to use the Tivoli Storage Manager archive client or application program interface (API) to archive and retrieve files.

IBM introduced an Index and Search feature within the appliance, enabling you to build an index of your information as it is archived, and then search that information for specific files, based on user-specified criteria. You can execute searches and associated retrievals without disrupting applications or archiving operations.

Document collections

IBM Information Archive uses collections to manage archived data (for details, see the next section “Archiving processing and data management” on page 215). A collection is a logical container used to store documents and define retention and access requirements. Information Archive provides two types of collections for maintaining security:

- The IBM Storage System Archive Manager (SSAM) software collection stores information from applications that implement the Tivoli Storage Manager API, the Tivoli Storage Manager archive client interface, and the Tivoli Storage Manager Hierarchical Storage Management (HSM) for Windows client.
- The File Archive collection stores information from applications that implement standard file interfaces and protocols, such as NFS and CIFS.

Archiving processing and data management

The data information life-cycle in IBM information Archive consists of three phases, as depicted in Figure 11-2.

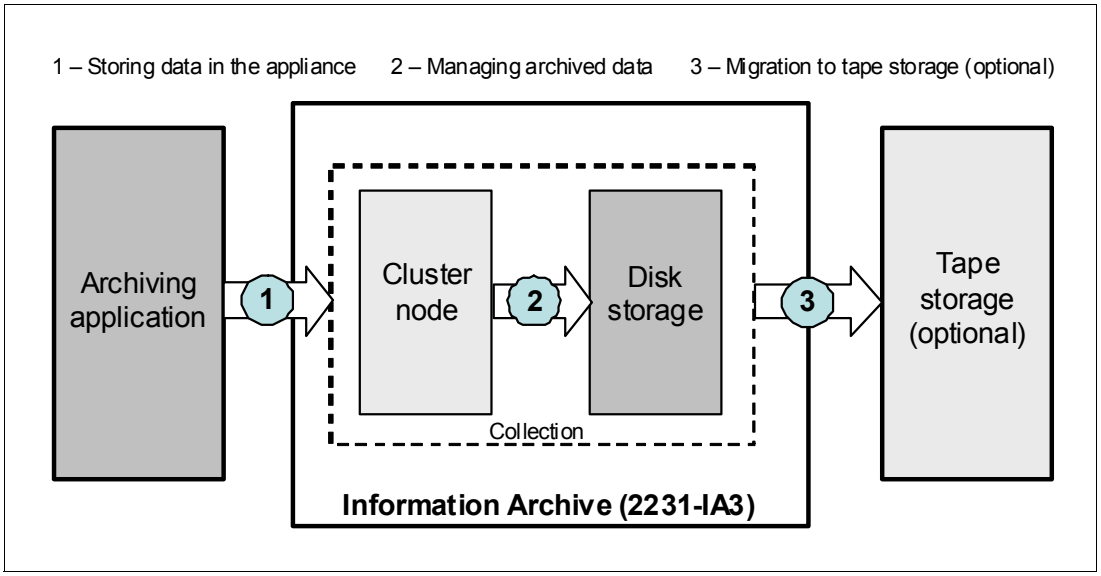

Figure 11-2 Data processing in the IBM Information Archive appliance

Here describe the three steps in Figure 11-2:

- *Storing data in the appliance:* Information Archive uses a logical object called a collection to manage archived data. Each collection is represented by a filesystem. To begin the process of storing data, an archiving application (for example, IBM Content Manager) writes a new file to the collection filesystem. The file is then committed to the archive by one of several available methods. Committing a file starts the process of ingestion. During ingestion, the file is added to the collection, along with a set of descriptive metadata. This metadata includes information about when the file was created, who owns the file, and other identifying information. You can extend the metadata by adding custom fields to the metafile schema. Files and their accompanying metadata are referred to as documents.
- *Managing archived data:* The Information Archive appliance includes up to three servers for ingesting and managing documents. These servers are called cluster nodes, because they use a clustered filesystem that provides workload balancing and high-availability through failover. Each collection is associated with a dedicated pool of disk-based storage. Each collection has its own policy domain, which includes rules for managing security, for setting document retention periods, and for managing storage resources. During the ingestion process, the retention policies that are defined for the collection are applied to the newly ingested document. These policies determine how long the document is retained in the archive, and whether the document can be deleted before its retention period expires. After a document has been ingested, it is migrated to the next level in the disk storage hierarchy. During the migration process, documents can be compressed and deduplicated to make the most efficient use of the available storage capacity.
- *Migration to tape storage (optional):* By default, all documents that are stored in the appliance are protected by a RAID6 storage configuration. A remote replication option is also available. To provide an additional level of disaster recovery protection, archived data can be automatically copied to tape storage and transferred to an off-site location. The Information Archive solution includes an integrated storage management application that provides policy-based management of this migrated data. Attaching a tape library to the appliance can significantly increase its overall storage capacity, and, at the same time, still provide the ability to access the documents through the Information Archive interface.

