Redbooks
ibm.com/redbooks

# IBM Private, Public, and Hybrid Cloud Storage Solutions


Larry Coyne

Joe Dain

Phil Gilmer

Patrizia Guaitani

Ian Hancock

Antoine Maille

Tony Pearson

Brian Sherman

Christopher Vollmar


Cloud

Storage

IBM

Redpaper

International Technical Support Organization

# IBM Private, Public, and Hybrid Cloud Storage Solutions

January 2017

REDP-4873-03

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

**Fourth Edition (January 2017)**

This document was created or updated on January 17, 2017.

# Contents

# Notices

This information was developed for products and services offered in the US. This material might be available from IBM in other languages. However, you may be required to own a copy of the product or product version in that language in order to access it.

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 grant you any license to these patents. You can send license inquiries, in writing, to:
*IBM Director of Licensing, IBM Corporation, North Castle Drive, MD-NC119, Armonk, NY 10504-1785, US*

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 jurisdictions 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 provide in any way it believes appropriate without incurring any obligation to you.

The performance data and client examples cited are presented for illustrative purposes only. Actual performance results may vary depending on specific configurations and operating conditions.

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.

Statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

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 actual people or business enterprises 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 programming 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. The sample programs are provided “AS IS”, without warranty of any kind. IBM shall not be liable for any damages arising out of your use of the sample programs.

# Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation, registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at “Copyright and trademark information” at http://www.ibm.com/legal/copytrade.shtml

The following terms are trademarks or registered trademarks of International Business Machines Corporation, and might also be trademarks or registered trademarks in other countries.

Active Cloud Engine®
AIX®
Aspera®
Bluemix®
Cloudant®
Cognos®
DB2®
developerWorks®
DS8000®
Easy Tier®
ECKD™
FlashCopy®
Global Technology Services®
GPFS™
HiperSockets™
HyperFactor®
HyperSwap®
IBM®
IBM Cloud Managed Services®
IBM Elastic Storage™
IBM FlashCore®
IBM FlashSystem®
IBM Flex System®
IBM SmartCloud®
IBM Spectrum™
IBM Spectrum Accelerate™
IBM Spectrum Archive™
IBM Spectrum Control™
IBM Spectrum Protect™
IBM Spectrum Scale™
IBM Spectrum Storage™
IBM Spectrum Virtualize™
IBM Watson™
IBM z Systems®
Insight™
Linear Tape File System™
LSF®
MicroLatency®
POWER8®
ProtecTIER®
PureSystems®
Real-time Compression™
Redbooks®
Redpaper™
Redbooks (logo) ®
Storwize®
System Storage®
Tivoli®
XIV®
z Systems®

The following terms are trademarks of other companies:

Accesser, Cleversafe, Dispersed Storage, Slicestor, and C device are trademarks or registered trademarks of Cleversafe, Inc., an IBM Company.

SoftLayer, and SoftLayer device are trademarks or registered trademarks of SoftLayer, Inc., an IBM Company.

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.

ITIL is a Registered Trade Mark of AXELOS Limited.

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

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.

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

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

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

# Preface

This IBM® Redpaper™ publication takes you on a journey that surveys cloud computing to answer several fundamental questions about storage cloud technology. What are storage clouds? How can a storage cloud help solve your current and future data storage business requirements? What can IBM do to help you implement a storage cloud solution that addresses these needs?

This paper shows how IBM storage clouds use the extensive cloud computing experience, services, proven technologies, and products of IBM to support a smart storage cloud solution designed for your storage optimization efforts. Clients face many common storage challenges and some have variations that make them unique. It describes various successful client storage cloud implementations and the options that are available to meet your current needs and position you to avoid storage issues in the future. IBM Cloud Services (IBM Cloud Managed Services® and IBM SoftLayer®) are highlighted as well as the contributions of IBM to OpenStack cloud storage.

This paper is intended for anyone who wants to learn about storage clouds and how IBM addresses data storage challenges with smart storage cloud solutions. It is suitable for IBM clients, storage solution integrators, and IBM specialist sales representatives.

## Authors

This paper was produced by a team of specialists from around the world working at the IBM Client Center Montpellier, France.

**Larry Coyne** is a Project Leader at the International Technical Support Organization, Tucson Arizona Center. He has 34 years of IBM experience with 23 in IBM storage software management. He holds degrees in Software Engineering from the University of Texas at El Paso and Project Management from George Washington University. His areas of expertise include client relationship management, quality assurance, development management, and support management for IBM Storage Management Software.

**Joe Dain** is a Senior Engineer and Master Inventor in Tucson, Arizona and works in the Storage and Software Defined Infrastructure CTO Office. He joined IBM in 2003 with a BS in Electrical Engineering. His areas of expertise include backup, restore, disaster recovery, object storage, data reduction techniques such as data deduplication and compression, and emerging storage technology trends. He is on his fourteenth IBM invention plateau with over 60 patents issued and pending worldwide, including 22 high-value patents.

**Phil Gilmer** is an Information Infrastructure Consultant in the IBM Systems and Technology Group Lab Consulting, in Tucson, Arizona, United States. He has 33 years of experience with IBM in various roles including Test Engineer, Systems Engineer, Network Integrator, Sales Representative, and PMP Certified Project/Program Manager. He holds a Bachelor’s degree in Electrical Engineering from Michigan State University. His current areas of expertise include business consulting and project management in information infrastructure architecture, business case development, information lifecycle, and availability management. He works with clients worldwide to perform cloud and storage IT optimization engagements. He has written numerous optimization reports, white papers, and IBM Redpaper publications related to storage virtualization and tiering.

**Patrizia Guaitani** is an Executive Infrastructure Architect and Manager of Client Technical Specialist based in Milan, Italy. Patrizia has more than 30 years of IT infrastructure experience spread on virtualization, high availability, cloud, and hybrid cloud project. She is mainly focused on helping customers build solutions using the IBM Software Defined Infrastructure offering, and develop private and hybrid storage cloud solutions using the IBM Spectrum™ Storage family and Converged Infrastructure solutions.

**Ian Hancock** is a Technical Sales specialist currently working with software-defined storage products for customers in the United Kingdom. He has over 37 years of experience in the IT industry, working for international vendors of IT systems in the UK and EMEA regions. He has worked in many fields from hardware development, services, consultancy, project management, and line management through to technical sales. Most of his career in IBM has been in storage software technical sales. However, he is also a qualified ITIL Manager, which has helped to bring a service management perspective to his approach for solving data management problems. His current area of interest is to help customers use the latest software-defined storage products and technologies to modernize their data protection environments by using innovative architectures and designs.

**Antoine Maille** is an IBM Certified Architect expert. Since 2002, he has been involved in planning and leading large distributed environments infrastructure projects. Initially, he worked as the benchmark manager responsible for testing and qualifying new products in real customer contexts. Currently, Antoine is one of the leaders of the storage design center at the IBM Client Center in Montpellier, France.

**Tony Pearson** is a Master Inventor and Senior Software Engineer in the IBM Tucson Executive Briefing Center, and is a subject matter expert for all IBM storage hardware and software solutions. He has worked on IBM Storage for more than 30 years, with 19 patents for storage solutions and technologies. He is known for his “Inside System Storage®” blog, one of the most popular blogs on IBM developerWorks®. He has a bachelor's degree in Computing Engineering and master's degree in Electrical Engineering, both from the University of Arizona.

**Brian Sherman** is an IBM Distinguished Engineer with over thirty years’ experience as an I/T Specialist since joining IBM from McMaster University in 1985 with a Mathematics and Computer Science Degree. Brian has been involved in Storage since joining IBM and has held various storage related roles including level 2 software support, storage implementation services and branch Systems Engineer in the Public and Financial Sectors. Brian currently is the technical lead for Software Defined Storage (SDS), Spectrum Storage Family, IBM DS8000®, and XIV/A9000 in the World Wide Advanced Technical Skills (ATS) organization. He also develops and provides World Wide technical education on new Storage Hardware and Software product launches and participates on several Storage Product Development Teams.

**Christopher Vollmar** is a Storage Solutions Specialist and Storage Client Technical Specialist based in Toronto, Ontario, Canada with the IBM Systems Group. Christopher is currently focused on helping customers build storage solutions using the IBM Spectrum Storage / Software Defined Storage family. He is also focused on helping customers develop private and hybrid storage cloud solutions using the IBM Spectrum Storage family and Converged Infrastructure solutions. Christopher has worked for IBM for over 14 years across a number of different areas of the IBM businessm, including System Integrators and System x. He has spent the past 7 years working with the IBM System Storage team in Toronto as a Client Technical Specialist and Solution Specialist. Christopher holds an honours degree in Political Science from York University and holds an IT Specialist-Storage certification from The Open Group.

Thanks to the following people for their contributions to this project:

Ann Lund
**International Technical Support Organization**

Erwan Auffray
Marc Bouzigues
Olivier Fraimbault
Benoit Granier
Joelle Haman
Thierry Huche
Hubert Lacaze
Marc Lapierre
Khanh Ngo
Christine O'Sullivan
Gauthier Siri
Olivier Vallod
**IBM Systems**

Thanks to the authors of the previous editions of this paper:

Larry Coyne
Mark Bagley
Gaurav Chhaunker
Phil Gilmer
Shivaramakrishnan Gopalakrishnan
Patrizia Guaitani
Tiberiu Hajas
Magnus Hallback
Mikael Lindström
Daniel Michel
John Sing
Hrvoje Stanilovic
Christopher Vollmar

## 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 papers to be as helpful as possible. Send us your comments about this paper 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

1

# What is cloud computing

Before focusing specifically on storage clouds, it is useful to describe the larger IT landscape for a general understanding of cloud computing concepts. The trade press, journals, and marketing collateral have generated substantial content about cloud computing, but differ widely in exactly what constitutes an IT cloud. A helpful way to think is in general terms of ownership (public and private clouds), and categorize the types of services that an IT cloud provides. These concepts are described in some detail in this chapter. Finally, the IBM Cloud Computing Reference Architecture is described as a definition of the basic elements of any cloud service environment.

This chapter includes the following sections:

- Cloud computing definition
- What is driving IT and businesses to cloud
- Introduction to cloud service models
- Introduction to cloud delivery models
- The IBM Cloud Computing Reference Architecture
- Hybrid Cloud Use Cases
- Cloud enabled data center journey
- Software-defined environment
- Role of OpenStack cloud software in cloud computing
- Storage cloud components within overall cloud

## 1.1 Cloud computing definition

The National Institute of Standards and Technology (NIST) provides the following definition[1] for cloud computing:

> "Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model is composed of five essential characteristics, three service models, and four deployment models."
>
> For a service to be considered a *cloud service*, NIST describes the following "Essential characteristics:
>
> **On-demand self-service**
>
> A consumer can unilaterally provision computing capabilities, such as server time and network storage, as needed automatically without requiring human interaction with each service provider.
>
> **Broad network access**
>
> Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms (e.g., mobile phones, tablets, laptops, and workstations).
>
> **Resource pooling**
>
> The provider's computing resources are pooled to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand. There is a sense of location independence in that the customer generally has no control or knowledge over the exact location of the provided resources but may be able to specify location at a higher level of abstraction (e.g., country, state, or data center). Examples of resources include storage, processing, memory, and network bandwidth.
>
> **Rapid elasticity**
>
> Capabilities can be elastically provisioned and released, in some cases automatically, to scale rapidly outward and inward commensurate with demand. To the consumer, the capabilities available for provisioning often appear to be unlimited and can be appropriated in any quantity at any time.
>
> **Measured service**
>
> Cloud systems automatically control and optimize resource use by leveraging a metering capability, typically this is done on a pay-per-use or charge-per-use basis, at some level of abstraction appropriate to the type of service (e.g., storage, processing, bandwidth, and active user accounts). Resource usage can be monitored, controlled, and reported, providing transparency for both the provider and consumer of the utilized service."

[1] See NIST Special Publication (SP) 800-145, A NIST Definition of Cloud Computing: http://dx.doi.org/10.6028/NIST.SP.800-145

Users interact with cloud computing environments with the services that the cloud environment provides. The following examples are of services that are provided by a cloud (cloud services):

- Virtual servers
- Database services
- Email applications
- Storage

A company can use cloud services that are provided by third parties, or it can build its own cloud. The company can then provide services from the cloud to internal company users, to selected business partners or customers, or to the world at large.

To provide these characteristics, the infrastructure that enables the cloud services takes advantage of two key enablers:

- Virtualization: Allows computing resources to be pooled and allocated on demand. It also enables pay-per-use billing to be implemented.
- Automation: Allows for the elastic use of available resources, and for workloads to be moved to where resources are available. It also supports provisioning and deprovisioning of service instances to support scalability.

Although these enablers are not part of any formal cloud definition, they are indispensable in delivering the essential cloud service characteristics.

Many traditional IT services are provisioned with some of the characteristics of a cloud service. So how do you know that you are providing a cloud service, or when you are using a cloud service? You know that you are *providing* a cloud service when your service exhibits the characteristics listed previously, typically provisioned by using the virtualization and automation enablers.

As the user of any service, whether it is being provisioned as a cloud service might be immaterial. However, you are likely to be *using* a cloud service when the service that you are using exhibits the characteristics that are listed previously. From a cloud user perspective, it is important that you are able to perform self-service activities to quickly provision new service instances and have resources that are elastically sized to meet your changing processing demands.

## 1.2 What is driving IT and businesses to cloud

Cloud computing has clearly moved beyond the hype and into the mainstream reality of today's IT environments. What are the drivers for this rapid adoption and disruption in the traditional IT world? In the past couple of years, this question has been thoroughly studied and documented by numerous sources. One example is the December 2014, KPMG Consulting report, "The Cloud Service Providers Survey", in which the reasons shown in Figure 1-1 for cloud adoption were derived.

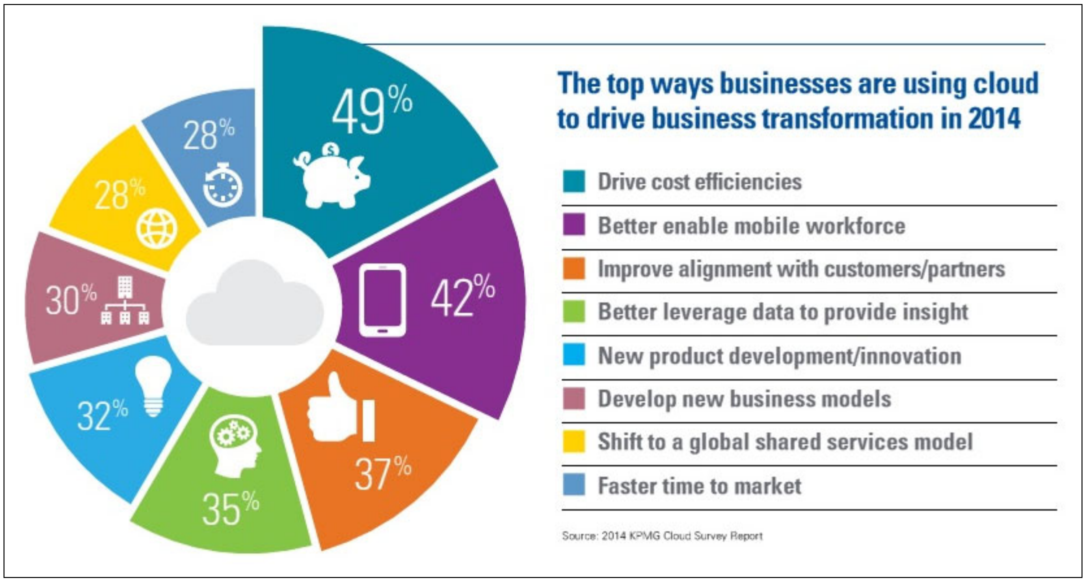



*Figure 1-1 Reasons to use cloud environments are business reasons*

As you can see, although cost savings are important, what is even more noteworthy is the number of business-related, time-to-market, competitive advantage, business revenue-related aspects of the move to the cloud. In aggregate, these business reasons are at least as important, if not more so, than the cost reduction. For more insight into the data behind what is driving cloud decisions, see the following website:

http://www.kpmginfo.com/EnablingBusinessInTheCloud/downloads/7397-CloudSurvey-Rev1-5-15.pdf

## 1.3 Introduction to cloud service models

When discussing cloud services (identified in 1.1, "Cloud computing definition" on page 2), a helpful approach is to organize service capabilities into groups. NIST formally describes a standard for grouping cloud services, referring to them as *service models*. The following sections describe the NIST service models.

### 1.3.1 Infrastructure as a service (IaaS)

The IaaS model is the simplest for cloud service providers to provision. It can include the following elements:

- Processing
- Storage
- Network

Each of these elements is provisioned in an elastic fashion. As an IaaS user, you can deploy and run your chosen software, including operating systems and applications. You do not need to manage or control the underlying cloud infrastructure, but you have control over the operating systems, storage, and deployed applications. You might also have limited control over select networking components such as host firewalls.

### Role of predefined IaaS offerings

Cloud IT preferred practices require different workflows and relationships among the functions of IT as compared to traditional IT. These practices require an IT reorganization like the cloud workflow to truly provide cloud IT services. These substantial required changes then create the following IT management and technical questions:

- How best to begin and accelerate the needed realignment of the IT organization?
- How best to redeploy existing skills and experienced personnel in this new cloud-oriented organization?
- What technologies and tools are available to address and implement the new, different cloud workflow and the newly required skill sets?

For almost any organization, the magnitude of effort that is required to construct internal custom-built answers to these questions from scratch is daunting and often is not feasible.

This concern is why proven, pre-built, pre-tested cloud workflow IaaS offerings are so popular for organizations that need to change quickly to stay competitive. IaaS offerings already implement the cloud preferred practices workflow, and good IaaS offerings come with a system of proven experience and proven users. By adopting a proven IaaS solution, an IT organization can obtain and implement a reliable template and toolset to create true cloud capabilities within the IT organization.

Examples of commercial implementations of IaaS include IBM SoftLayer, IBM Cloud Managed Services, IBM Cloud Managed Backup, Amazon Elastic Compute Cloud (EC2), and Rackspace.

## 1.3.2 Platform as a service (PaaS)

The PaaS model includes services that build on IaaS services. They add value to the IaaS services by providing a platform on which the cloud users can provision their own applications, or conduct application development activities. The user does not need to manage the underlying cloud infrastructure (network, storage, operating systems), but can control configuration of the provisioned platform services. The following services are provisioned in PaaS models:

- Middleware
- Application servers
- Database servers
- Portal servers
- Development runtime environments

Examples of commercial implementations of PaaS environments include IBM Bluemix®, IBM Cloudant®, Amazon Relational Database Service, and Microsoft Azure.

### 1.3.3 Software as a service (SaaS)

The SaaS model provides software services that are complete applications that are ready to use. The cloud user simply connects to the application, which is running at a remote location. The user might not know where the system is located. The cloud service provider is responsible for managing the cloud infrastructure, the system on which the application is running, and the application itself. This approach eliminates the need for the users to install and run the application on their own computers, significantly reducing the need for maintenance and support.

SaaS is sometimes referred to as *applications as a service* because SaaS essentially provides applications as a service, rather than just software in general. SaaS also includes content services (for example, video on demand) and higher value network services (for example VoIP) than typically encountered in communication service provider scenarios.

Examples of commercial implementations of SaaS environments include IBM Watson™ Analytics, IBM API Management on Cloud, IBM Payment Systems, SalesForce, and NetSuite.

### 1.3.4 Other service models

Since the publishing of the NIST Cloud Computing definition, various new service delivery models have been coined:

- Business process as a service (BPaaS)
- Storage as a service (STaaS)
- Disaster recovery as a service (DRaaS)

The BPaaS model combines software and workflow elements to deliver end-to-end business processes as a service. Many business processes have the potential to be delivered through, vertical markets (such as healthcare and insurance). BPaaS allows businesses to pass on some of their day-to-day operating costs to service providers by using a fee-for-service model so that the businesses can focus on their core competencies.

Examples of commercial implementations of BPaaS include IBM Source to Pay on Cloud, IBM Customer Experience on Cloud, IBM Watson Business Solutions, ADP HR, and Google AdSense.

## 1.3.5 Cloud service model layering

Figure 1-2 illustrates how the service models described previously can be layered. It also contrasts the level of effort required of the service provider with that of the service user through the service model layers. As you travel *up* the service model layers, the service provider is responsible for providing more effort as the level of functionality increases. By contrast, as you travel *down* the service layers, the service user must provide more effort in terms of environment customization. For more information about service providers, service users, and other roles, see 1.5.2, "Cloud service roles" on page 13.

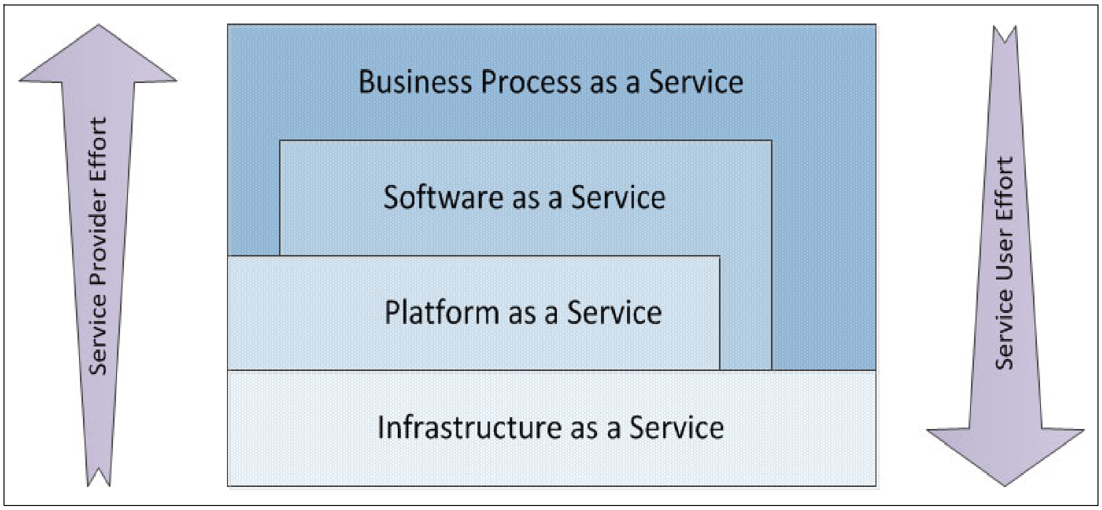



*Figure 1-2 Cloud service model layering*

Table 1-1 lists the functions that are provided by the cloud service provider and the cloud service user for each service model. For any service model, the service provider also provides the functions that are listed in the service models below it. The cloud user provides the functions listed in the service models above it, if required, as indicated by the arrows in the table.

*Table 1-1 Cloud service provider and service user responsibilities by service model*

| **Service model** | **Cloud service provider delivered functions** | **Cloud user delivered functions** |
|---|---|---|
| Business process as a service | Business process ↓ | Business process configuration ↑ |
| Software as a service | Applications ↓ | Application configuration ↑ |
| Platform as a service | Languages<br>Libraries<br>Tools<br>Middleware<br>Application servers<br>Database servers ↓ | Applications ↑ |
| Infrastructure as a service | Processing<br>Storage<br>Network | Languages<br>Libraries<br>Tools<br>Middleware<br>Application servers<br>Database servers |

## 1.4 Introduction to cloud delivery models

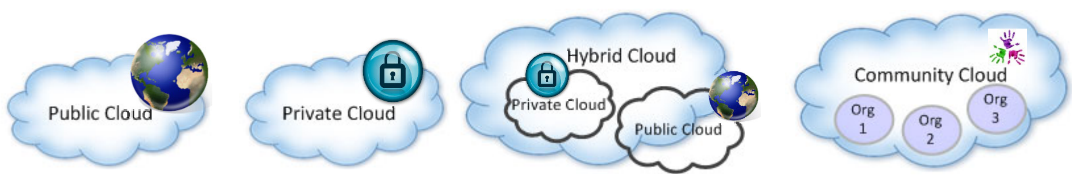
Cloud delivery models refer to how a cloud solution is used by an organization, where the data is located, and who operates the cloud solution. Cloud computing supports multiple delivery models that can deliver the capabilities needed in a cloud solution.

The cloud delivery models are as follows:

- Public cloud
- Private cloud
- Hybrid cloud
- Community cloud

These delivery models provide services in line with the service models described in 1.3, “Introduction to cloud service models” on page 4. You can integrate them with existing IT systems and other clouds.

Figure 1-3 illustrates these cloud delivery models, and identifies some of their characteristics in terms of roles, users, and accessibility.




| | | | | |
|---|---|---|---|---|
| **Host** | Provider | Enterprise, 3rd Party | Enterprise, 3rd Party | Community, 3rd Party |
| **Owner** | Provider | Enterprise | Enterprise | Community |
| **Access** | Internet | Intranet, VPN | Intranet, VPN | Intranet, VPN |
| **Users** | Public Individuals Organizations | Business Units | Business Units | Community Members |

*Figure 1-3 Cloud delivery models*

### 1.4.1 Public clouds

A public cloud is one in which the cloud infrastructure is made available to the general public or a large industry group over the Internet. The infrastructure is not owned by the user, but by an organization that provides cloud services. Services can be provided either at no cost, as a subscription, or as a pay-as-you-go model.

Examples of public clouds include IBM SoftLayer, Amazon Elastic Compute Cloud (EC2), Google AppEngine, and Microsoft Azure App Service.

### 1.4.2 Private clouds

A private cloud refers to a cloud solution where the infrastructure is provisioned for the exclusive use of a single organization. The organization often acts as a cloud service provider to internal business units that obtain all the benefits of a cloud without having to provision their own infrastructure. By consolidating and centralizing services into a cloud, the organization benefits from centralized service management and economies of scale.

A private cloud provides an organization with some advantages over a public cloud. The organization gains greater control over the resources that make up the cloud. In addition, private clouds are ideal when the type of work being done is not practical for a public cloud because of network latency, security, or regulatory concerns.

A private cloud can be owned, managed, and operated by the organization, a third party, or a combination. The private cloud infrastructure is usually provisioned on the organization's premises, but it can also be hosted in a data center that is owned by a third party.

IBM uses the term *Local* when referring to on-premises private clouds that are owned, managed, and operated by the organization, and the term *Dedicated* when referring to off-premise third-party managed private clouds.

### 1.4.3 Hybrid clouds

A hybrid cloud, as the name implies, is a combination of various cloud types (public, private, and community). For more information, see 1.4.4, "Community clouds" on page 9. Each cloud in the hybrid mix remains a unique entity, but is bound to the mix by technology that enables data and application portability.

The hybrid approach allows a business to take advantage of the scalability and cost-effectiveness of off-premise third-party resources without exposing applications and data beyond the corporate intranet. A well-constructed hybrid cloud can service secure, mission-critical processes, such as receiving customer payments (a private cloud service), and secondary processes such as employee payroll processing (a public cloud service).

The challenge for a hybrid cloud is the difficulty in effectively creating and governing such a solution. Services from various sources must be obtained and provisioned as though they originated from a single location, and interactions between on-premises and off-premise components make the implementation even more complicated.

### 1.4.4 Community clouds

A community cloud shares the cloud infrastructure across several organizations in support of a specific community that has common concerns (for example, mission, security requirements, policy, and compliance considerations). The primary goal of a community cloud is to have participating organizations realize the benefits of a public cloud, such as shared infrastructure costs and a pay-as-you-go billing structure, with the added level of privacy, security, and policy compliance that is usually associated with a private cloud.

The community cloud infrastructure can be provided on-premises or at a third party's data center, and can be managed by the participating organizations or a third party.

The following guidance is provided from NIST.GOV:

"Carefully plan the security and privacy aspects of cloud computing solutions before engaging them. Public cloud computing represents a significant paradigm shift from the conventional norms of an organizational data center to a de-perimeterized infrastructure open to use by potential adversaries. As with any emerging information technology area, cloud computing should be approached carefully with due consideration to the sensitivity of data. Planning helps to ensure that the computing environment is as secure as possible and in compliance with all relevant organizational policies and that privacy is maintained. It also helps to ensure that the agency derives full benefit from information technology spending.

The security objectives of an organization are a key factor for decisions about outsourcing information technology services and, in particular, for decisions about transitioning organizational data, applications, and other resources to a public cloud computing environment. Organizations should take a risk-based approach in analyzing available security and privacy options and deciding about placing organizational functions into a cloud environment. The information technology governance practices of the organizations that pertain to the policies, procedures, and standards used for application development and service provisioning, as well as the design, implementation, testing, use, and monitoring of deployed or engaged services, should be extended to cloud computing environments.

To maximize effectiveness and minimize costs, security and privacy must be considered throughout the system lifecycle from the initial planning stage forward. Attempting to address security and privacy issues after implementation and deployment is not only much more difficult and expensive, but also exposes the organization to unnecessary risk.

Understand the public cloud computing environment offered by the cloud provider. The responsibilities of both the organization and the cloud provider vary depending on the service model. Organizations consuming cloud services must understand the delineation of responsibilities over the computing environment and the implications for security and privacy. Assurances furnished by the cloud provider to support security or privacy claims, or by a certification and compliance review entity paid by the cloud provider, should be verified whenever possible through independent assessment by the organization.

Understanding the policies, procedures, and technical controls used by a cloud provider is a prerequisite to assessing the security and privacy risks involved. It is also important to comprehend the technologies used to provision services and the implications for security and privacy of the system. Details about the system architecture of a cloud can be analyzed and used to formulate a complete picture of the protection afforded by the security and privacy controls, which improves the ability of the organization to assess and manage risk accurately, including mitigating risk by employing appropriate techniques and procedures for the continuous monitoring of the security state of the system.

Ensure that a cloud computing solution satisfies organizational security and privacy requirements. Public cloud providers' default offerings generally do not reflect a specific organization's security and privacy needs. From a risk perspective, determining the suitability of cloud services requires an understanding of the context in which the organization operates and the consequences from the plausible threats it faces. Adjustments to the cloud computing environment may be warranted to meet an organization's requirements. Organizations should require that any selected public cloud computing solution is configured, deployed, and managed to meet their security, privacy, and other requirements.

Non-negotiable service agreements in which the terms of service are prescribed completely by the cloud provider are generally the norm in public cloud computing. Negotiated service agreements are also possible. Similar to traditional information technology outsourcing contracts used by agencies, negotiated agreements can address an organization's concerns about security and privacy details, such as the vetting of employees, data ownership and exit rights, breach notification, isolation of tenant applications, data encryption and segregation, tracking and reporting service effectiveness, compliance with laws and regulations, and the use of validated products meeting federal or national standards (e.g., Federal Information Processing Standard 140). A negotiated agreement can also document the assurances the cloud provider must furnish to corroborate that organizational requirements are being met.

Critical data and applications may require an agency to undertake a negotiated service agreement in order to use a public cloud. Points of negotiation can negatively affect the economies of scale that a non-negotiable service agreement brings to public cloud computing, however, making a negotiated agreement less cost effective. As an alternative, the organization may be able to employ compensating controls to work around identified shortcomings in the public cloud service. Other alternatives include cloud computing environments with a more suitable deployment model, such as an internal private cloud, which can potentially offer an organization greater oversight and authority over security and privacy, and better limit the types of tenants that share platform resources, reducing exposure in the event of a failure or configuration error in a control."

For more information, see the following website:

http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-144.pdf

# 1.5 The IBM Cloud Computing Reference Architecture

This section introduces the IBM Cloud Computing Reference Architecture (CCRA), and describes the cloud service roles that are defined within it.

## 1.5.1 Introduction to the CCRA

A *reference architecture* is a proven template solution for architecture within a domain, in this case the cloud computing domain. A reference architecture is important to have because it offers these benefits:

- Delivers best practices in a standardized, methodical way
- Ensures consistency and quality across the development and delivery processes
- Mitigates risk by taking an asset-based approach to solution development

CCRA is an IBM-defined reference architecture for the cloud computing domain. It is an evolving architecture that is based on real-world input from many cloud implementations around the globe, and was submitted to the Open Group Cloud Architecture Project.

The IBM CCRA is designed around a set of architectural principles that establish the framework within which architectural decisions are made. CCRA has these architectural principles:

- Design for cloud-scale efficiencies
- Support lean service management
- Identify and use commonalities
- Define and manage cloud services generically during their lifecycle

As shown in Figure 1-4, the IBM CCRA defines basic elements of any cloud service environment. You can use it to identify the physical components of a cloud implementation such as network, storage, virtualization, and also the software components that are required to run and manage the cloud environment. In addition, it defines governance policies that are tailored for the environment or organization.

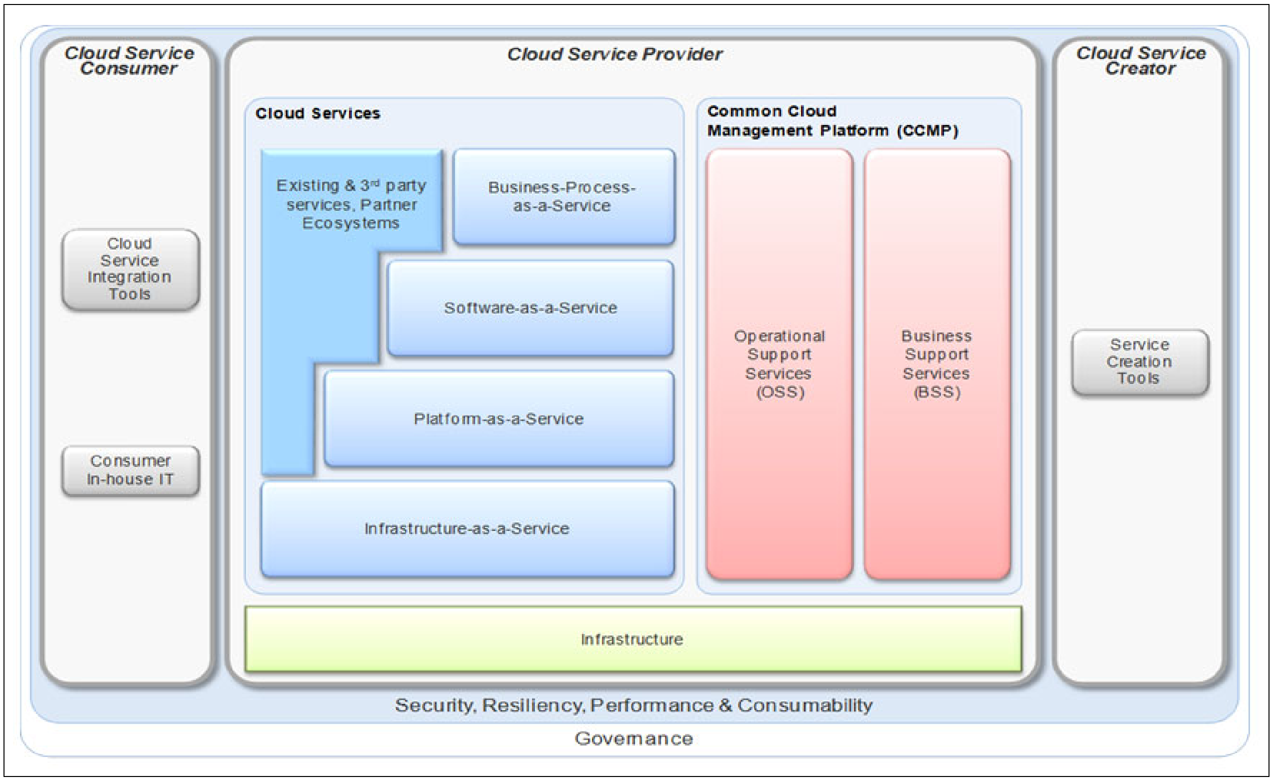



*Figure 1-4 IBM Cloud Computing Reference Architecture*

For more information about IBM CCRA, see the following website:

https://ibm.biz/BdEWLz

The roles that are defined by the CCRA are described, at a high level, in 1.5.2, “Cloud service roles” on page 13.

CCRA categorizes the cloud business models and corresponding architecture by the following cloud adoption patterns:

- Cloud enabled data center (IaaS)
- Platform as a service (PaaS) adoption pattern
- Software as a service (SaaS)
- Cloud service providers
- Mobile
- Analytics
- Government - Cloud

For each cloud adoption pattern, CCRA identifies these patterns:

- Common architecture patterns that describe the business drivers, the use cases, and the technologies that underlie each type of cloud computing implementation.
- Common architecture patterns for items that cut across all the adoption patterns, which include security, resiliency, and performance.

Figure 1-5 shows the cloud adoption patterns.

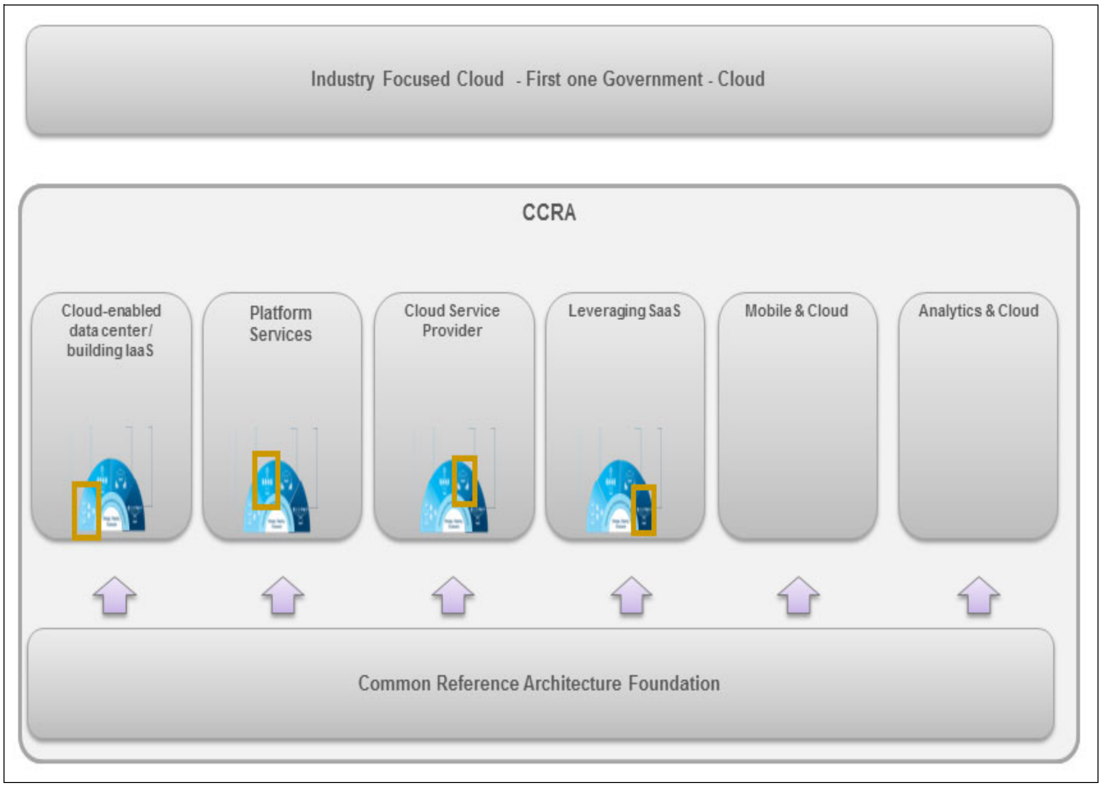



*Figure 1-5 Cloud adoption patterns*

## 1.5.2 Cloud service roles

As shown in Figure 1-4 on page 12, the IBM CCRA defines the following interrelated roles:

- Cloud Service Creator
- Cloud Service Provider
- Cloud Service Consumer

These roles are interrelated in that a *Cloud Service Creator* is responsible for creating a cloud service, which can be run by a *Cloud Service Provider*, and exposed to *Cloud Service Consumers*. Multiple roles can be fulfilled by the same organization or person.

### Cloud Service Creator

The Cloud Service Creator is responsible for creating a cloud service. The creator can be an individual or an organization that designs, implements, and maintains runtime and management artifacts that are specific to a cloud service. Typically, Cloud Service Creators build their cloud services by using functions that are exposed by a Cloud Service Provider.

Also typical is that the operations staff, who are responsible for operating a cloud service, are closely integrated with the development organization that develops the service. This integration is commonly referred to as *DevOps*. This close integration helps to achieve the delivery efficiency that is expected from cloud services because it allows a short feedback loop to implement changes in the cloud service.

### Cloud Service Provider

The Cloud Service Provider has the responsibility of providing cloud services to Cloud Service Consumers. The provider sets up the cloud service, and manages the effective running of the service, which can include the following tasks:

- Determine performance service levels and management strategies
- Monitor performance of virtualization infrastructure and service level agreements (SLAs)
- Manage long-term capacity and performance trends
- Analyze how to prevent costly service quality problems
- Ensure alignment of business and operational support systems
- Track performance against the provider business plan

A Cloud Service Provider might be a link within a chain of service providers and service consumers, with each provider adding some value to the service within the chain. In this case, each service provider needs to establish a partnership with their service providers to be able to ensure service levels to their clients. This chain is illustrated in Figure 1-4 on page 12 by the shaded segment named “Existing and Third Party Services, Partner Ecosystems.”

### Cloud Service Consumer

A Cloud Service Consumer is the user of a cloud service. The consumer might be an organization, a human being, or an IT system that requests, uses, and manages instances of a cloud service. Managing a service can include performing activities such as changing quotas for users, changing CPU capacity assigned to a virtual machine (VM), or increasing the maximum number of seats for a web conference. The service consumer can be billed for all (or a subset) of its interactions with the cloud service and the provisioned service instances.

Within the Cloud Service Consumer role, more specific roles can exist. The consumer organization might require a technical role responsible for making service consumption work from a technical perspective. There might also be a business person on the consumer side who is responsible for the financial aspects of consuming the service. In simple public cloud scenarios, all of these consumer roles can be collapsed into a single person.

The Cloud Service Consumer browses the service offering catalog and triggers service instantiation and management from there. Interaction with the service delivery catalog can be tightly embedded within the actual cloud service. In particular, these cases are common for SaaS and BPaaS cloud services where application-level virtualization is implemented.

## 1.6 Hybrid Cloud use cases

IBM is focusing its strategic investment and resources on helping clients realize potential with hybrid cloud. IBM intends to provide greater visibility, control, and security when it comes to the combination of traditional IT, private and public cloud resources.

Figure 1-6 shows seven use cases for Hybrid Cloud.