Appliance administration

Information Archive includes a web-based user interface called the *administrative interface*. Use this interface to configure and manage the appliance. The administrative interface provides wizards to guide you through common tasks. Properties notebooks allow you to modify settings and perform advanced management tasks. The administrative interface is only used to configure and manage the appliance, and to download audit logs. This interface is not used to archive or retrieve documents.

The Information Archive administrative interface is installed as a component of the IBM Integrated Solutions Console, which offers a single point of management for the majority of key IBM Tivoli products, such as Tivoli Storage Manager.

11.2.2 Enhanced remote mirroring

The optional Enhanced Remote Mirroring feature allows data synchronization between two Information Archive appliances installed at primary and secondary sites. The secondary Information Archive appliance can take over I/O responsibility when the primary appliance becomes unavailable. Enabling the feature reduces the possibility of data loss and system downtime.

Remote replication is enabled when two Information Archive appliances are manufactured, during which time one appliance is specified as the primary appliance, and the other, the secondary appliance. The Enhanced Remote Mirroring feature must be installed on each of the DS4200 storage controllers installed in the appliance.

The detailed description and step-by-step procedure for configuring and enabling Enhanced Remote Mirroring between two sites with IBM Information Archives, is available in *IBM Information Archive Architecture and Deployment*, SG24-7843.

11.2.3 IBM Information Archive additional features and benefits

There are many additional features and benefits of the IBM Information Archive. The following list summarizes the most significant features:

- IBM Information Archive is designed to offer a fast time-to-value, as it is quickly and easily configurable.
- The product is easy to use:
  - You can manage both appliances from a single point of management.
  - You can manage more data with fewer resources.
- Information protection levels and retention policies address unique data retention needs with flexible information protection levels:
  - Data can be stored in up to three collections per system for which various levels of protection can be assigned, and there are retention policies for various types of data.
  - The product allows compliant and non-compliant storage within the same appliance footprint.
  - Enhanced Tamper protection is offered.
  - The product can accept retention policies from other applications, or can create them,
  - Time-based and event-based information retention policies are accepted to protect information from intentional or accidental deletions and modifications.
  - The product accepts retention-hold and deletion-hold policies that set an indefinite period of retention for a file, such as during legal discovery.

- The product grows with your needs; you can add storage to increase capacity and nodes for performance.
- Uses industry standard interfaces, thus eliminating the need for customized API or add on feature requirements
- Data deduplication and compression are provided:
  - These help to lower total cost of ownership by supporting embedded deduplication and compression capabilities.
  - The product is designed to optimize storage capacity and to improve productivity.
- Tiered storage management:
  - The product is designed to allow a cost-effective mix of nearline storage and offline storage (disk and tape).
  - Contains costs and utilizes storage technology for optimum usage
- Universal storage repository for all types of content (structured and unstructured, compliant or non-compliant) are provided.

11.2.4 More information

For additional information about the IBM Information Archive, see these websites:

- IBM Information Archive:

  http://www.ibm.com/systems/storage/disk/archive
- IBM Information Archive V1.2 Information Center:

  http://publib.boulder.ibm.com/infocenter/tivihelp/v37r1/index.jsp
- *IBM Information Archive Architecture and Deployment*, SG24-7843:

  http://www.redbooks.ibm.com/abstracts/sg247843.html
- *IBM Information Archive V1.2 - Introduction and Planning Guide*, SC27-2324-02:

  http://www.ibm.com/support/docview.wss?rs=1328&uid=ssg1S7003200
- *IBM Information Archive, V1.2 - Installing and Configuring IBM Information Archive*, GC27-2326-02:

  http://www.ibm.com/support/docview.wss?rs=1328&uid=ssg1S7003240
- *IBM Information Archive V1.2 - User's Guide*, SC27-2325-02:

  http://www.ibm.com/support/docview.wss?rs=1328&uid=ssg1S7003239
- *IBM Information Archive V1.2 - Service Guide*, SC27-2327-02:

  http://www.ibm.com/support/docview.wss?rs=1328&uid=ssg1S7003241

11.3 IBM Enterprise tape libraries with LTO5

Although tape libraries are typically configured for backup/recovery scenarios, they are also used for long-term archiving, or as a combination of both, as it is possible with the IBM Enterprise TS3500 Tape Library (formerly known as IBM TotalStorage 3584).