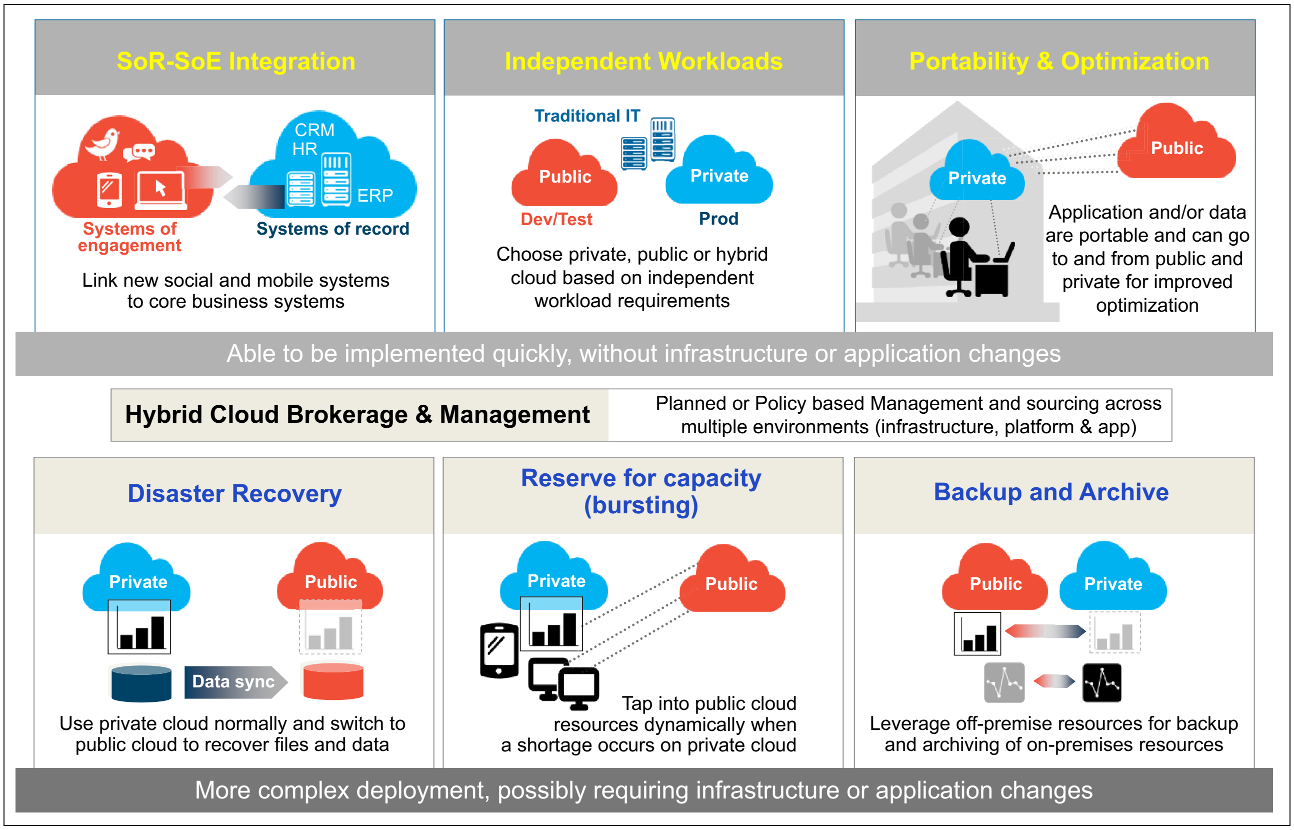



*Figure 1-6 Hybrid Cloud use cases*

## 1.6.1 Systems integration

Customer-interaction-oriented models are often referred to as Systems of Engagement, whereas traditional backend structures are called Systems of Record. Together they form IBM Systems of Insight™ as shown in Figure 1-7.

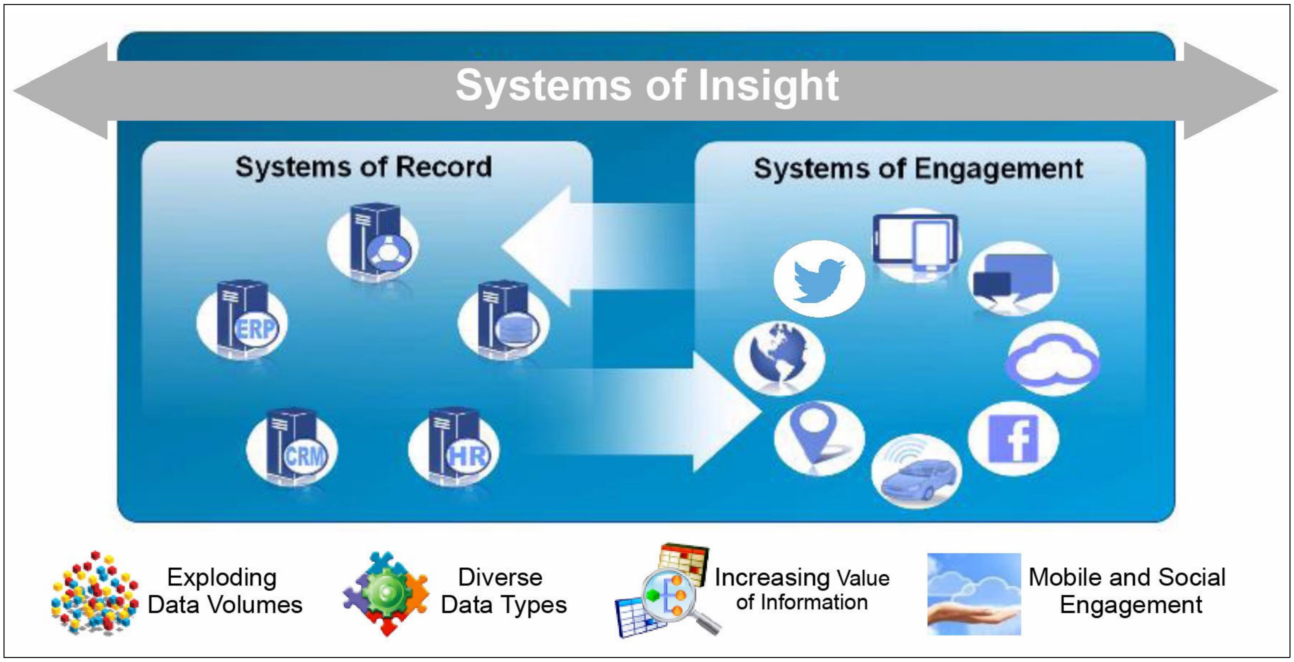



*Figure 1-7 Systems of Insight*

Supporting Systems of Insight workloads is driving IT to reshape itself to accommodate these new business needs while integrating them with traditional applications and infrastructure. Meeting these increasingly challenging business requirements is the primary goal of software-defined architecture.

For more information about the impact of Systems of Engagement, see:

https://www.ibm.com/software/ebusiness/jstart/sna

## 1.6.2 Independent workloads

For Independent workloads, select a private or hybrid cloud based on independent workload requirements. For example, a company might run production in a local or dedicated private cloud, and run development and testing on the public cloud to ensure that production is not affected by sharing resources with other workloads.

## 1.6.3 Portability and optimization

Application and data are portable in a cloud, and can be moved between private and public cloud to optimize workloads.

### 1.6.4 Hybrid cloud brokerage and management

Hybrid cloud brokerage and management can help you manage mixed cloud environments. It is planned or policy-based management and sourcing across multiple environments. Cloud brokerage is the answer for all cloud implementations. It provides visibility and control on the usage of cloud. It also reduces so-called shadow IT (business users using existing cloud services from the market on their own, bypassing central IT). The cloud brokerage function provides a one-stop-shop for the users using self-service portals to order their environment.

With a brokerage solution, you can plan, purchase, and manage IT services across all cloud models from multiple supplies. It helps you choosing the right cloud.

### 1.6.5 Disaster recovery

Cloud can help you set up and make available a parallel environment off-premises without building a second data center yourself.

Cloud-based disaster recovery as a service (DRaaS) has emerged rapidly as both small and large organizations look for a cost-effective way to ensure that data is protected and business activities can continue during a data center disruption. The evolution of today's leading DRaaS offerings centers around traditional managed storage and collocation service models. Some organizations have evolved solutions from either backup and recovery (B/R) software or cloud-related compute and storage services offerings.

### 1.6.6 Capacity bursting

Third-party managed private or public clouds offer an opportunity to use additional storage resources for large jobs, high-performance computing, and big data analytics batch jobs. Resources to handle extra requirements that are caused by seasonal workloads can be obtained from the public cloud while the off-seasonal load is optimized on a private cloud.

### 1.6.7 Backup and archive

You can use off-premise resources to back up or archive your on-premises data by connecting over the network. This configuration offers an alternative to writing to physical tapes that are then taken off-site by truck or other means. IBM offers backup and archive software that can be run in various configurations to support this use case.

## 1.7 Cloud enabled data center journey

Given the information presented so far in this chapter, what does the journey and industry best practices toward a true cloud-enabled infrastructure and a storage cloud look like? And, where is the *storage cloud* positioned within this overall cloud infrastructure? If you look at the Cloud Adoption Patterns in Figure 1-8 on page 18, the Storage Cloud fits under IaaS - Cloud Enabled Data Center (CEDC). The CEDC pattern can be implemented in three different models:

- On-premise-hybrid
- Off-premise-hybrid
- Full off-premise

IBM has published two cloud papers that answer these important questions in detail:

- *IBM SmartCloud®: Building a Cloud Enabled Data Center*, REDP-4893
- *IBM SmartCloud: Becoming a Cloud Service Provider,* REDP-4912

The cloud data center journey is illustrated in Figure 1-8.

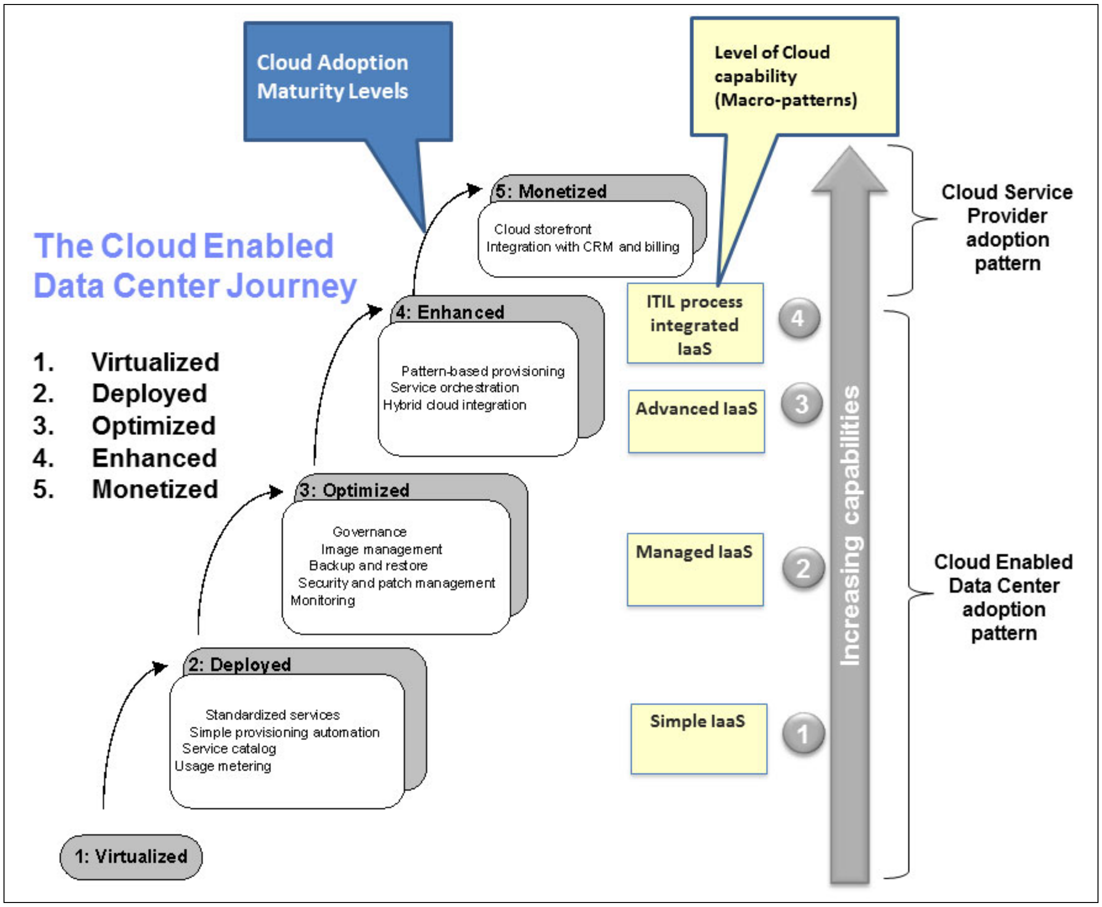



*Figure 1-8 The cloud-enabled data center journey*

Figure 1-8 shows the preferred practices steps that help ensure a successful cloud deployment.

The specific *relationship patterns and workflows* that the successful cloud must implement are not depicted in this progression. The next section summarizes those workflows so that you can exactly see what best practices and organizational structure are required to provide true cloud services, and where the storage cloud is positioned within the overall cloud.

## 1.7.1 Cloud Macro patterns

The *IBM Cloud Computing Reference Architecture 4.0* shows the overall cloud preferred practices macro-patterns and use cases of a CEDC adoption pattern. See Figure 1-9.

1. Simple IaaS: Start to embrace the cloud model by delivering simple VMs by using a service catalog
2. Managed IaaS: Manage the cloud infrastructure to increase SLA/QoS, security, and reliability
3. Advanced IaaS: Deliver more sophisticated and composite cloud services
4. ITIL process integrated IaaS: Integrate IaaS with existing enterprise ITIL processes

At the bottom is the Software Defined Infrastructure layer that underpins the adoption patterns. Software Defined Infrastructure is discussed in 1.8, "Software-defined environment" on page 20.

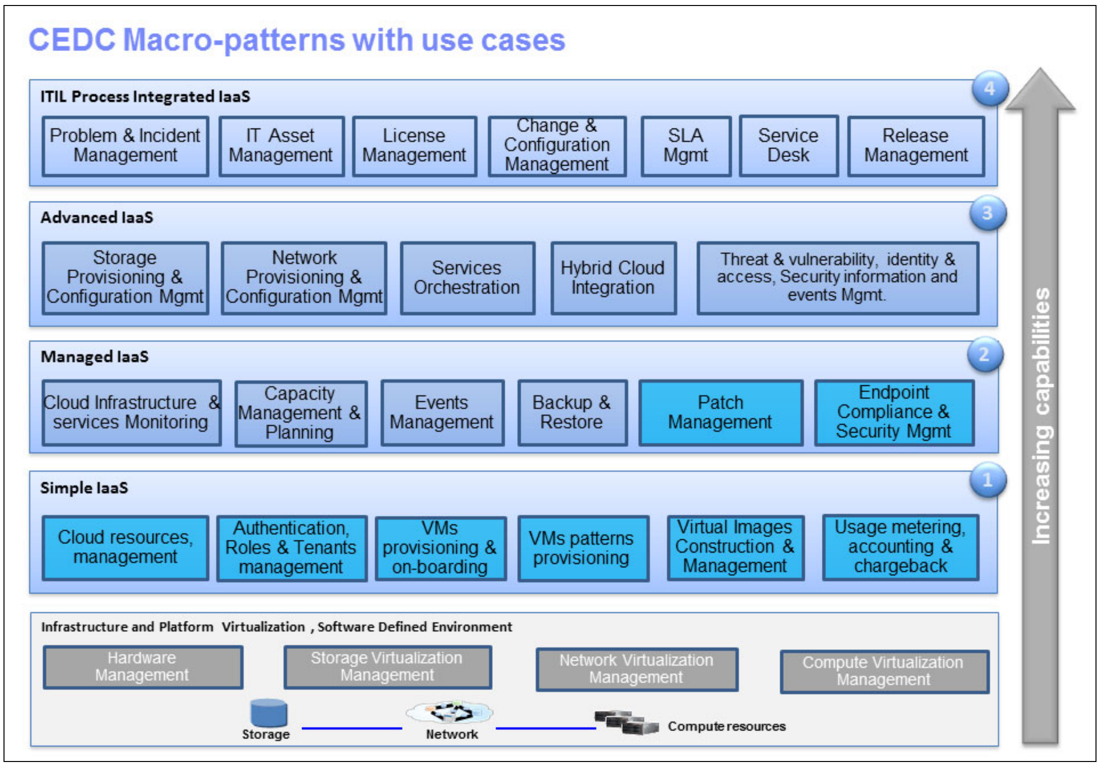



Figure 1-9 Cloud preferred practices Macro-patterns

The following sections detail the use cases that make up each of these macro-patterns.

## 1.7.2 Cloud IT organizational structure

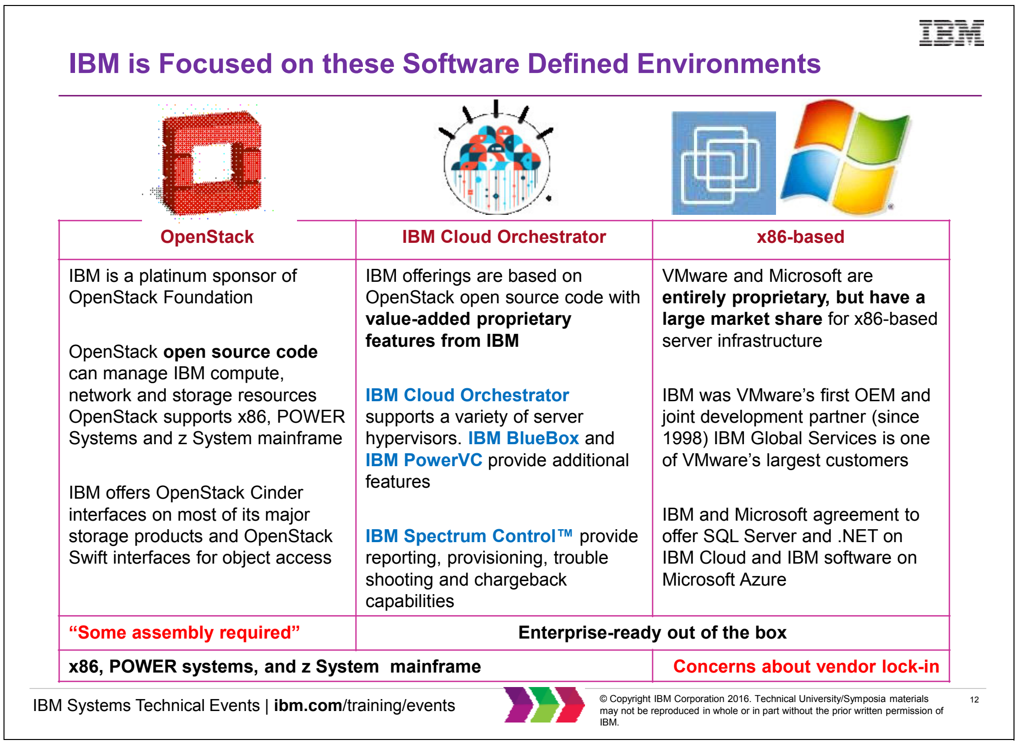
To deliver true cloud services, a traditional IT organization is unfortunately almost certainly *not* organized like the workflows shown in Figure 1-9. As a result, a traditional IT organization often finds it difficult to deliver truly elastic modern cloud IT capabilities.

Therefore, to be able to truly deliver elastic, scalable, automated cloud IT services, an IT reorganization from the existing traditional structure is *required*. Most IT organizations will need to realign based on the workflows that make up the preferred practices cloud infrastructure.

The cloud workflows also give you an organization template for how your IT organization should and *must* be reorganized over time to deliver true cloud services.

## 1.8 Software-defined environment

A software-defined environment (SDE) combines workload automation and orchestration with a software-defined infrastructure. IBM has decided to focus on three software defined environments, as shown in Figure 1-10.


IBM is Focused on these Software Defined Environments

| OpenStack | IBM Cloud Orchestrator | x86-based |
|---|---|---|
| IBM is a platinum sponsor of OpenStack Foundation<br><br>OpenStack **open source code** can manage IBM compute, network and storage resources OpenStack supports x86, POWER Systems and z System mainframe<br><br>IBM offers OpenStack Cinder interfaces on most of its major storage products and OpenStack Swift interfaces for object access | IBM offerings are based on OpenStack open source code with **value-added proprietary features from IBM**<br><br>**IBM Cloud Orchestrator** supports a variety of server hypervisors. **IBM BlueBox** and **IBM PowerVC** provide additional features<br><br>**IBM Spectrum Control™** provide reporting, provisioning, trouble shooting and chargeback capabilities | VMware and Microsoft are **entirely proprietary, but have a large market share** for x86-based server infrastructure<br><br>IBM was VMware's first OEM and joint development partner (since 1998) IBM Global Services is one of VMware's largest customers<br><br>IBM and Microsoft agreement to offer SQL Server and .NET on IBM Cloud and IBM software on Microsoft Azure |
| "Some assembly required" | **Enterprise-ready out of the box** | |
| **x86, POWER systems, and z System mainframe** | | Concerns about vendor lock-in |

IBM Systems Technical Events | **ibm.com**/training/events

© Copyright IBM Corporation 2016. Technical University/Symposia materials may not be reproduced in whole or in part without the prior written permission of IBM. 12

*Figure 1-10 Software-defined environments*

## 1.9 Software-defined infrastructure

This section introduces software-defined infrastructure (SDI), which includes new terminology and concepts that might differ from traditional approaches.

Fundamentally, SDI is an IT advancement that is aimed at enabling automation of infrastructure configuration to support rapid deployment. It is aligned to real-time application requirements that are long standing goals of IT systems optimization. SDI is evolving technology that has been made feasible by the abstraction of infrastructure component interfaces delivered through the virtualization of server, storage, and network infrastructures.

A key driver of SDI development and deployment is cloud configuration automation requirements. And although SDI is finding widespread application in cloud implementations, it can provide substantial agility and utilization improvements across IT environments, especially those with rapidly changing application infrastructure support requirements.

### 1.9.1 New and Traditional Workloads

SDI targets new business models that use tighter interactions with customers such as big data, analytics, social business, and mobile. It can also be used with traditional IT business workloads like enterprise resource planning (ERP), human resources (HR), and customer relationship management (CRM) systems that continue to be important within an integrated infrastructure.

### 1.9.2 SDI Components

SDI is an excellent framework for creating and implementing optimized IT infrastructures that can help enterprises attain competitive advantage by delivering higher value and profitability through speed and efficiency in provisioning IT services. Most enterprise IT architectures already use virtualization to manage growth and improve agility. Virtualization and the abstraction of IT component interfaces that it provides allow integrated software definition across the entire infrastructure.

Standardized software interfaces support the automation of infrastructure administrative tasks like configuration, monitoring, and provisioning in real time in response to changing application and business requirements. It is increasingly critical to quickly and efficiently deliver resources in support of not only traditional business workloads, but also enable cloud, big data and analytics, mobile, and social-business services. The SDI approach helps an enterprise fulfill its business requirements and respond to business requests faster and more effectively.

Software-defined storage is a key component that supports the SDI framework along with software-defined compute and software-defined network constructs as shown in Figure 1-11. Although each of these constructs can be used separately, substantial synergy and value results from an integrated approach.

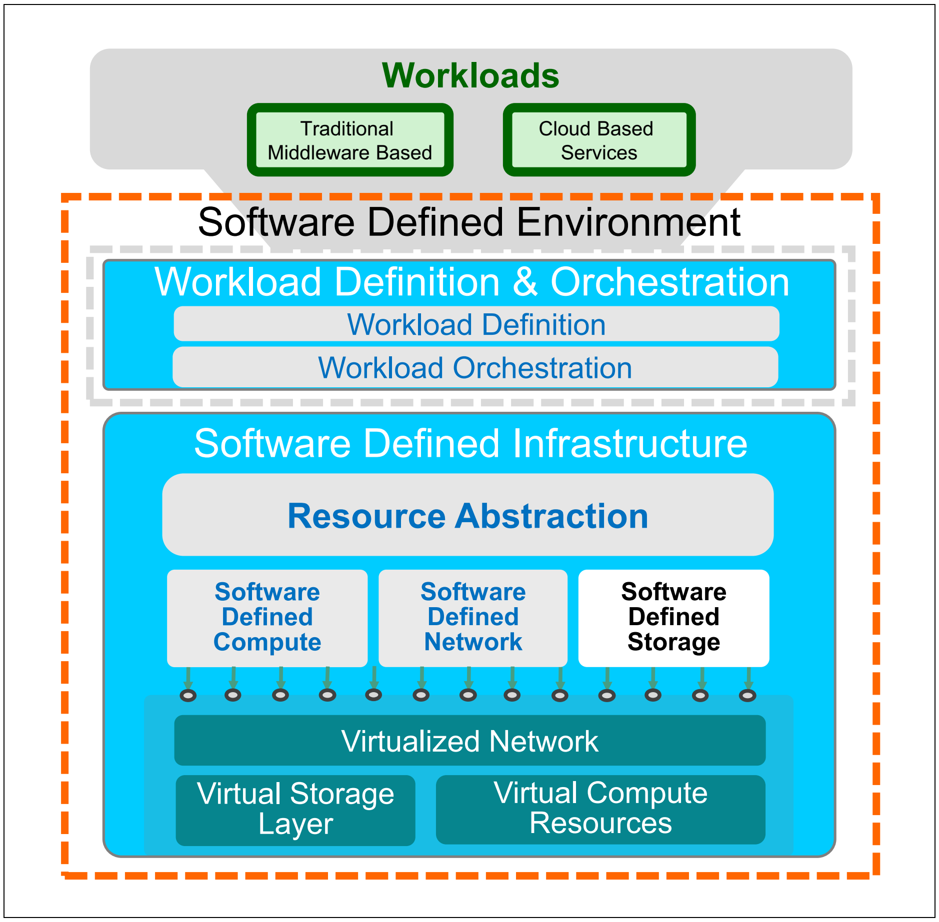



*Figure 1-11 Software-defined infrastructure framework*

SDI drives efficiency by optimizing the connections between workloads and resources based on the needs of the business. Workloads are defined and orchestrated based on usage patterns, and resources are managed and deployed according to business rules and policies. An SDI offers several core advantages that provide enterprises that employ an SDI approach with improvements within several processes that IT operations, until now, have traditionally handled manually.

SDI automatically coordinates infrastructure resources (compute, storage, networks, and management) to programmatically (that is, by using software programming) meet workload requirements in real time. From this point of view, the following are some of the major attributes of SDI:

- Agility: IT resource customers expect to use infrastructure resources on demand based on immediate business requirements. The agility of IT resource allocation and consumption needs to be made near instantaneous to support emerging workloads.
- Standardization: Consumers are less interested in the specific infrastructure components and are more concerned with ensuring that the appropriate service level characteristics that are needed for their applications are in place. SDI brings uniformity by automating, standardizing, and integrating IT infrastructures.

- Provisioning and Orchestration: Rather than building unique infrastructure systems of server, network, and storage components for applications, IT providers need to configure pools of resources and put them together in a way that can be dynamically delivered programmatically (that is, by using software) with service-level-oriented interfaces appropriate to IT consumers.

An SDI requires hardware to provide resources to support the server, storage, and network infrastructure. The essential characteristic requirement for an SDI is that these hardware components be dynamically configurable to support real-time service level requirements.

It is important to consider that SDI by itself will not provide infrastructure that is aligned to business IT service level requirements unless the proper software definable components are in place. High performance, availability, and security service levels require software definable components that can be configured to meet these business requirements.

Similarly, lower level (for example, best effort) service levels should generally be configured with software definable components cost aligned to these business requirements. SDI architectures that need to support varied service levels will still require appropriate performance and capacity planning across higher performance components and differentiation of availability requirements for cost optimization. SDI supports the optimization of infrastructure service levels to available component resources in real time. Implementing an SDI framework supports the transformation from static infrastructure into dynamic, continually optimized, workload-aware, and virtualized resources that allow line-of-business users to better use IT as needed. This system enables far greater business agility.

The deployment velocity requirements of Systems of Engagement (SoE) demand this new interaction between the consumer and the infrastructure provider to define workloads in a way that enables the infrastructure to dynamically respond to the needs of those workloads.

Analytics processing, for example, typically needs to rapidly access required data, efficiently process that data, and then release resources when the analysis completes. SDI is an ideal IT infrastructure implementation approach in this scenario. Similarly, SDI supports efficient deployment of rapidly growing and dynamically evolving transactional applications that support the increasing number of mobile devices that IT now manages. SDI value is even more apparent in hybrid scenarios like social analytics that are employed in sentiment analysis used to determine customer opinions and thinking, or develop macro-level understanding of worldwide events to create opportunity out of the data. Without SDI, the ability to react in a timely fashion to Cloud, Analytics, Mobile, Security and Social (CAMSS) workloads requirements is limited and inhibits IT's ability to expeditiously meet the dynamic infrastructure requirement for these workloads. As a result, these applications are often delayed and less effectively deployed, resulting in under-realized or missed business opportunities.

Although SDI provides a compelling framework to address the challenge of dynamic infrastructure configuration and provisioning, important architectural design considerations still must be addressed despite the claims of some over-zealous visionaries. These concerns range from network data latency to infrastructure component interoperability within a specific software defined implementation. The following are some examples:

- Data locality and latency along with network bandwidth and data scale considerations can be more readily addressed within SDI, but certainly must be properly planned for during the design process.
- Infrastructure diversity in terms of public, hybrid, and private cloud along with legacy infrastructure support are benefits of SDI implementations, but still require careful planning around security, availability, and relative cost parameters.

- Component selection, fit for purpose, and web-scale (custom commoditization) require attention to relative cost/performance (as previously mentioned) especially during SDI transition.
- Application and component interoperability and open standards conformity typically drive reductions of component and administrative support cost, but there are design scenarios where custom/vendor proprietary solution continue making business sense through SDI transition and likely beyond.
- Transition and legacy support must be handled in an evolutionary way to minimize effects on business and customers.

The business value of SDI is too great to ignore and is maximized when these design parameters are given proper consideration while planning and deploying software defined infrastructures. IBM is investing heavily in developing offerings across the spectrum of the software defined universe, from building block components to integrated cloud offerings, as well as implementing and supporting open API standards and architectures. These objects help businesses achieve improved agility and competitiveness, and produce outstanding customer satisfaction. Figure 1-12 shows the building blocks required to support new IT infrastructure, highlighting software-defined storage (SDS).

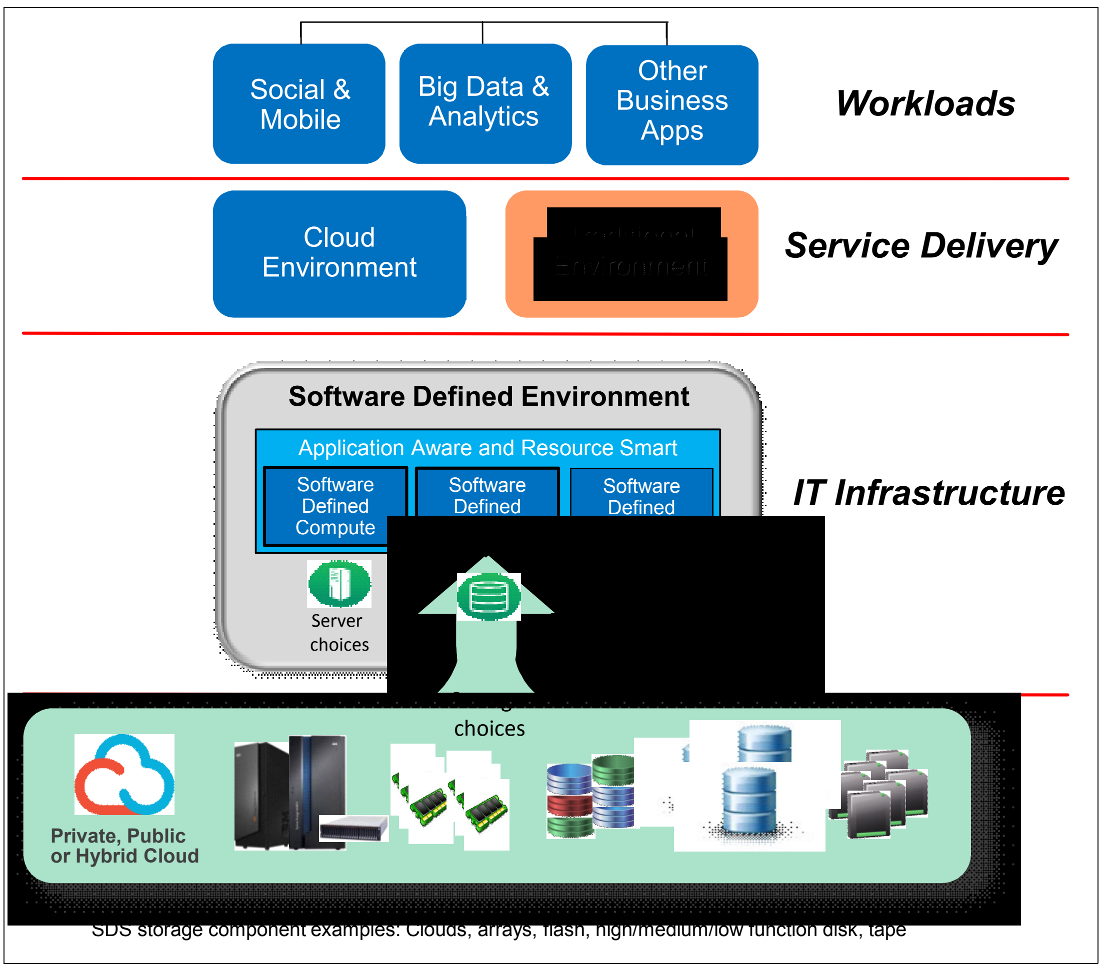



*Figure 1-12 SDS building block of SDI for support of new IT business requirements*

For more information about SDS, see *IBM Software-Defined Storage Guide*, REDP-5121.

## 1.9.3 Role of OpenStack cloud software in cloud computing

In just the past two years, the open source OpenStack cloud software has become one of the most popular and fastest growing of the generally available cloud IaaS offerings. See Figure 1-13.

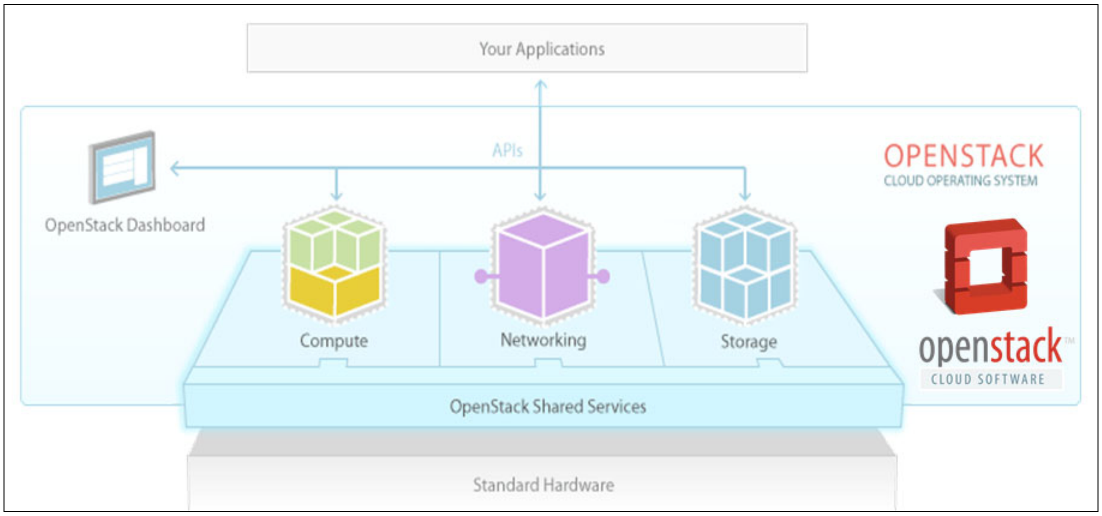



*Figure 1-13 OpenStack architecture high-level overview*

The OpenStack cloud software was initially released into open source in 2010 by Rackspace Cloud and the US National Aeronautics and Space Administration (NASA). The OpenStack architecture goal is to provide an open source cloud operating system IaaS platform for creating and managing large groups of virtual private servers in a cloud computing environment. OpenStack cloud software is an open source IaaS cloud operating system that is released under the terms of the Apache 2.0 license. The design goals of OpenStack cloud software are scale and elasticity, share nothing, and distribute everything. OpenStack cloud software and offerings like it provide a means for traditional IT to quickly adopt newer cloud computing workflows and best practices.

By adopting and using offerings such as OpenStack cloud software, the IT organization can organize, develop skill sets, and deploy cloud computing around proven offerings that already implement industry cloud computing best practices.

The OpenStack cloud software has experienced rapid adoption in just the past four years, and has become largest, fastest growing open source IaaS offering. OpenStack community currently has more than 507 companies who have joined and actively contribute to the OpenStack software and the OpenStack Foundation project, including IBM (Figure 1-14).


*Figure 1-14 OpenStack community status and growth as of April 2016*[2]

[2] Source: OpenStack Foundation

## 1.9.4 OpenStack architecture overview: An open source cloud IaaS platform

OpenStack architecture implements the complete collection of cloud infrastructure as a service workflow, which is shown in Figure 1-9 on page 19.

OpenStack architecture is modular and includes the major components that are required for a cloud infrastructure. The OpenStack architecture diagram shows the various OpenStack components, already implemented in a cloud best practices workflow (Figure 1-15).

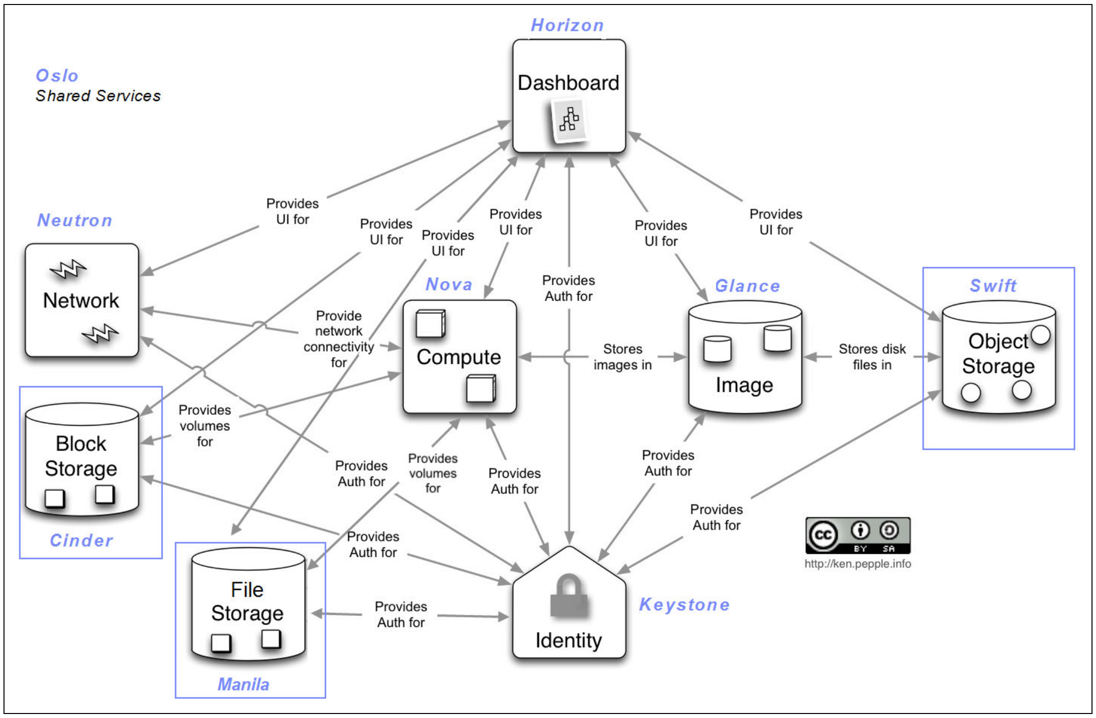



*Figure 1-15 OpenStack architecture with storage components*

OpenStack includes these modular components:

- Compute (Nova)
- Block Storage (Cinder)
- Object Storage (Swift)
- File Storage (Manila)
- Virtual Machine Boot Image Service (Glance)
- OpenStack Identity Management (Keystone)
- User Interface Dashboard (Horizon)
- Telemetry (Ceilometer)
- Orchestration (Heat)
- Bare Metal Provisioning (Ironic)

OpenStack IaaS solution provides a full cloud best practices solution for implementing cloud IT services. In addition, the most important aspect of OpenStack cloud software is its large and rapidly growing open source community of contributors worldwide. In a similar manner to Linux, which has become a reliable, mission-critical capable technology for operating systems, OpenStack technology seems like it will play a similar role for cloud operating systems.

### 1.9.5 IBM participation in OpenStack Foundation

IBM believes that an open source approach to cloud is the most beneficial strategic means for clients to enter and take advantage of the benefits of cloud computing. As such, IBM is investing in supporting the OpenStack Foundation as a Platinum member:

- One of the top contributors
- Over 210 IBM OpenStack contributors
- Over 500 dedicated IBM professionals

For more information about IBM participation, see the OpenStack website at:

http://www.openstack.org/foundation/companies/profile/ibm

IBM views support of OpenStack cloud software as a strategic and key component of IBM participation in providing cloud computing capability.

Also, IBM has Cloud OpenStack Services, which can help reduce your need to invest in up-front capital resources for an in-house private cloud infrastructure. The IBM hosted private cloud runs on dedicated, high-performing SoftLayer bare metal servers that are housed in global data centers designed to meet stringent industry and regulatory compliance requirements.

Features such as physical infrastructure isolation for compute and storage, network gateways, and a virtual private network connection with an encrypted tunnel can help you feel more confident that your data is being protected with the same rigor as an on-premises solution.

For more information about IBM OpenStack Cloud Services, see:

- IBM Cloud OpenStack Services

  http://www.ibm.com/common/ssi/ShowDoc.wss?docURL=/common/ssi/rep_ca/8/877/ENUSZS14-0048/index.html&lang=en&request_locale=en

- IBM Cloud

  https://www.ibm.com/cloud-computing

## 1.10 Storage cloud components within overall cloud

Having now shown the overall cloud picture, the remainder of this paper describes the storage-specific portions of the cloud journey. It focuses on the role that storage plays in the cloud workflow, and storage cloud best practices.

The paper reviews what a storage cloud is; what the storage features that enable a storage cloud are; key technology aspects such as storage efficiency, automation, and management; and security and data protection. It provides an overview of storage key enablers of a cloud IaaS, including a description of OpenStack storage components. It also highlights specific IBM products that participate in the storage cloud workflow.

2

# What is a storage cloud

Cloud data storage is a critical component in the cloud computing model. Without cloud storage, there can be no cloud service. As stated in Chapter 1, “What is cloud computing” on page 1, a specific definition of what constitutes a storage cloud is not always clear in this emerging paradigm. The growing interest in cloud storage together with cloud computing is explained in terms of the challenges that traditional IT presents. This chapter explores how these challenges can be addressed in the various storage cloud models that are aligned to cloud computing constructs (that is, public, private, and hybrid clouds).

This chapter includes the following sections:

- Storage cloud overview
- Traditional storage versus storage cloud
- Benefits and features of storage cloud
- Storage classes for cloud
- Storage cloud delivery models
- The storage cloud journey

## 2.1 Storage cloud overview

A storage cloud provides storage as a service (SaaS) to storage consumers. It can be delivered in any of the previously described cloud delivery models (public, private, hybrid, and community). A storage cloud can be used to support a diverse range of storage needs, including mass data stores, file shares, backup, archive, and more. Implementations range from public user data stores to large private storage area networks (SAN) or network-attached storage (NAS), hosted in-house or at third-party managed facilities. The following examples are publicly available storage clouds:

- IBM Cloud offers various storage options, including archive, backup, and object storage.
- Skydrive from Microsoft allows the public to store and share nominated files on the Microsoft public storage cloud service.
- Email services, such as Hotmail, Gmail, and Yahoo, store user email and attachments in their respective storage clouds.
- Facebook and YouTube allow users to store and share photos and videos.

Storage cloud capability can also be offered in the form of *storage as a service*, where you pay based on the amount of storage space used. A storage cloud can be used in various ways, based on your organization's specific requirements.

Figure 2-1 shows how various electronic or portable devices can access storage through the Internet without necessarily knowing the explicit details of the type or location of storage that is used underneath. Although the devices can access SAN or NAS storage, SAN or NAS storage can itself use storage cloud for backup or other purposes.

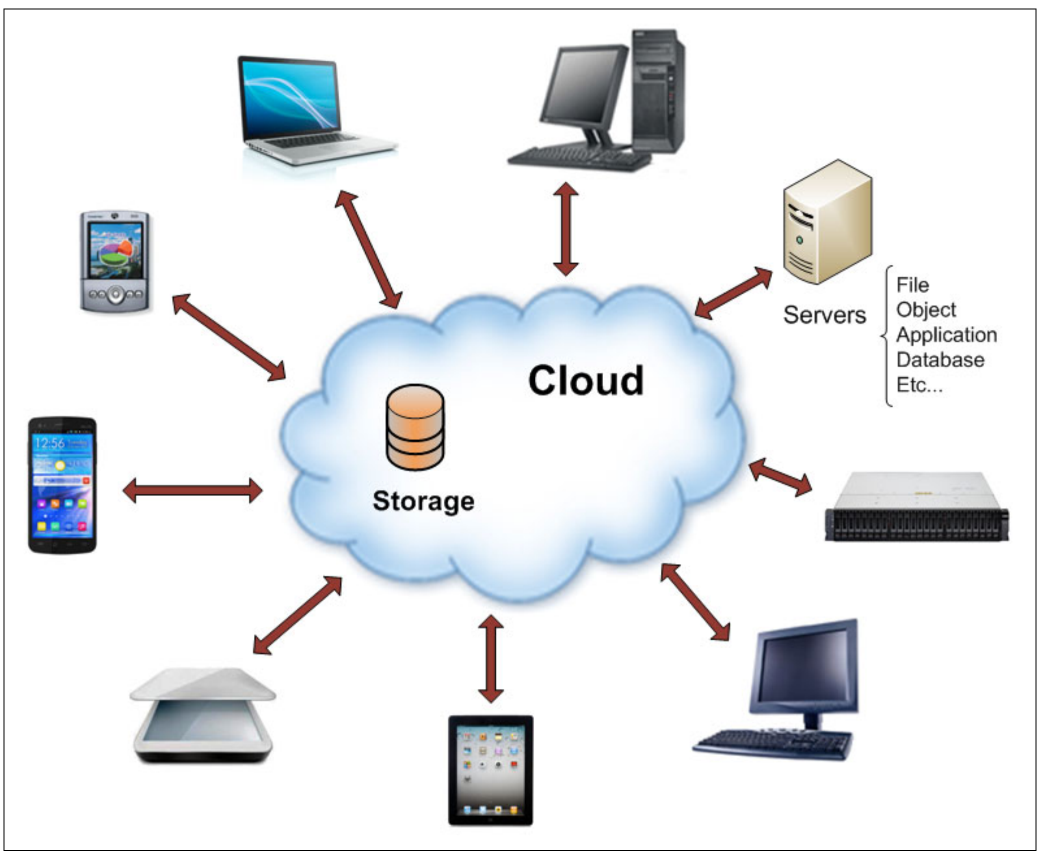



*Figure 2-1 Overview of storage cloud*

## 2.1.1 Storage usage differences within a storage cloud infrastructure

Within a cloud infrastructure, a useful distinction can be made between how storage capacity is used. This distinction is similar to the difference that exists in traditional IT between system data (files, libraries, utilities, and so on), and application data and user files. This distinction becomes important for storage allocation in virtual server implementations.

### Storage cloud

Storage cloud is the storage capacity service that is provided for client data and the primary focus of this paper. A storage cloud exhibits the characteristics that are essential to any cloud service (self-service provisioning, Internet and intranet accessibility, pooled resources, elastic, and metered). It is a cloud environment on which the offered services can store and retrieve data on behalf of computing processes that are not part of the storage cloud service. A storage cloud can be used in combination with a compute cloud, a private compute facility, or as storage for a computing device. Storage in a storage cloud can be categorized as follows:

- Hosted storage

  This category is primary storage for block or file data that can be written and read on demand, and is provisioned as generally higher performance and availability storage.

- Reference storage

  This category is fixed content storage to which blocks or files are typically written to once, and read from many times. Examples of data typically on reference storage include multimedia, archival data, medical imaging, surveillance data, and log files.

### Storage for cloud

*Storage for cloud* is a general name that is applied to the type of storage environment, implemented in cloud computing that is required to provision cloud computing services. For example, when a virtual server is created, some storage capacity is required. This storage is provisioned as part of the virtual machine creation process to support the operating system and runtime environment for the instance. It is not delivered by a storage cloud. However, it can be provisioned from the same storage infrastructure as a storage cloud. The types of storage provisioned for a cloud service can be categorized as follows:

- Ephemeral storage

  This storage is required only while a virtual machine is running. It is freed from use and made available to the storage pool when the virtual machine is shut down. Examples of this category of storage include boot volumes, page files, and other temporary data.

- Persistent storage

  This storage is required across virtual machine restarts. It is retained even when a virtual machine is shut down. It includes “gold” (master template) images, systems customization, and user data.

Figure 2-2 illustrates the categories of storage found in cloud computing.

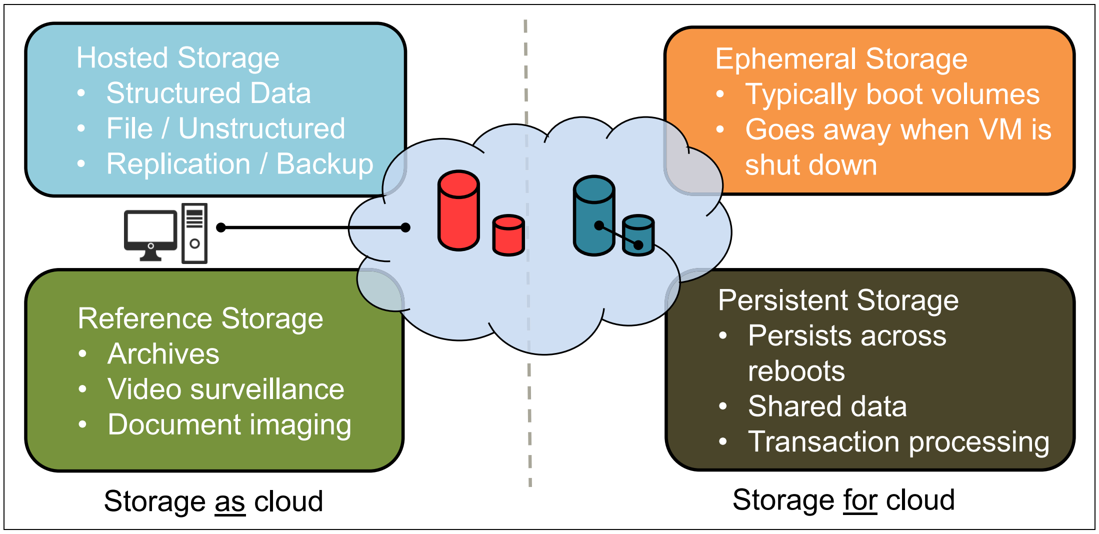



*Figure 2-2 Storage categories used in cloud*

## 2.2 Traditional storage versus storage cloud

This section compares the various challenges of traditional and cloud storage, outlines the advantages of cloud storage, and explains key implementation considerations for potential storage cloud infrastructure deployments.

### 2.2.1 Challenges of traditional storage

Before exploring the advantages and benefits of storage cloud, this section lists several limitations of current IT infrastructure that businesses deal with daily. This categorization is from a high level. Challenges in one category can sometimes be applicable to other categories.

#### Constrained business agility

The time that is required to provision storage capacity for new projects or unexpectedly rapid growth affects an organization’s ability to quickly react to changing business conditions. This situation can often negatively affect the ability to develop and deliver products and services within competitive time-to-market targets. The following constraints are examples:

- Time that is required to deploy new or upgraded business functions
- Downtime that is required for data migration and technology refresh
- Unplanned storage capacity acquisitions
- Staffing limitations

Substantial reserve capacity is often required to support growth, which requires planning and investment far in advance of the actual need to store data. The reason is because the infrastructure cannot easily scale up the needed extra capacity as a result of an inability to seamlessly add required storage resources. This key issue makes it more difficult to cope with rapidly changing business environments, adversely affecting the ability to make better decisions more rapidly and proactively optimize processes with more predictable outcomes.

The following additional issues can affect business agility:

- Inability to meet demand for data availability, and therefore not being able to access the correct data at the correct time to make better business decisions
- Inability to support unplanned acquisitions and staffing limitations

### Suboptimal utilization of IT resources

The variation in workloads and the difficulty in determining future requirements typically results in IT storage capacity inefficiencies:

- Difficulty in predicting future capacity and service-level needs
- Peaks and valleys in resource requirements
- Over and under provisioning of IT resources

Extensive capacity planning effort is needed to plan for future storage capacity and service level requirements. Capacity is often underutilized because the storage infrastructure requires reserve capacity for unpredictable future growth requirements, and therefore cannot be easily scaled up or down. Compounding these issues is the frequent inability to seamlessly provision more storage capacity without impacting application uptime.

### Organizational constraints

Another barrier to efficient use of resources can be traced to artificial resource acquisition, ownership, and operational practices:

- Project-oriented infrastructure funding
- Constrained operational budgets
- Difficulty implementing resource sharing
- No chargeback or showback mechanism as incentive for IT resource conservation

The limited ability to share data across the enterprise, especially in the context of interdepartmental sharing, can degrade overall use of IT resources including storage capacity. Parallel performance requirements in existing storage systems result in one node supporting one disk, leading to multiplication of nodes and servers.

### IT resource management

Efficient IT support is based on cost-effective infrastructure and service-level management to address business needs:

- Rapid capacity growth
- Cost control
- Service-level monitoring and support (performance, availability, capacity, security, retention, and so on)
- Architectural open standardization

The continued growth of resource management complexity in the storage infrastructure is often based on a lack of standardization and high levels of configuration customization. For example, adjusting storage performance through multiple RAID settings and manual tuning the distribution of I/O loads across various storage arrays consumes valuable staff resources.

Sometimes, the desire to avoid vendor lock-in because of proprietary protocols for data access also creates tremendous pressure on storage resource management. Other challenges are related to managing and meeting stringent service level agreement (SLA) requirements and lack of enough in-house expertise to manage complex storage infrastructures. New service levels, adjusting existing SLAs to align IT disaster recovery, business resilience requirements, and high-availability solutions are also factors.

Duplicate data that exists in the form of copies across organizational islands within the enterprise leads to higher costs for data storage and backup infrastructure. Compounding all of these limitations are tight operational and project budgets, and lack of dynamic chargeback or showback models as incentives for IT resource conservation.

### 2.2.2 Advantages of a storage cloud

Storage cloud has redefined the way storage consumers can do business, especially those who have seasonal or unpredictable capacity requirements, and those requiring rapid deployment or contraction of storage capacity. Storage cloud can help them focus more on their core business and worry less about supporting a storage infrastructure for their data. Storage cloud offers these advantages:

- Facilitates rapid capacity provisioning that supports business agility
- Improves storage utilization by avoiding unused capacity
- Supports storage consolidation and storage virtualization functions
- Chargeback and showback accounting of usage as incentive to conserve resources

Storage cloud helps companies to become more flexible and agile, and supports their growth. Improvement in quality of service (QoS), by automating provisioning and management of underlying complex storage infrastructure, helps improve the overall efficiency of IT storage. Cloud features such as data deduplication, compression, automatic tiering, and data migration capabilities are generally built-in options, and also support the optimizing of storage costs by implementing tiered storage.

Often the growth in file-based systems is restricted to approximately a few terabytes (TB). This restriction can be easily overcome with storage cloud. Ubiquitous access to data over the Internet, intranet, or both, provides location-independent access. This configuration can provide a single management platform to manage hundreds of nodes, with data flowing from all the nodes to all the storage arrays.

Capital expenditure can be reduced with a cloud operational-based, pay-as-you-go model. Storage clouds can be tailored and services acquired to support key storage operations such as backup and recovery, remote site disaster recovery, archive, and development and test operations.

Figure 2-3 shows layers that provide unique benefits in the storage cloud.

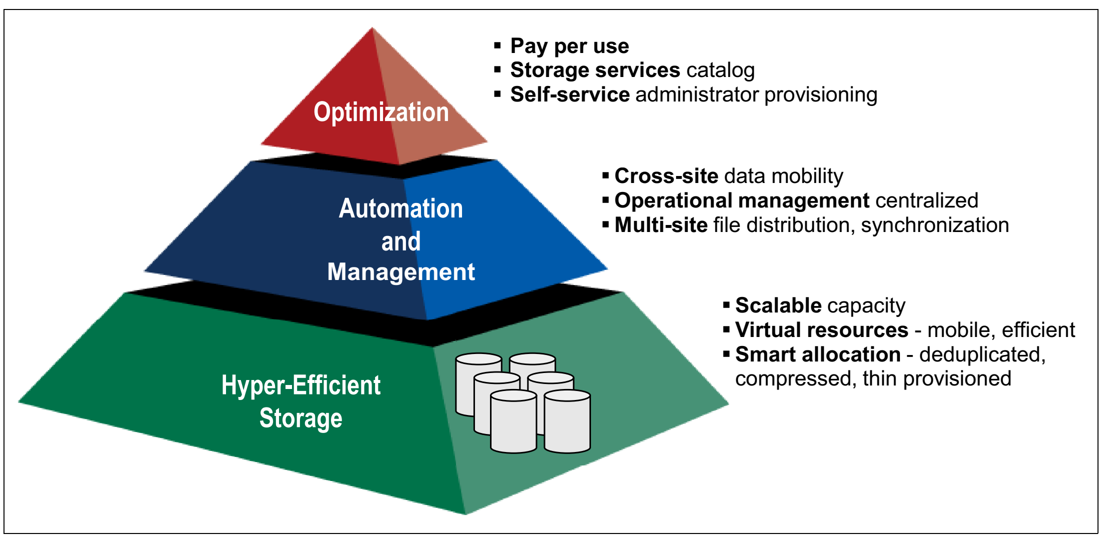



*Figure 2-3 Storage cloud characteristics*

### 2.2.3 Implementation considerations for storage cloud

Storage cloud is still an emerging paradigm. Although it offers many advantages, you need to be aware of these challenges:

- You need a reliable and robust network infrastructure for remote data access. Because the storage is accessed over the Internet or intranet, a good network connection is essential. The reliability of network providers such as Internet service providers (ISPs) is an important factor because in some parts of the globe, the Internet is not up to current standards.
- Security is an important factor. Beyond user name and password, consider encryption for sensitive data.
- You need to maintain security and control of data that is stored off-site, especially at third-party locations. Data can be encrypted when transmitted from an on-premises data center to an off-premises cloud service provider.
- Ensure that regulatory compliance is preserved for various standards such as the Health Insurance Portability and Accountability Act (HIPAA), Payment Card Industry Data Security Standard (PCI-DSS) and the Sarbanes-Oxley Act (Sarbox or SOX).
- Because standards are still evolving, avoiding vendor lock-in should be part of a selection process. Focus on cloud service providers who adopt open standards and participate in open source communities.
- Know the overall reliability of the cloud storage provider. Are SLAs required and will providers offer adequate assurance of service delivery? Will the provider remain viable in the future?
- Multitenancy (isolation) can be critical. Data needs to be protected from other clients who share cloud storage resources, security threats, viruses, and so on, because data is stored on a common shared storage infrastructure.
- Difficulty in applying policies across many independent file systems in an enterprise can cause operational problems.
- Determine whether the cloud storage provider can scale to your capacity requirements and maintain the required performance service levels.
- Be able to manage complexity of separate hardware from multiple vendors. Standardization can simplify management for heterogeneous storage devices. Storage virtualization across SAN arrays such as with IBM SAN Volume Controller, or Global Namespace solutions such as IBM Spectrum Scale™ can provide solutions to this issue.

## 2.3 Benefits and features of storage cloud

The overall benefits of storage cloud vary significantly based on the underlying storage infrastructure. Storage cloud can help businesses achieve more effective functionality at a lower cost while improving business agility and reducing project scheduling risk. Figure 2-4 identifies basic differences between the traditional IT model and a storage cloud model.

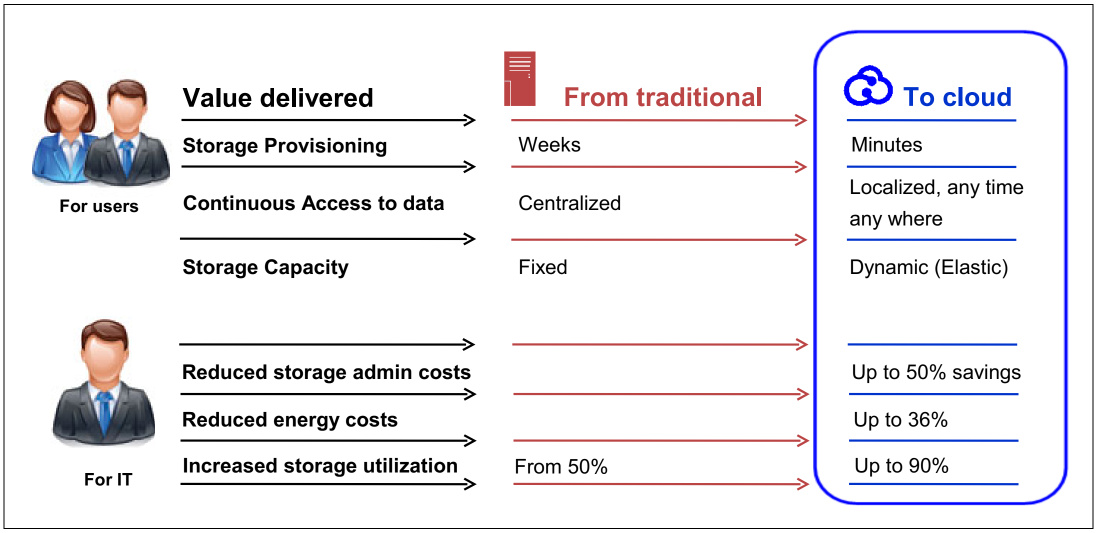



*Figure 2-4 Benefits of moving to storage cloud from traditional IT infrastructure*

### 2.3.1 Dynamic scaling and provisioning (elasticity)

One of the key advantages of storage cloud is dynamic scaling, also known as *elasticity*. Elasticity means that storage resources can be dynamically allocated (scaled up) or released (scaled down) based on business needs. Traditional IT storage infrastructure administration most often acquires capacity that is needed within the next year or two, which necessarily means this reserve capacity will be idle or underutilized for some period or time. A storage cloud can start small and grow incrementally with business requirements, or even shrink to lower costs if appropriate to capacity demands. For this key reason, storage cloud can support a company's growth while reducing net capital investment in storage.

### 2.3.2 Faster deployment of storage resources

New enterprise storage resources can be provisioned and deployed in minutes compared to less optimized traditional IT, which typically takes more time, sometimes days or even months.

### 2.3.3 Reduction in TCO and better ROI

Enterprise storage virtualization and consolidation lower infrastructure total cost of ownership (TCO) significantly, with centralized storage capacity and management driving improved usage and efficiency. It generally provides a significantly higher return on investment (ROI) through storage capacity cost avoidance. In addition, savings can be gained because of reduced floor space, energy required for cooling, labor costs, and also support and maintenance. This gain can be important where storage costs grow faster than revenues and directly affect profitability.

### 2.3.4 Reduce cost of managing storage

Virtualization helps in consolidating storage capacity and helps achieve much higher utilization, significantly reducing the capital expenditure on storage and its management. Storage virtualization is explained further in 3.2.1, “Virtualization” on page 50.

### 2.3.5 Dynamic, flexible chargeback model (pay-per-use)

By implementing storage cloud, an organization pays only for the amount of storage that is actually used rather than paying for spare capacity that remains idle until needed. This model can provide an enterprise with enormous benefits financially. Savings can also be realized from hardware and software licensing for functions such as replication and point-in-time copy.

### 2.3.6 Self-service user portal

A self-service user portal that is based on a service catalog empowers clients to automatically provision based on predefined templates. You can manage IT infrastructure that is based on the user’s needs.

### 2.3.7 Integrated storage and service management

The storage cloud infrastructure usually includes integrated management software, which helps to manage the complete storage infrastructure from a single console, without having to buy proprietary management software from multiple vendors. This technique saves time and helps reduce spending on management software.

### 2.3.8 Improved efficiency of data management

Consolidation and standardization of storage resources facilitates less infrastructure complexity, which is intrinsically simpler to manage. Consistent policies and processes with integrated management tools support geographically diverse infrastructure requirements that are driven by performance or availability considerations.

### 2.3.9 Faster time to market

Automation, self-service portals, rapid deployment, dynamic scaling, and centralized storage management enhance business agility by facilitating significant improvements, such as decreased time-to-market for new products. Businesses can focus on building their core products and competencies instead of worrying about the management of their IT infrastructure.

## 2.4 Storage classes for cloud

Enterprises with optimized storage infrastructures use storage tiers with characteristics that are aligned to business process operational requirements. These tiers support granular service levels for performance, resiliency, availability, security, retention, and so on, as defined for various workloads, which are outlined in Table 2-1.

*Table 2-1 Typical storage service level requirements for various workloads*

| Service level | Mission-critical: OLTP | Business-vital: OLTP | Business-vital: Data warehouse | Business-vital: File service | Business-important: File service |
|---|---|---|---|---|---|
| Availability | | | | | |
| ► Planned uptime | Five 9s + | Four 9s + | Four 9s | Four 9s | Three 9s |
| ► Redundant local disk | Double | Single, double | Single | Single | Single |
| ► Remote replication | Yes | Yes | Possibly | Possibly | No |
| ► Snapshot | Multiple | Multiple | Yes | Multiple | Yes |
| Performance | | | | | |
| ► Sequential I/O latency | Best | Better | Best | Better | Better |
| ► Random I/O latency | Best | Best | Good | Better | Good |
| ► I/O throughput | Best | Best | Good | Better | Good |
| Recovery | | | | | |
| ► RPO | 5 minutes | 4 hours | 24 hours | 4 hours | 24 hours |
| ► RTO | 2 hours | 4 hours | 6 hours | 4 hours | 24 hours |
| ► Disaster resources | Tier 1 | Tier 2 | Tier 2 | Tier 2 | Tier 3 |
| Storage class | Enterprise | Enterprise, mid-range | Mid-range | Mid-range | Mid-range, low cost |

Table 2-2 shows classes of storage and their characteristics.

*Table 2-2 Types of storage and their features or requirements*

| Types of data | Typical features or requirements |
|---|---|
| Structured, transactional, or both | ► Storage to support runtime computations of a compute cloud, for example, database indexing, which are considered “tier one”<br>► Must be co-located with the computation<br>► Has the most stringent latency, I/O operations per second (IOPS), and data protection requirements<br>► Is the least sensitive to cost and is the smallest quantity of storage |
| File, unstructured, or both | ► Storage that allows a customer to flexibly increase file storage capacity, for example productivity, web content<br>► Must be relatively close to customer data center<br>► Has intermediate latency and IOPS requirements<br>► Has immediate sensitivity to cost |
| Fixed Content | ► Contains objects that are written once and never modified but can be replaced, for example records, images<br>► Can accept some latency in access to first byte and is not focused on IOPS<br>► Has high sensitivity to cost and is the largest quantity |

## 2.5 Storage cloud delivery models

The cloud delivery model that is described in 1.3, “Introduction to cloud service models” on page 4 can be extended to include storage cloud as outlined in the following descriptions.

### 2.5.1 Public storage cloud

Data is stored on the premises of the cloud storage service provider and is accessed through network services. All the management tasks that are associated with storage, such as upgrading and replacing, are carried out by the storage service provider. You just pay for the amount of storage space that is consumed. Typically, this storage capacity is somewhat inexpensive because of economies of scale, with different levels of performance and availability at different price points. For data stored in the public storage cloud, security and multitenancy are major areas of concern that need to be evaluated in accordance with business requirements. Storage resources can be scaled up or down to meet the user requirements. In this model, the bulk of capital expenditures (CAPEX) to acquire storage capacity is shifted to operational expense because the storage cloud service provider purchases the resources and therefore incurs the CAPEX.

### 2.5.2 Dedicated private storage cloud

Data is stored on the premises of the cloud storage service provider and accessed over the client's intranet. The management can be done either by the client or can be outsourced to the service provider. Like the public storage cloud, different levels of performance and availability can be provided at different price points.

Unlike the public model, data is comparatively secure behind enterprise firewalls on dedicated hardware. Because the storage space is not shared by other organizations, security and multitenancy concerns are similar to traditional IT. In this model, the client might also save significantly with storage consolidation and virtualization.

### 2.5.3 Local private storage cloud

Data is stored on the client's own premises and accessed within the client's intranet. The management can be done either by the client or can be outsourced to a service provider. Like the dedicated private cloud storage model, data is comparatively secure behind enterprise firewalls. Because the storage space is not shared by other organizations, security and multitenancy concerns are the same as in traditional IT. In this model, the client can also save significantly with storage consolidation and virtualization.

### 2.5.4 Hybrid storage cloud

As the name implies, hybrid storage data is provisioned in a mixed local, dedicated, and public environment. For example, business-critical data (payroll processing, HR, finance) can be stored in a dedicated or local private cloud (to provide security and control over the data) and relatively less important data can be maintained in public cloud storage.

### 2.5.5 Community storage cloud

A community storage cloud limits access to a cloud infrastructure to organizations within a specific "community" that has common requirements and concerns (for example, mission, security requirements, policy, and compliance considerations). The participating organizations realize the benefits of a storage cloud, such as shared infrastructure costs and a pay-as-you-go billing structure, with added levels of privacy, security, and policy compliance that are usually associated with a private cloud. The community cloud infrastructure can be delivered on premises or at a third party's data center, and can be managed by the participating organizations or a third party.

## 2.6 The storage cloud journey

The journey to storage cloud starts at different places for different organizations. This section, describes an effective path to transition from a traditional IT infrastructure to a cloud-based storage infrastructure.

Figure 2-5 shows the typical journey from a traditional model to cloud-based model.

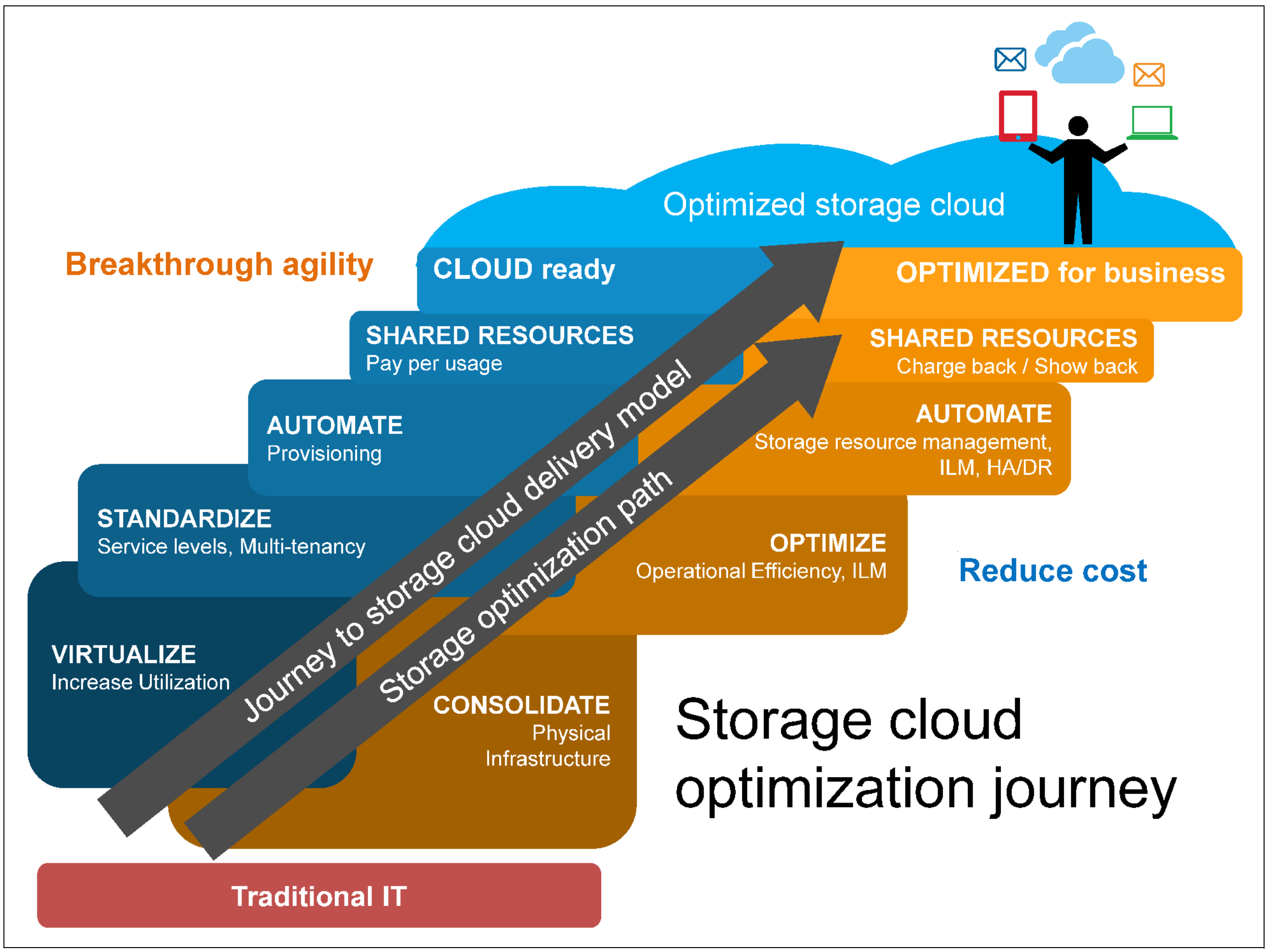



*Figure 2-5 The overall cloud journey from traditional IT to storage cloud*

Storage cloud offers a path to IT optimization by implementing common key practices such as virtualization, standardization, and automation. An optimized storage infrastructure aligns IT resources to business requirements through managed service levels that are usually defined in a service catalog, which is supported within a storage cloud implementation. The journey takes the following path:

- Traditional IT

  Evaluate the current IT infrastructure (servers, storage, networking, and so on) and identify where servers and storage can be consolidated for better performance and utilization, and operational efficiency.

- Consolidate

  Inventory the storage capacity by location, identifying opportunities to combine capacity where feasible to drive inherent economies of scale and usage improvement.

- Virtualize

  Virtualize storage capacity for better utilization and performance.

- Optimize

  Optimization aligns business requirements with cost-effective infrastructure through service-level management. Tiering, archiving, and space reclamation are key practices in achieving an optimized storage infrastructure.

- Automate

  Automate the storage administrative processes, such as the movement of data, by using policies across different storage tiers,. This system enables faster access to the most frequently used data, and also ensures that the data is stored in the correct place.

- Shared resources

  After consolidating and virtualizing the storage resources, the infrastructure is ready to be shared across the global enterprise.

- Cloud-ready

  Although all of these practices are not mandatory, they are all instrumental for deploying an optimized infrastructure within a storage cloud implementation for your enterprise. Consolidation of servers and storage with virtualization technologies improves utilization, while standardizing infrastructure and processes improves operational efficiency. Automation facilitates flexible delivery while enabling client self-service. Establishing common workloads on shared resources allows clients to provision new workloads in a dynamic fashion to achieve a true cloud-enabled environment.

**Solutions:** IBM offers a comprehensive set of solutions geared toward enabling a cloud infrastructure for clients, from small and medium businesses to global enterprises. See Chapter 4, “IBM Storage Systems for SDS” on page 59 for a survey of industry-leading, enterprise-ready IBM storage offerings for cloud.

## 2.6.1 Example: Distributed computing environment

From an IT management perspective, centralizing data is an often repeated mantra because it results in reduced capital expenditure, management costs, and security risks. However, for many organizations, centralizing data storage, although a laudable goal, might not be achievable, perhaps because of technology limitations, or a business operational model. The use case presented here illustrates how an organization, operating in a distributed computing environment, can benefit from the introduction and use of a storage cloud.

## Environment description

Figure 2-6 illustrates a typical topology of an organization that is operating within a distributed computing environment model.

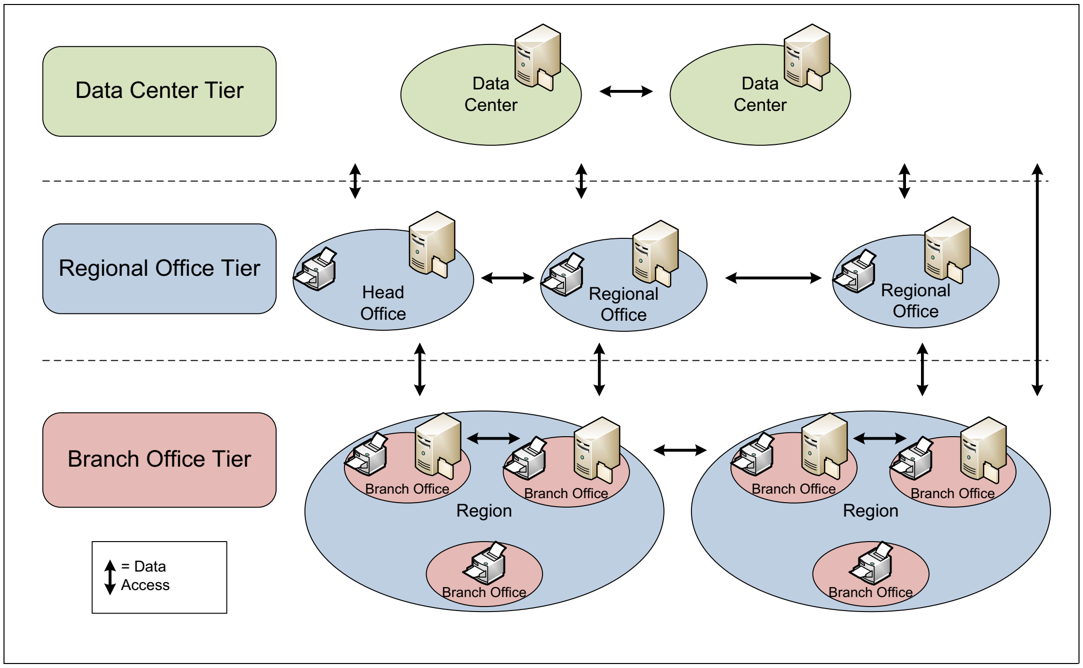



*Figure 2-6 Distributed Computing Environment tiering model*

The following examples are types of organizations that might operate within the distributed computing model:

- Financial institutions
- Government departments
- Retail organizations

The following sections describe the tiering structure that is shown in Figure 2-6, and some of the operational characteristics of an organization that is operating within the distributed computing model.

### *Data center tier*

The organization has one or more data centers. The data centers host a high concentration of the corporate IT infrastructure and data. Where more than one data center exists, data replication is usually required between the data centers to meet business continuity and high availability requirements. A data center typically does not house any users, and can be operated in a “lights out” (remotely, automatically operated) fashion. Some data that is held within the data center (typically high-value transactional data) is accessed only over the organization’s wide area network (WAN). Other data held there can be accessed at another tier to avoid WAN latency and contention issues. Backup data is typically held there also.

### *Regional office tier*

Regional offices are large corporate offices that host IT infrastructure in support of the personnel who are at the office. A regional office might also provide services to branches within the region. A head office can act as a regional office in this tier. A regional office can be co-located with a data center, and therefore share the data center infrastructure. In some organizations, this tier might be small, or omitted entirely.

Read-only data that is held in this tier includes IT support data, such as standard operating environment images, client-side application packages, updates, and patches. Other read-only data can include corporate policy documents and other reference material. Although this type of data is often accessed through web technologies, where manipulation or printing is required, the data might be better placed locally to reduce the impact of WAN traffic.

Read/write data that is held in this tier includes a user's personal data, and data shared among co-workers within a team. Teams might be spread across regional offices within this tier.

Although most users who are operating within this tier are normally dedicated to a single regional office, users in management roles might roam across the regional offices.

#### *Branch office tier*

Branches often represent the public face of an organization. It is here that much of the transactional data is initiated. Branches can vary widely in terms of size and numbers of users. Some can be so small that the presence of significant local IT infrastructure cannot be cost-justified. In this case, the branch can be serviced out of the closest regional office, or directly from the data center. Data requirements for a branch are often identical to a regional office, including read-only and read/write data.

For some organizations, branch users are not dedicated to a single branch, but roam among branches within a region. Regional managers might also spend time at the branches for which they are responsible.

### Benefits of a storage cloud implementation

For a Distributed Computing Environment, a storage cloud provides significant benefits for the accessibility, replication, and hierarchical storage management of data.

#### *Data accessibility*

One of the features of a storage cloud is its ability to consolidate multiple disparate data islands into a single data repository, accessible to anyone from anywhere throughout an organization (if security permits it). This single view of data is helpful in a distributed computing environment, where data islands are prevalent. Users and administrators can take advantage of this consolidated view to more easily access the data across the organization.

#### *Data replication*

Data replication is the key to enabling effective user roaming within and across the Distributed Computing Environment tiers. It can reduce WAN congestion and improve operational performance by having users access copies of data that is on their local area network (LAN) rather than across the WAN.

Branch staff can have their personal data replicated to branches within their region. Regional managers can have their personal data replicated to all of the branches within their region. Inter-region managers can have their personal data replicated to all regional offices. Teams that operate across regions can have their shared data replicated to their own regional office.

Each tier can have data replicated to its parent to facilitate high availability at the originating tier, and also to enhance the efficiency of the enterprise backup strategy. Corporate data can be replicated out to the branches for local manipulation, including printing.

IT infrastructure data can be replicated to all locations to facilitate IT-related tasks, such as workstation builds, software distribution, and patch management.

Although data replication is the key enabler for solving the data distribution dilemma, a smart storage cloud solution enhances the process by supporting automated management functions. These functions include features such as caching to reduce the amount of WAN traffic when accessing remote files, checking file “staleness” to ensure that the current version of a file is always used, delta updates to minimize network traffic for updated files, and multiuser access management to eliminate update conflicts. These features are provided by the IBM Active Cloud Engine®. For further information, see “Active File Management” on page 96.

#### *Hierarchical storage management*

Data islands, as encountered within a distributed computing environment, exacerbate the issue of organization data growth. As requirements for data storage grow, pressure is placed on each individual data island for capacity expansion. This pressure can be relieved by smart management of the data on each island. Because a storage cloud has a consolidated view of the storage environment across the enterprise, it is able to make intelligent decisions about whether data should be stored at a particular location based on the metadata. Data that is infrequently used can be migrated to a central location. Inactive data can be archived and retrieved on demand. These features allow for storage optimization, without the need for administrators to individually manage each storage repository.

## 2.6.2 Example: Development and test environment

The software development lifecycle poses many challenges to a development organization. These challenges include enabling an agile development environment that supports the short time-to-market goals of the business, managing version control, meeting dynamically changing requirements, and keeping ahead of the competition.

### Business description and challenge

Consider the example of a web development and hosting business unit within an IT services company. Such a unit is likely to experience peaks and troughs in demand, based on the business activity cycles of the clients that they service. These demand-based fluctuations lead to bursts of intense activity, requiring access to large amounts of human and technology resources. At the completion of the development tasks, the resource requirements diminish significantly. Similarly, web hosting services add to the resource demand fluctuations as website traffic changes dynamically based on user-access patterns that are driven by market forces, some of which are predictable, and some are not.

A traditional technology infrastructure presents the following challenges to an organization that is operating this type of business unit:

- Cost of provisioning and managing separate infrastructure for the differing business units
- Forecasting infrastructure requirements
- Provisioning adequate infrastructure for demand peaks

  Lead times for procuring and commissioning hardware are relatively lengthy, resulting in capacity that does not meet demand at critical project times, or resulting in the business being unable to compete for business opportunities.
- Wasting capital investment

  After hardware is provisioned, it might remain idle for lengthy periods of time, wasting capital investment.
- Determining pricing models, based on apportionment of infrastructure resource utilization

- Cost of upgrading

  Infrastructure investment can leave the business behind in terms of current technology because the cost of upgrading diverse infrastructure can be prohibitive.

Figure 2-7 shows the current IT structure of Organization ABC's currently isolated IT structures.

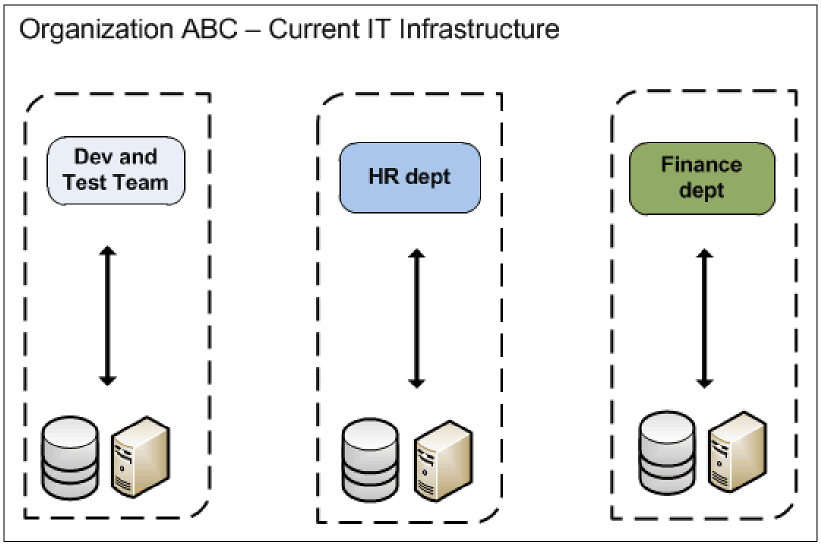



*Figure 2-7 Various teams' dedicated access makes sharing hardware resources difficult*

Figure 2-8 shows how Organization ABC is now better prepared to adapt to changing demands.

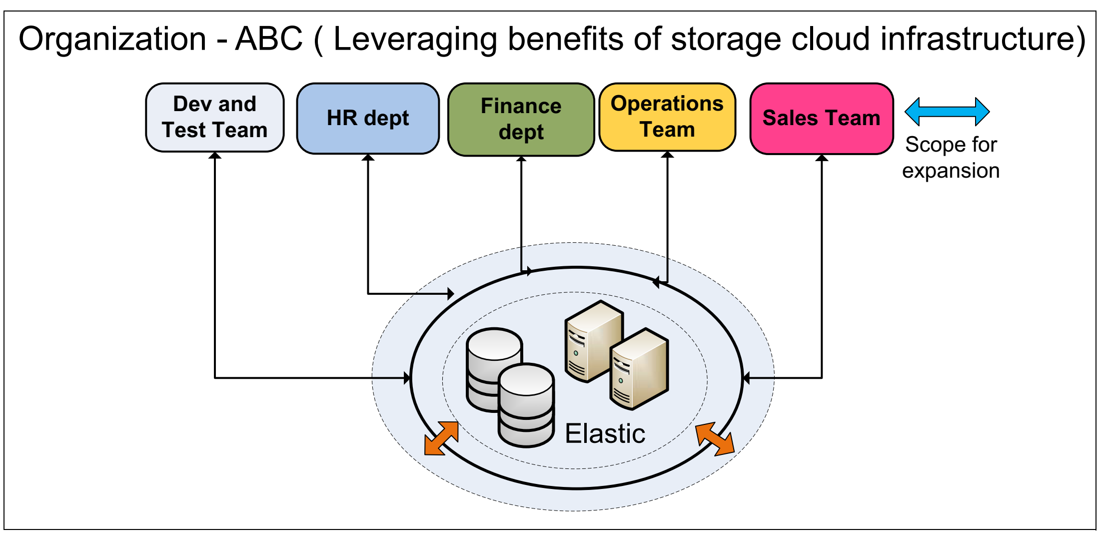



*Figure 2-8 Storage and compute resources can be scaled up or down to meet new demands*

## Benefits of a storage cloud implementation

A storage cloud can help the business units become more agile and dynamic to meet the fluctuating resource demands of their clients. Storage cloud also helps the larger organization to implement a pay-per-use model to accurately track and recover infrastructure costs across the business units.