Standard backup solution requires the file, folder, or database to be recovered in case of accidental deletion by a user, either when the user requests previous versions of an existing file (such as document, spreadsheet, financial records, etc.), or if the file is corrupted due to a hardware or software malfunction. The backup solution (for example, Tivoli Storage Manager

with enterprise tape library) is, in this case, configured to retain a given number of versions of each file for a specific time period, generally days, weeks, or months. Longer retention policies of one or more years are rare and not likely.

In contrast, the archive solution can retain data for many years, with a given requirement to keep track of each file stored in the system. Typically, the file is removed from the source system after it is archived. After an amount of time, when using a retention product *without* a transparent content management process, the user is not able to recognize the original location of the file or perhaps even the file name.

In the following section, we briefly discuss Linear Tape-Open (LTO) Ultrium 5 tape technology and its benefits, together with the IBM System Storage TS3500 tape library.

11.3.1 IBM LTO Ultrium5 tape drive

The fifth generation of the widely-used, industry-proven LTO Ultrium5 technology enhances the benefits and performance of all previous generations, especially generation four. It provides nearly double the capacity, increased read/write access time to the tape, and enhanced encryption functionality.

Table 11-1 describes the technical and functional details of all available LTO generations and the planned parameters for future generations represented by the LTO product roadmap.

Table 11-1 Functional aspects of LTO Ultrium generations

Attribute	LTO1	LTO2	LTO3	LTO4	LTO5	LTO6	LTO7	LTO8
Release date	2000	2003	2005	2007	2010	TBA	TBA	TBA
Native Capacity [GB]	100	200	400	800	1500	3200	6400	12800
Max Speed [MBps]	15	40	80	120	140	270	315	472
Compression Capable?	Yes 2:1					Planned 2.5:1		
WORM Capable?	No		Yes			Planned		
Encryption Capable?	No			Yes		Planned		
Partition Capable?	No				Yes	Planned		

For more details about the LTO roadmap, which provides a clearly defined eight-generation development path for scalability and growth, see this website:

http://www.ultrium.com/technology/roadmap.html

LTO5 technology introduces a new feature named *tape partitioning*. At present, IBM offers only one filesystem or product which benefits from this functionality, the IBM Long Term File System (LTFS).

LTFS is a low-cost, self-describing, tape filesystem that provides an easy storage and distribution solution without the additional database application needs. LTFS provides the following features:

- A filesystem that is implemented on dual partitions of the LTO5 media
- Fast filesystem access to tape data (drag and drop files to and from tape)
- Files and directories that appear on desktop, directory listing
- Data exchange supported
- A simple, one-time intuitive installation

The latest available generation of IBM Ultrium LTO5 tape drives also incorporates additional features to increase reliability and consistence of the data on tape:

- *Data integrity:* The drive performs a read after write, for verification.
- *Power loss:* No recorded data is lost as a result of normal or abnormal power loss when the drive is reading or writing data.
- *Error correction:* Two levels of error correction that can provide recovery from longitudinal media scratches.
- *Integrated head cleaner:* During the load process, a brush integrated into the drive mechanism cleans the head before it is used. Additionally, IBM tape libraries introduce dedicated slots for special LTO cleaning cartridges (increasing the data life-cycle), together with the automatic cleaning functions of tape drives and libraries.
- *Surface control guiding:* This guides the tape along the tape path, using the surface of the tape, rather than the tape edges, to control tape motion.
- *Flat lap head:* Improves the contact between the read and write recording elements and the tape, producing higher quality recording and reading of the data.
- *Statistical Analysis and Reporting System (SARS):* Only IBM LTO drives provide this level of preventive diagnostic reporting to assist in isolating failures between media and hardware.

For more information, see the website at this website:

`http://www.ibm.com/servers/storage/tape/lto/`

11.3.2 IBM System Storage TS3500 tape library

The IBM System Storage TS3500 tape library (formerly known as IBM TotalStorage 3584) is the only IBM tape automation product available on the market that adheres to the strict classification of the enterprise tape library. The TS3500 is a highly scalable, automated tape library for backup and archive of mainframe and open systems, from midrange to enterprise environments. It has the following features:

- Extensive scalability up to 16 frames with 192 LTO5 tape drives and 20,000 cartridges
- Logical partitioning of the single physical tape library
- High availability of the library, with dual tape accessor robotic units
- High availability to the host systems, benefiting from Data Path and Control Path Failover
- High density of cartridge storage slots using tiered layers in high-density frames
- Advanced Library Management System for transparent library management
- Compatibility with mainframe and open systems
- Supports IBM tape virtualization engine TS7740 for System z
- Capacity-on-demand is managed transparently by license keys as data grows
- Supports a virtual I/O station to easily insert and remove a larger number of tapes
- A remote, easy-to-use web-based management interface with an IBM standard layout
- Application-managed and hardware-managed data encryption support

Figure 11-3 shows the maximum capable configuration of TS3500 tape library containing 16 frames.