### Cost reduction

The business unit can provision storage to its clients at a significantly reduced cost because the infrastructure costs are shared across multiple customers and other business units, rather than paid solely by the client.

By consolidating its storage infrastructure, the organization is able to provide a single storage infrastructure over a broader client base. This way provides both economies of scale, and the potential to even out demand peaks and troughs.

Pooling of storage resources means that the organization can allocate storage from anywhere to where it is the most effective in meeting a client needs.

### Elasticity

Client resource demands can be met with agility because a storage cloud enables resources to be provisioned in an elastic fashion, dynamically as demands dictate.

Internal resource peak and trough demands for resources can also be met by provisioning a storage cloud. After activities, such as testing, are completed, the virtual machines and the attached storage that is associated with these activities can be released and added back to the virtual storage pool to be used later, or by other business units.

### Rapid provisioning

A storage cloud allows for rapid provisioning of resources by providing a consolidated view of resources and automation of the overall storage provisioning process through a storage cloud portal.

Automation and self-provisioning also helps the temporary workforce, in terms of providing the test setup in minutes rather than weeks. This feature means that personnel can be productive on startup, rather than being delayed by infrastructure provisioning workflows.

Standard deployment templates, which can be customized for differing environments, ensure that the provisioned environments are more stable and less error-prone, improving the quality of deliverables.

### Faster time to market

As a result of the reduction in time spent for manual provisioning processes, the business unit is able to focus on its core competencies, rather than being distracted by storage infrastructure administration. Less administrative complexity provides benefits like faster time to market for new products and services.

3

# What enables a storage cloud

As outlined in the first two chapters, certain functions are vital to implementing a cloud architecture. Other capabilities enhance the overall infrastructure to make it a more effective and more efficient implementation that is optimized to business requirements. This chapter describes the key features and capabilities that enable a storage cloud.

This chapter includes the following sections:

- Automation and management
- Storage efficiency
- Monitoring and metering
- Resiliency and data protection
- Security and audit
- Scalability and elasticity
- WAN optimization
- Hybrid cloud enablement

## 3.1 Automation and management

As storage vendors seek to reduce the costs of configuring and managing their products, they are increasingly turning to automation to eliminate repetitive tasks, freeing up administrators for more productive activities. In storage clouds, vendors are introducing smarter automation that optimizes the storage by taking into account the capacity and workload requirements in a multi-tenant environment.

OpenStack and VMware technologies are key enablers of cloud infrastructure as a service (IaaS) capability. Both OpenStack and VMware provide overall cloud preferred practices, and automated provisioning, workflow, and orchestration capabilities.

### 3.1.1 Block storage automation and management

Automation and management of block storage in cloud environments is achieved by integrating with APIs provided by the automation and orchestration layer. This section describes the storage APIs that are used by OpenStack and VMware cloud automation software to reduce the cost of managing storage environments.

### 3.1.2 Block storage support for OpenStack

The OpenStack Cinder component enables the creation, provisioning, and management of block storage for server images in a seamless fashion. When provisioning servers, the Nova component can talk directly to the Cinder component to request storage. The Cinder component is able to perform any requested operations on storage systems, and return the results to Nova for attachment to the compute server images.

Storage vendors implement Cinder drivers by using the Cinder API, which provides the following functions:

- Creating, modifying, extending, and deleting block storage LUNs
- Attaching/detaching these LUNs to server images created in the Nova component
- Performing block storage functions such as snapshots (create, delete, list)
- Providing support for backup and restore of volumes (volume from snapshot/image)
- Protection functions such as Volume Replications and Consistency Groups
- Migrating volumes from one system to another

### 3.1.3 Block storage support automation and management for VMware

IBM Block storage solutions have deep integration with VMware. The methods for using those APIs with an IBM SDS Block Storage based solution are detailed in "IBM Spectrum Control Base Edition" on page 64. The data path based APIs such as the Storage Resource Management are supported directly on the IBM SDS offering. The management-based APIs are supported for the IBM SDS offerings through the IBM Spectrum Control™ Base tool, including these items:

- vSphere Storage APIs for Storage Awareness (VASA): With the VASA v1.0 provider, a VMware vCenter Web Client Administrator can view IBM Storage capabilities and optimize VM storage placement automatically (VMware Storage DRS). With the VASA 2.0 provider, a VMware vCenter Web Client Administrator can offload VM-granular snapshots and cloning to IBM Storage, automate IBM storage provisioning by workload-aware policy, and apply VM-granular backup and in-place restore based on IBM Storage snapshots. With the VASA v2.0 provider, a storage administrator can define and publish to vCenter workload-specific storage services and does not need to manage VVoLs or pre-allocate large capacity for data stores.

- VMware VWC (WebClient): With the VWC VMware, a vCenter administrator can discover data store relationships with IBM Storage volumes, view native IBM Storage array, pool, and volumes/shares properties as well as define self-provision volumes and file shares from delegated pools.

vRealize Suite for vSphere based APIs use these applications through IBM Spectrum Control Base (Management):

- VMware vRA/vRO: The VMware (vRO) administrator can apply simple IBM Storage discovery and provisioning in custom automated workflows and can easily develop IBM storage-based workload (PaaS) and storage (SaaS) blueprints. Application owners can self-provision IBM-Storage based workloads.
- vROps: The VMware vROps operator can be notified about unexpected IBM Storage behavior (trend analysis, alerts, and events), easily traverse relations between VM and IBM Storage components (resolve root cause from impacted workload), view trends of a rich set of IBM Storage statistics and apply ready or custom thresholds for notification, and centrally view IBM Storage alerts and events.

VMware Storage based APIs used directly through the IBM software-defined storage offerings for IBM Spectrum Accelerate™ and IBM Spectrum Virtualize™ block offerings provide these features:

- vStorage APIs for Array Integration (VAAI) is a feature that provides hardware acceleration functions. It enables your host to offload specific virtual machine and storage management operations to compliant storage hardware. With the storage hardware assistance, your host performs these operations faster and consumes less CPU, memory, and storage fabric bandwidth.
- vSphere Virtual Volumes (VVoL), in concert with the VASA v2.0, provides storage abstraction, and delivers easy automated provisioning with tenant domains, policy-compliant service, snapshot and cloning offloading, and instant space reclamation. The Virtual Volume model eliminates the complexity of managing the storage infrastructure. It introduces a new control plane in Storage Policy-Based Management.
- VMware vCenter Site Recovery Manager (SRM) is the disaster recovery management product that ensures simple and reliable disaster protection for all virtualized applications. Site Recovery Manager can use storage-based replication to provide centralized management of recovery plans, enable nondisruptive testing, and automate site recovery and migration processes through the VMware Storage Replication Adapter (SRA) for the individual software-defined storage (SDS) block storage offerings. The SRA enables the communication with vSphere SRM to enable the awareness of storage-based replication.

For more information about the IBM Storage Integration points on VMware, see these resources:

- *IBM FlashSystem V9000 and VMware Best Practices Guide*, REDP-5247
- *IBM Spectrum Virtualize and SAN Volume Controller Enhanced Stretched Cluster with VMware*, SG24-8211

### 3.1.4 File storage automation and management

Similar to block storage, automation and management of file storage in cloud environments is also achieved by integrating with APIs provided by the automation and orchestration layer. This section describes the storage APIs that are used by OpenStack and VMware to provide this capability.

### 3.1.5 File storage support for OpenStack

The OpenStack Manila component provides file storage that allows coordinated access to shared or distributed file systems. Although the primary consumption of shares is OpenStack compute instances, the service is also intended to be accessed independently, based on the modular design established by OpenStack services. Storage vendors implement Cinder drivers by using the Manila API, which allows you to create, delete, and list file shares.

### 3.1.6 Object storage automation and management

Object storage provides RESTful APIs that enable automation for the following management tasks:

- Adding and removing users and groups
- Allocating and de-allocating vaults
- Creating, modifying, and deleting capacity quotas
- Modify role-based access and ACLs
- Manage object storage software and hardware components
- Capacity / node provisioning / de-provisioning
- Creating, modifying, and deleting notifications such as SNMP and email alerts

## 3.2 Storage efficiency

The insatiable desire for increased data storage space has led to significant innovations in storage efficiency. This section describes these innovations and the ways in which storage clouds are using them to provide users with a better return on their storage investment.

### 3.2.1 Virtualization

Storage virtualization refers to the abstraction of storage systems from applications and servers. It is a foundation for the implementation of other technologies, such as thin provisioning, tiering, and data protection, that are transparent to the server. It is one of the key enablers for storage cloud environments where several cloud services typically share one common infrastructure. Storage virtualization abstracts storage from multiple sources into a single storage pool. It helps you to manage the rapid information growth by using your storage equipment and data center space more effectively. The increase in storage utilization reduces power costs and keeps the footprint of your storage hardware small.

### 3.2.2 Compression

The amount of stored data continues to grow exponentially every year, which creates tremendous strain on the IT infrastructure, especially on storage systems. Additional storage systems can help to meet these storage growth requirements in the near-term. However, shrinking IT budgets are pressuring IT managers to increase the lifetime of existing storage systems. More storage systems lead to higher energy costs, and available floor space in data centers is often a considerable limitation. Compression provides an innovative approach that is designed to overcome these challenges.

Online compression immediately reduces the physical storage across all storage tiers. It allows storage administrators to gain back free disk space in the existing storage system without the need to change any administrative processes or requiring users to clean up or archive data. The benefits to the business are immediate because the capital expense of upgrading the storage environment is delayed. Compression also reduces the environmental requirements per unit of storage. After compression is applied to stored data, the required power and cooling per unit of storage are reduced because more logical data is stored in the same physical space.

## 3.2.3 Data deduplication

Data deduplication is a key technology to dramatically reduce the amount of, and the cost associated with, storing large amounts of data by consolidating redundant copies of a file or file subset. Incoming or existing data is standardized into “chunks” that are then examined for redundancy. If duplicates are detected, pointers are shifted to reference a single copy of the chunk and the extraneous duplicates are then released.

## 3.2.4 Thin provisioning

Traditional storage provisioning pre-allocates and dedicates physical storage space for use by applications or hosts. However, the total requested capacity is usually not required from the beginning when the assignment is made, but it needs to be physically available already. Furthermore, estimating the exact amount of required space for a new application, which can lead to over-provisioning, is sometimes difficult or even impossible. This challenge results in wasted space, which is known as white space, and inefficient use of the physical storage. Figure 3-1 illustrates the advantages of thin provisioning in terms of storage allocation.

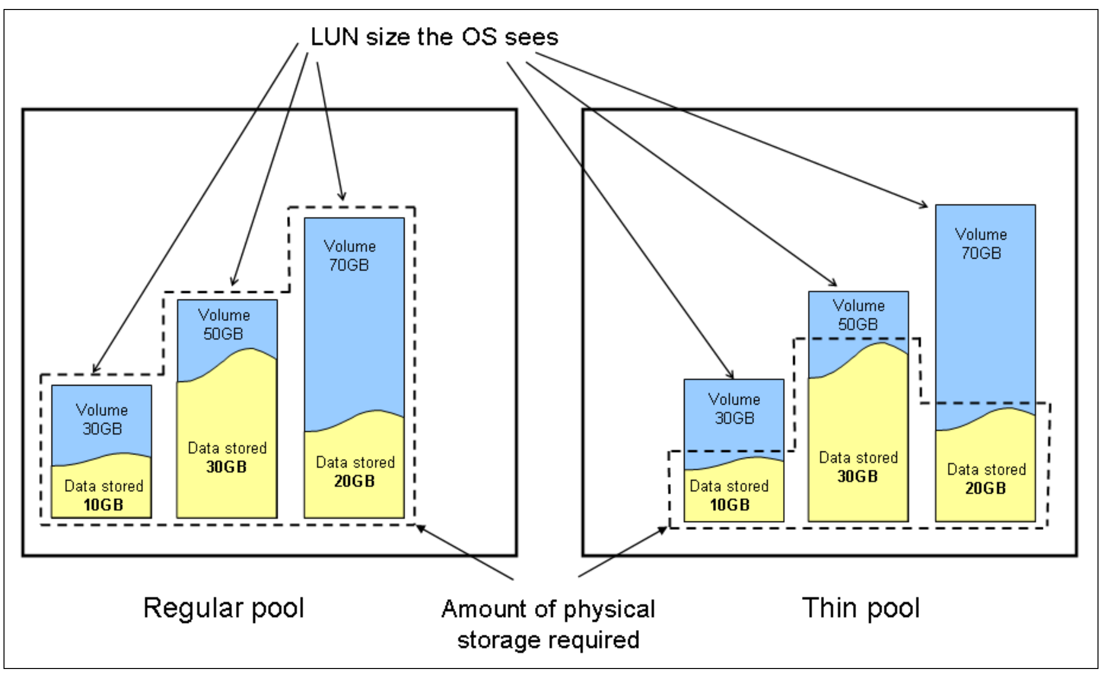



*Figure 3-1 Advantages of thin provisioning over regular storage provisioning*

Thin provisioning allows applications and servers to see logical volume sizes that are larger than the physical capacity that is actually dedicated to the volumes on the storage system. Physically, capacity is allocated for the volumes only as needed. This method allows a higher storage utilization, which in turn leads to a reduction in the amount of storage that is needed, lowering capital expenses. Furthermore, the usage of thin-provisioned storage postpones the need to invest in more storage. Thin provisioning also simplifies your 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 some volumes are capacity constrained while others have capacity to spare. In this way, your storage environment becomes more agile.

### 3.2.5 Automated tiering

In modern and complex application environments, the increasing and often unpredictable demands for storage capacity and performance lead to related issues in terms of planning and optimization of storage resources. Determining the amount of I/O activity on storage and what data to move to which storage tier is usually complex. As a manual process, these corrective actions are expensive in terms of hardware resources and labor, and are critical to service availability.

*Automated tiering* refers to the automated migration of data between storage tiers based on real-time analysis of access patterns. This continuously ongoing process consists of these steps:

1. The workload on the storage is monitored by the storage system.
2. After a certain period, the storage system evaluates the historical information to identify “hot spots,” which means data with a high I/O density.
3. The storage system creates a migration plan for moving this “hot spot” data to a higher tier storage that can provide the required performance.

Data whose I/O density has dropped off is moved back to a lower tier. Automated tiering helps you to more precisely plan and manage both storage costs and application performance.

## 3.3 Monitoring and metering

Centralized storage management gives you insights into storage usage and allocation across your heterogeneous storage infrastructures and allows you to manage infrastructure proactively rather than reactively. It leads to cost savings through improved operational efficiency, allows more intelligent business decisions, and enables chargeback and show-back in a multi-tenant environment. A centralized GUI along with RESTful APIs provided by the underlying storage infrastructure are used to provide monitoring and metering capabilities.

## 3.4 Resiliency and data protection

Organizations today demand that their data is protected from corruption and loss, whether by accident or intent. Despite rapid data growth, data protection and retention systems are expected to maintain service levels and data governance policies. Data has become integral to business decision-making and basic operations, from production to sales and customer management.

This section highlights the following data protection mechanisms available within storage clouds, and their relevance to providing the data integrity, which businesses have come to expect:

- Backup and restore
- Disaster recovery
- Archive
- Continuous Data Availability

## 3.4.1 Backup and restore

Backups protect data by creating an extra copy of the data and a backup window defines how long it takes to complete a backup of the storage environment. The backup window is influenced by the amount of data that must be protected as well as the number of files, objects, and volumes that must be protected.

Restores are performed to recover the loss or corruption of data from operational issues such as inadvertent or malicious delete, localized hardware failures, and software issues. The recovery time objective (RTO) defines how long it takes to restore the data and the recovery point objective (RPO) defines the amount of time that elapses between backup operations. The amount of data that is created and modified after the last backup but before the next backup is at risk of being lost because it has not been protected yet. For example, with a nightly backup, the amount of data that is created and modified during the day after the last nightly backup, but before the next nightly backup is at risk of being lost.

## 3.4.2 Disaster recovery

Disaster recovery provides protection against catastrophic site disasters such as earthquakes and floods. To provide disaster recovery protection, backup data at a primary site is replicated to a different geographic location, often referred to as the disaster recovery (DR) site. During a catastrophic failure, data is restored at the DR site, enabling business operations to resume. The amount of time that is required to replicate the backup data to the DR site must be considered as part of the RPO because data that is not fully replicated to the DR site is at risk of being lost during a catastrophic failure at the primary site.

## 3.4.3 Archive

Archiving retains inactive data that has long-term data retention requirements, either for compliance or business purpose. It does so by providing secure and cost effective solutions with automated process for retention policies and data migration to low-cost storage.

## 3.4.4 Continuous data availability

Continuous data availability ensures uninterrupted access to data for critical business systems, reducing the risk of downtime during failure conditions, including site failures. Replication across geographic boundaries and geo-dispersed erasure coding are key functions that provide this capability.

Continuous data availability generally adheres to either a strong consistency or eventual consistency model. Strong consistency ensures that the copy of the data that is returned to the application is always the most up-to-date version, even if the data is written in one location and accessed from another, resulting in an RPO of zero. Eventual consistency does not guarantee that the data written at one geographic location is immediately visible at a different geographic location and is often used to replicate data over long distances. With eventual consistency, the delay in visibility is small, on the order of seconds and is often referred to as a near-zero RPO.

### Data replication

Data replication creates multiple copies of data in different geographic locations to protect against site failure. Synchronous replication provides a strong consistency model with an RPO of zero, ensuring that the data is identical at the different geographic locations. It is often used for mission critical application data, is deployed over metro distances, and is sometimes referred to as mirroring between the two locations. Asynchronous replication of data provides eventual consistency with a near zero RPO, and is typically deployed between two or more geographic locations.

### Geo-dispersed erasure coding

Geo-dispersed erasure coding provides a strong consistency model with an RPO of zero, while reducing the need for storage capacity. Sophisticated algorithms slice a single copy of the data into multiple chunks and distribute them across geographic locations. An operator-defined subset of slices is needed to retrieve data perfectly in real time. The level of resiliency is fully customizable, resulting in a massively reliable and efficient way to store data at scale as opposed to RAID and replication techniques.

# 3.5 Security and audit

Security and audit functions are critical functions of storage clouds. According to the Storage Networking Industry Association (SNIA), data security and audit in the context of storage systems is responsible for safeguarding the data against theft, prevention of unauthorized disclosure of data, and prevention of data tampering and accidental corruption. These features ensure accountability, authenticity, business continuity, and regulatory compliance.

This section contains the following security and audit topics:

- Multitenancy
- Identity management and role-based access
- Encryption
- Audit

## 3.5.1 Multitenancy

The term *multitenancy* refers to an architecture that is typically used in cloud environments. Instead of providing each cloud service consumer (tenant) a separate, dedicated infrastructure (single-tenancy architecture), all consumers share one common environment. Shared layers must behave as though they were set up in a dedicated fashion in terms of customization, isolation, and so on.

A cloud environment has two primary technology stacks where multitenancy is relevant:

- The management environment (cloud management stack)
- The managed environment (infrastructure, platform, or application that is provided as a service)

Depending on the service model, the level and degree of shared infrastructure varies as illustrated in Figure 3-2. For IaaS, typically hypervisors are installed on the managed hardware infrastructure. For platform as a service (PaaS), a multitenancy-enabled middleware platform is used, and for software as a service (SaaS), the multitenancy-enabled software application is divided into virtual partitions.

Multitenancy in cloud service models implies a need for policy-driven enforcement, segmentation, isolation, service levels, and chargeback billing models, because multiple service consumers are using a shared infrastructure. Service consumers can be either distinct organizations in a public cloud service or separate business units in a private cloud service. All cloud service consumers want to ensure that, although from a physical perspective they are sharing infrastructure, from a logical perspective they are isolated without risk to their sensitive data or their workloads.

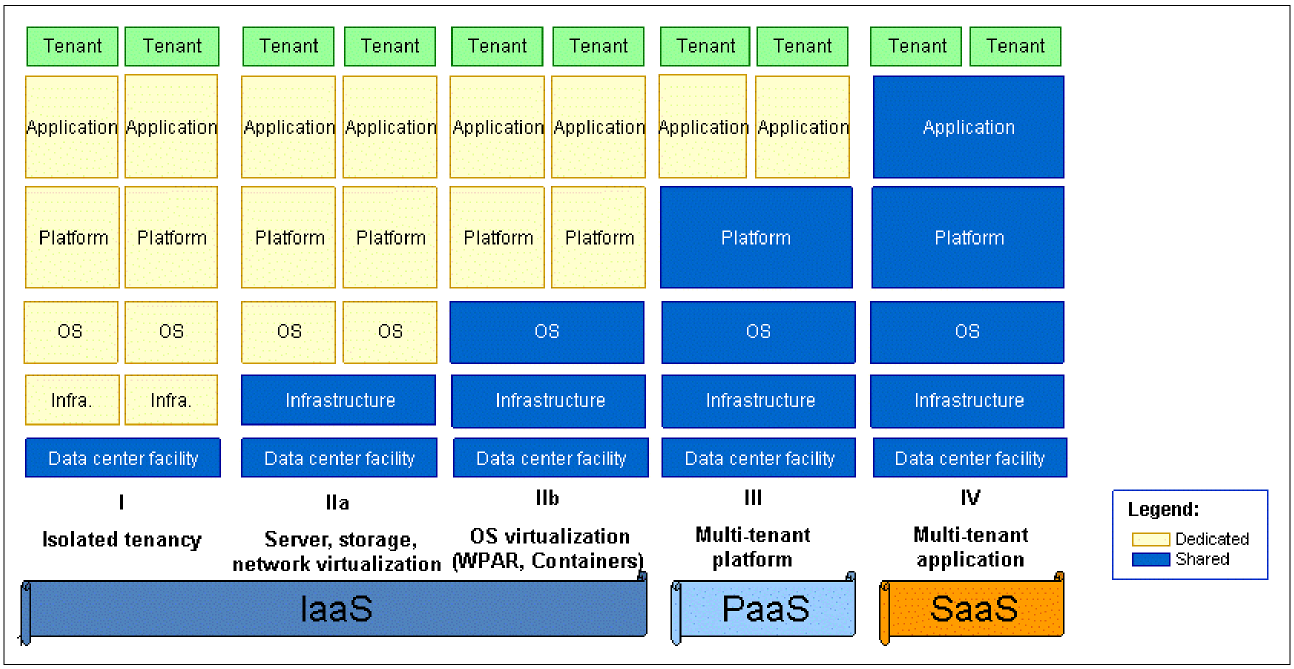



*Figure 3-2 Multitenancy in cloud environments*

Multitenancy offers several main benefits:

- Can quickly scale to more tenants
- Is cost-effective because the infrastructure is shared by all tenants
- Requires less management effort than a virtualized or mediated approach
- Requires less storage

## 3.5.2 Identity management

Authentication is the process of validating the identity of an entity or individual. Authorization is the process of verifying that an entity or individual is allowed to access or alter a resource. With role-based access control (RBAC) roles are defined, permissions are defined for each role, and users are assigned to one or more roles.

Access control lists (ACLs) provide fine grain control of which resources individual users and groups can access. For example, in the context of object storage, ACLs indicate which users and groups are able to access individual buckets or vaults. In the context of file-based storage, ACLs define which users are able to access individual files and directories.

Identity management services are either provided locally by the storage cloud or externally with Active Directory (AD), Lightweight Directory Access Protocol (LDAP), or Keystone. LDAP is a set of protocols that are used to access centrally stored information over a network. User lists and groups within an organization can be consolidated into a central repository accessible from anywhere on the network. Keystone is an OpenStack service that provides API client authentication, service discovery, and distributed multi-tenant authorization for services in the OpenStack family by implementing OpenStack's identity API. Active Directory is a directory service that was developed for Microsoft Windows domain networks that authenticates and authorizes all users and computers by assigning and enforcing security policies.

## 3.5.3 Encryption

Encryption is a technique that is used to encode data with an encryption key so that the information content of the data can be decoded only with knowledge of a decryption key. Data that is encrypted is referred to as *ciphertext*. Data that is not encrypted is referred to as *plaintext* or *cleartext*. With an appropriately derived encryption key and an appropriate encryption algorithm, guessing the decryption key is prohibitively difficult. Data that is encrypted into ciphertext is considered secure from anyone who does not have possession of the decryption key. This section describes the following encryption considerations for storage clouds:

- Encryption of data in motion
- Encryption of data at rest
- Secure data deletion
- Encryption considerations for public storage clouds

### Encryption of data in motion

Encryption of data in motion refers to encrypting data at one endpoint, sending it over a communication wire, and decrypting it at another endpoint, regardless of whether the data is already encrypted. The Transport Layer Security (TLS) is the best protocol that provides this capability and so all communication in a storage cloud should use TLS. Additionally, storage clouds use REST APIs such as OpenStack Swift and Amazon S3 as well as proprietary management APIs over the HTTPS protocol, which uses TLS to secure the base HTTP protocol.

### Encryption of data at rest

Self-encrypting drives (SEDs) perform the data encryption and decryption operations on a dedicated crypto-processor that is part of the drive controller. SEDs protect data at rest by preventing unauthorized access to the storage device by using user-defined authentication credentials when the host system is powered on. If the proper credentials are provided, the drive is unlocked and the user has full access to the drive's decrypted data. Thus, this encryption method protects against attacks targeting the disks, such as theft or acquisition of improperly discarded disks.

Alternatively, encryption can be provided by the storage software and does not require the use of SED drives. In addition to protecting against theft of drives, storage-software-based encryption protects against attacks by unprivileged users of a multi-tenant system.

Encryption key management is the administration of tasks involved with protecting, storing, backing up, and organizing encryption keys, and is a critical component of managing encryption of data at rest. Keys can be managed locally by the storage cloud infrastructure or might be managed externally by using dedicated encryption key management infrastructure.

### Secure data deletion

Secure data deletion uses encryption and key management to ensure erasure of files beyond the physical and logical limitations of normal deletion operations. If data is encrypted, and the master key (or keys) required to decrypt it have been deleted from the key server, that data is effectively no longer retrievable.

### Encryption considerations for public clouds

When using public storage clouds, such as object storage, it is preferable to encrypt data before sending it to the public cloud and keep the encryption keys stored and protected locally on-premises. Because the data is encrypted before it is sent to a public cloud and the keys are not provided, the data in the public cloud cannot be easily deciphered.

## 3.5.4 Audit

Referencing NIST SP-800-14, audit trails maintain a record of system activity by system or application processes and by user activity. Audit trails provide the following advantages:

- Individual accountability: The audit trail supports accountability by providing a trace of user actions. While users cannot be prevented from using resources to which they have legitimate access authorization, audit trail analysis can be used to examine their actions.
- Reconstruction of events: An organization should use audit trails to support investigations of how, when, and why normal operations ceased.
- Intrusion detection: If audit trails have been designed and implemented to record appropriate information, they can help intrusion detection. Intrusions can be detected in real time by examining audit records as they are created or after the fact, by examining audit records in a batch process.
- Problem identification: Audit trails can also be used as online tools to help identify problems other than intrusions as they occur. This feature is often referred to as real-time auditing or monitoring.

The Cloud Auditing Data Federation (CADF) open standard defines a full event model that anyone can use to complete the essential data needed to certify, self-manage, and self-audit application security in cloud environments.

For more information about audit capabilities, see the following links:

- NIST standards

  `http://csrc.nist.gov/publications/nistpubs/800-14/800-14.pdf`

- CADF standards

  `https://www.dmtf.org/standards/cadf`

# 3.6 Scalability and elasticity

The ability to non-disruptively add capacity and remove it as needed in a global namespace is a key function for storage clouds. A global namespace aggregates disparate storage infrastructure, potentially across geographical boundaries, to provide a consolidated file or object view that simplifies administration.

## 3.7 WAN acceleration

WAN acceleration provides efficiency gains and cost savings for storage clouds. WAN acceleration enables data transfer over any network at maximum speed for on-premises, public, and hybrid cloud storage systems, regardless of distance and network conditions.

## 3.8 Hybrid cloud enablement

The following functions enable hybrid cloud capability in storage clouds:

- Extension of a global namespace spanning on-premises and public cloud infrastructure. Data can be accessed on-premises and in a public cloud transparently to the application.
- Tiering to object storage enables the usage of object storage as a tier in the cloud storage infrastructure transparently to applications. Data can be moved to and from object storage transparent to applications to maximize efficiency and optimize cost by placing colder data on low-cost object storage.
- Peer to peer sharing of data between multiple storage systems using object storage. In this model, data is exported from one storage system to object storage and imported into a different storage system from the object storage.
- Using object storage as a backup pool and optionally using it as part of a disaster recovery procedure.

# 4

# IBM Storage Systems for SDS

This chapter is a general overview of IBM Storage Systems, which provide broad functions and can be combined to meet the customers' business requirements for their software-defined storage (SDS) environments. Some IBM Storage Systems come with SDS software that can make deployment even easier. Some fit within the SDS because of the functions that are provided to meet specific critical business requirements.

The requirements for a customer's SDS environment can include response times for *hot* and *cold* data. A FlashSystem V9000 or FlashSystem A9000/A9000R, with Spectrum Scale, Spectrum Archive, and an IBM TS4500 library might be the answer to the requirement for fast access for hot data. The FlashSystem is integrated with active archive requirements for cold data that can be stored on tape. The storage systems can be seamlessly integrated with Spectrum Scale and Spectrum Archive.

SDS is part of the software-defined infrastructure (SDI). A successful implementation can be traced back to the appropriate planning needed to produce the architecture that meets the customer's business requirements.

IBM disk, hybrid disk, and all flash systems provide storage efficiency solutions such as inline Real-time Compression, inline data deduplication, automated tiering, virtualization, and thin provisioning. These storage solutions increase the data storage optimization opportunities for organizations of all sizes to boost system performance and lower IT costs.

This chapter includes the following sections:

- Overview
- SDS Control Plane
- SDS Data Plane
- IBM storage support of OpenStack components
- IBM storage supporting the data plane
- IBM cloud services

# 4.1 Overview

The Spectrum Storage family of products are organized by their functions within the SDS control plane or SDS data plane. The control plane is the software layer that manages administrative functions (for example, configuration, monitoring, replication, policy automation, and provisioning) for software-defined storage resources while data is processed and stored in the data plane.

Figure 4-1 shows the IBM SDS architecture with a mapping of the Spectrum Storage family of products across the SDS control plane and data plane.

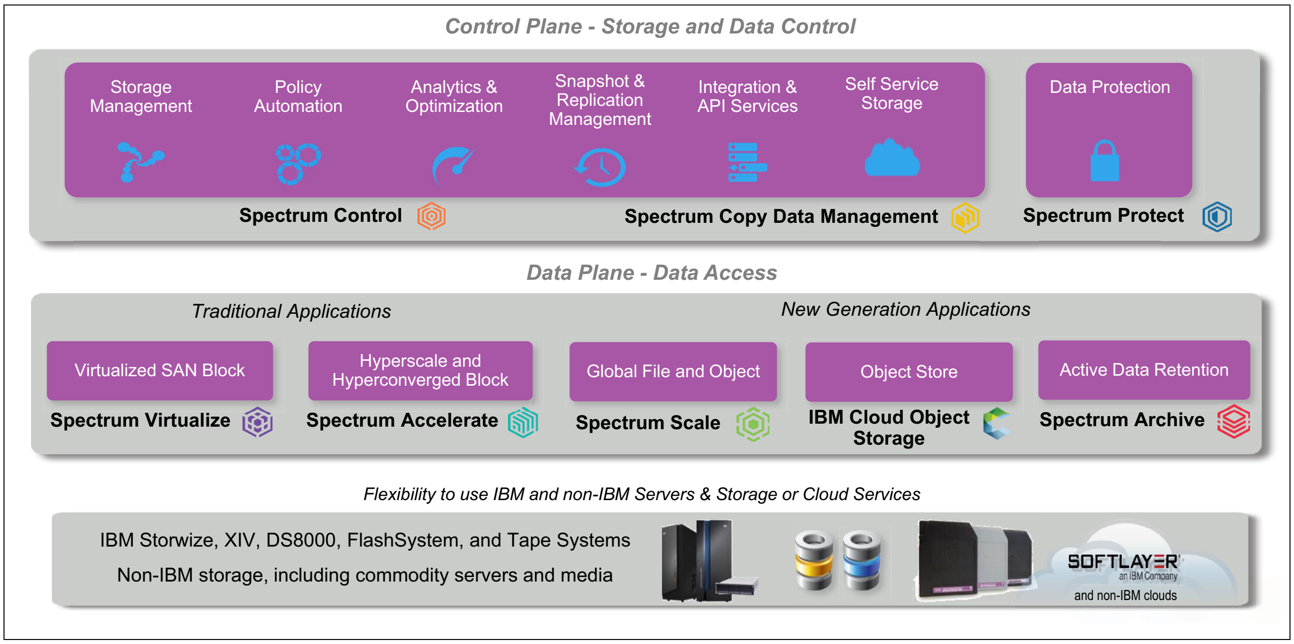



*Figure 4-1 IBM Spectrum Storage family mapped to SDS Control Plane and Data Plane*

Table 4-1 on page 61 is an overview of the IBM Spectrum Storage family including high-level descriptions with the products that provide those functions.

*Table 4-1 IBM Spectrum Storage Family descriptions*

| | Spectrum Storage family member | Description | Former name |
|---|---|---|---|
| | Spectrum Storage Suite | A single software license for all your changing software-defined storage needs. Straightforward per-TB pricing for the entire Spectrum Storage suite. | |
| **SDS Control Plane** | | | |
| | Spectrum Control | Automated control and optimization of storage and data infrastructure | IBM Tivoli® Storage Productivity Center, management layer of IBM Virtual Storage Center, and IBM Storage Integration Server |
| | Copy Services Manager | Automated control and optimization of storage replication features | IBM Tivoli Storage Productivity Center for Replication |
| | IBM Spectrum Protect | Optimized data protection for client data through backup and restore capabilities | IBM Tivoli Storage Manager Suite for Unified Recovery |
| | IBM Spectrum Protect Snapshot | Integrated application-aware point-in-time copies | IBM Tivoli FlashCopy® Manager |
| | Spectrum Copy Data Management | Automate creation and use of copy data snapshots, vaults, clones, and replicas on existing storage infrastructure | |
| **SDS Data Plane** | | | |
| | Spectrum Virtualize | Core SAN Volume Controller function is virtualization that frees client data from IT boundaries | IBM SAN Volume Controller software |
| | Spectrum Accelerate | Enterprise storage for cloud that is deployed in minutes instead of months | IBM XIV® software |
| | Spectrum Scale | Storage scalability to yottabytes and across geographical boundaries | IBM General Parallel File System (GPFS™) |
| | Spectrum Archive | Enables long-term storage of low activity data | IBM Linear Tape File System™ Enterprise Edition, Library Edition, and Single Drive Edition |

# 4.2 SDS Control Plane

The control plane is a software layer that manages the virtualized storage resources. It provides all the high-level functions that are needed by the customer to run the business workload and enable optimized, flexible, scalable, and rapid provisioning storage infrastructure capacity. These capabilities span functions like storage virtualization, policy automation, analytics and optimization, backup and copy management, security, and integration with the API services, including other cloud provider services.

This section describes the IBM software product offerings that provide the building blocks for the SDS control plane:

- IBM Spectrum Control
- IBM Spectrum Protect

## 4.2.1 IBM Spectrum Control

IBM Spectrum Control provides efficient infrastructure management for virtualized, cloud, and software-defined storage by reducing the complexity associated with managing multi-vendor infrastructures and helps businesses optimize provisioning, capacity, availability, protection, reporting, and management for today's business applications without having to replace existing storage infrastructure. With support for block, file, and object workloads, Spectrum Control enables administrators to provide efficient management for heterogeneous storage environments.

### Key capabilities

Spectrum Control helps organizations transition to new workloads and updated storage infrastructures by providing these advantages to significantly reduce total cost of ownership:

- A single management console that supports IBM Spectrum Virtualize, IBM Spectrum Accelerate, IBM Cloud Object Storage, and IBM Spectrum Scale environments, enabling holistic management of physical and virtual block, file, and object systems storage environments.
- Insights that offer advanced, detailed metrics for storage configurations, performance, and tiered capacity in an intuitive web-based user interface with customizable dashboards so that the most important information is always accessible.
- Performance monitoring views that enable quick and efficient troubleshooting during an issue with simple threshold configuration and fault alerting for high availability.

### Benefits

Spectrum Control can help reduce the administrative complexity of managing a heterogeneous storage environment, improve capacity forecasting, and reduce the amount of time spent troubleshooting performance-related issues. Spectrum Control provides these key values:

- Transparent mobility across storage tiers and devices for IBM Spectrum Virtualize based designs
- Centralized management that offers visibility to block, file, and object workloads as well as control and automation of block storage volumes

Figure 4-2 shows the Spectrum Control dashboard window where all the managed resources in your data server are presented in an aggregated view.

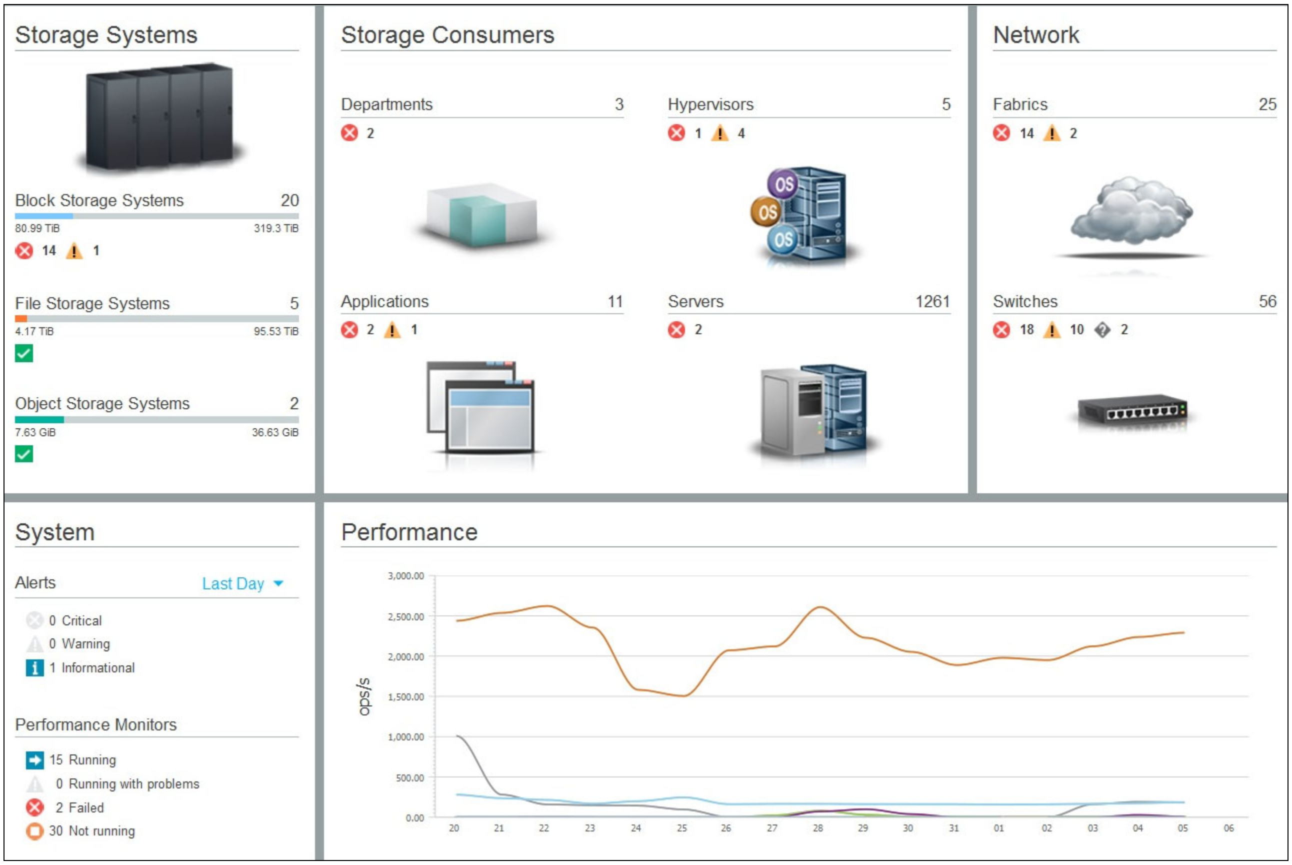



*Figure 4-2 Single dashboard for monitoring all storage components*

## IBM Data and Storage Management Solutions features

Spectrum Control solutions provide improved visibility, simplified administration, and greater scalability. This section describes the features of the specific products that provide the functions for Spectrum Control.

**Note:** The Management Layer of VSC is now called IBM Spectrum Control Advanced Edition.

Figure 4-3 shows the IBM Spectrum Control offerings.

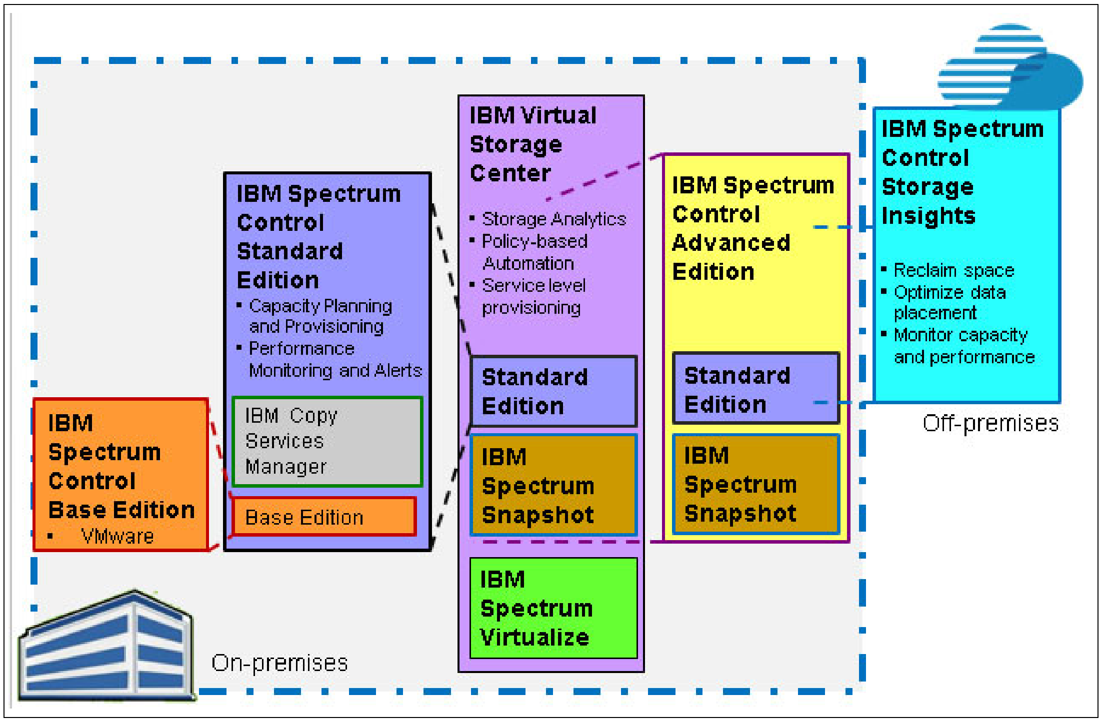



*Figure 4-3 IBM Spectrum Control offerings*

## IBM Spectrum Control Base Edition

IBM Spectrum Control Base Edition is a centralized server system that consolidates a range of IBM storage provisioning, automation, and monitoring solutions through a unified server platform.