Figure 11-3 Full configuration of IBM System Storage TS3500 with 16 frames

Seven types of frames are currently available to build a TS3500 tape library. Each frame is identified by a three character model number (L23, L53, D23, D53, S24, S54, or HA1) that describes the nature of the frame. Libraries are built of modules as follows:

- Every library requires a base frame (model L*xx*), to which optional expansion frames (model D*xx* or model S*xx*) can be added. Only one base frame is permitted per library.
- Base and expansion frames support either of these combinations:
  - LTO1 ... LTO 5 tape drives (*x*53 frames)
  - IBM System Storage 3592 model E and model J1A (*x*23 frames)
- An optional second accessor is made available by the addition of model HA1 frames in conjunction with a Service Bay frame.

IBM System Storage TS3500 tape library offers a wide scale of capacities and storage solutions, based on the concrete frame configuration. Here are the highlights:

- Number of tape cartridges:
  - 3592: L23 - up to 260; D23 - up to 400; S24 - up to 1000; (total > 15000)
  - LTO*x*: L53 - up to 287; D53 - up to 440; S54 - up to 1320; (total > 20000)
- Storage capacity:
  - 3592: up to 15 PB native data (45 PB using 3:1 compression)
  - LTO5: up to 30 PB native data (60 PB using 2:1 compression)
- Functional parameters:
  - Maximum of 16 library frames, including one base frame L*x*3
  - Maximum of 192 tape drives in the library (12 per L*x*3 or D*x*3 frame)
  - Maximum of 224 I/O slots in maximally 4 I/O station doors (minimum 16 I/O slots)

11.3.3 Advanced Library Management System

The Advanced Library Management System is an available feature that supports dynamic storage management, allowing the user to dynamically create and change logical libraries and configure any drive into any logical library. This product is required when attaching to mainframe environments and when integrating high-density frames into a library configuration. Based on capacity requirements, this library management system can be implemented as an entry level, intermediate, or full Capacity-on-Demand level in the TS3500 tape library. The Advanced Library Management System virtualizes the locations of cartridges in the TS3500. Logical libraries can then consist of unique drives and ranges of volume serial numbers instead of fixed locations.

This library management system is an extension of the IBM-patented, multi-path architecture. With this product, the TS3500 is the industry's first standards-based tape library to virtualize the locations of cartridges (called SCSI element addresses) and, at the same time, maintain the native SAN attachment for the tape drives. When you enable the library management system with its license key, you can assign tape drives to any logical library by using the Tape Library Specialist web interface. Logical libraries can also be added, deleted, or easily changed without disruption. Storage capacity can be changed without impact to the host applications.

The Advanced Library Management System offers dynamic management of cartridges, cartridge storage slots, tape drives, and logical libraries. It enables the TS3500 to achieve unprecedented levels of integration for functionality through dynamic partitioning, storage slot pooling, and flexible drive assignment. This product also reduces downtime when you add expansion frames, add or remove tape drives, or change the allocation of logical drives.

11.3.4 More information

For more information, see the following references:

- To become more familiar with the IBM System Storage TS3500 tape library, download *IBM System Storage TS3500 Tape Library Introduction and Planning Guide IBM 3584 Tape Library*, available at this website:

  `http://www.ibm.com/support/docview.wss?uid=ssg1S7001738`

- For the latest specifications, features and functions, see this website:

  `http://www.ibm.com/systems/storage/tape/ts3500/index.html`

- The enterprise tape library TS3500 is also discussed in *IBM System Storage Tape Library Guide for Open Systems*, SG24-5946.

- If you are interested in mainframe backup and archive environments, study the *IBM TS3500 Tape Library with System z Attachment A Practical Guide to Enterprise Tape Drives and TS3500 Tape Automation*, SG24-6789.

11.4 IBM System Storage N series SnapLock

Today's business relies on the frequent use of Write Once Read Many (WORM) data storage devices to meet regulatory compliance requirements or simply to add another layer to the data protection roadmap. This section discusses the integration of N series into environments that require WORM data storage for retention and compliance.

The use of SnapLock addresses issues faced by growing business requirements for WORM data storage, and it alleviates issues inherent with traditional WORM storage solutions. The N series SnapLock product allows clients to implement the data permanence functionality of traditional WORM storage in an easier-to-manage, faster-to-access, lower-cost, disk-based solution.