As shown in Figure 4-4, IBM Spectrum Control Base Edition provides a single-server back-end location and enables centralized management of IBM storage resources for the use of independent software vendor (ISV) platforms and frameworks. These frameworks currently include VMware vCenter Server, VMware vSphere Web Client, and VMware vSphere Storage APIs for Storage Awareness (VASA). Spectrum Control Base Edition is available for no extra fee to storage-licensed clients.

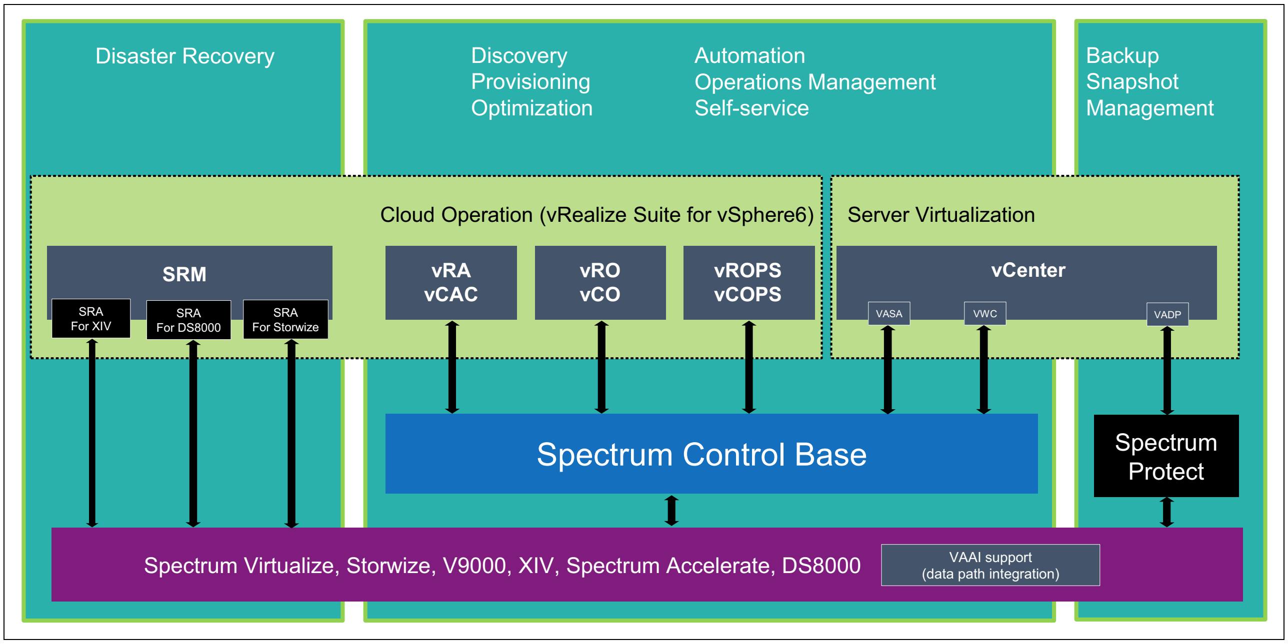



*Figure 4-4 Spectrum Control Base Edition*

In Figure 4-3 on page 85, IBM Spectrum Control Base Edition is not in the data path. Spectrum Control Base Edition runs in the control plane as shown in Figure 4-1 on page 60. Spectrum Control Base Edition provides integration between IBM Block Storage and VMware. Clients utilize Spectrum Control Base Edition if they are or plan on using the VMware Web Client (VWC), VMware Virtual Volumes (VVol) or the vRealize Automation Suite from VMware.

Spectrum Control Base Edition provides common services like authentication, high availability, and storage configuration for IBM Block Storage in homogeneous and heterogeneous multiple target environments. Spectrum Control Base Edition manages IBM XIV Storage System, A9000, A9000R, IBM DS8000 series, IBM SAN Volume Controller, the IBM Storwize® family, and third party storage subsystems.

Figure 4-5 shows IBM Storage connectivity to VMware through IBM Spectrum Control Base Edition.

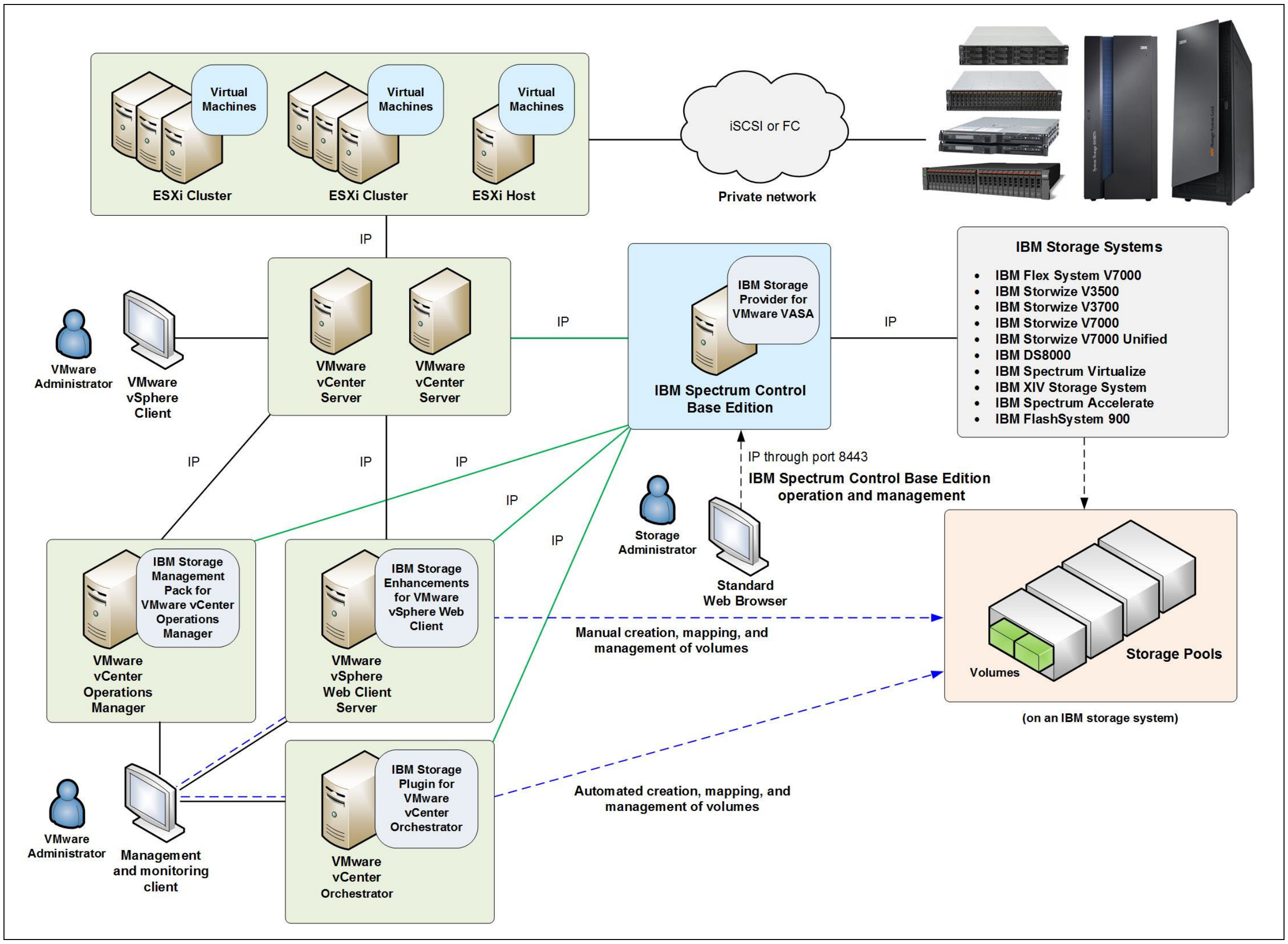



*Figure 4-5 IBM Storage connectivity to VMware through IBM Spectrum Control Base Edition*

For more information, see IBM Spectrum Control Base Edition at:

https://www.ibm.com/support/knowledgecenter/STWMS9/landing/IBM_Spectrum_Control_Base_Edition_welcome_page.html

## IBM Spectrum Control Standard Edition

IBM Spectrum Control Standard Edition is designed to provide storage infrastructure and data management capabilities for traditional and software-defined storage environments. IBM Spectrum Control Standard Edition has these primary features:

- Capacity visualization and management
- Performance reporting and troubleshooting
- Health and performance alerting
- Data Path view
- Department and Application grouping
- Hypervisor integration with VMware
- Management of IBM Replication features with Copy Services Manager (CSM)

### IBM Spectrum Control Advanced Edition

IBM Spectrum Control Advanced Edition includes all of the features of Spectrum Control Standard Edition, and adds the following advanced capabilities:

- Tiered storage optimization with intelligent analytics for IBM Spectrum Virtualize
- Service catalog with policy-based provisioning
- Self-service provisioning with restricted use logins
- Application-based snapshot management from IBM Spectrum Protect™ Snapshot

Advanced Edition has built-in efficiency features that help users avoid complicated integration issues or the need to purchase add-ons or additional licenses:

- Simplified user experience: Virtual Storage Center provides an advanced GUI and a VMware vCenter plug-in to reduce administration complexity. Administrators can perform common tasks consistently over multiple storage systems, including those from different vendors. The IBM storage GUI enables simplified storage provisioning with intelligent presets and embedded best practices, and integrated context-sensitive performance management.
- Near-instant, application-aware backup and restore: To reduce downtime in high-availability virtual environments, critical applications such as mission critical databases or executive email requiring near-instant backups must have little or no impact on application performance. Application-aware snapshot backups can be performed frequently throughout the day to reduce the risk of data loss. Virtual Storage Center simplifies administration and recovery from snapshot backups.

  IBM Spectrum Protect Snapshot, previously known as IBM Tivoli Storage FlashCopy Manager, is designed to deliver data protection for business-critical applications through integrated application snapshot backup and restore capabilities. These capabilities are achieved through the utilization of advanced storage-specific hardware snapshot technology to help create a high-performance, low-impact, application data protection solution. It is designed for easy installation, configuration, and deployment, and integrates with various traditional storage systems and software-defined storage environments.
- IBM Tiered Storage Optimizer: Virtual Storage Center uses performance metrics, advanced analytics, and automation to enable storage optimization on a large scale. Self-optimizing storage adapts automatically to workload changes to optimize application performance, eliminating most manual tuning efforts. It can optimize storage volumes across different storage systems and virtual machine vendors. The Tiered Storage Optimizer feature can reduce the unit cost of storage by as much as 50 percent, based on deployment results in a large IBM data center.

Spectrum Control Advanced Edition is data and storage management software for managing heterogeneous storage infrastructures. It helps to improve visibility, control, and automation for data and storage infrastructures. Organizations with multiple storage systems can simplify storage provisioning, performance management, and data replication.

Spectrum Control Advanced Edition simplifies the following data and storage management processes:

- A single console for managing all types of data on disk, flash, file, and object storage systems.
- Simplified visual administration tools that include an advanced web-based user interface, a VMware vCenter plug-in, and IBM Cognos® Business Intelligence with pre-designed reports.
- Storage and device management to give you fast deployment with agent-less device management.
- Intelligent presets that improve provisioning consistency and control.

- Integrated performance management features end-to-end views that include devices, SAN fabrics, and storage systems. The server-centric view of storage infrastructure enables fast troubleshooting.
- Data replication management that enables you to have remote mirror, snapshot, and copy management, and supports Windows, Linux, UNIX, and IBM z Systems® data.

Spectrum Control enables multi-platform storage virtualization, and data and storage management. It supports most storage systems and devices by using the Storage Networking Industry Association (SNIA) Storage Management Initiative Specification (SMI-S), versions 1.0.2, 1.1, and 1.5 and later.

Hardware and software interoperability information is provided on the IBM Support Portal for Spectrum Control. The interoperability matrix can be found at:

http://www.ibm.com/support/docview.wss?uid=swg27047049

Advanced Edition enables you to adapt to the dynamic storage needs of your applications by providing storage virtualization, automation, and integration for cloud environments with features that include the following:

- OpenStack cloud application provisioning: Advanced Edition includes an OpenStack Cinder volume driver that enables automated provisioning using any of the heterogeneous storage systems that are controlled by IBM Cloud Orchestrator or Virtual Storage Center. OpenStack cloud applications can access multiple storage tiers and services without adding complexity.
- Self-service portal: Advanced Edition can provide provisioning automation for self-service storage portals, which enables immediate responses to service requests while eliminating manual administration tasks.
- Pay-per-use invoicing: Advanced Edition now includes a native chargeback tool. This tool allows customers to create chargeback or showback reports from the native GUI, or work with more advanced reporting as part of the embedded Cognos engine that is also included for building custom reports.

IBM Cognos-based reporting helps create and integrate custom reports on capacity, performance, and utilization. Spectrum Control provides better reporting and analytics with no additional cost through integration with Cognos reporting and modeling. Some reporting is included. Novice users can rapidly create reports with the intuitive drag function. Data abstraction and ad hoc reporting makes it easy to create high-quality reports and charts. You can easily change the scaling and select sections for both reporting and charting. Reports can be generated on schedule or on demand in multiple distribution formats, including email.

Spectrum Control provides better user management and integration with external user repositories, like Microsoft Active Directory. Enhanced management for virtual environments provides enhanced reporting for virtual servers (VMware). Tiered Storage Optimization provides integration with the existing storage optimizer and storage tiering reporting. Tiered Storage Optimization is policy-driven information lifecycle management (ILM) that uses virtualization technology to provide recommendations for storage relocation. It provides recommendations for workload migration based on user-defined policy that is based on file system level data, performance, and capacity utilization. This feature ensures that only the highest performing workloads are allocated to the most expensive storage.

### Spectrum Control in an OpenStack environment

The Spectrum Control OpenStack Cinder driver enables your OpenStack powered cloud environment to use your Spectrum Control installation for block storage provisioning.

Spectrum Control provides block storage provisioning capabilities that a storage administrator can use to define the properties and characteristics of storage volumes within a particular service class. For example, a block storage service class can define RAID levels, tiers of storage, and various other storage characteristics.

### Spectrum Control in a VMware environment

The Spectrum Control supports the following functions in VMware:

- Both act as a control plane function
- Decision around where you are delivering storage services like provisioning versus just performance monitoring
- Base as the control plane aggregation layer

## 4.2.2 IBM Virtual Storage Center

Organizations need to spend less of their IT budgets on storage capacity and storage administration so that they can spend more on new, revenue-generating initiatives. Virtual Storage Center (VSC) delivers an end-to-end view of storage with the ability to virtualize Fibre Channel block storage infrastructures, helping you manage your data with more confidence with improved storage utilization and management efficiency. It combines IBM Spectrum Control Advanced features with IBM Spectrum Virtualize capabilities to deliver an integrated infrastructure to transform your block storage into an agile, efficient, and economical business resource.

IBM Virtual Storage Center is a virtualization platform and a management solution for cloud-based and software-defined storage. It is an offering that combines both IBM Spectrum Control Advanced Edition with IBM Spectrum Virtualize, including SAN Volume Controller, members of the IBM Storwize family, and FlashSystem V9000. VSC helps organizations transition to new workloads and update storage infrastructures. It enables organizations to monitor, automate, and analyze storage. It delivers provisioning, capacity management, storage tier optimization, and reporting. VSC helps standardize processes without replacing existing storage systems, and can also significantly reduce IT costs by making storage more user and application oriented.

Cloud computing is all about agility. Storage for clouds needs to be as flexible and service-oriented as the applications it supports. IBM Virtual Storage Center can virtualize existing storage into a private storage cloud with no “rip and replace” required.

## 4.2.3 IBM Spectrum Control Storage Insights

IBM Spectrum Control Storage Insights is an analytics-driven, storage resource management solution that is delivered over the cloud. The solution uses cloud technology to provide visibility into on-premises storage with the goal of helping clients optimize their storage environments in today's data-intense world. This software as a solution (SaaS) solution running on SoftLayer can deploy in as little as 5 minutes and show actionable insights in 30 minutes.

Storage Insights is a cloud data and storage management service that is deployed in a secure and reliable cloud infrastructure that provides the following features:

- Accurately identify and categorize storage assets
- Monitor capacity and performance from the storage consumer's view, including server, application, and department-level views
- Increase capacity forecasting precision by using historical growth metrics
- Reclaim unused storage to delay future purchases and improve utilization
- Optimize data placement based on historical usage patterns that can help lower the cost

Figure 4-6 shows an example of the Storage Insights dashboard.

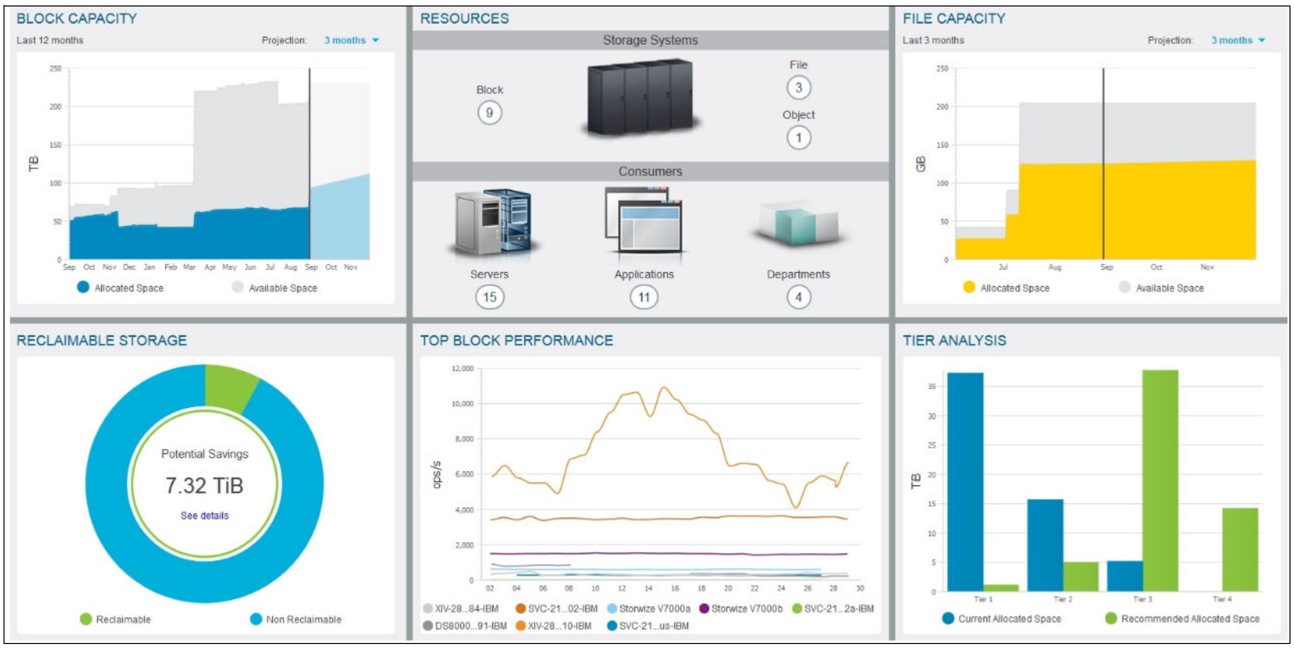



*Figure 4-6 Storage Insights dashboard*

For more information, see these websites about IBM Storage Insights:

http://www.ibm.com/systems/storage/spectrum/insights
http://www.ibm.com/marketplace/cloud/analytics-driven-data-management/us/en-us

## 4.2.4 IBM Copy Services Manager

IBM CSM replication management tool set (formerly in IBM Tivoli Storage Productivity Center) is included in IBM Spectrum Control. This replication management solution delivers central control of your replication environment by using simplified and automated complex replication tasks. Using the CSM functions within Spectrum Control, you can coordinate copy services on IBM Storage, including DS8000, DS6000TM, SAN Volume Controller, Storwize V7000, Spectrum Accelerate, and XIV. You can also help prevent errors and increase system continuity by using source and target volume matching, site awareness, disaster recovery testing, and standby management. Copy services include IBM FlashCopy, Metro Mirror, Global Mirror, and Metro Global Mirror.

You can use Copy Services Manager to complete the following data replication tasks and help reduce the downtime of critical applications:

- Plan for replication when you are provisioning storage
- Keep data on multiple related volumes consistent across storage systems during a planned or unplanned outage
- Monitor and track replication operations
- Automate the mapping of source volumes to target volumes
- Practice disaster recovery procedures

The IBM Copy Services Manager family of products consists of the following products:

- Copy Services Manager provides high availability and disaster recovery for multiple sites
- Copy Services Manager for z Systems provides high availability and disaster recovery for multiple sites
- Copy Services Manager Basic Edition for z Systems provides high availability for a single site if a disk storage system failure occurs

Figure 4-7 shows the Copy Services Manager overview window.

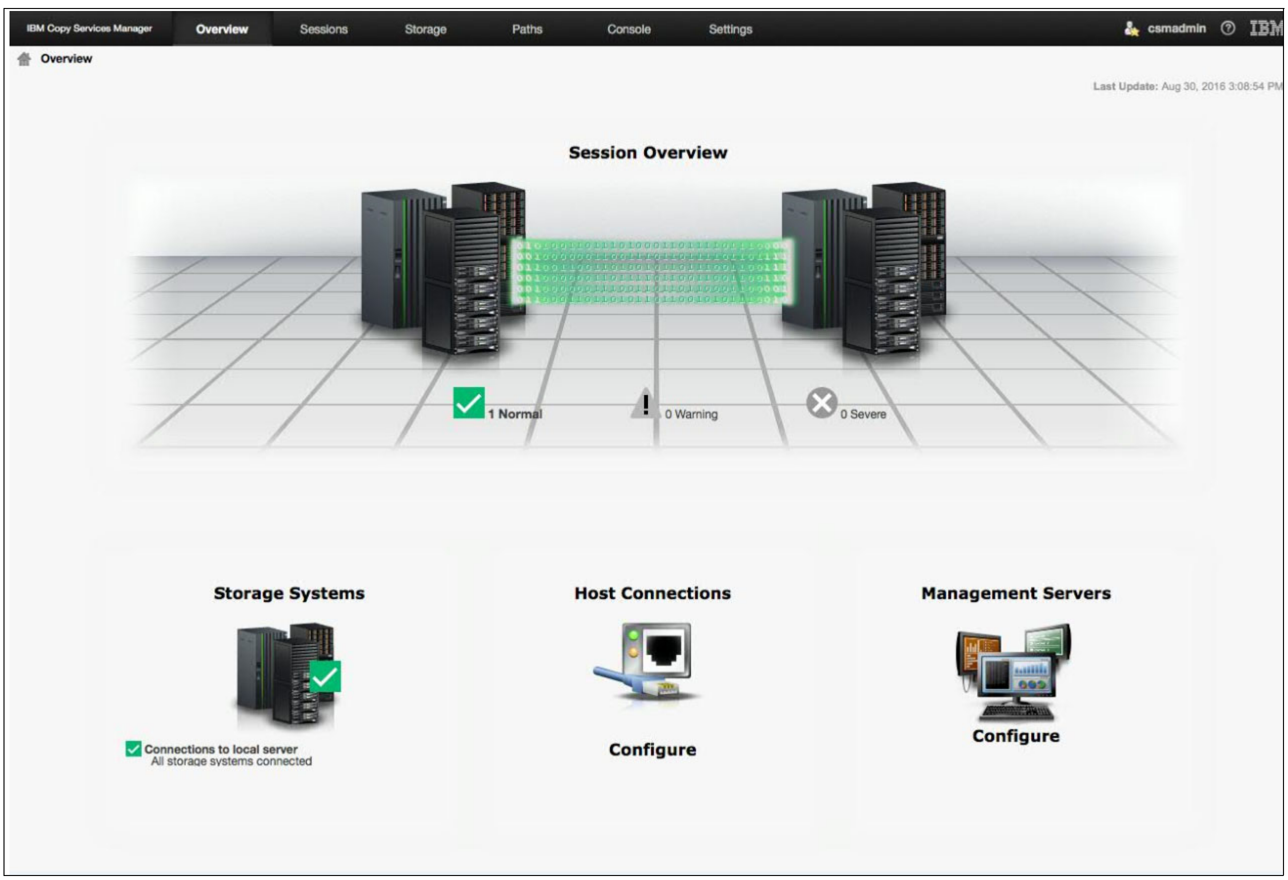



*Figure 4-7 Copy Services Manager Overview window*

## 4.2.5 IBM Spectrum Protect

IBM Spectrum Protect is an intuitive, intelligent, and transparent software that provides a set of product features that allow you to design adaptive and comprehensive data protection solutions. It is a comprehensive data protection and recovery solution for virtual, physical, and cloud data. IBM Spectrum Protect provides backup, snapshot, archive, recovery, space management, bare machine recovery, and disaster recovery capabilities.

### Key capabilities

Here are a few capability highlights:

- Protects virtual, physical, and cloud data with one solution
- Reduces backup and recovery infrastructure costs
- Delivers greater visualization and administrator productivity
- Simplifies backups by consolidating administration tasks
- Space Management moves less active data to less expensive storage, such as tape or cloud
- Provides long-term data archive for data retention, such as for compliance with government regulations

### Benefits

The following are highlights of the benefits of IBM Spectrum Protect:

- Application-aware and VM-aware data protection for any size organization
- Simplified administration
- Built-in efficiency features: Data deduplication, incremental 'forever' backup
- Integrated multi-site replication and disaster recovery
- Multi-site data availability with active-active replication-based architecture and heterogeneous storage flexibility using disk, tape, or cloud

Whatever your data type and infrastructure size, IBM Spectrum Protect scales from a small environment, consisting of 10 to 20 machines to a large environment with thousands of machines to protect. The software product consists of two basic functional components:

- IBM Spectrum Protect server with IBM DB2® database engine

  The IBM Spectrum Protect server provides backup, archive, and space management services to the IBM Spectrum Protect clients, and manages the storage repository. The storage repository can be implemented in a hierarchy of storage pools using any combination of supported media and storage devices. These devices must be directly connected to the IBM Spectrum Protect server system or be accessible through a SAN.

- IBM Spectrum Protect clients with application programming interfaces (APIs)

IBM Spectrum Protect enables data protection from failures and other errors by storing backup, archive, space management, and "bare-metal" restore data, and also compliance and disaster-recovery data in a hierarchy of auxiliary storage. IBM Spectrum Protect can help protect computers that run various operating systems, on various hardware platforms and connected together through the Internet, wide area networks (WANs), local area networks (LANs), or storage area networks (SANs). It uses web-based management, intelligent data move-and-store techniques, and comprehensive policy-based automation that work together to increase data protection and potentially decrease time and administration costs.

The progressive incremental methods that are used by IBM Spectrum Protect back up only new or changed versions of files, greatly reducing data redundancy, network bandwidth, and storage pool consumption as compared to traditional methods.

### Backup and recovery

Despite rapid data growth, data protection and retention systems are expected to maintain service levels and data governance policies. Data has become integral to business decision-making and basic operations, from production to sales and customer management. Data protection and retention are core capabilities for their role in risk mitigation and for the amount of data involved.

The storage environment offers three functions that improve the efficiency and effectiveness of data protection and retention:

- Backup and recovery: Provides cost-effective and efficient backup and restore capabilities, improving the performance, reliability, and recovery of data that is aligned to business required service levels. Backups protect current data, and are unlikely to be accessed unless data is lost or corrupted.
- Archiving: Stores data that has long-term data retention requirements, either for compliance or business purposes, by providing secure and cost effective solutions with automated process for retention policies and data migration to different storage media.
- Continuous data availability: Ensures uninterrupted access to data for critical business systems, reducing the risk of downtime by providing the capability to fail over transparently and as instantaneously as possible to an active copy of the data. The total mirroring strategy needs to be automated to ensure automated failover and then an appropriate automated fail-back.

Optimizing all of these areas helps an organization deliver better services with reduced application downtime. Data protection and retention, archiving, and continuous data availability can improve business agility by ensuring that applications have the correct data when needed, while inactive data is stored in the correct places for the correct length of time. This method requires that the data protection functions must be application aware.

### Tool set

IBM Spectrum Protect is a family of tools that helps manage and control the “information explosion” by delivering a single point of control and administration for storage management needs. It provides a wide range of data protection, recovery management, movement, retention, reporting, and monitoring capabilities by using policy-based automation.

> **Products:** For an updated list of the available products in the IBM Spectrum Protect family, see the following website:
>
> `http://www.ibm.com/software/products/en/spectrum-protect`
>
> See the IBM Spectrum Protect Knowledge Center for information about the most recent releases:
>
> `https://www.ibm.com/support/knowledgecenter/SSGSG7/landing/welcome_ssgsg7.html`

Table 4-2 lists the main features, functions, and benefits offered by the IBM Spectrum Protect family.

*Table 4-2 Main features, functions, and benefits of IBM Spectrum Protect*

| Feature | Function | Benefits |
|---|---|---|
| Backup and recovery management | Intelligent backups and restores using a progressive incremental backup and restore strategy, where only new and used files are backed up | Centralized protection based on smart-move and smart-store technology, which leads to faster backups and restores with fewer network and storage resources needed |
| Hierarchical storage management | Policy-based management of file backup and archiving | Ability to automate critical processes related to the media on which data is stored while reducing storage media and administrative costs associated with managing data |
| Archive management | Managed archives | Ability to easily protect and manage documents that need to be kept for a designated length of time |
| Advanced data reduction | Combines incremental backup, source inline, and target data deduplication, compression, and tape management to provide data reduction | Reduces the costs of data storage, environmental requirements, and administration |

## IBM Spectrum Protect Operations Center

IBM Spectrum Protect Operations Center is a graphical user interface (GUI), with new features (as shown in Figure 4-8). It provides an advanced visualization dashboard, built-in analytics, and integrated workflow automation features that dramatically simplify backup administration.

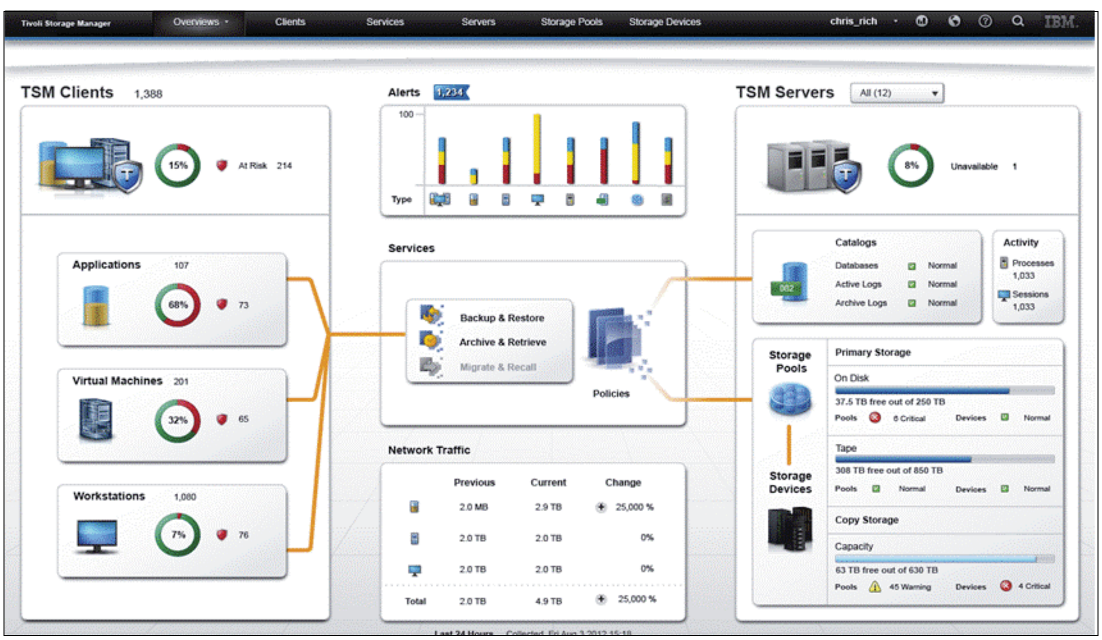



*Figure 4-8 IBM Spectrum Protect Operations Center*

### IBM Spectrum Protect cloud architectures

IBM Spectrum Protect has multiple cloud architectures to meet various requirements. Figure 4-9 shows several IBM Spectrum Protect cloud architectures for storing IBM Spectrum Protect cloud-container storage pools.

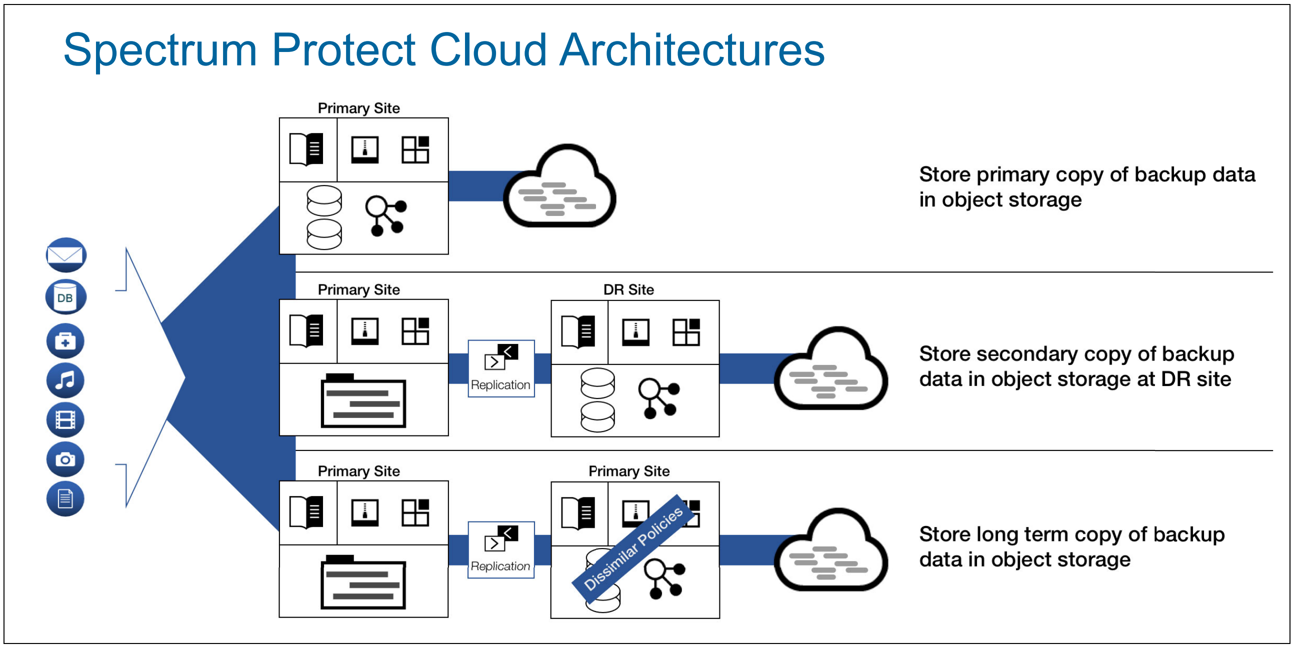



*Figure 4-9 IBM Spectrum Protect Cloud Architectures*

IBM Spectrum Protect supports the following cloud providers:

- IBM SoftLayer
- OpenStack Swift with Keystone Versions 1 and 2

IBM Spectrum Protect also supports IBM Cloud Object Storage (Cleversafe®) dsNet as a storage system within Amazon Simple Storage Service (S3) protocol. In addition, IBM Spectrum Protect lets you configure cloud-container storage pools to use Amazon Web Services (AWS) using the Amazon S3 protocol.

Also, IBM Spectrum Protect can protect data that is hosted in an OpenStack environment, and can use the OpenStack (Swift) environment as a repository for backup and archive objects. The IBM Spectrum Protect cloud architecture in Figure 4-10 shows an IBM Spectrum Protect agent that is deployed within a VM guest.

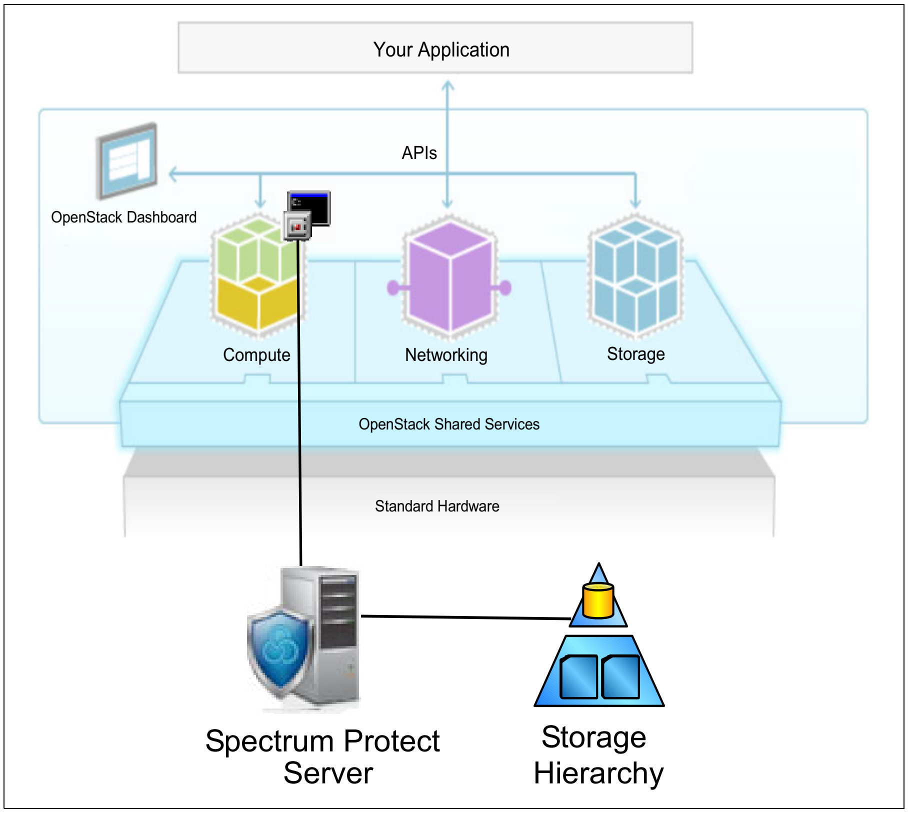



*Figure 4-10 IBM Spectrum Protect protecting data hosted in an OpenStack environment*

### *Data privacy considerations*

Although the security of sensitive data is always a concern, data that you store off-premises in a cloud computing system should be considered particularly vulnerable. Data can be intercepted during transmission, or a weakness of the cloud computing system might be used to gain access to the data.

To guard against these threats, define a cloud-container storage pool to be encrypted. When you do, the server encrypts data before it is sent to the storage pool. After data is retrieved from the storage pool, the server decrypts it so is understandable and usable again. Your data is protected from eavesdropping and unauthorized access when it is outside your network because it can be understood only when it is back on premises.

## IBM Spectrum Protect for Virtual Environments

IBM Spectrum Protect for Virtual Environments simplifies data protection for virtual and cloud environments. It protects VMware and Microsoft Hyper-V virtual machines by offloading backup workloads to a centralized IBM Spectrum Protect server for safe keeping. Administrators can create backup policies or restore virtual machines with just a few clicks.

IBM Spectrum Protect for Virtual Environments enables your organization to protect data without the need for a traditional backup window. It allows you to reliably and confidently safeguard the massive amounts of information that virtual machines generate.

IBM Spectrum Protect for Virtual Environments provides these benefits:

- Improves efficiency with data deduplication, incremental “forever” backup, and other advanced IBM technology to help reduce costs
- Simplifies backups and restores for VMware with an easy-to-use interface that you can access from within VMware vCenter or vCloud Director
- Enables VMware vCloud Director and OpenStack cloud backups
- Enables faster, more frequent snapshots for your most critical virtual machines
- Flexible recovery and copy options from image-level backups give you the ability to perform recovery at the file, mailbox, database object, volume, or VM image level by using a single backup of a VMware image
- Eliminates processor usage caused by optimized virtual machine backup by supporting VMware vStorage APIs for Data Protection and Microsoft Hyper-V technology, which simplify and optimize data protection

### 4.2.6 IBM Spectrum Protect Snapshot

In today’s business world, where application servers are operational 24 hours a day, the data on these servers must be fully protected. You cannot afford to lose any data, but you also cannot afford to stop these critical systems for hours so you can protect the data adequately. As the amount of data that needs protecting continues to grow exponentially and the need to keep the downtime associated with backup to an absolute minimum, IT processes are at their breaking point. Data volume snapshot technologies such as IBM Spectrum Protect Snapshot can help minimize the effect caused by backups and provide near instant restore capabilities.

Although many storage systems are now equipped with volume snapshot tools, 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 extra snapshot process to ensure that all parts of a data transaction are flushed from memory and committed to disk before the snapshot. This process is necessary to ensure that you have a usable, consistent copy of the data.

IBM Spectrum Protect Snapshot helps deliver the highest levels of protection for mission-critical IBM DB2, SAP, Oracle, Microsoft Exchange, and Microsoft SQL Server applications using integrated, application-aware snapshot backup and restore capabilities. This protection is achieved by using advanced IBM storage hardware snapshot technology to create a high performance, low impact application data protection solution.

The snapshots captured by IBM Spectrum Protect Snapshot can be retained as backups on local disk. With optional integration with IBM Spectrum Protect, customers can use the full range of advanced data protection and data reduction capabilities such as data deduplication, progressive incremental backup, hierarchical storage management, and centrally managed policy-based administration as shown in Figure 4-11.

Because a snapshot operation typically takes much less time than the time for a tape backup, the window during which the application must be aware of a backup can be reduced. This advantage 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.

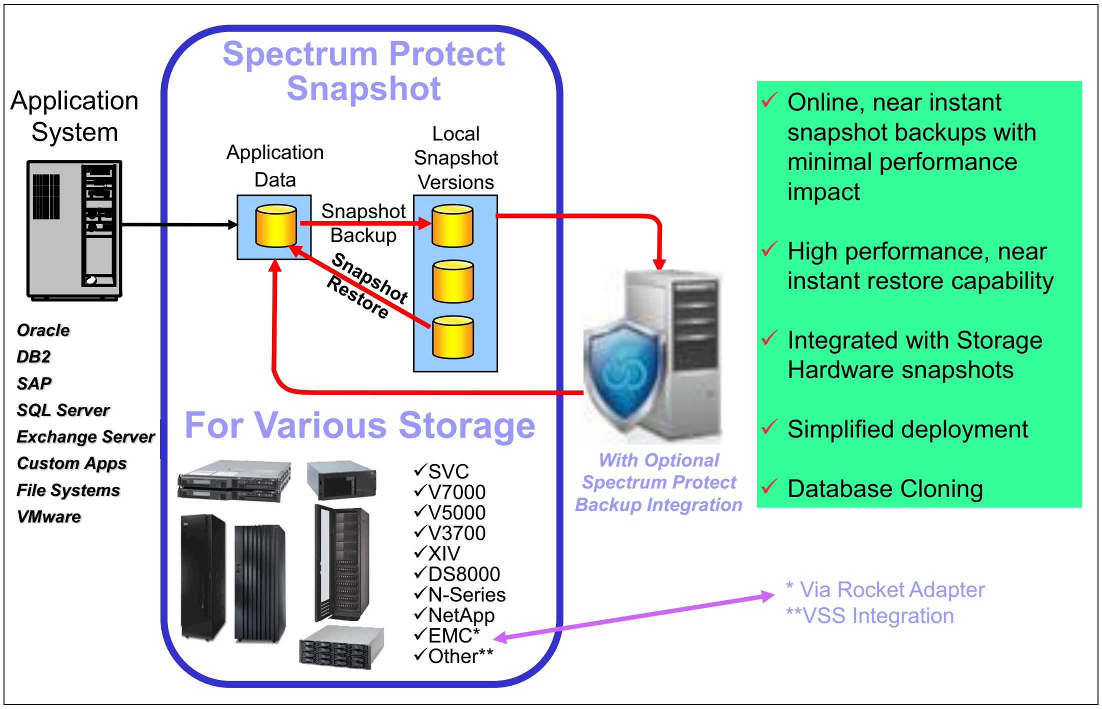



*Figure 4-11 IBM Spectrum Protect Snapshot storage snapshot capabilities*

Application availability is also significantly improved due to the reduction of the load on the production servers. IBM Spectrum Protect Snapshot uses storage snapshot capabilities to provide high speed, low impact, application-integrated backup and restore functions for the supported application 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 way to use and quick to deploy data protection solution that enables the most stringent database recovery time requirements to be met.