11.4.1 SnapLock Compliance and SnapLock Enterprise

The SnapLock Compliance and SnapLock Enterprise software products both provide nonerasable, nonrewritable, WORM data permanence functionality, utilizing high-throughput, magnetic disk drives in a cost-efficient, highly available RAID configuration. From a data protection perspective, the process of committing data to WORM status on either SnapLock product can be thought of in the same manner as storing data on an optical platter.

Like an optical platter burned with data, both SnapLock software products protect data that is committed to WORM status from any possible alteration or deletion until their retention period has expired. Although SnapLock Compliance and SnapLock Enterprise data permanence is analogous to traditional optical WORM media, comparisons end there. SnapLock offers performance and reliability improvements over traditional WORM storage, and reduces both maintenance impact and total cost of ownership.

SnapLock Compliance and SnapLock Enterprise are implemented as add-on licenses to Data ONTAP (version 7.3.1. and higher). Both SnapLock software products run on N series. These storage systems feature lower cost, higher capacity Advanced Technology Attachment-(ATA-)based drives on the EXN1000 or EXN3000 models. The higher-performance EXN4000 features Fiber-Channel attached disk drives.

SnapLock Compliance

SnapLock Compliance is designed to assist organizations in implementing a comprehensive archival solution for meeting Security Exchange Commission or governmental regulations for data retention. Records and files committed to WORM storage on a SnapLock Compliance volume cannot be altered or deleted before the expiration of their retention period. Moreover, a SnapLock Compliance volume cannot be deleted until all data stored on it has passed its retention period and has been deleted by the archival application or another process.

SnapLock Enterprise

SnapLock Enterprise is geared towards assisting organizations with meeting self-regulated and best practice guidelines for protecting digital assets with WORM-type data storage. Data stored as WORM on a SnapLock Enterprise volume is protected from alteration or modification with one main difference from SnapLock Compliance. As the data being stored is not for regulatory compliance, a SnapLock Enterprise volume can be deleted, including the data it contains, by an administrator with system privilege.

A comparison of SnapLock Compliance and SnapLock Enterprise is shown in Figure 11-4.

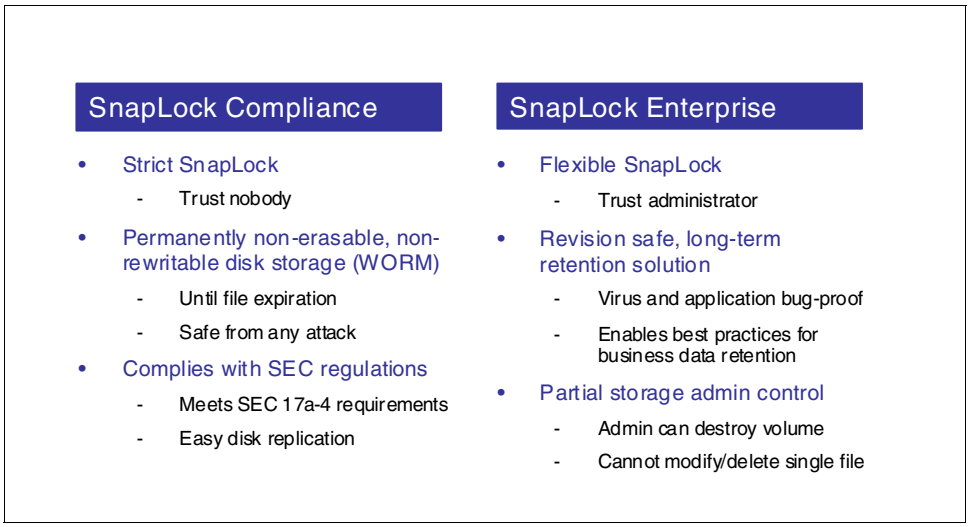

Figure 11-4 Differences between the SnapLock Compliance and SnapLock Enterprise

11.4.2 SnapLock volume usage

SnapLock WORM protection works at the individual file level. However, creating a file or copying a read-only file to a SnapLock volume does not trigger the WORM commit operation required for SnapLock data protection.

Rather, committing a file to WORM on a SnapLock volume for both Microsoft and UNIX environments requires a two-step sequence:

1. Using a CIFS share or NFS mount, copy or create a file with write permission on a SnapLock volume.
2. Change the permission on the file to read-only.

The two-step process for triggering a WORM commit can be manual for testing purposes, or scripted for in-house archival application development. In most enterprises, applications drive the archival process, and those certified with SnapLock interoperability perform the WORM commit process.

The SnapLock commit action occurs after a transition from a writable state to a read-only state no matter who owns the file or which user carried out the SnapLock trigger operation. After the commit of a file to the SnapLock state occurs, any modification or deletion attempt will fail (even from traditional superuser accounts) until the file retention period has expired.