For more information, see the following website:

http://www.rocketsoftware.com/resource/rocket-device-adapter-pack-ibm-tivoli-flash copy-manager-overview

## 4.2.7 IBM Spectrum Copy Data Management

IBM Spectrum Copy Data Management makes copies available to data consumers when and where they need them, without creating unnecessary copies or leaving unused copies on valuable storage. It catalogs copy data from across local, hybrid cloud, and off-site cloud infrastructure, identifies duplicates, and compares copy requests to existing copies. This process ensures that the minimum number of copies are created to service business requirements. Data consumers can use the self-service portal to create the copies they need when they need them, creating business agility. Copy processes and work flows are automated to ensure consistency and reduce complexity. IBM Spectrum Copy Data Management rapidly deploys as an agentless VM as it helps manage snapshot and FlashCopy images made to support DevOps, data protection, disaster recovery, and Hybrid Cloud computing environments.

This member of the IBM Spectrum Storage family automates the creation and catalogs the copy data on existing storage infrastructure, such as snapshots, vaults, clones, and replicas. One of the key use cases centers around use with Oracle, Microsoft SQL server, and other databases that are often copied to support application development, testing, and data protection.

The IBM Spectrum Copy Data Management software is an IT modernization technology that focuses on using existing data in a manner that is efficient, automated, scalable, and easy to use to improve data access. IBM Spectrum Copy Data Management (Figure 4-12), with IBM storage arrays, delivers in-place copy data management that modernizes IT processes and enables key use cases with existing infrastructure.

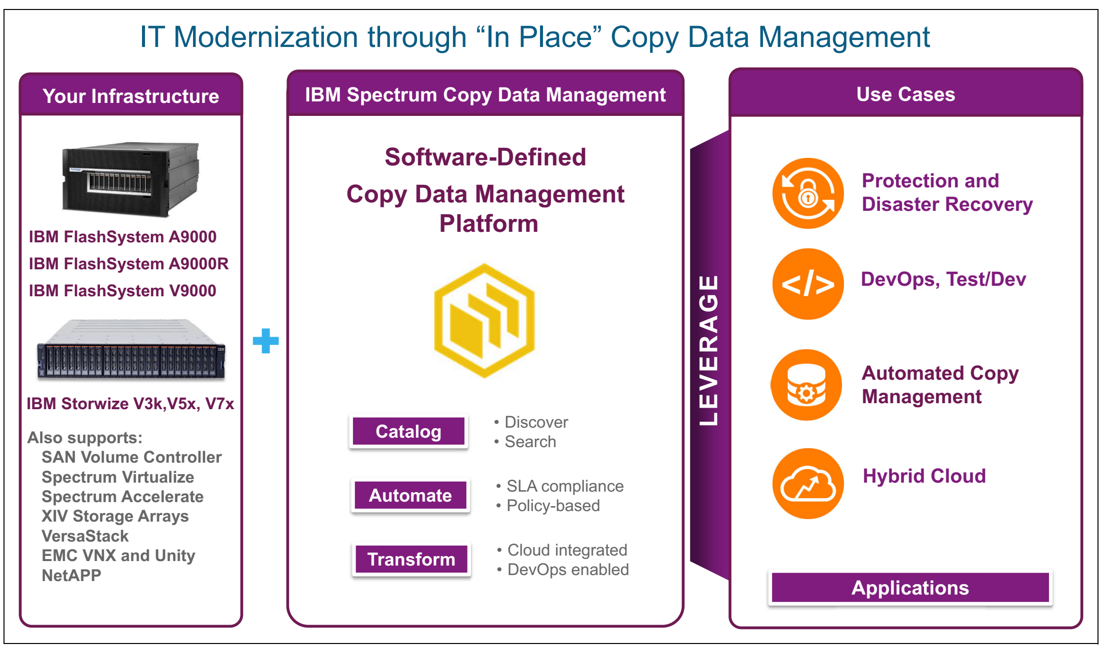



*Figure 4-12 Software-Defined IBM Spectrum Copy Data Management Platform*

IBM Spectrum Copy Data Management includes support for the following copy data management use cases:

- Automated Copy management
- Development and operations
- Data protection and disaster recovery
- Test and development
- Hybrid cloud computing

### Automated copy management

IT functions that rely heavily on copies or ‘snapshots’ are typically managed by using a complex mix of scripts, tools, and other products, none of which are optimized for copy management. With IBM Spectrum Copy Data Management, organizations have a holistic, simplified approach that greatly reduces cycle time and frees staff to manage more productive projects.

IT teams can use the core policy engine, catalog, and reporting of IBM Spectrum Copy Data Management to dramatically improve IT operations that rely on copies of data, including disaster recovery, testing and development, business analytics, and local recovery. IBM Spectrum Copy Data Management improves operations by using automated, service-level based copy policies that are consistent, reliable, and easily repeatable. This feature provides huge savings in operating expenses.

### Development and operations (DevOps)

Organizations are increasingly moving toward DevOps for faster delivery of new applications to market. IBM Spectrum Copy Data Management enables IT teams to use their existing storage infrastructure to enable DevOps, helping to meet the needs of the development teams for rapid deployment of the infrastructure. IBM Spectrum Copy Data Management templates define the policies for infrastructure deployment. The whole system is accessible through the REST API. Rather than following legacy processes to requisition IT resources, developers include the infrastructure deployment commands directly within their development systems, such as Chef, Puppet, or IBM Bluemix. Predefined scripts and plug-ins for popular DevOps tools simplify implementation.

### Next-generation data protection and disaster recovery

Through its template-based management and orchestration of application-aware copies, IBM Spectrum Copy Data Management can support next-generation data protection and recovery workflows. IBM Spectrum Copy Data Management enables IT to mount and instantly access copies that are already in the production storage environment. IBM Spectrum Copy Data Management catalogs all snapshots and replicas, and alerts you if a snap or replication job was missed or failed. Disaster recovery can be fully automated and tested nondisruptively. In addition, IBM Spectrum Copy Data Management can coordinate sending data through the AWS Storage Gateway to the AWS storage infrastructure. This feature provides a simplified, low-cost option for longer term or archival storage of protection copies.

### Automated test and development

The speed and effectiveness of test and development processes are most often limited by the time it takes to provision IT infrastructure. With IBM Spectrum Copy Data Management, test and development infrastructure can be spun-up in minutes, either on an automated, scheduled basis, or on-demand basis.

### Hybrid cloud computing

IBM Spectrum Copy Data Management is a powerful enabler of the hybrid cloud, enabling IT to take advantage of cloud compute resources. IBM Spectrum Copy Data Management not only helps customers move data to the cloud, it also enables IT organizations to create live application environments that can use the less expensive, elastic compute infrastructure in the cloud. Being able to spin up workloads and then spin them back down reliably helps maximize the economic benefit of the cloud by only using and paying for infrastructure as needed.

IBM Spectrum Copy Data Management is a software platform that is designed to use the existing infrastructure in the IT environment. It works directly with hypervisor and enterprise storage APIs to provide the overall orchestration layer that uses the copy services of the underlying infrastructure resources. IBM Spectrum Copy Data Management also integrates with AWS S3 for cloud-based data retention, as well as Puppet, IBM Bluemix, and others.

### Database-specific functionality

IBM Spectrum Copy Data Management allows the IT team to easily create and share copies of all popular database management systems by integrating key database management system (DBMS) tasks within well-defined policies and work-flows. The solution also includes application-aware integration for Oracle and Microsoft SQL Server platforms, providing a deeper level of coordination with the DBMS.

### Secure multi-tenancy

Secure multi- tenancy meets the needs of both managed service providers and large organizations that need to delegate resources internally. Individual “tenants” can be created within a single IBM Spectrum Copy Data Management instance, allowing each tenant its own set of resources and the ability to support administration within the tenancy to create users, define jobs, and perform other functions.

### Policy templates for automation and self-service

Template-based provisioning and copy management provides easy self- service access for internal customers to request the resources that they need, when they need them. Templates are pre- defined by the IT team, and they are accessible through a self- service portal interface or through API calls.

### Compatibility

IBM Spectrum Copy Data Management is a simple-to-deploy software platform that is designed to use the existing IT infrastructure. It works directly with hypervisor and storage APIs to provide the overall orchestration layer that uses the copy services of the underlying infrastructure resources. It also integrates with Amazon Web Services S3 for cloud-based data retention.

IBM Spectrum Copy Data Management delivers the following benefits:

- Automate the creation and use of copy data on existing storage infrastructure, such as snapshots, vaults, clones, and replicas
- Reduce time that is spent on infrastructure management while improving reliability
- Modernize existing IT resources by providing automation, user self-service, and API-based operations without the need for any additional hardware
- Simplify management of critical IT functions such as data protection and disaster recovery
- Automate test and development infrastructure provisioning, drastically reducing management time
- Drive new, high-value use cases, such as using hybrid cloud compute
- Catalog and track IT objects, including volumes, snapshots, virtual machines, data stores, and files

## 4.3 SDS Data Plane

The data plane encompasses the infrastructure where data is processed. It consists of all basic storage management functions such as virtualization, RAID protection, tiering, copy services (remote, local, synchronous, asynchronous, and point-in-time), encryption, and data deduplication that can be started and managed by the control plane. The data plane is the interface to the hardware infrastructure where the data is stored. It provides a complete range of data access possibilities, spanning traditional access methods like block I/O (for example, iSCSI) and File I/O (POSIX compliant), to object-storage and Hadoop Distributed File System (HDFS).

Block, file, and object are simply different approaches to accessing data. Figure 4-13 is a high-level view of these differences.

The section describes the following IBM software product offerings (organized by block, file, and object support) that provide the building blocks for the SDS data plane:

- IBM Spectrum Virtualize
- IBM Spectrum Accelerate
- IBM Spectrum Scale
- IBM Spectrum Archive™

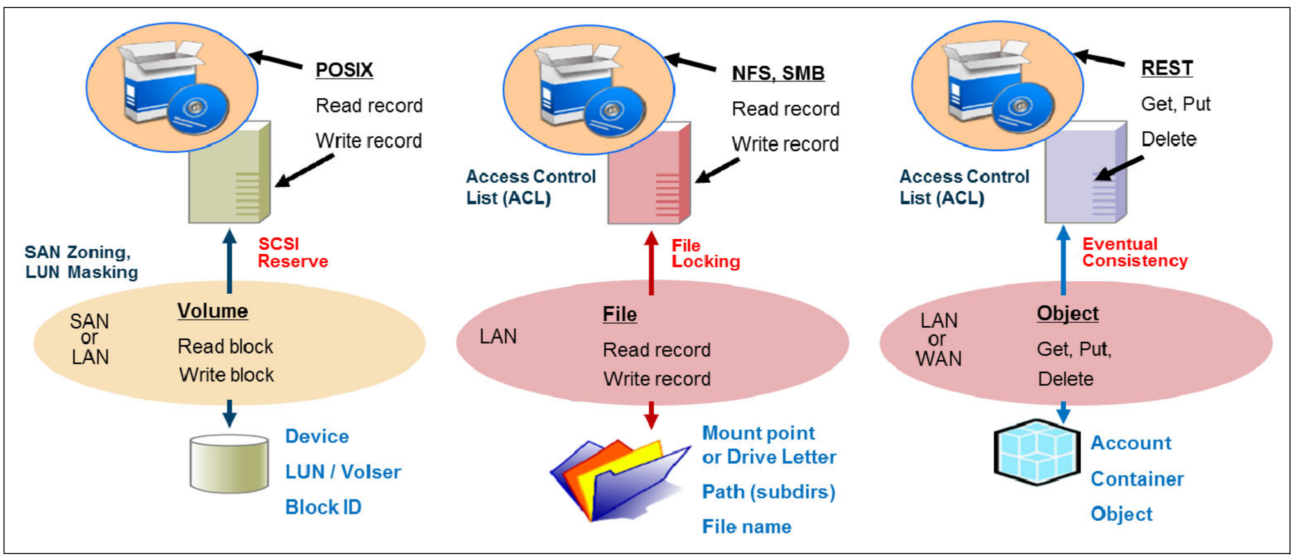



*Figure 4-13 High-level view of data access differences between file, block, and object storage*

### 4.3.1 Block Storage

Block storage offerings are differentiated by speed/throughput (as measured in IOPS) and segmented by lifecycle of the disk. Data is split into evenly sized chunks or blocks of data, each with its own unique address. How blocks of data are accessed is up to the application. Few applications access the blocks directly. Rather, a Portable Operating System Interface (POSIX) file system is used in a hierarchical way of organizing files so that an individual file can be located by describing the path to that file.

## 4.3.2 File Storage

File storage uses protocols to access individual directories and files over various protocols, including NFS, CIFS/SMB, and so on. Certain file attributes can describe a file and its contents, such as its owner, who can access the file, and its size. This metadata is stored along with the data or related directory structure.

## 4.3.3 Object storage

With object storage, data is written into self-contained entities called objects. Unlike file systems, an object storage system gives each object a unique ID, which is managed in a flat index. There are no folders and subfolders. Unlike files, objects are created, retrieved, deleted or replaced in their entirety, rather than being updated or appended in place.

Object storage also introduces the concept of "eventual consistency." If one user creates an object, a second user might not see that object listed immediately. Eventually, all users will be able to see the object.

When a user or application needs access to an object, the object storage system is provided with a unique ID. This flat index approach provides greater scalability, enabling an object storage system to support faster access to a massively higher quantity of objects or files as compared to traditional file systems.

## 4.3.4 IBM block storage solutions

This section describes IBM block storage solutions.

### IBM Spectrum Virtualize

IBM Spectrum Virtualize software is at the heart of IBM SAN Volume Controller, IBM Storwize family, IBM FlashSystem® V9000, and VersaStack. It enables these systems to deliver better data value, security, and simplicity through industry-leading virtualization. This virtualization transforms existing and new storage and streamlines deployment for a simpler, more responsive, scalable, and cost efficient IT infrastructure.

IBM Spectrum Virtualize systems provide management of storage from entry and midrange up to enterprise disk systems, and enable hosts to attach through SAN, FCoE, or iSCSI to existing Ethernet networks. IBM Spectrum Virtualize is easy to use, enabling existing staff to start working with it rapidly. IBM Spectrum Virtualize uses virtualization, thin provisioning, and compression technologies to improve storage utilization and meet changing needs quickly and easily. In this way, Spectrum Virtualize products are the ideal complement to server virtualization strategies.

### Key Capabilities

IBM Spectrum Virtualize software capabilities are offered across various platforms, including SAN Volume Controller (SVC), Storwize V7000 (Unified), Storwize V5000, and FlashSystem V9000. IBM Spectrum Virtualize products are designed to deliver the benefits of storage virtualization and advanced storage capabilities in environments from large enterprises to small businesses and midmarket companies:

- IBM Real-time Compression™ for inline, real-time compression
- Stretched Cluster and IBM HyperSwap® for a high-availability solution
- IBM Easy Tier® for automatic and dynamic data tiering
- Distributed RAID for better availability and faster rebuild times

- Encryption for internal and external virtualized capacities
- FlashCopy snapshots
- Remote data replication

## Benefits

The sophisticated virtualization, management, and functions of Spectrum Virtualize provide these storage benefits:

- Improves storage utilization up to 2x
- Supports up to 5x as much data in the same physical space
- Simplifies management of heterogeneous storage systems
- Enables rapid deployment of new storage technologies for greater ROI
- Improves application availability with virtually zero storage-related outages

The SAN Volume Controller combines software and hardware into a comprehensive, modular appliance that uses symmetric virtualization.

Symmetric virtualization is achieved by creating a pool of managed disks (MDisks) from the attached storage systems. Those storage systems are then mapped to a set of volumes for use by the attached host systems. System administrators can view and access a common pool of storage on the SAN. This function helps administrators to use storage resources more efficiently and provides a common base for advanced functions.

Figure 4-14 shows the Spectrum Virtualize functions.

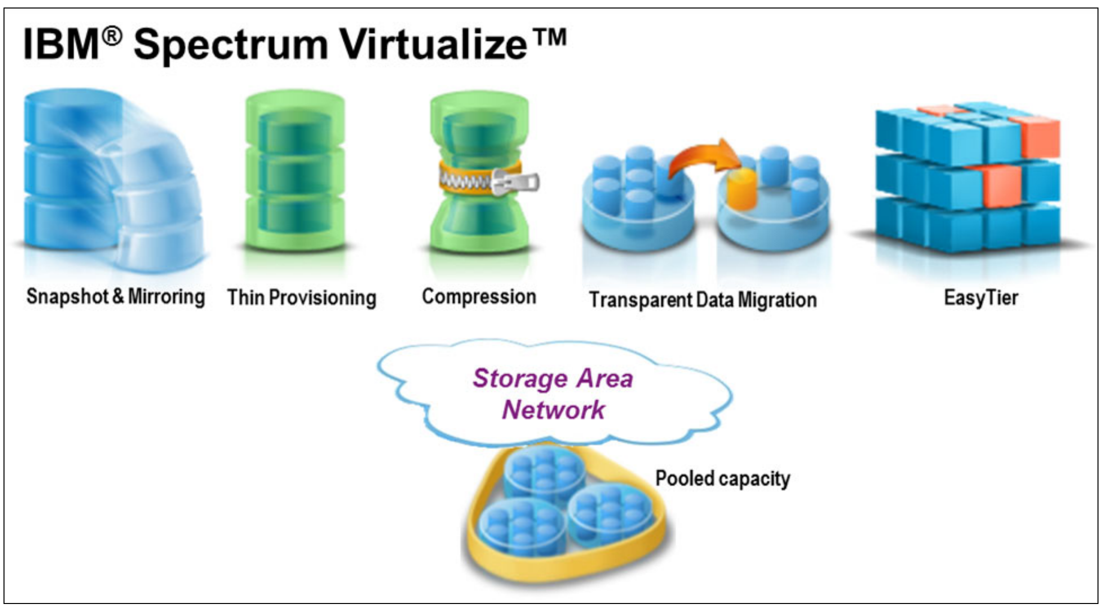



*Figure 4-14 IBM Spectrum Virtualize functions*

Table 4-3 describes Spectrum Virtualize features and benefits.

*Table 4-3 Spectrum Virtualize features and benefits*

| Feature | Benefits |
|---|---|
| Single point of control for storage resources | ► Designed to increase management efficiency<br>► Designed to help support business application availability |
| Pools the storage capacity of multiple storage systems on a SAN | ► Helps you manage storage as a resource to meet business requirements and not just as a set of boxes<br>► Helps administrators better deploy storage as required beyond traditional “SAN islands”<br>► Can help increase use of storage assets<br>► Insulates applications from physical changes to the storage infrastructure |
| Clustered pairs of IBM SAN Volume Controller data engines | ► Highly reliable hardware foundation<br>► Designed to avoid single points of hardware failure |
| IBM Real-time Compression | ► Increases effective capacity of storage systems up to five times, helping to lower costs, floor-space requirements, and power and cooling needs<br>► Can be used with a wide range of data, including active primary data, for dramatic savings<br>► Hardware compression acceleration helps transform the economics of data storage |
| Innovative and tightly integrated support for flash storage | ► Designed to deliver ultra-high performance capability for critical application data<br>► Move data to and from flash storage without disruption; make copies of data onto hard disk drive (HDD) |
| Support for IBM FlashSystem | Enables high performance for critical applications with IBM MicroLatency®, coupled with sophisticated functions |
| Easy-to-use IBM Storwize family management interface | ► Single interface for storage configuration, management, and service tasks regardless of storage vendor<br>► Helps administrators use their existing storage assets more efficiently |
| IBM Storage Mobile Dashboard | Provides basic monitoring capabilities to securely check system health and performance |
| Dynamic data migration | ► Migrate data among devices without taking applications that are using that data offline<br>► Manage and scale storage capacity without disrupting applications |
| Manage tiered storage | Helps balance performance needs against infrastructure costs in a tiered storage environment |
| Advanced network-based copy services | ► Copy data across multiple storage systems with IBM FlashCopy<br>► Copy data across metropolitan and global distances as needed to create high-availability storage solutions |
| Integrated Bridgeworks SANrockIT technology for IP replication | ► Optimize use of network bandwidth<br>► Reduce network costs or speed replication cycles, improving the accuracy of remote data |

| Feature | Benefits |
|---|---|
| Enhanced stretch cluster configurations | ► Provide highly available, concurrent access to a single copy of data from data centers up to 300 km apart<br>► Enable nondisruptive storage and virtual machine mobility between data centers |
| Thin provisioning and “snapshot” replication | ► Dramatically reduce physical storage requirements by using physical storage only when data changes<br>► Improve storage administrator productivity through automated on-demand storage provisioning |
| Hardware snapshots integrated with IBM Spectrum Protect Snapshot Manager | ► Performs near-instant application-aware snapshot backups, with minimal performance impact for IBM DB2, Oracle, SAP, VMware, Microsoft SQL Server, and Microsoft Exchange<br>► Provides advanced, granular restoration of Microsoft Exchange data |

Virtualizing storage with SAN Volume Controller helps make new and existing heterogeneous storage arrays more effective by including many functions that are traditionally deployed within disk array systems. By including these functions in a virtualization system, SAN Volume Controller standardizes functions across virtualized storage for greater flexibility and potentially lower costs.

Figure 4-15 shows how SAN Volume Controller stretch virtual volume standardizes heterogeneous storage across data centers.

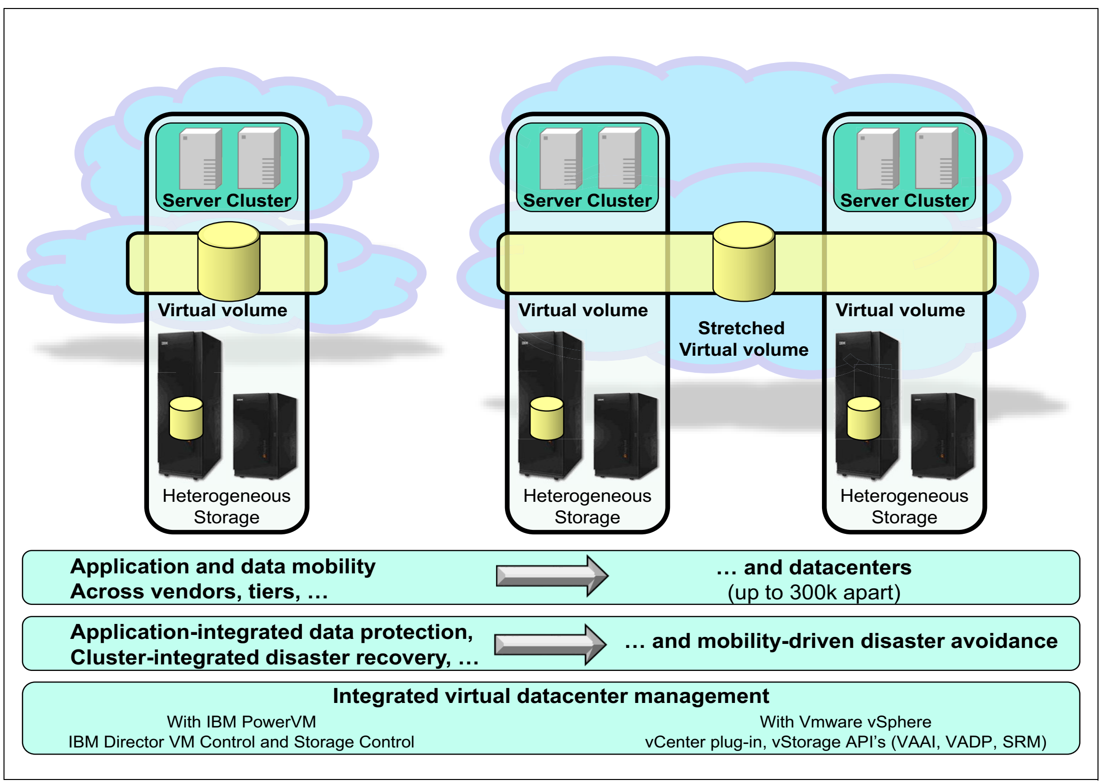



*Figure 4-15 Stretching virtual volume across data centers with heterogeneous storage*

SAN Volume Controller functions benefit all virtualized storage. For example, IBM Easy Tier optimizes use of flash memory, and Real-time Compression enhances efficiency even further by enabling the storage of up to five times as much active primary data in the same physical disk space[1]. Finally, high-performance thin provisioning helps automate provisioning. These benefits can help extend the useful life of existing storage assets, reducing costs.

Integrating these functions into SAN Volume Controller also means that they are designed to operate smoothly together, reducing management effort:

- Storage virtualization: Virtualization is a foundational technology for software defined infrastructures that enables software configuration of the storage infrastructure. Without virtualization, networked storage capacity utilization averages about 50 percent, depending on the operating platform. Virtualized storage enables up to 90 percent utilization by enabling pooling across storage networks with online data migration for capacity load balancing. Virtual Storage Center supports a virtualization of storage resources from multiple storage systems and vendors (that is, heterogeneous storage). Pooling storage devices enables access to capacity from any networked storage system, which is a significant advantage over the limitations inherent in traditional storage arrays.
- IBM Easy Tier: Virtual Storage Center helps optimize flash memory with automated tiering for critical workloads. Easy Tier helps make the best use of available storage resources by automatically moving the most active data to the fastest storage tier, which helps applications and virtual desktop environments run up to three times faster.
- Thin provisioning: Thin provisioning helps automate provisioning and improve productivity by enabling administrators to focus on overall storage deployment and utilization, and on longer-term strategic requirements, without being distracted by routine storage-provisioning requests.
- Remote mirroring: IBM Metro Mirror and Global Mirror functions automatically copy data to remote sites as it changes, enabling fast failover and recovery. These capabilities are integrated into the advanced GUI, making them easy to deploy.
- IBM Real-time Compression: Real-time Compression is patented technology that is designed to reduce space requirements for active primary data. It enables users to store up to five times as much data in the same physical disk space, and can do so without affecting performance.

## IBM Spectrum Accelerate

IBM Spectrum Accelerate is a highly flexible, software-defined storage solution that enables rapid deployment of block data storage services for new and traditional workloads, both on and off premises. It is a key member of the IBM Spectrum Storage portfolio. Spectrum Accelerate allows you to run hotspot-free, grid-scale software that runs on the XIV Storage System Gen3 enterprise storage platform in your existing data center infrastructure or in a cloud provider such as IBM SoftLayer. It offers proven grid- scale technology, mature features, and ease of use, and is already deployed on over 100,000 servers worldwide.

IBM Spectrum Accelerate delivers predictable, consistent storage performance, management scaling to more than 68 petabytes usable, and a rich feature set that includes remote mirroring and granular multi-tenancy. It deploys on premises on x86 commodity servers and on the optimized XIV Storage System, and off-premises as a public cloud service on SoftLayer. You can manage all your Spectrum Accelerate instances, wherever they are deployed, in a single, intuitive interface. Hardware-independent, transferable licensing offers superb operational flexibility and cost benefits.

[1] Compression data based on IBM measurements. Compression rates vary by data type and content.

IBM Spectrum Accelerate also allows customers to deploy a Hyperconverged solution across their on-premises and off-premises deployments to help meet the unpredictability of today's cloud world. It runs as a virtual machine on the VMware vSphere ESXi hypervisor. It converges compute and storage, enabling customer-built, hyper -converged solutions based on proven XIV technology. The ability to run application workload VMs on the same servers as the storage enables customers to rapidly provision and decommission workloads in a dynamic fashion.

Spectrum Accelerate delivers a single management experience across software-defined storage infrastructure using IBMs HyperScale Manager, which can manage both Spectrum Accelerate instances, IBM XIV, and the IBM A9000 all flash solution. This combination helps cut costs through reduced administration effort and training, reduces procurement costs, standardizes data center storage hardware operations and services, and provides licensing flexibility that enables cost-efficient cloud building.

Figure 4-16 shows how straightforward scaling is by building a storage grid with Spectrum Accelerate.

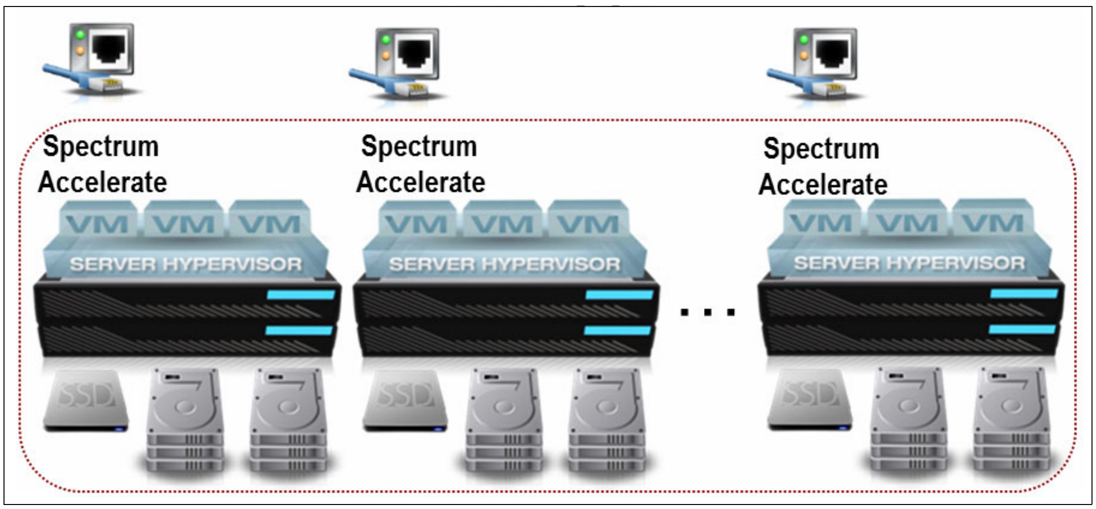



*Figure 4-16 Spectrum Accelerate iSCSI storage grid*

## Key capabilities

Spectrum Accelerate gives organizations these capabilities:

- Enterprise cloud storage in minutes, using commodity hardware
- Hotspot-free performance and QoS without any manual or background tuning needed
- Advanced remote replication, role-based security, and multi-tenancy
- Deploy on-premises or on the cloud (also as a service on SoftLayer)
- Hyper-scale management of dozens of petabytes
- Best in class VMware and OpenStack integration
- Run IBM Spectrum Accelerate and other application virtual machines on the same server

Spectrum Accelerate runs as a virtual machine on vSphere ESXi hypervisor, enabling you to build a server-based SAN from commodity hardware that includes x86 servers, Ethernet switches, solid-state drives (SSDs), and direct-attached, high-density disks. Spectrum Accelerate essentially acts as an operating system for your self-built SAN storage, grouping virtual nodes and spreading the data across the entire grid.

IBM Spectrum Accelerate release 11.5.3 manages up to 15 nodes in a grid. It provides a single point of management of up to 144 grids connected through Hyper-Scale Manager, up to 2,160 nodes.

Spectrum Accelerate allows you to deploy storage services flexibly across different delivery models, including customer-choice hardware, existing data center infrastructure, and IBM storage systems.

Spectrum Accelerate includes these benefits:

- Cost reduction by delivering hotspot-free storage to different deployment models on- and off-premises, enabling organizations to pay less overall for the same capacity by optimizing utilization, acquiring less hardware, and minimize administrative overhead
- Increased operational agility through easy cloud building, faster provisioning, small capacity increments, and flexible, transferable licensing
- Rapid response through enterprise-class storage availability, data protection, and security for the needs of new and traditional workload in the data center and other sites, while flexibly balancing capital and operational expenses
- Scaleout across 144 virtual systems and seamless management across IBM Spectrum Accelerate instances on- and off-premises and the XIV storage system

Table 4-4 describes Spectrum Accelerate features with their associated benefits.

*Table 4-4 Spectrum Accelerate features and benefits*

| Feature | Benefit |
|---|---|
| Performance | ► Ensures even data distribution through massive parallelism and automatic load balancing including upon capacity add<br>► Distributed cache |
| Reliability and Availability | ► Grid redundancy maintains two copies of each 1-MB data partition with each copy being on a different VM, proactive diagnostics, fast and automatic rebuilds, event externalization<br>► Advanced monitoring; network monitoring; disk performance tracking/reporting; data center monitoring; shared monitoring for some components; data and graphical reports on I/O, usage, and trends<br>► Self-healing, whcih minimizes the rebuild process by rebuilding only actual data<br>► Automated load balancing across components; minimized risk of disk failure due to rapid return to redundancy |
| Management | Intuitive GUI: Scales to up to 144 virtual arrays and up to more than 45 PB with IBM Hyper-Scale Manager; extensive CLI; RESTful API; mobile app support with push notifications; multi-tenancy with quality of service by tenant, pool, or host |
| Cloud automation and Self-service | OpenStack; VMware vRealize Orchestrator through IBM Spectrum Control Base |
| Snapshot management | Space efficient snapshots: Writable, snapshot of snapshot, restore from snapshot, snapshots for consistency groups, mirroring |
| Thin provisioning; space reclamation | Thin provisioning per pool, thick-to-thin migration; VMware, Microsoft, Symantec space reclamation support |
| Mirroring | Synchronous/asynchronous; volumes and consistency groups, recovery point objective (RPO) of seconds; online/offline initialization; failover/failback; mirroring across platforms including with XIV Storage System |

| Feature | Benefit |
| --- | --- |
| Security | Role-based access management, multi-tenancy, iSCSI Challenge Handshake Authentication Protocol (CHAP) and auditing; integrates with Lightweight Directory Access Protocol (LDAP) and Microsoft Active Directory servers |

### *OpenStack device support for IBM XIV*

IBM has built and contributed the OpenStack Cinder block storage driver for XIV to the OpenStack community. This driver allows Spectrum Accelerate to be the first enterprise class storage system to have OpenStack software support. This feature allows ease of use and fast time to implementation characteristics to be magnified by being able to be automatically managed and provisioned within the OpenStack environment.

The IBM Storage Driver for OpenStack Cinder component added support starting with the Folsom release as shown in Figure 4-17, and then expanded the support for the Grizzly and Havana releases. The driver enables OpenStack clouds to be able to directly access and use IBM Spectrum Accelerate Storage System Gen3.

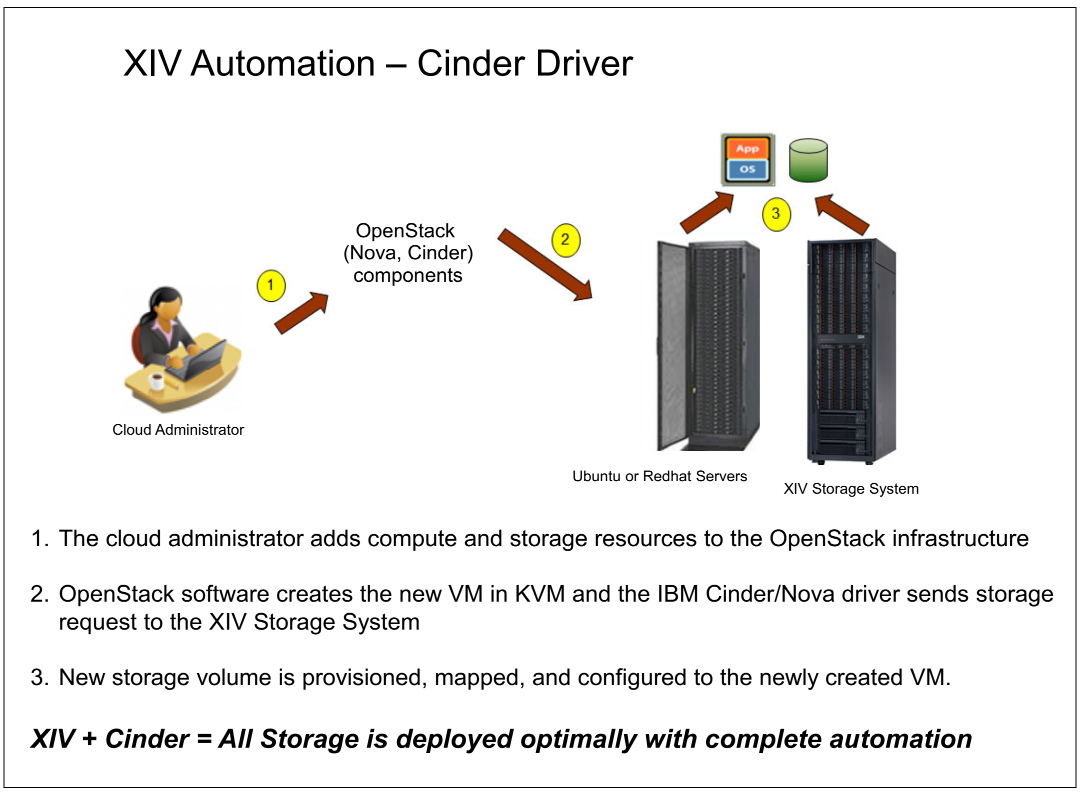



*Figure 4-17 OpenStack Cinder support for XIV*

### *Hyperconverged Flexible deployment*

IBM Spectrum Accelerate provides customers the capability to create Hyperconverged solutions running both Compute and the Storage services on the same physical x86 servers wherever they are deployed. Using the VMware ESX hypervisor, additional resources that are not used by the Spectrum Accelerate instances, such as memory and processors, can be provisioned to additional guest workloads. The Spectrum Accelerate instance can be managed and administered from its native GUI or through the IBM HyperScale Manager option. This capability includes actions such as creating pools, replication structures, hardware component replacement, and firmware updates. See Figure 4-18.

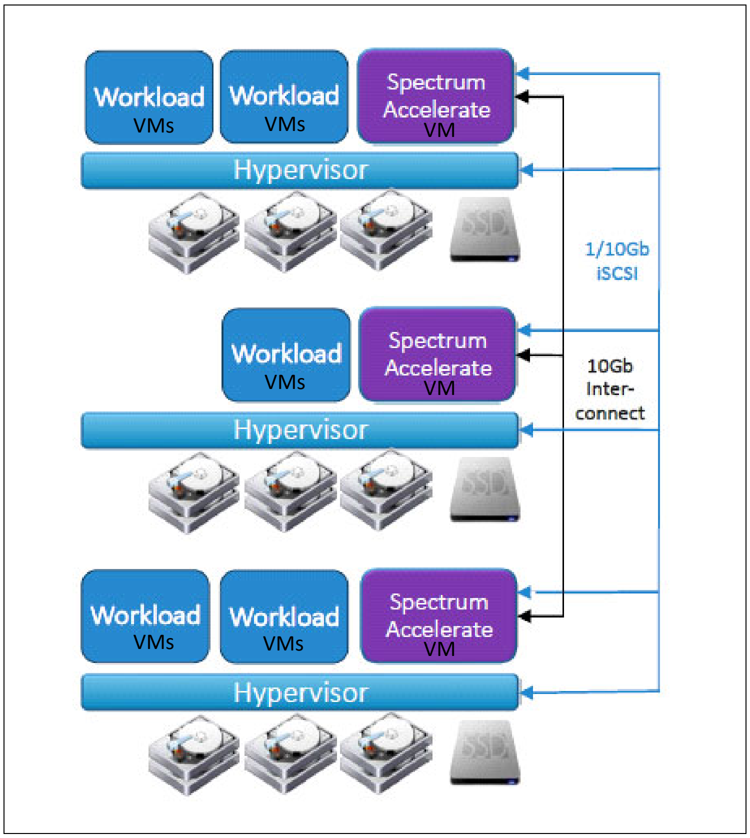



*Figure 4-18 Spectrum Accelerate in a Hyperconverged infrastructure*

### *Hyperconverged Orchestration*

Using IBM Spectrum Control Base (see Figure 4-5 on page 66, IBM Storage VMware Integration), orchestration of the compute layer, provisioning from predefined Spectrum Accelerate pools, and replication can all be done through the VMware integration points. Therefore, the Hyperconverged solution can be controlled through the vRealize suite and can use vCenter and VMware SRM-based APIs.

### *Spectrum Accelerate as a pre-configured Hyperconverged Solution*

Customers looking to benefit from Spectrum Accelerate's Hyperconverged capabilities and its ability to work as a storage system that can replicate to the IBM XIV, but want the ability to deploy it as a pre-integrated solution can order it from Supermicro. IBM and Supermicro have jointly designed an appliance deliverable that combines IBM Spectrum Accelerate software with Supermicro hardware. This product is delivered as a pre-configured, preinstalled, and pre-tested solution that is ready to be integrated into customer networks. The three basic building blocks are *Small*, *Medium*, and *Large*, and those building blocks can be customized for based on customer-specific requirements.

For more information, you can visit:

https://www.supermicro.com/solutions/spectrum-accelerate.cfm

### IBM XIV Storage System Gen3

IBM Spectrum Accelerate is the common software defined layer inside the IBM XIV Storage System Gen3.

> **Note:** For more information about Spectrum Accelerate, see these IBM publications:
>
> - *IBM Spectrum Accelerate Deployment, Usage, and Maintenance*, SG24-8267
> - *Deploying IBM Spectrum Accelerate on Cloud*, REDP-5261
> - *IBM Spectrum Accelerate Reference Architecture*, REDP-5260

### IBM FlashSystem A9000 and A9000R

IBM Spectrum Accelerate is the common software defined layer across the IBM FlashSystem A9000 and A9000R all-flash arrays.

> **Note:** For more information about IBM FlashSystem A9000 and A9000R, see these IBM publications:
>
> - *IBM FlashSystem A9000 and IBM FlashSystem A9000R Architecture, Implementation, and Usage*, SG24-8345
> - *IBM FlashSystem A9000 Product Guide*, REDP-5325

## 4.3.5 IBM File Storage Solutions

This section describes IBM file storage solutions.

### IBM Spectrum Scale

IBM Spectrum Scale is a proven, scalable, high-performance file management solution that is based on IBM's General Parallel File System (GPFS). IBM Spectrum Scale provides world-class storage management with extreme scalability, flash accelerated performance, and automatic policy-based storage tiering from flash to disk, then to tape. IBM Spectrum Scale reduces storage costs up to 90% while improving security and management efficiency in cloud, big data, and analytics environments.

First introduced in 1998, this mature technology enables a maximum volume size of 8 YB, a maximum file size of 8 EB, and up to 18.4 quintillion (two to the 64th power) files per file system. IBM Spectrum Scale provides simplified data management and integrated information lifecycle tools such as software-defined storage for cloud, big data, and analytics. It introduces enhanced security, flash accelerated performance, and improved usability. It also provides capacity quotas, access control lists (ACLs), and a powerful snapshot function.