Tip: An important consideration for SnapLock storage is how the directories are treated. Directories, after being created on a SnapLock volume, cannot be renamed, regardless of the access permissions of the directory.

Although directories cannot be renamed, they can be deleted as long, as no files committed to WORM state are contained within their hierarchy.

11.4.3 Setting retention period with SnapLock

In keeping with the open protocol design of SnapLock, support for record retention periods was implemented without the need for the use of proprietary APIs or protocols. Record retention dates can be set and queried programmatically through standard system call interfaces supplied by most operating systems or interactively through standard command line tools.

Operations for setting retention dates occur over standard network filesystem interfaces, such as NFS and CIFS. This flexibility allows applications to utilize SnapLock from compiled code or scripts, without the requirement for any libraries or software being installed on the client systems.

The retention date for WORM records on a SnapLock volume is stored in the last access time stamp of the record file metadata. To set a retention date for a WORM record, the application must explicitly set the file last access time to the desired retention date before setting the file to read-only and engaging the WORM commit operation. After being committed to the WORM state, the access time of the file is immutable, with the only exception being for extending the record retention period.

11.4.4 SnapLock disaster protection

The N series offers multiple methods to ensure data protection and security against disaster. Considerations regarding cost, criticality, location, and type of data determine which one is best for you.

Replication to remote site

Compliance with data retention rules and regulatory agencies might require keeping a second copy of archived data at a remote site. The most straightforward and natural way to comply with this requirement is to replicate data from a primary N series to a secondary N series at a separate location. See Figure 11-5.

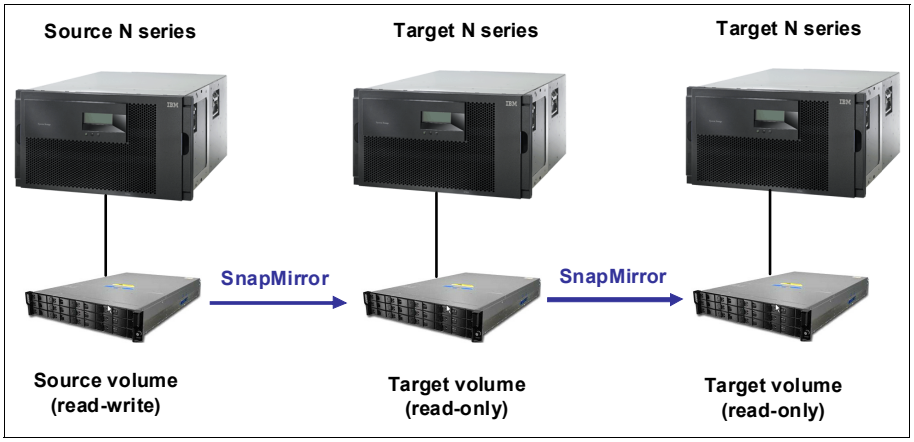

Figure 11-5 SnapMirror of SnapLock volumes to a separate N series

There are two integrated N series solutions available to seamlessly perform data replication:

- The easiest and most robust solution is using SnapMirror to replicate data to a remote location. SnapMirror, in either synchronous or asynchronous mode, replicates SnapLock data to a remote appliance, and maintains all aspects of the original WORM file, such as the date and time stamp and file name, including the path.
- The second solution, **`ndmpcopy`**, is a no-cost utility and is already bundled with Data ONTAP. Like SnapMirror, **`ndmpcopy`** maintains WORM aspects of the original files in the replicated copy. SnapMirror can replicate at either the volume or quota tree (also known as the qtree) level.

Tape backup

An N series offers substantial performance improvement and storage capabilities for near-line data storage over optical or tape-based storage. Even so, tape backups, including off-site tape rotation, are still a valuable part of an overall data protection strategy for enterprises. See Figure 11-6.