## Key capabilities

IBM Spectrum Scale adds elasticity with the following capabilities:

- Global namespace with high-performance access scales from departmental to global
- Automated tiering, data lifecycle management from flash (6x acceleration) to tape (10x savings)
- Enterprise ready with data security (encryption), availability, reliability, large scale
- POSIX compliant
- Integrated with OpenStack components and Hadoop

## Benefits

IBM Spectrum Scale provides these benefits:

- Improves performance by removing data-related bottlenecks
- Automated tiering, data lifecycle management from flash (acceleration) to tape (savings)
- Enables sharing of data across multiple applications
- Reduces cost per performance by placing data on most applicable storage (flash to tape or cloud)

IBM Spectrum Scale is part of the IBM market-leading software-defined storage family:

- As a Software-only solution: Runs on virtually any hardware platform and supports almost any block storage device. IBM Spectrum Scale runs on Linux (including Linux on IBM z Systems), IBM AIX®, and Windows systems.
- As an integrated IBM Elastic Storage™ Server solution: A bundled hardware, software, and services offering that includes installation and ease of management with a graphical user interface. Elastic Storage Server provides unsurpassed end-to-end data availability, reliability, and integrity with unique technologies that include IBM Spectrum Scale RAID.
- As a cloud service: IBM Spectrum Scale delivered as a service provides high performance, scalable storage, and integrated data governance for managing large amounts of data and files in the IBM SoftLayer cloud.

IBM Spectrum Scale features enhanced security with native encryption and secure erase. It can increase performance by using server-side flash cache to increase I/O performance up to six times. IBM Spectrum Scale provides improved usability through data replication capabilities, data migration capabilities, Active File Management (AFM), transparent cloud tiering (TCT), File Placement Optimizer (FPO), and IBM Spectrum Scale Native RAID.

Figure 4-19 shows an example of the Spectrum Scale architecture.

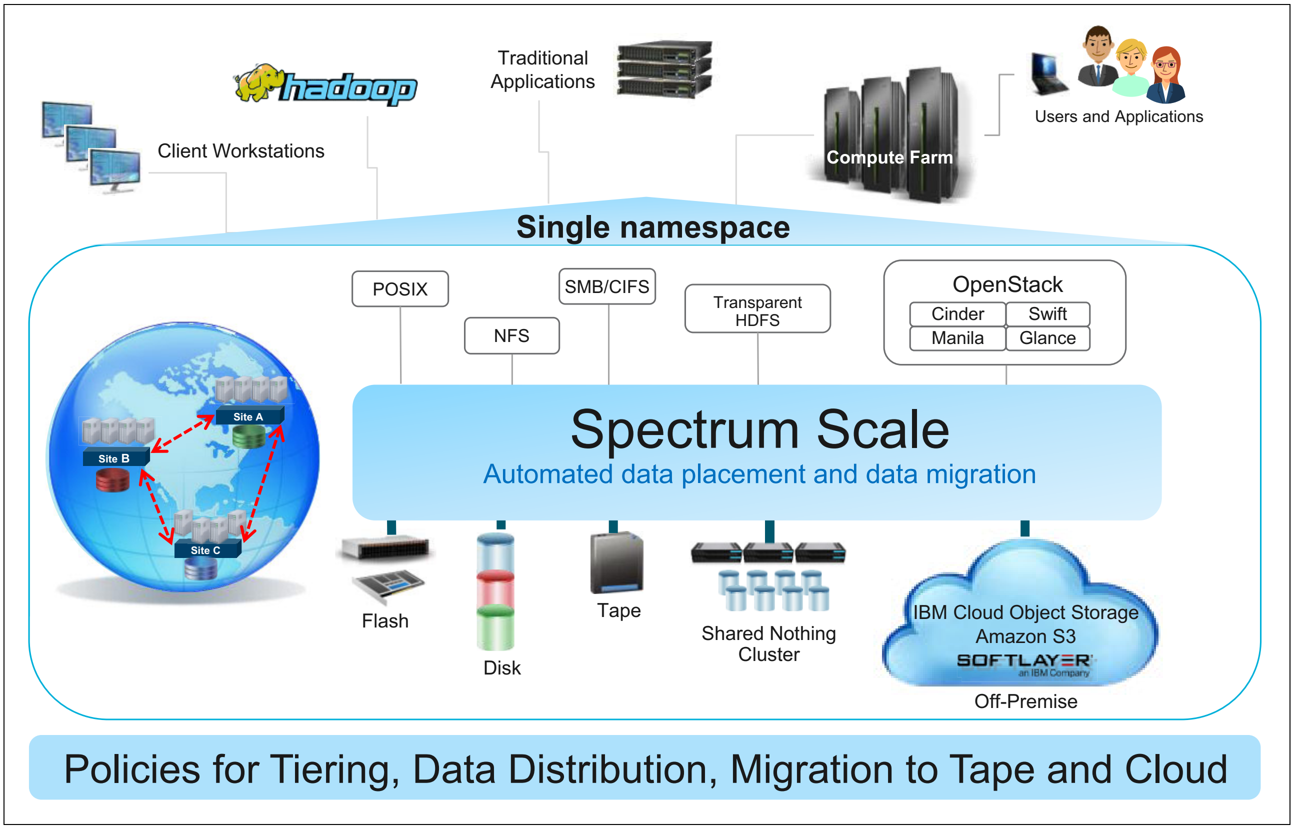



*Figure 4-19 Spectrum Scale architecture*

IBM Spectrum Scale is based around the following concepts:

- Storage pools
- File sets
- Policy engine
- Mirroring, replication, and migration capabilities
- Active File Management
- File Placement Optimizer
- Licensing

### *Storage pools*

A storage pool is a collection of disks or arrays with similar attributes. It is an organizational structure that allows the combination of multiple storage locations with identical characteristics. There are three different types of storage pools:

**System Pool** One system pool is needed per file system. The system pool contains file system metadata and can be used to store data.

**Data Pool** A data pool is used to store file data. A data pool is optional.

**External Pool** An external pool is used to attach auxiliary storage, such as tape to IBM Spectrum Scale. An external pool is optional.

### *File sets*

IBM Spectrum Scale creates a single name space, so there are tools that provide a fine grained management of the directory structure. A file set acts like a partition of a file system, a subdirectory tree. File sets can be used for operations such as quotas or used in management policies. It is a directory tree that behaves like a “file system” within a file system:

- Part of the global namespace.
- Can be linked and unlinked (like mount/unmount).
- Policy scan can be restricted to only scan file sets. This setting can be helpful when the file system has billions of files.
- A file set can be assigned to a storage pool.

There are two kinds of file sets:

**Dependent file sets** A dependent file set allows for a finer granularity of administration. It shares the inode space with another file set.

**Independent file set** An independent file set has a distinct inode space. An independent file set allows file set level snapshots, independent file scans, and enabled advanced features like AFM.

### *Policy engine*

The policy engine uses an SQL style syntax to query or operate on files based on file attributes. Policies can be used to migrate all data that has not been accessed in 6 months (for example) to less expensive storage or used to query the contents of a file system. Management policies support advanced query capabilities, though what makes the policy engine most useful is the performance. The policy engine is capable of scanning billions of objects as shown in Table 4-5.

*Table 4-5 Speed comparison for GPFS policy engine*

| **Search through** 1000000000 (1 billion) files | |
|---|---|
| `find` | ~ 47 hours |
| GPFS policy engine | ~ 5 hours |

Table 4-5 shows the power of the GPFS policy engine. Although an average **`find`** across 1 billion files took ~ 47 hours, the GPFS policy engine can satisfy the request within five hours. The GPFS policy engine can also create a candidate list for backup applications to use to achieve a massive reduction in candidate identification time.

IBM Spectrum Scale has next generation availability with features that include rolling software and hardware upgrades. You can add and remove servers to adapt the performance and capacity of the system to changing needs. Storage can be added or replaced online, and you can control how data is balanced after storage is assessed.

### *Mirroring, replication, and migration capabilities*

In IBM Spectrum Scale, you can replicate a single file, a set of files, or the entire file system. You can also change the replication status of a file at any time by using a policy or command. Using these capabilities, you can achieve a replication factor of two, which equals mirroring, or a replication factor of three.

A replication factor of two in Spectrum Scale means that each block of a replicated file is in at least two failure groups. A failure group is defined by the administrator and contains one or more disks. Each storage pool in a file system contains one or more failure groups. Failure groups are defined by the administrator and can be changed at any time. So when a file system is fully replicated, any single failure group can fail and the data remains online.

For migration, IBM Spectrum Scale provides the capability to add storage to the file system, migrate the existing data to the new storage, and remove the old storage from the file system. All of this can be done online without disruption to your business.

### *Active File Management*

AFM enables the sharing of data across unreliable or high latency networks. With AFM, you can create associations between IBM Spectrum Scale clusters and define the location and flow of file data. AFM allows you to implement a single name space view across clusters, between buildings, and around the world.

AFM operates at the file set level. This configuration means that you can create hundreds of AFM relationships in each file system. AFM is a caching technology though inode. File data in a cache file set is the same as an inode and file data in any IBM Spectrum Scale file system. It is a “real” file that is stored on disk. The job of the cache is to keep the data in the file consistent with the data on the other side of the relationship.

AFM can be implemented in five different modes:

- Read-Only (ro)
- Local-Update (lu)
- Single-Writer (sw)
- Independent Writer (iw)
- Asynchronous DR

These modes can be used to collect data at a remote location (single-writer), create a flash cache for heavily read data (read-only), provide a development copy of data (local-update), create a global interactive name space (independent-writer), and create asynchronous copies of file data (asynchronous DR).

### *Transparent Cloud Tiering*

Data in the enterprise is growing at an alarming rate led by growth in unstructured data, leading to a capacity crisis. Cooler and cold data constitutes a large proportion of data in the enterprise. Migrating cooler and cold data to lower-cost cloud object storage provides cost savings.

Transparent cloud tiering is a new feature of IBM Spectrum Scale 4.2.1 that provides hybrid cloud storage capability. This software defined capability enables usage of public, private, and on-premises cloud object storage as a secure, reliable, transparent storage tier that is natively integrated with Spectrum Scale without introducing additional hardware appliances or new management touch points. It uses the existing ILM policy language semantics available in IBM Spectrum Scale. the semantics allow administrators to define policies for tiering cooler and cold data to a cloud object storage such as the following:

- IBM Cloud Object Storage (Cleversafe)
- Amazon Web Services S3
- OpenStack Swift

This configuration frees up storage capacity in higher-cost storage tiers that can be used for more active data.

Figure 4-20 highlights the Spectrum Scale transparent cloud tiering feature.

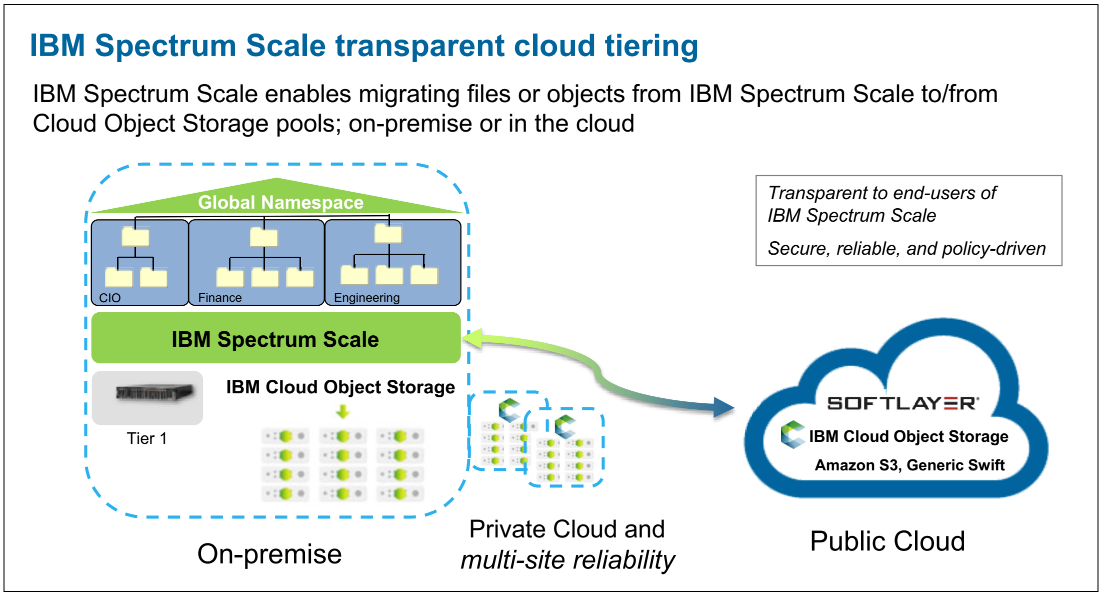



*Figure 4-20 Spectrum Scale transparent cloud tiering feature highlights*

For more information, see *Enabling Hybrid Cloud Storage for IBM Spectrum Scale Using Transparent Cloud Tiering*, REDP-5411.

### *IBM Spectrum Scale Management GUI*

The IBM Spectrum Scale Management GUI (Graphical User Interface) can be used in conjunction with the existing command line interface. The GUI is meant to support common administrator tasks, such as provisioning additional capacity, which can be accomplished faster and without knowledge of the command-line interface. System health, capacity, and performance displays can be used to identify trends and respond quickly to any issues that may arise. The GUI is available to Spectrum Scale Clusters running at or above the 4.2 release for the Standard Edition and Advanced Edition. See Figure 4-21 on page 98.

The IBM Spectrum Scale management GUI provides an easy way to configure and manage various features that are available with the IBM Spectrum Scale system. You can perform the following important tasks through the IBM Spectrum Scale management GUI:

- Monitoring the performance of the system based on various aspects
- Monitoring system health
- Managing file systems
- Creating file sets and snapshots
- Managing Objects and NFS and SMB data exports
- Creating administrative users and defining roles for the users
- Creating object users and defining roles for them
- Defining default, user, group, and file set quotas
- Monitoring the capacity details at various levels such as file system, pools, file sets, users, and user groups


*Figure 4-21 Spectrum Scale management GUI dashboard*

### *IBM DeepFlash 150 for IBM Spectrum Scale*

IBM DeepFlash 150 provides an essential big data building block for petabyte-scale, cost-constrained, high-density, and high-performance storage environments. It delivers the response times of an all-flash array with extraordinarily competitive cost benefits. DeepFlash 150 is an ideal choice to accelerate systems of engagement, unstructured data, big data, and other workloads that require low latency, high performance, and sustained throughput.

The DeepFlash 150 all-flash storage array is a building block for IBM Spectrum Scale infrastructures. It is primarily targeted at big data, media and entertainment streaming, virtual desktop, high-speed database, and hyper-scale environments. DeepFlash 150 provides highly scalable capacity and performance plus highly competitive economics, in both capital expenditures and operational expenditures, surpassing that of conventional enterprise-grade storage systems.

DeepFlash 150 does not use conventional SSDs. Instead, this innovative new storage system relies on a larger, systems-level approach that enables organizations to manage much larger data sets without having to manage individual storage modules. The DeepFlash 150 system comes complete with the hardware necessary for enterprise and hyper-scale storage, including up to 64 purpose-engineered flash cards in a 3U chassis and 12-Gbps SAS connectors for up to eight servers. The flash modules have a capacity of 8 TB each. DeepFlash 150 comes preinstalled with 16, 32 or 64 Board Solid State Drives (BSSDs).

### *File Placement Optimizer*

FPO allows Spectrum Scale to use locally attached disks on a cluster of servers that communicate by using the network, rather than the regular case of using dedicated servers for shared disk access (such as using SAN). Spectrum Scale FPO is suitable for workloads like SAP HANA, and IBM DB2 with Database Partitioning Feature. It can be used as an alternative to Hadoop Distributed File System (HDFS) in big data environments. The use of FPO extends the core Spectrum Scale architecture, providing greater control and flexibility to use data location, reduce hardware costs, and improve I/O performance. The following are some benefits when using FPO:

- Allows your jobs to be scheduled where the data is located (locality awareness)
- Metablocks that allow large and small block sizes to coexist in the same file system
- Write affinity that allows applications to dictate the layout of files on different nodes, maximizing write and read bandwidth
- Pipelined replication to maximize use of network bandwidth for data replication
- Distributed recovery to minimize the effect of failures on ongoing computation

More information about IBM Spectrum Scale FPO can be found in the *GPFS V4.1: Advanced Administration Guide*, SC23-7032.

### *IBM Spectrum Scale Native RAID*

IBM Spectrum Scale Native RAID provides next generation performance and data security. Using IBM Spectrum Scale native RAID, just a bunch of disks (JBOD) are directly attached to the systems running IBM Spectrum Scale software. This technology uses declustered RAID to minimize performance degradation during RAID rebuilds and provides extreme data integrity by using end-to-end checksums and version numbers to detect, locate, and correct silent disk corruption. An advanced disk hospital function automatically addresses storage errors and slow performing drives so that your workload is not affected.

IBM Spectrum Scale native RAID is available with the IBM Power8 architecture in the IBM Elastic Storage Server (ESS) offering.

### *Licensing*

IBM Spectrum Scale V4.2 offers different editions so you only pay for the functions that you need:

- Express Edition contains the base IBM Spectrum Scale function
- Standard Edition includes the base function plus ILM, AFM, and integrated multi-protocol support, which includes NFS, SMB, and Object
- Advanced Edition includes encryption of data at rest, secure erase, asynchronous multisite disaster recovery, and all the features of Standard Edition

For each of these editions, you can choose an IBM Spectrum Scale license for Server, Client, and FPO.

For more information, see:

- IBM Spectrum Scale:

  `http://www.ibm.com/systems/storage/spectrum/scale/index.html`

- IBM Spectrum Scale (IBM Knowledge Center):

  `http://www.ibm.com/support/knowledgecenter/STXKQY/ibmspectrumscale_welcome.html`

- IBM Spectrum Scale Wiki:

  `https://ibm.biz/BdFPR2`

- IBM Elastic Storage Server

  http://www.ibm.com/systems/storage/spectrum/ess

### *IBM Spectrum Scale for Linux on IBM z System*

The IBM Spectrum Scale for Linux on IBM z System implements the IBM Spectrum Scale Software-based delivery model in the Linux on IBM z System environment. The highlights of IBM Spectrum Scale for Linux on IBM z System include the following features:

- Supports both extended count key data (IBM ECKD™) DASD disks and Fibre Channel Protocol attached SCSI disks
- Supports IBM HiperSockets™ for communication within one z System

For more information, see *Getting started with IBM Spectrum Scale for Linux on z Systems*:

http://www.ibm.com/common/ssi/cgi-bin/ssialias?htmlfid=ZSW03272USEN

## IBM Elastic Storage Server

The IBM Elastic Storage Server (ESS) is a modern implementation of software-defined storage, combining IBM Spectrum Scale software with IBM servers and disk arrays. This technology combines the CPU and I/O capability of the IBM POWER8® architecture and matches it with 2U and 4U storage enclosures. This architecture allows the IBM Spectrum Scale Native RAID software capability to actively manage all RAID functionality that is formerly accomplished by a hardware disk controller. RAID rebuild times are reduced to a fraction of the time that is needed with hardware RAID.

The IBM building block solution for high performance storage allows you to accomplish these tasks:

- Deploy petascale class high-speed storage quickly with pre-assembled and optimized servers, storage, and software
- Host multiple tenants, adjust resource allocation, and scale as your needs evolve
- Experience higher performance and scalability with lower cost
- Rebuild failed disks faster with IBM developed de-clustered RAID technology

IBM has implemented Elastic Storage Server configurations for various workloads, from high-velocity import through high-density cloud storage usage models, deploying the latest SSD, serial-attached SCSI (SAS), and Nearline SAS drives. For performance-oriented workloads, the system's affordable building block configurations start at 24 or 48 drives. For high-capacity storage, IBM offers configurations that can support almost 2 petabytes of usable, deployable storage in a single industry-standard 42U rack. For mixed workloads, the server supports varied configurations of building blocks, with placement rules for the creation and management of all data on the appropriate storage tier.

For more information, see these websites:

http://www.ibm.com/systems/storage/spectrum/ess
http://www.ibm.com/support/knowledgecenter/SSYSP8_2.5.0/sts25_welcome.html

Newly developed RAID techniques from IBM use this CPU and I/O power to help overcome the limitations of current disk drive technology.

Elastic Storage Server is a building block that provides these benefits:

- Deploy petascale class high-speed storage quickly with pre-assembled and optimized servers, storage, and software
- Host multiple tenants, adjust resource allocation, and scale as your needs evolve
- Experience higher performance and scalability with lower cost
- Achieve superior sustained streaming performance
- Rebuild failed disks faster with IBM developed de-clustered RAID technology

For more information, see the Elastic Storage Server in the IBM Knowledge Center at:

http://www.ibm.com/support/knowledgecenter/POWER8/p8ehc/p8ehc_storage_landing.htm

Also, see IBM Elastic Storage Server at:

http://www.ibm.com/systems/storage/spectrum/ess/

### IBM DeepFlash Elastic Storage Server

The IBM DeepFlash Elastic Storage Server is a software-defined flash storage solution that is composed of high-density, cost- effective IBM DeepFlash 150 all-flash storage integrated with the IBM Spectrum Scale ESSs massively scalable architecture. IBM DeepFlash 150 provides a big data building block for exabyte- scale, cost- constrained, high- density and high-performance storage environments. An IBM DeepFlash Elastic Storage Server implementation is constructed by deploying multiple IBM Elastic Storage Server nodes. Each node runs IBM Spectrum Scale software, shares storage management duties, and is connected to IBM DeepFlash 150 all-flash storage. This building block or grid architecture approach offers great resiliency to the overall IBM Spectrum Scale implementation, and other advantages such as linear scaling of both capacity and performance as more IBM Elastic Storage Server and IBM DeepFlash 150 building blocks are added.

Built from the ESS DeepFlash Storage drawers and Power ESS data servers, it enables big data storage to dramatically expand scale-out deployments with high density and performance. DeepFlash Elastic Storage Server V5.0 supports 256 TB or 512 TB configurations that use one or two flash storage drawers deployed with Spectrum Scale for DeepFlash Elastic Storage Server. Two ESS Data Servers are connected to this storage through 12 Gb high-speed SAS adapters to provide maximum data throughput. The data servers come preinstalled with Spectrum Scale for DeepFlash Elastic Storage Software to provide a complete solution that has been optimized for workloads requiring low latency, high capacity, and sustained throughput.

ESS DeepFlash (Figure 4-22 on page 102) is designed for hyperscale and cloud workloads that are also evolving and demanding scalability, high levels of availability, and agility in how IT resources are allocated and used. New performance levels bring the advantages of ESS DeepFlash Storage to a wider variety of applications, including high-performance computing workloads.

The offering includes these features:

- Improved performance and rack density
- DeepFlash storage with either 256 TB or 512 TB of flash storage
- IBM PowerLinux servers with up to 20 processor cores and high-performance I/O bandwidth for optimal performance

- IBM Spectrum Scale for IBM DeepFlash Elastic Storage Server, which delivers the industry-leading, software-defined storage solution of IBM Spectrum Scale, as well as IBM Spectrum Scale RAID

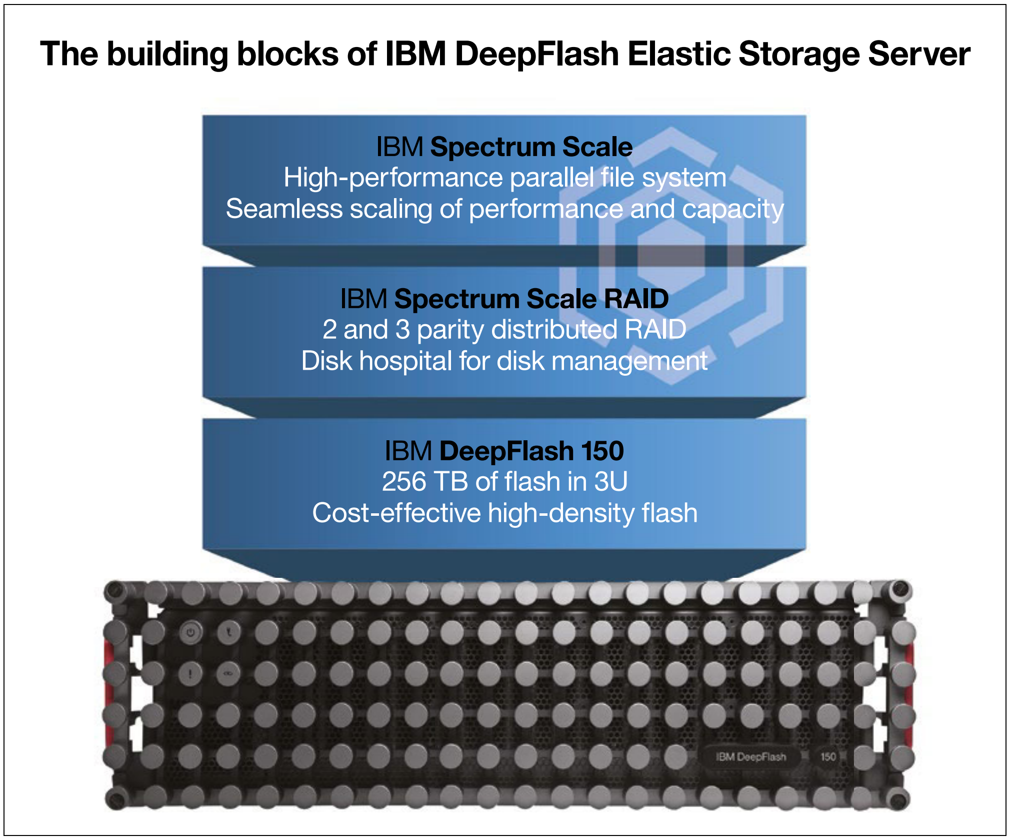



*Figure 4-22 IBM DeepFlash Elastic Storage Server*

### IBM Spectrum Archive

IBM Spectrum Archive, a member of the IBM Spectrum Storage family, enables direct, intuitive, and graphical access to data stored in IBM tape drives and libraries by incorporating the Linear Tape File System (LTFS) format standard for reading, writing, and exchanging descriptive metadata on formatted tape cartridges. Spectrum Archive eliminates the need for extra tape management and software to access data.

Spectrum Archive offers three software solutions for managing your digital files with the LTFS format:

- Single Drive Edition (SDE)
- Library Edition (LE)
- Enterprise Edition (EE)

With Spectrum Archive Enterprise Edition and Spectrum Scale, tape can now add savings as a low-cost storage tape tier. Being able to use a tier of tape for active but “cold” data enables enterprises to look at new ways to cost optimize their unstructured data storage. They are able to match the value of the data, or the value of the copies of data to the most appropriate storage media. In addition, the capability to store the data at the cost of tape storage allows customers to build their cloud environments to take advantage of this new cost structure. Spectrum Archive provides enterprises with the ability to store cold data at costs that can be cheaper than some public cloud provider options. To understand the potential costs with large-scale cold data storage and retention, IBM has created a Tape TCO Calculator, which is available at:

http://www.ibm.com/systems/storage/tape/tco-calculator

Network attached unstructured data storage with native tape support using LTFS delivers the best mix of performance and lowest cost storage.

### Key capabilities

Spectrum Archive options can support small, medium, and enterprise businesses with these advantages:

- Seamless virtualization of storage tiers
- Policy-based placement of data
- Single universal namespace for all file data
- Security and protection of assets
- Open, non-proprietary, cross platform interchange
- Integrated functionality with IBM Spectrum Scale

### Benefits

IBM Spectrum Archive enables direct, intuitive, and graphical access to data stored in IBM tape drives and libraries by incorporating the LTFS format standard for reading, writing, and exchanging descriptive metadata on formatted tape cartridges. Spectrum Archive eliminates the need for additional tape management and software to access data.

Spectrum Archive takes advantage of the low cost of tape storage while making it easy to use. Spectrum Archive provides these benefits:

- Access and manage all data in stand-alone tape environments as easily as though it were on disk
- Enable easy-as-disk access to single or multiple cartridges in a tape library
- Improve efficiency and reduce costs for long-term, tiered storage
- Optimize data placement for cost and performance
- Enable data file sharing without proprietary software
- Scalable and low cost

### Linear Tape File System

IBM developed LTFS and then contributed it to SNIA as an open standard so that all tape vendors can participate. LTFS is the first file system that works with Linear Tape-Open (LTO) generation 7, 6, and 5 tape technology (or IBM TS1150 and TS1140 tape drives) to set a new standard for ease of use and portability for open systems tape storage. With this application, accessing data that is stored on an IBM tape cartridge is as easy and intuitive as using a USB flash drive. Tapes are self-describing, and you can quickly recall any file from a tape without having to read the whole tape from beginning to end.

Furthermore, any LTFS-capable system can read a tape that is created by any other LTFS-capable system (regardless of the operating system and platform). Any LTFS-capable system can identify and retrieve the files that are stored on it. LTFS-capable systems have the following characteristics:

- Files and directories are displayed to you as a directory tree listing.
- More intuitive searches of cartridge and library content are now possible due to the addition of file tagging.

- Files can be moved to and from LTFS tape by using the familiar drag-and-drop metaphor common to many operating systems.
- Many applications that were written to use files on disk can now use files on tape without any modification.
- All standard File Open, Write, Read, Append, Delete, and Close functions are supported.

### Spectrum Archive Editions

As shown in Figure 4-23, Spectrum Archive is available in different editions that support small, medium, and enterprise businesses.

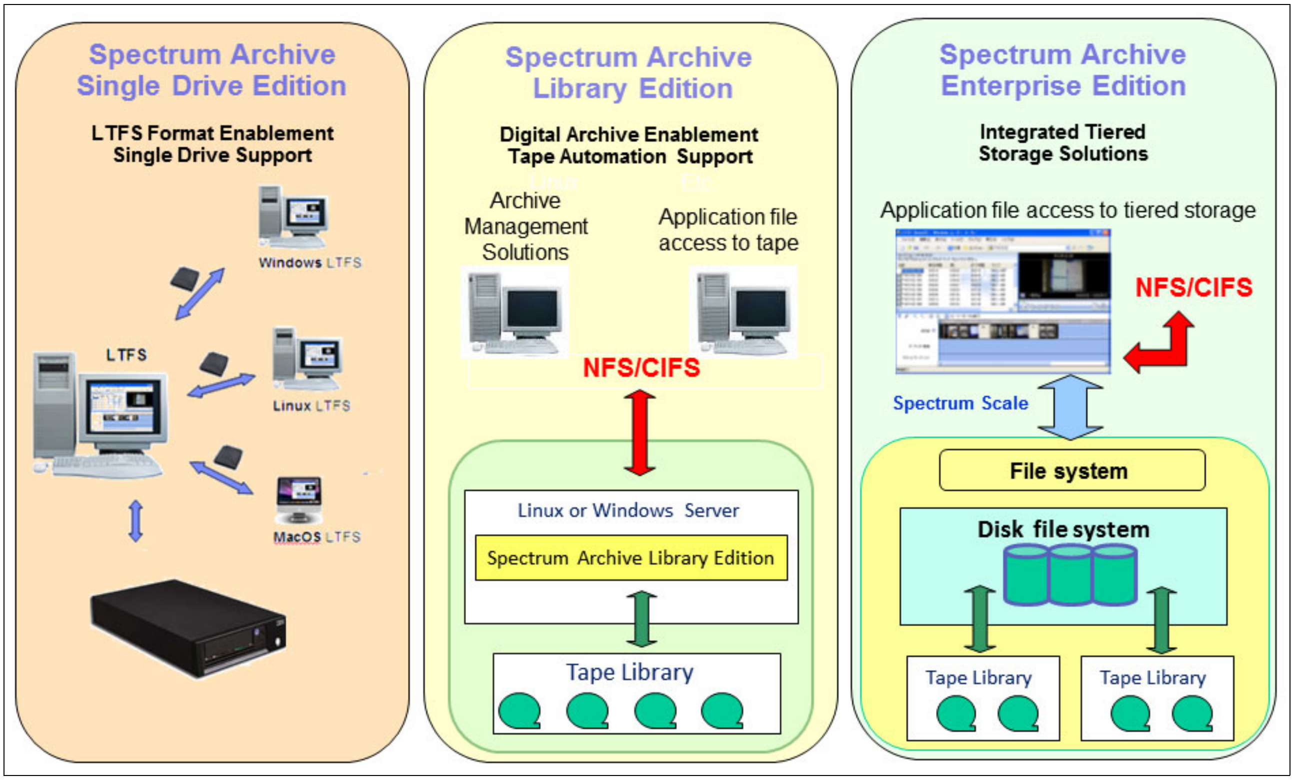



*Figure 4-23 Spectrum Archive SDE, LE, and EE implementations*

#### *IBM Spectrum Scale Single Drive Edition*

The IBM Spectrum Archive Single Drive Edition implements the LTFS Format and allows tapes to be formatted as LTFS Volumes. These LTFS Volumes can then be mounted by using LTFS to allow users and applications direct access to files and directories that are stored on the tape. No integration with tape libraries exists in this edition. You can access and manage all data in stand-alone tape environments as simply as though it were on disk.

#### *IBM Spectrum Archive Library Edition*

IBM Spectrum Archive Library Edition extends the file management capability of the IBM Spectrum Archive SDE. Spectrum Archive LE is introduced with Version 2.0 of LTFS. It enables easy-as-disk access to single or multiple cartridges in a tape library.

LTFS is the first file system that works with IBM System Storage tape technology to optimize ease of use and portability for open-systems tape storage. It manages the automation and provides operating system-level access to the contents of the library. Spectrum Archive LE is based on the LTFS format specification, enabling tape library cartridges to be interchangeable with cartridges that are written with the open source SDE version of Spectrum Archive. IBM Spectrum Archive LE supports most IBM tape libraries:

- TS2900 tape autoloader
- TS3100 tape library
- TS3200 tape library
- TS3310 tape library
- TS3500 tape library
- TS4500 tape library

IBM TS1150 and IBM TS1140 tape drives are supported on the IBM TS4500 and IBM TS3500 tape libraries only.

Spectrum Archive LE enables the reading, writing, searching, and indexing of user data on tape and access to user metadata. *Metadata* is the descriptive information about user data that is stored on a cartridge. Metadata enables searching and accessing of files through the GUI of the operating system. Spectrum Archive LE supports both Linux and Windows.

Spectrum Archive LE provides the following product features:

- Direct access and management of data on tape libraries with LTO Ultrium 7 (LTO-7), LTO Ultrium 6 (LTO-6), LTO Ultrium 5 (LTO-5), and TS1150 and TS1140 tape drives
- Tagging of files with any text, allowing more intuitive searches of cartridge and library content
- Exploitation of the partitioning of the media in LTO-5 tape format standard
- One-to-one mapping of tape cartridges in tape libraries to file folders
- Capability to create a single file system mount point for a logical library that is managed by a single instance of LTFS and runs on a single computer system
- Capability to cache tape indexes, and to search, query, and display tape content within an IBM tape library without having to mount tape cartridges

The IBM Spectrum Archive LE offers the same basic capabilities as the SDE with additional support of tape libraries. Each LTFS tape cartridge in the library appears as an individual folder within the file space. The user or application can navigate to these folders to access the files that are stored on each tape. The Spectrum Archive LE software automatically controls the tape library robotics to load and unload the necessary LTFS Volumes to provide access to the stored files.

### *IBM Spectrum Archive Enterprise Edition*

IBM Spectrum Archive Enterprise Edition (EE) gives organizations an easy way to use cost-effective IBM tape drives and libraries within a tiered storage infrastructure. By using tape libraries instead of disks for Tier 2 and Tier 3 data storage (data that is stored for long-term retention), organizations can improve efficiency and reduce costs. In addition, Spectrum Archive EE seamlessly integrates with the scalability, manageability, and performance of IBM Spectrum Scale, an IBM enterprise file management platform that enables organizations to move from simply adding storage to optimizing data management.

Here are some of the Spectrum Archive Enterprise Edition highlights:

- Simplify tape storage with the IBM LTFS format, which is combined with the scalability, manageability, and performance of IBM Spectrum Scale
- Help reduce IT expenses by replacing tiered disk storage (Tier 2 and Tier 3) with IBM tape libraries
- Expand archive capacity by simply adding and provisioning media without affecting the availability of data already in the pool

- Add extensive capacity to IBM Spectrum Scale installations with lower media, floor space, and power costs
- Support for attaching up to two tape libraries to a single Spectrum Scale cluster

Spectrum Archive EE for the IBM TS4500, IBM TS3500, and IBM TS3310 tape libraries provides seamless integration of Spectrum Archive with Spectrum Scale by creating an LTFS tape tier. You can run any application that is designed for disk files on tape by using Spectrum Archive EE. Spectrum Archive EE can play a major role in reducing the cost of storage for data that does not need the access performance of primary disk. This configuration improves efficiency and reduces costs for long-term, tiered storage.

With Spectrum Archive EE, you can enable the use of LTFS for the policy management of tape as a storage tier in a Spectrum Scale environment and use tape as a critical tier in the storage environment. Spectrum Archive EE supports IBM LTO Ultrium 7, 6, and 5, IBM System Storage TS1150, and TS1140 tape drives that are installed in TS4500, TS3500, or LTO Ultrium 7, 6, and 5 tape drives that are installed in the TS3310 tape libraries.

The use of Spectrum Archive EE to replace disks with tape in Tier 2 and Tier 3 storage can improve data access over other storage solutions. It improves efficiency and streamlines management for files on tape. Spectrum Archive EE simplifies the use of tape by making it transparent to the user and manageable by the administrator under a single infrastructure.

Figure 4-24 shows the integration of Spectrum Archive EE archive solution with Spectrum Scale.

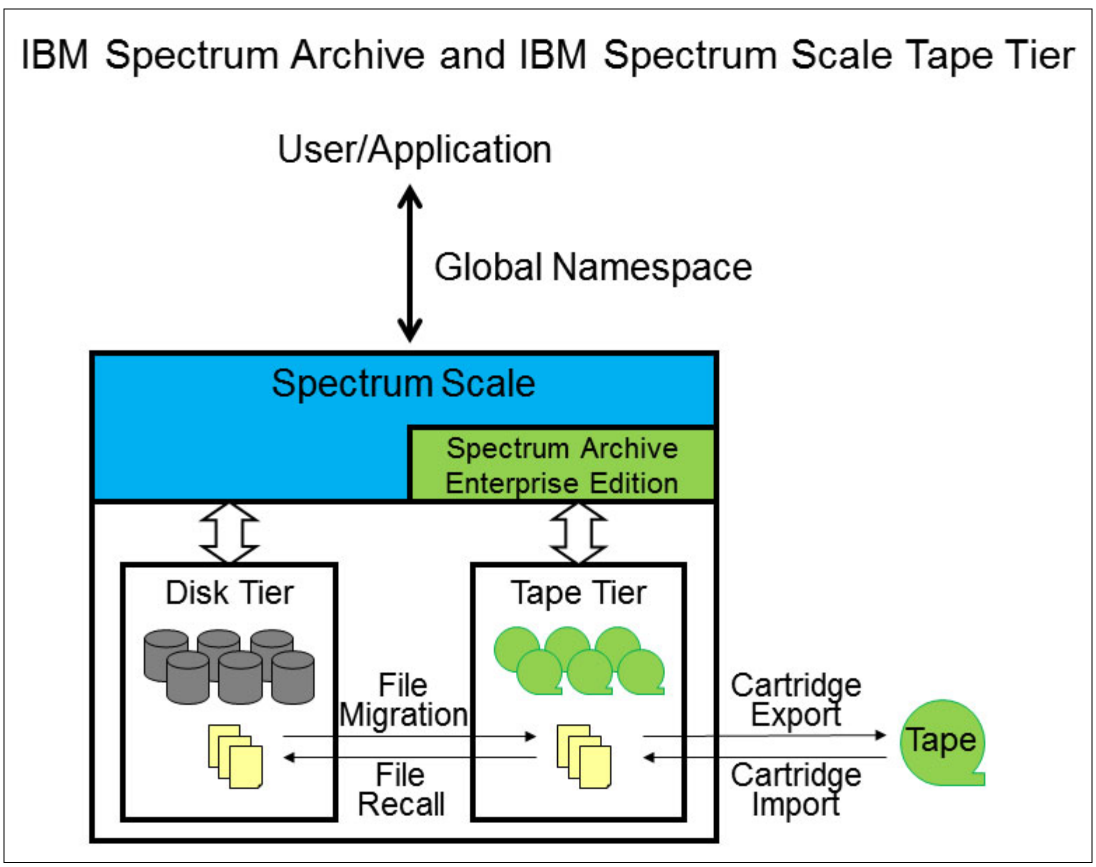



*Figure 4-24 Integration of Spectrum Scale and Spectrum Archive Enterprise Edition*

The seamless integration offers transparent file access in a continuous name space. It provides file level write and read caching with disk staging area, policy-based movement from disk to tape, creation of multiple data copies on different tapes, load balancing, and high availability in multi-node clusters. It also offers data exchange on LTFS tape by using import and export functions, fast import of file name space from LTFS tapes without reading data, built-in tape reclamation and reconciliation, and simple administration and management.

For more information, see IBM Spectrum Archive at:

http://www.ibm.com/systems/storage/tape/ltfs

IBM Spectrum Archive Enterprise Edition can also be used to provide object storage by using OpenStack Swift. This configuration means that Objects can be stored in the file system and can exist on either disk or tape tiers within the enterprise. More information on creating an object storage "Active Archive" with IBM Spectrum Scale and Spectrum Archive can be found in *Active Archive Implementation Guide with IBM Spectrum Scale Object and IBM Spectrum Archive*, REDP-5237.

## 4.3.6 IBM object storage solutions

This section describes IBM object storage solutions.

### IBM Cloud Object Storage

The IBM Cloud Object Storage (COS) system is a breakthrough cloud platform that helps solve petabyte and beyond storage challenges for companies worldwide. Clients across multiple industries use IBM Cloud Object Storage for large-scale content repository, backup, archive, collaboration, and SaaS.

The Internet of Things (IoT) allows every aspect of life to be instrumented through millions of devices that create, collect, and send data every second. These trends are causing an unprecedented growth in the volume of data being generated. IT organizations are now tasked with finding ways to efficiently preserve, protect, analyze, and maximize the value of their unstructured data as it grows to petabytes and beyond. Object storage is designed to handle unstructured data at web-scale.

The IBM Cloud Object Storage portfolio gives clients strategic data flexibility, simplified management, and consistency with on-premises, cloud, and hybrid cloud deployment options. See Figure 4-25.

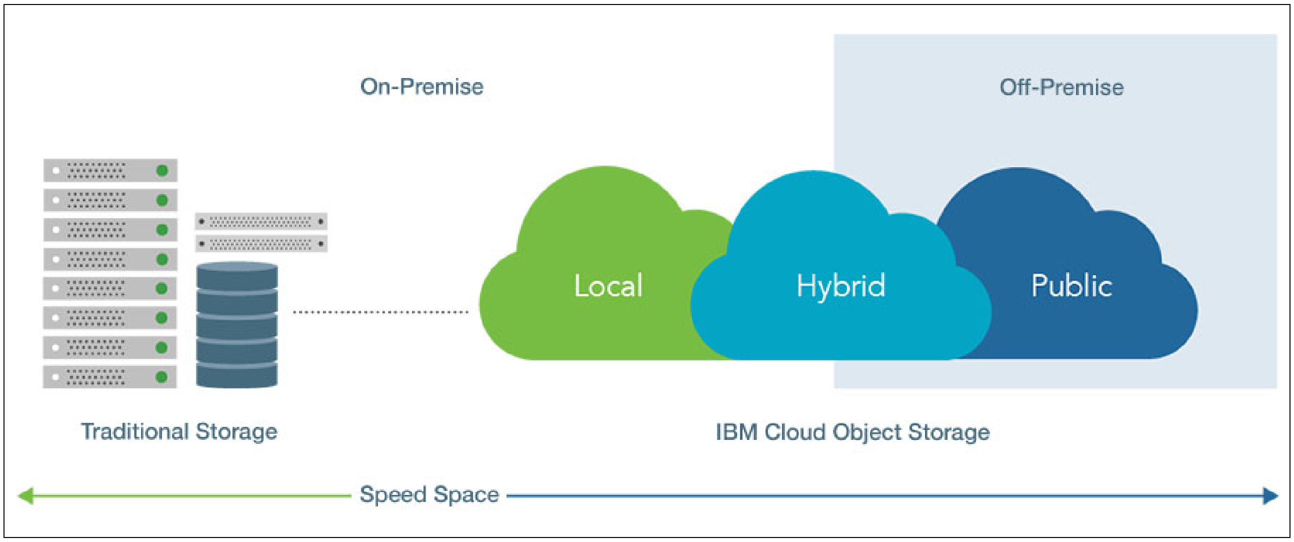



*Figure 4-25 Cloud Object Storage offers flexibility for on-premises, cloud, and hybrid cloud deployment options*

IBM Cloud Object Storage's Dispersed Storage® Network (dsNet) solutions enhances on-premises storage options for clients and service providers with low-cost, large-scale active archives and unstructured data content stores. The solutions complement the IBM software defined Spectrum Storage portfolio for data protection and backup, tape archive, and a high-performance file and object solution where the focus is on response time.