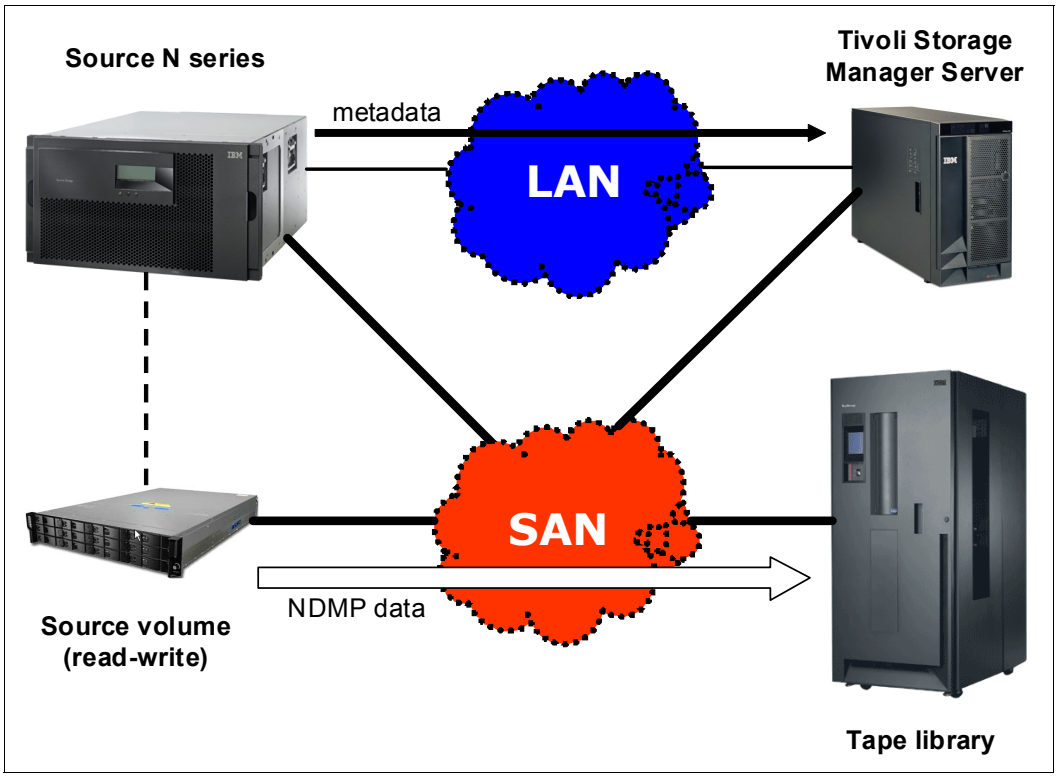

Figure 11-6 Backup of SnapLock data to tape, using SAN infrastructure and Tivoli Storage Manager

Industry-standard NDMP-generated backups (see “Replication to remote site” on page 224) of SnapLock data include all file WORM metadata for each file, so that a subsequent NDMP restore to a new SnapLock volume puts the data back into its original WORM state.

11.4.5 More information

For more information, see the following documentation:

- Further comprehensive information about the N series and the SnapLock feature is located in *IBM System Storage N series Software Guide*, SG24-7129.

  http://www.redbooks.ibm.com/abstracts/sg247129.html

- Hardware related information about the N series is described in the *IBM System Storage N Series Hardware Guide*, SG24-7840.

  http://www.redbooks.ibm.com/abstracts/sg247840.html

- Tape backup of the N series is discussed in *Using the IBM System Storage N series with IBM Tivoli Storage Manager*, SG24-7243.

  http://www.redbooks.ibm.com/abstracts/sg247243.html

- N series SnapMirror and its complete configuration are described in *IBM System Storage N Series SnapMirror*, SG24-7260.

  http://www.redbooks.ibm.com/abstracts/sg247260.html

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this book.

IBM Redbooks publications

The following IBM Redbooks publications provide additional information about the topic in this document. Note that some publications referenced in this list might be available in softcopy only.

- *IBM System Storage Solutions Handbook*, SG24-5250
- *Deployment Guide Series: IBM TotalStorage Productivity Center for Data*, SG24-7140
- *Designing an IBM Storage Area Network*, SG24-5758
- *IBM b-type Data Center Networking: Design and Best Practices Introduction*, SG24-7786
- *IBM b-type Data Center Networking: Product Introduction and Initial Setup*, SG24-7785
- *IBM j-type Data Center Networking Introduction*, SG24-7820
- *Implementation of IBM j-type Ethernet Switches and Routers*, SG24-7882
- *Implementation of IBM j-type Ethernet Appliances*, SG24-7883
- *IBM/Cisco Multiprotocol Routing: An Introduction and Implementation*, SG24-7543
- *IBM Midrange System Storage Implementation and Best Practices Guide*, SG24-6363
- *IBM System Storage DS Storage Manager Copy Services Guide*, SG24-7822
- *IBM Midrange System Storage Hardware Guide*, SG24-7676
- *IBM SAN Survival Guide*, SG24-6143
- *IBM System Storage DS3000: Introduction and Implementation Guide*, SG24-7065
- *IBM System Storage DS3500 Introduction and Implementation Guide*, SG24-7914
- *IBM System Storage DS4000 and Storage Manager V10.30*, SG24-7010
- *IBM System Storage DS8000: LDAP Authentication*, REDP-4505
- *IBM System Storage DS8000: Remote Pair FlashCopy (Preserve Mirror)*, REDP-4504
- *IBM System Storage DS8000 Series: IBM FlashCopy SE*, REDP-4368
- *IBM System Storage DS8700 Disk Encryption*, REDP-4500
- *IBM System Storage DS8800: Architecture and Implementation*, SG24-8886
- *IBM System Storage DS8700 Easy Tier*, REDP-4667
- *IBM SAN Volume Controller and IBM Tivoli Storage FlashCopy Manager*, REDP-4653
- *IBM System Storage Tape Library Guide for Open Systems*, SG24-5946
- *IBM System Storage TS7650,TS7650G and TS7610*, SG24-7652
- *TS7680 Deduplication ProtecTIER Gateway for System z*, SG24-7796
- *IBM Information Archive Architecture and Deployment*, SG24-7843
- *IBM System Storage Data Encryption*, SG24-7797

- *Using IBM Tivoli Key Lifecycle Manager: Business Benefits and Architecture Overview*, REDP-4529
- *IBM Tivoli Key Lifecycle Manager for z/OS*, REDP-4472
- *IBM Tivoli Storage Productivity Center V4.1 Release Guide*, SG24-7725
- *IBM TotalStorage Virtual Tape Server: Planning, Implementing, and Monitoring*, SG24-2229
- *IBM TotalStorage 3494 Tape Library: A Practical Guide to Tape Drives and Tape Automation*, SG24-4632
- *IBM Virtualization Engine TS7500: Planning, Implementation, and Usage Guide*, SG24-7520
- *IBM Virtualization Engine TS7510: Tape Virtualization for Open Systems Servers*, SG24-7189
- *IBM Virtualization Engine TS7700 with R1.7*, SG24-7712
- *IBM XIV Storage System: Architecture, Implementation, and Usage*, SG24-7659
- *Implementing an IBM/Cisco SAN*, SG24-7545
- *Implementing the IBM System Storage SAN Volume Controller V6.1*, SG24-7933
- *Implementing the IBM Storwize V7000*, SG24-7938
- *Introduction to Storage Area Networks*, SG24-5470
- *IP Storage Networking: IBM NAS and iSCSI Solutions*, SG24-6240
- *IBM System Storage N Series Hardware Guide*, SG24-7840
- *IBM Scale Out Network Attached Storage: Architecture, Planning, and Implementation Basics*, SG24-7875
- *The IBM TotalStorage NAS Integration Guide*, SG24-6505
- *The IBM Virtualization Engine TS7510: Getting Started with i5/OS and Backup Recovery and Media Services*, SG24-7510
- *An Introduction to Fibre Channel over Ethernet, and Fibre Channel over Convergence Enhanced Ethernet*, REDP-4493

Online resources

These websites are also relevant as further information sources:

- IBM System Storage: Hardware, Software, and Service Solutions:

  `http://www.ibm.com/systems/storage/`
- IBM Product Information centers:

  `http://www.ibm.com/support/publications/us/library/`
- Storage Networking Industry Association:

  `http://www.snia.org`

Help from IBM

- IBM Support and downloads:
  http://www.ibm.com/support
- IBM Global Services:
  http://www.ibm.com/services

Index

D

E

F

G

H

I

L

M

N

O

P

R

S

T

U

IBM System Storage Solutions for Smarter Systems

(0.5” spine)
0.475”<->0.873”
250 <-> 459 pages

IBM System Storage Solutions for Smarter Systems

Get an overview of IBM storage technologies, products, and solutions

Learn about the different aspects of storage efficiency

Understand the latest concepts in data protection

IBM offers storage solutions that address the key needs of storage efficiency and data protection and retention. This IBM Redbooks publication shows that IBM has the technology, products, and services to help keep your storage systems enabled and your business flourishing.

The information presented is structured as follows:

Storage efficiency:

- Virtualization
- Tiering
- Thin provisioning
- Flexible delivery
- Compression
- Deduplication
- Visibility

Data protection:

- Backup and restore
- Continuous operations
- Retention and compliance

This book has been written for those most concerned about efficient, protected, and available data, along with maintaining their IT environment at a competitive price. The audience for this book is that of IT professionals, such as CEOs, CIOs, CFOs, practitioners, information architects, and current and prospective IBM clients and Business Partners.

INTERNATIONAL TECHNICAL SUPPORT ORGANIZATION

BUILDING TECHNICAL INFORMATION BASED ON PRACTICAL EXPERIENCE

IBM Redbooks are developed by the IBM International Technical Support Organization. Experts from IBM, Customers and Partners from around the world create timely technical information based on realistic scenarios. Specific recommendations are provided to help you implement IT solutions more effectively in your environment.

For more information:
ibm.com/redbooks

SG24-7945-00 ISBN 073843583X