IBM Cloud Object Storage can be deployed as an on-premises, public cloud, or hybrid solution, providing you unprecedented choice, control, and efficiency:

**On-Premise solutions**

Deploy IBM Cloud Object Storage on premises for optimal scalability, reliability, and security. The software runs on industry standard hardware for flexibility and simplified management.

**Cloud Solutions** Easily deploy IBM Cloud Object Storage on the IBM SoftLayer public cloud.

**Hybrid Solutions** For optimal flexibility, deploy IBM Cloud object storage as a hybrid solution to support multiple sites across your enterprise (on-premises and in the public cloud) for agility and efficiency.

### *Access methods*

The IBM Cloud Object storage pool can be shared and is jointly accessible by multiple access protocols:

- Object-based access methods: The Simple Object interface is accessed with a HTTP/REST API. Simple PUT, GET, DELETE, and LIST commands allow applications to access digital content, and the resulting object ID is stored directly within the application. The IBM COS Accesser® does not require a dedicated appliance because the application can talk directly to the IBM COS Slicestor® using object IDs. See Figure 4-26.
- REST API access to storage: REST is a style of software architecture for distributed hypermedia information retrieval systems such as the World Wide Web. REST style architectures consist of clients and servers. Clients send requests to servers. Servers process those requests and return associated responses. Requests and responses are built around the transfer of various representations of the resources. The REST API works in way that is similar to retrieving a Universal Resource Locator (URL). But instead of requesting a web page, the application references an object.
- File-based access methods: Dispersed storage can also support the traditional NAS protocols (SMB/CIFS and NFS) through integration with third-party gateway appliances. Users and storage administrators are able to easily transfer, access, and preserve data assets over standard file protocol.

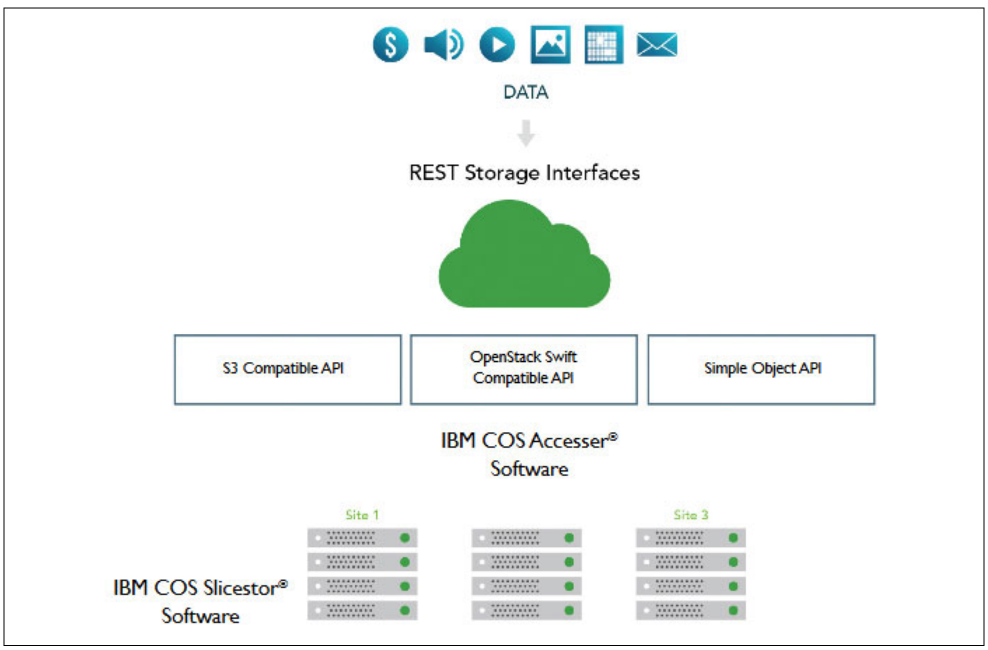



*Figure 4-26 REST APIs accessing objects using object IDs with IBM COS Slicestor*

The IBM COS System is deployed as a cluster that combines three types of nodes as shown in Figure 4-27. Each node consists of IBM COS software running on an industry-standard server. IBM COS software is compatible with a wide range of servers from many sources, including a physical or virtual appliance. In addition, IBM conducts certification of specific servers that customers want to use in their environment to help ensure a quick initial installation, long-term reliability, and predictable performance.

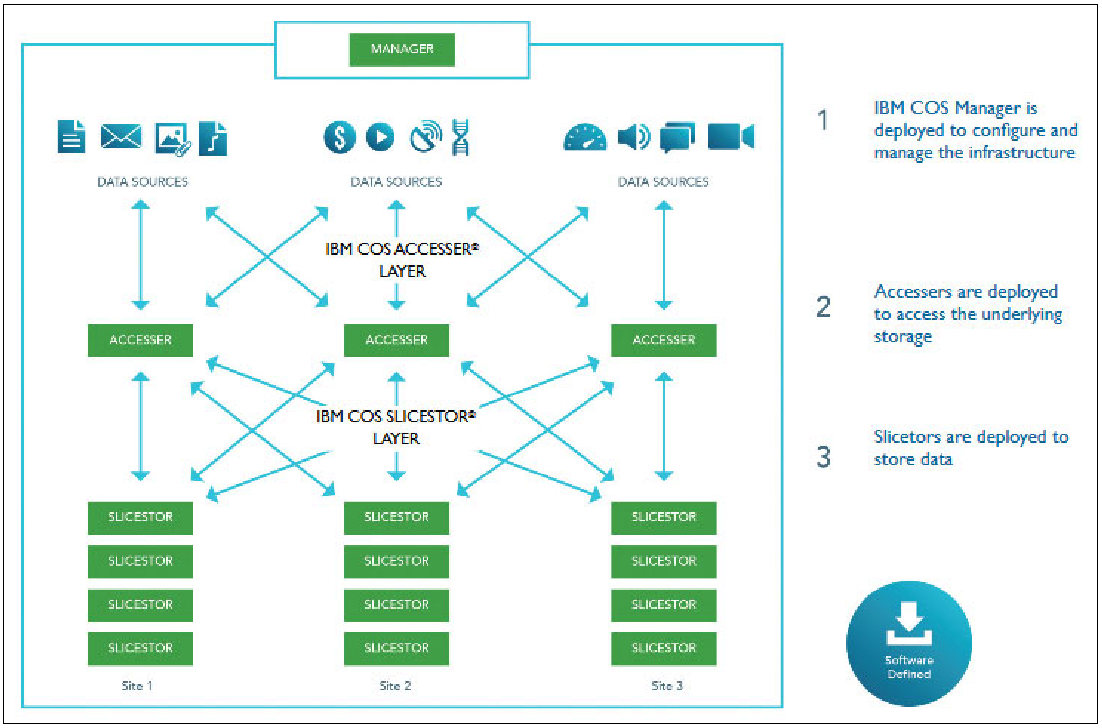



*Figure 4-27 IBM COS System deployed as a cluster combining three types of nodes*

The following three types of nodes are available:

- IBM Cloud Object Storage Manager
- IBM Cloud Object Storage Accesser
- IBM Cloud Object Storage Slicestor

Each IBM COS System include the following nodes:

- A single Manager node, which provides out-of-band configuration, administration and monitoring capabilities
- One or more Accesser nodes, which provide the storage system endpoint for applications to store and retrieve data
- One or more Slicestor nodes, which provide the data storage capacity for the IBM COS System

The Accesser is a stateless node that presents the storage interface of the IBM COS System to client applications and transforms data using an Information Dispersal Algorithm (IDA). Slicestor nodes receive data to be stored from Accesser nodes on ingest and return data to Accesser nodes as required by reads.

The IDA transforms each object written to the system into a number of slices such that the object can be read bit-perfectly by using a subset of those slices. The number of slices created is called the IDA Width (or Width) and the number required to read the data is called the IDA Read Threshold (or Read Threshold). The difference between the Width and the

Read Threshold is the maximum number of slices that can be lost or temporarily unavailable while still maintaining the ability to read the object. For example, in a system with a width of 12 and threshold of seven, data can be read even if five of the 12 stored slices cannot be read.

Storage capacity is provided by a group of Slicestor nodes, which are referred to as a storage pool. In the diagram in Figure 4-27 on page 110, 12 Slicestor nodes are grouped in a storage pool. A single IBM COS System can have one or multiple storage pools.

A Vault is not part of the physical architecture, but is an important concept in an IBM COS System. A Vault is a logical container or a virtual storage space, upon which reliability, data transformation options (for example, IBM COS SecureSlice and IDA algorithm), and access control policies can be defined. Multiple vaults can be provisioned on the same storage pool.

The Information Dispersal Algorithm combines encryption and erasure-coding techniques that are designed to transform the data in a way that enables highly reliable and available storage without making copies of the data as would be required by traditional storage architectures.

### *Information Dispersal*

At the foundation of the IBM COS System is a technology called *information dispersal*. Information dispersal is the practice of using erasure codes as a means to create redundancy for transferring and storing data. An erasure code is a Forward Error Correction (FEC) code that transforms a message of k symbols into a longer message with n symbols such that the original message can be recovered from a subset of the n symbols (k symbols).

Erasure codes use advanced deterministic math to insert “extra data” in the “original data” that allows a user to need only a subset of the “coded data” to re-create the original data.

An IDA can be made from any Forward Error Correction code. The additional step of the IDA is to split the coded data into multiple segments. These segments can then be stored on different devices or media to attain a high degree of failure independence. For example, using forward FEC alone on files on your computer is less likely to help if your hard disk drive fails. However, if you use an IDA to separate pieces across machines, you can now tolerate multiple failures without losing the ability to reassemble that data. See Figure 4-28.

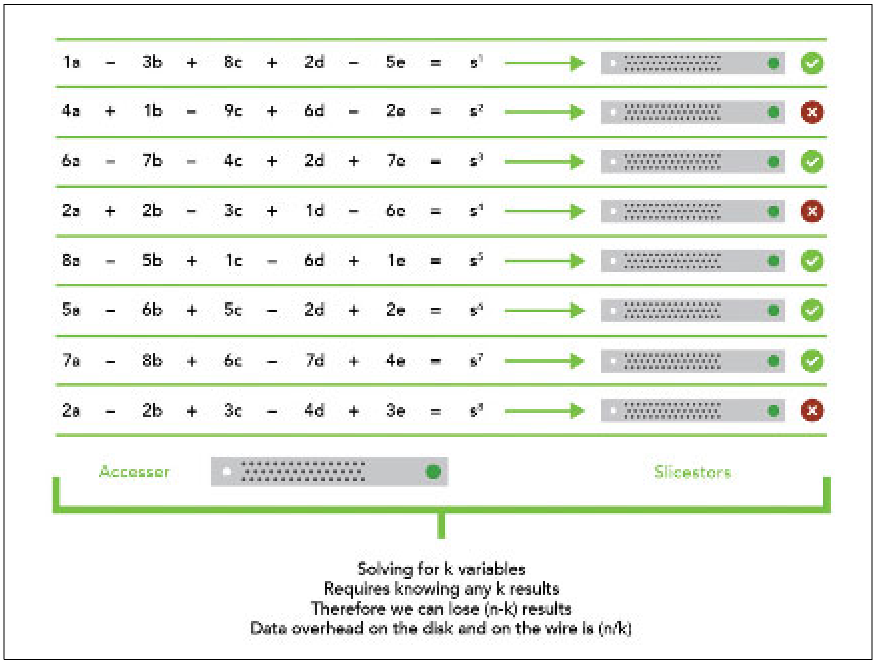



*Figure 4-28 Example of calculations to illustrate how information dispersal works*

As you can see in Figure 4-28, there are five variables (a through e) and eight different equations that use these variables, with each yielding a different output. To understand how information dispersal works, imagine the five variables are bytes. Following the eight equations, you can compute eight results, each of which is a byte. To solve for the original five bytes, you can use any five of the resulting eight bytes. This process is how information dispersal can support any value for k, and n- k is the number of variables, and n is the number of equations.

### *How the Storage Dispersal and Retrieval works*

At a basic level, the IBM COS System uses three steps for slicing, dispersing, and retrieving data. See Figure 4-29 on page 113.

1. Data is virtualized, transformed, sliced, and dispersed by using IDAs. In the example in Figure 4-29 on page 113, the data is separated into 12 slices. So the “width” (n) of the system is 12.
2. Slices are distributed to separate disks, storage nodes, geographic locations, or some combination of these three. In this example, the slices are distributed to three different sites.
3. The data is retrieved from a subset of slices. In this example, the number of slices that are needed to retrieve the data is 7. Therefore, the “threshold” (k) of the system is 7. Given a width of 12 and a threshold of 7, this example can be called a “7 of 12” (k of n) configuration.

The configuration of a system is determined by the level of reliability required. In a “7 of 12” configuration, five slices can be lost or unavailable and the data can still be retrieved because the threshold of seven slices has been met. With a “5 of 8” configuration, only three slices can be lost, so the level of reliability is lower. Conversely, with a “20 of 32” configuration, 12 slices can be lost, so the level of reliability is higher.

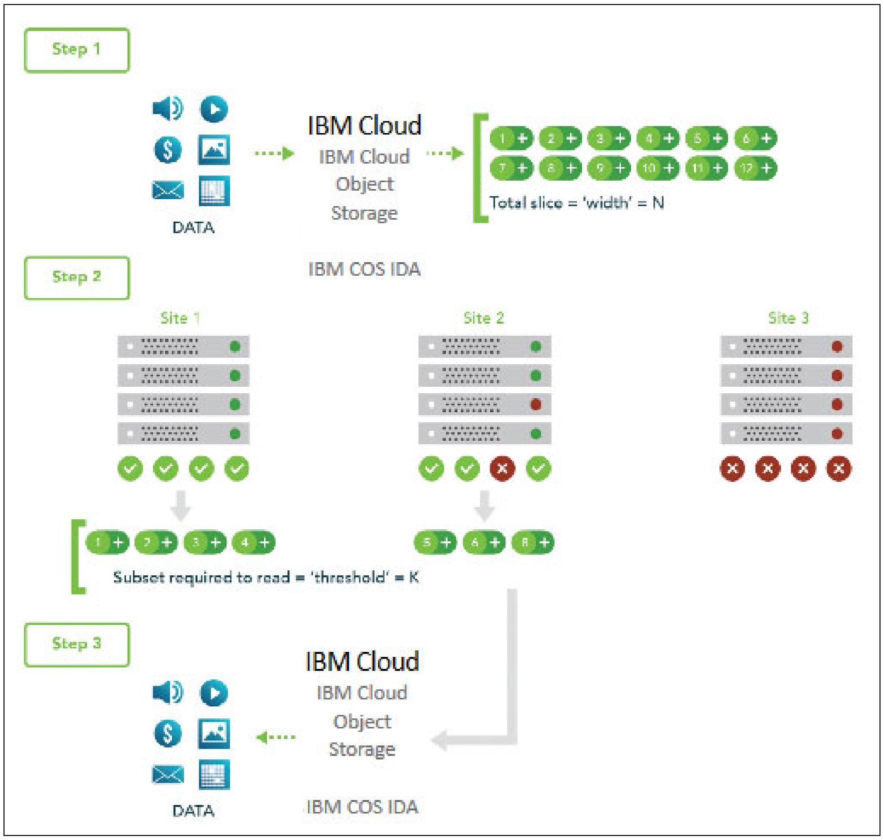



*Figure 4-29   COS System’s three steps for slicing, dispersing, and retrieving data*

For more information about IBM Cloud Object Storage, see:

https://www.ibm.com/cloud-computing/infrastructure/object-storage

### IBM Spectrum Scale Object support

Spectrum Scale supports file and object solutions. See “IBM Spectrum Scale” on page 92 for Spectrum Scale information about object support.

## 4.4 IBM storage support of OpenStack components

OpenStack environment is a cloud operating system that controls large pools of compute, storage, and networking resources throughout a data center. The resources are managed through a dashboard that gives administrators control while empowering their users to provision resources through a web interface. OpenStack Juno is the tenth release of the open source software for building public, private, and hybrid clouds. It has nearly 342 new features to support software development, big data, analysis, and application infrastructure. The Juno release adds enterprise features such as storage policies, a new data processing service that provisions Hadoop and Spark, and lays the foundation for OpenStack Software to be the platform for Network Functions Virtualization.

The OpenStack community continues to grow and attract developers and experts. The Mitaka release was designed and built by an international community of 2,336 developers, operators, and users from 345 organizations.

Because OpenStack software design and development is done in the open, public documentation is available regarding the development status of the current release and decisions made at each Design Summit. You can review this information in technical detail at the following link:

https://releases.openstack.org/mitaka

## 4.4.1 Global collaboration for OpenStack storage components

OpenStack technology is a key enabler of cloud infrastructure as a service (IaaS) capability. OpenStack architecture provides an overall cloud preferred practices workflow solution that is readily installable, and supported by a large ecosystem of worldwide developers in the OpenStack open source community.

Within the overall cloud workflow, specific OpenStack components support storage. The following OpenStack components support storage:

- IBM Cinder storage drivers
- Swift (object storage)
- Manila (file storage)

Figure 4-30 shows the OpenStack storage components Cinder, Manila, and Swift.

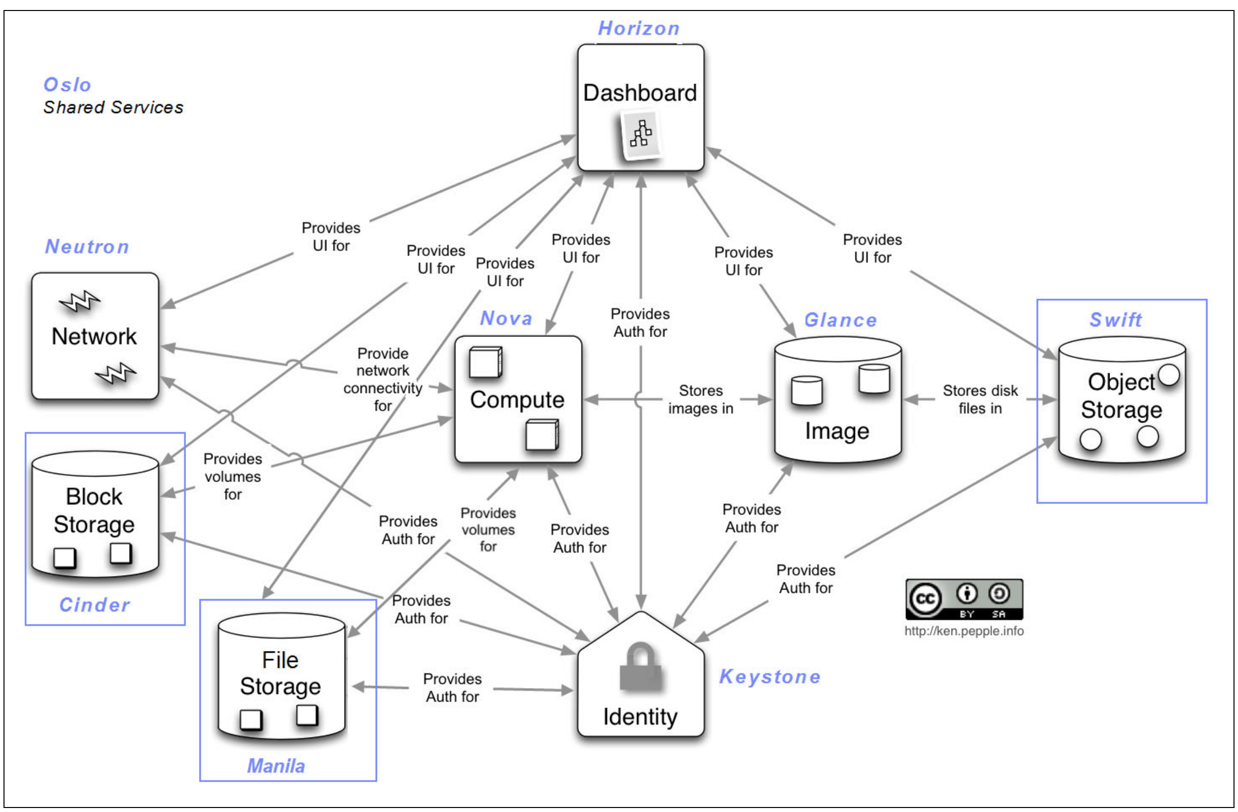



*Figure 4-30 OpenStack storage components Cinder (block), Swift (object), and Manila (file)*

OpenStack architecture is one implementation of a preferred practices cloud workflow. Regardless of the cloud operating system environment that is used, the following key summary points apply:

- Cloud operating systems provide the necessary technology workflow to provide truly elastic, pay per use cloud services
- OpenStack cloud software provides a vibrant open source cloud operating system that is growing quickly
- OpenStack storage components

### 4.4.2 Cinder

Cinder is an OpenStack project to provide *block storage as a service* and provides an API to users to interact with different storage backend solutions. Cinder component provides support, provisioning, and control of block storage. The following are standards across all drivers for Cinder services to properly interact with a driver.

Icehouse updates for Cinder are block storage added backend migrations with tiered storage environments, allowing for performance management in heterogeneous environments. Mandatory testing for external drivers now ensures a consistent user experience across storage platforms, and fully distributed services improve scalability.

### 4.4.3 Swift

The OpenStack Object Store project, which is known as OpenStack Swift, offers cloud storage software so that you can store and retrieve lots of data with a simple API. It is built for scale and optimized for durability, availability, and concurrency across the entire data set. Swift is ideal for storing unstructured data that can grow without bound.

**Note:** Do not confuse OpenStack Swift with Apple Swift, a programming language. In this paper, the term “Swift” always refers to OpenStack Swift.

### 4.4.4 Manila

The OpenStack Manila (File) component provides file storage, which allows coordinated access to shared or distributed file systems. Although the primary consumption of shares would be OpenStack compute instances, the service is also intended to be accessed independently, based on the modular design established by OpenStack services.

Manila has the following capabilities:

- Shared file system services for VMs
- Vendor-neutral API for NFS/CIFS and other network file systems
- IBM Spectrum Scale Manila (in Kilo):
  - Extends Spectrum Scale data plane into VM
  - Supports both kNFS and Ganesha 2.0
  - Create/list/delete Shared and Snapshots
  - Allow/deny access to a share based on IP address
  - Multi-tenancy

For more information about OpenStack technology, see the following site:

http://www.openstack.org

The following sections highlight the IBM SDS products that have interfaces to OpenStack components:

- The IBM Storage Driver for OpenStack environments: The IBM Storage Driver for OpenStack environments is a software component that integrates with the OpenStack cloud environment. It enables the usage of storage resources that are provided by the following IBM storage systems:
  - DS8880: This storage system can offer a range of capabilities that enable more effective storage automation deployments in private or public clouds. Enabling the OpenStack Cinder storage component with DS8880 allows for storage to be made available whenever it is needed without the traditional associated cost of highly skilled administrators and infrastructure. For more information, see *Using IBM DS8870 in an OpenStack Environment*, REDP-5220.
  - IBM XIV: Remote cloud users can issue requests for storage resources from the OpenStack cloud. These requests are transparently handled by the IBM Storage Driver. The IBM Storage Driver communicates with the XIV Storage System and controls the storage volumes on it. With the release of Version 11.5 software, the XIV introduced support for multi-tenancy. *Multi-tenancy* enables cloud providers to divide and isolate the XIV resources into logical domains, which can then be used by tenants without any knowledge of the rest of the system resources. For more information, see *Using XIV in OpenStack Environments*, REDP-4971.
  - IBM Storwize family/SAN Volume Controller: The volume management driver for the Storwize family and SAN Volume Controller provides OpenStack Compute instances with access to IBM Storwize family or SAN Volume Controller storage systems.

    Storwize and SAN Volume Controller support fully transparent live storage migration in OpenStack Havana:
    - No interaction with the host is required: All advanced Storwize features are supported and exposed to the Cinder system.
    - Real-time Compression with EasyTier supports iSCSI + FC attachment.
  - IBM FlashSystem (Kilo release): The volume driver for FlashSystem provides OpenStack Block Storage hosts with access to IBM FlashSystems.
- IBM Spectrum Scale: As of OpenStack Juno Release, Spectrum Scale combines the benefits of Spectrum Scale with the most widely used open source object store today, OpenStack Swift. Spectrum Scale provides enterprise ILM features. OpenStack Swift provides a robust object layer with an active community that is continuously adding innovative new features. To ensure compatibility with the Swift packages over time, no code changes are required to either Spectrum Scale or Swift to build the solution. For more information, see *A Deployment Guide for IBM Spectrum Scale Object*, REDP-5113.
- IBM Spectrum Protect: IBM data protection and data recovery solutions provide protection for virtual, physical, cloud, and software-defined infrastructures as well as core applications and remote facilities. These solutions fit nearly any size organization and recovery objective. They deliver the functions of IBM Spectrum Protect.

  IBM Spectrum Protect enables software-defined storage environments by delivering automated data protection services at the control plane for file, block, and object backup.

  IBM Spectrum Protect enables cloud data protection with OpenStack and VMware integration, cloud portal, and cloud deployment options.

For more information, see “IBM Spectrum Protect for Virtual Environments” on page 76. Also, see *Protecting OpenStack with Tivoli Storage Manager for Virtual Environments* at:

https://ibm.biz/BdXZmY

- IBM Spectrum Control integration with OpenStack (Cinder/Manila): IBM Spectrum Control has support for Cinder driver, which enables cloud storage to be powered by OpenStack software.

  IBM Spectrum Control provides block storage capabilities that the storage administrator can use to define the properties and characteristics of storage volumes with a particular service class. A block storage service class can define the RAID levels, tiers of storage, and various other characteristics.

  IBM Spectrum Control Cinder driver is provided at no extra charge, and can be downloaded at:

  https://www.ibm.com/developerworks/servicemanagement/sm/spectrum_control/downloads.html

  To install IBM Spectrum Control Cinder driver, make sure to have access to Virtual Storage Center license with IBM Spectrum Control version 5.2.2, OpenStack (Havana release). The OpenStack Cinder node must run either Red Hat Enterprise 6.4 or higher, or Ubuntu 12.04.

**Note:** For more information about the IBM storage drivers and functions that are supported in the various OpenStack releases, see the following wiki:

https://wiki.openstack.org/wiki/CinderSupportMatrix

# 4.5 IBM storage supporting the data plane

This section describes the IBM storage that supports the software defined storage products in the data plane.

## 4.5.1 IBM FlashSystem family

For businesses to act based on insight from data and transform it into competitive advantage, the data driven applications must operate at high availability and provide sufficient performance. IBM FlashSystem delivers extreme performance to provide measurable economic value across the data architecture, including servers, software, applications, and storage. IBM offers a comprehensive flash portfolio with the IBM FlashSystem family, supporting SDS solutions as a concept, as well as flash-optimized XIV, Storwize V7000, and DS8000 storage.

The IBM FlashSystem family allows you to take advantage of best-in-breed solutions that provide extreme performance, macro efficiency, and microsecond response times. The IBM FlashSystem V9000 Enterprise Performance Solution, the IBM FlashSystem A9000, and the IBM FlashSystem 900 members of the FlashSystem family are described in this section. IBM FlashSystem A9000 is the newest addition to the FlashSystem family of storage systems. You can consider IBM FlashSystem as a major Tier for SDS.

### FlashSystem benefits

Flash technology has fundamentally changed the paradigm for IT systems, enabling new use cases and unlocking the scale of enterprise applications. Flash technology enhances the performance, efficiency, reliability, and design of essential enterprise applications and solutions. It does this by addressing the bottleneck in the IT process (data storage), enabling truly optimized information infrastructure. IBM FlashSystem shared flash memory systems offer affordable, high-density, ultra low-latency, high reliability, and scalable performance in a storage device that is both space and power efficient.

IBM Flash products can either augment or replace traditional hard disk drive storage systems in enterprise environments. They empower applications to work faster and scale further. In addition to optimizing performance, the IBM FlashSystem family brings enterprise reliability and macro efficiency to the most demanding data centers, allowing businesses to receive the following benefits:

- Reduce customer complaints by improving application response time
- Service more users with less hardware
- Reduce I/O wait and response times of critical applications
- Simplify solutions
- Reduce power and floor space requirements
- Speed up applications, enhancing the pace of business
- Improve utilization of existing infrastructure
- Complement existing infrastructure
- Eliminate storage bottlenecks

From the client business perspective, IBM FlashSystem provides focus benefits and value in these essential areas:

- Extreme Performance: Enable business to unleash the power of performance, scale, and insight to drive services and products to market faster.
- MicroLatency: Achieve competitive advantage through applications that enable faster decision making due to microsecond response times.
- Macro Efficiency: Decrease costs by getting more from efficient use of the IT staff, IT applications, and IT equipment due to the efficiencies that flash brings to the data center.
- Enterprise Reliability: Durable and reliable designs that use enterprise class flash and patented data protection technology.

## IBM FlashCore technology

IBM FlashCore® technology refers to the IBM innovations that enable FlashSystem storage to deliver extreme performance, IBM MicroLatency, enterprise-grade reliability, and a wide range of operational and cost efficiencies. These technologies and innovations are represented in the FlashCore hardware-accelerated architecture and IBM MicroLatency modules. They are also in other advanced flash management features and capabilities that are used in the IBM FlashSystem 900, A9000, and V9000:

- Hardware Accelerated Architecture: By using an all-hardware data path, FlashSystem arrays minimize the amount of software interaction during I/O activity, resulting in the highest performance and lowest latency for all-flash storage arrays.
- IBM MicroLatency Modules: By using IBM-designed, purpose-engineered flash memory modules, FlashSystem delivers extreme performance, greater density, unlimited scalability, and mission-critical reliability.
- Advanced Flash Management: Unique, patented IBM hardware and software innovations enable FlashSystem to provide the most reliable, feature-rich, and highly available flash data storage.

### FlashSystem 900

Flash memory gives organizations the ability to deliver fast, reliable, and consistent access to critical data. With IBM FlashSystem 900, you can make faster decisions based on real-time insights and unleash the power of the most demanding applications. These applications include online transaction processing and analytics databases, virtual desktop infrastructures, technical computing applications, and cloud environments. FlashSystem 900 can also lower operating costs and increase the efficiency of IT infrastructure by using much less power and space than traditional hard disk drive (HDD) and SSD solutions. Here are some reasons to consider FlashSystem 900 for implementation:

- When speed is critical: IBM FlashSystem 900 is designed to accelerate the applications that drive business. Powered by IBM FlashCore Technology, FlashSystem 900 delivers high performance at lower cost:
  - 90 μ/155 μ read/write latency.
  - Up to 1.1 million random read 4 K IOPS.
  - Up to 10 GBps read bandwidth.
- High capacity business needs: IBM FlashSystem 900 has 12 hot-swappable IBM MicroLatency storage modules: 1.2 TB. 2.9 TB, and 5.7 TB.
- Provides higher density: IBM FlashSystem 900 employs 20 nm multi-level cell (MLC) chips with IBM-enhanced Micron MLC technology for higher storage density and improved endurance.
- Highly scalable: FlashSystem 900 is configurable with 2.4 - 57 TB of capacity for increased flexibility.
- Easy to integrate into VMware environments: FlashSystem 900 is easy to integrate with VMware VASA by using Spectrum Control Base Edition to use the following features:
  - Greater communication between vSphere and FlashSystem.
  - Ability of vSphere to monitor and directly manage FlashSystem, allowing greater efficiency.
  - Integration of VASA Unmap for greater storage efficiency.

Through SAN Volume Controller, the FlashSystem 900 has support for OpenStack cloud environments.

For more information, see these IBM publications:

- *FlashSystem 900 Product Guide*, TIPS1261
- *Implementing IBM FlashSystem 900*, SG24-8271

### IBM FlashSystem V9000

IBM FlashSystem V9000 is a comprehensive all-flash enterprise storage solution. FlashSystem V9000 delivers the full capabilities of IBM FlashCore technology plus a rich set of storage virtualization features.

FlashSystem V9000 offers the advantages of software-defined storage at the speed of flash. These all-flash storage systems deliver the full capabilities of the hardware-accelerated I/O provided by FlashCore Technology. FlashSystem V9000 also delivers the enterprise reliability of MicroLatency modules and advanced flash management. These features are coupled with a rich set of the features that are found in the most advanced software-defined storage solutions. These features include Real-time Compression, dynamic tiering, thin provisioning, snapshots, cloning, replication, data copy services, and high-availability configurations.

V9000 now also supports physical expansions receiving SSD drives and NearlLine SAS drives offering three tiers of storage.

Figure 4-31 shows the V9000 connectivity to VMware through IBM Spectrum Control Base Edition 2.1.1.

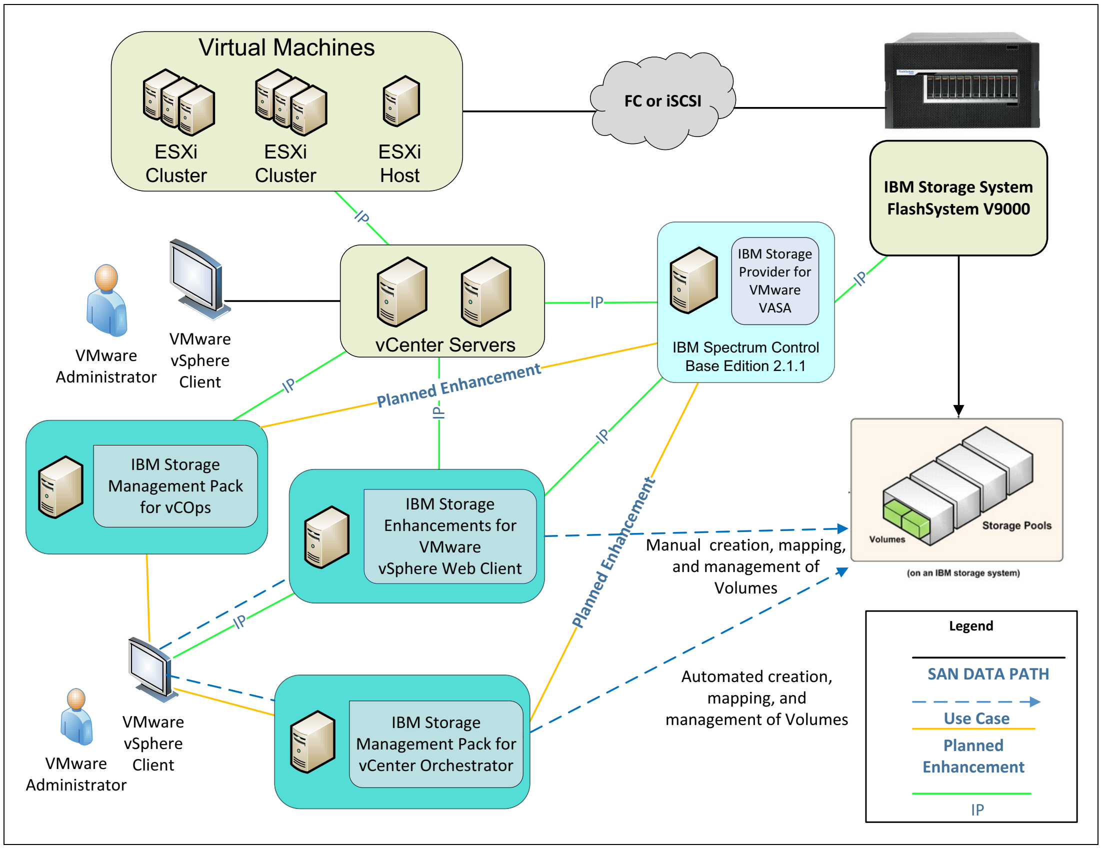



*Figure 4-31 V9000 connectivity with IBM Spectrum Base Edition and VMware*

FlashCore technology plus a rich set of storage virtualization features allow FlashSystem V9000 to deliver industry-leading value to enterprises in scalable performance, enduring economics, and agile integration:

- Fast: Optimize your infrastructure with scale-up, scale-out capabilities of fast FlashSystem performance.
- Cost-effective: Powerful virtualized storage enables you to realize immediate and long-term economic benefits.
- Easy: Unlike conventional storage, FlashSystem is easy to deploy, can virtualize legacy systems, and delivers value in hours.

The FlashSystem V9000 has connectivity to OpenStack cloud environments through the Cinder driver.

For more information, see the following IBM publications:

- *IBM FlashSystem V9000 Product Guide*, TIPS1281
- *IBM FlashSystem V9000 and VMware Best Practices Guide*, REDP-5247

### IBM FlashSystem A9000/A9000R

IBM FlashSystem A9000 is a new comprehensive all-flash enterprise storage solution. It delivers the full capabilities of IBM FlashCore technology combined with advanced reduction mechanism and all the features of the IBM Spectrum Accelerate software stack.

As a cloud optimized solution, IBM FlashSystem A9000 suits the requirements of public and private cloud providers who require features, such as inline data deduplication, multi-tenancy, and quality of service. It also uses powerful software-defined storage capabilities from IBM Spectrum Accelerate, such as Hyper-Scale technology and VMware integration:

- An enhanced management interface simplifies storage administration
- Data reduction: Pattern removal, data deduplication, and compression
- VMware vStorage API for Array Integration (VAAI)
- Multi-tenancy
- Host Rate Limiting: QoS
- Fibre Channel and iSCSI support
- Snapshots
- Synchronous and asynchronous remote mirroring
- Data Migration
- Hyper-Scale Mobility
- Encryption
- Authentication by using Lightweight Directory Access Protocol (LDAP)
- OpenStack and REST support
- VMware synergy

IBM FlashSystem A9000 is a fixed storage solution to provide up to 300 TB of effective capacity by using 8U of rack space.

IBM FlashSystem A9000R is a scalable storage solution into a single rack from 300 TB to 1800 TB of effective capacity.

Both FlashSystem A9000 and FlashSystem A9000R use the same firmware, and both offer onsite setup and service that are provided by IBM. They also share a feature set.

**For more information:**

*IBM FlashSystem A9000 Product Guide*, REDP-5325

*IBM FlashSystem A9000 and IBM FlashSystem A9000R Architecture, Implementation, and Usage*, SG24-8345

## 4.5.2 IBM ProtecTIER

This section describes the IBM TS7650G ProtecTIER® Deduplication Gateway solutions. IBM ProtecTIER provides global data deduplication for multiple attached systems even in heterogeneous environments.

Data deduplication is a key technology to dramatically reduce the amount of, and the cost associated with, storing large amounts of data by consolidating redundant copies of a file or file subset. Incoming or existing data is standardized into “chunks” that are then examined for redundancy. If duplicates are detected, pointers are shifted to reference a single copy of the chunk and the extraneous duplicates are then released.

As mentioned, business data is growing at an exponential rate and backup windows are typically shrinking. ProtecTIER helps you to back up and recover data quickly. The ProtecTIER application is available as part of a gateway-based solution and as an appliance with integrated storage. The cornerstone of ProtecTIER is IBM HyperFactor®, an IBM technology that deduplicates data inline as it is received from the source. HyperFactor is based on a series of algorithms that identify and filter out the elements of a data stream that were previously stored by ProtecTIER. This search is quick because it uses a small and efficient memory-resident index. Over time, HyperFactor considerably increases the usable capacity of your physical storage. With ProtecTIER native replication, the data reduction value of HyperFactor is extended to bandwidth and storage savings for disaster recovery operations.

IBM ProtecTIER offers two different types of interface for front end connectivity:

- Virtual tape library (VTL) mode
- File System Interface (FSI) mode

With VTL mode, ProtecTIER provides true dual-node active-active clustering for higher availability and higher performance.

VTL mode allows the emulation of different types of tape libraries and tape drives, most prominently the IBM TS3500 tape library with IBM LTO3 drives. These libraries and drives can perfectly interface into any backup application that supports this mode.

FSI offers CIFS and NFSv3 connectivity at the same time. Using IP-based connectivity, ProtecTIER can either be used as a target for backup applications or can be directly addressed as a backup target by applications like Oracle.

As schematically shown in Figure 4-32, ProtecTIER native IP replication uses smart algorithms to maximize the use of available network resources to replicate data between different remote sites. Even small portions of data are replicated with maximum efficiency.

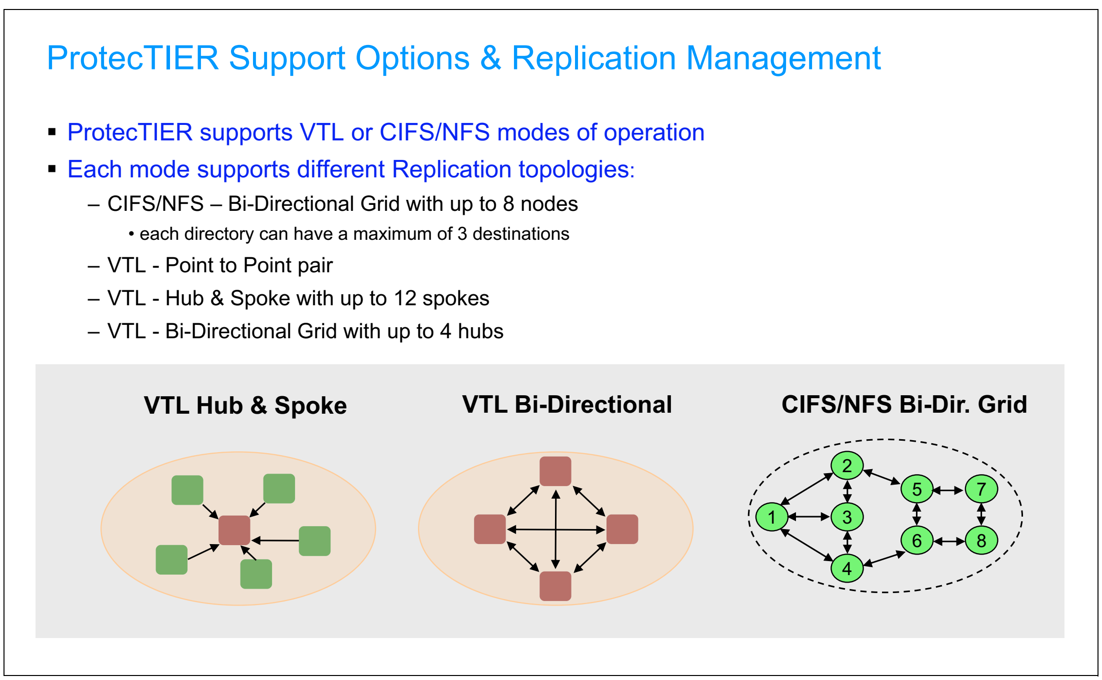



*Figure 4-32   ProtecTIER implementation concept with native IP replication*

The ProtecTIER native replication algorithm is designed to make the best use of unreliable low-bandwidth IP connections.

For more information, see IBM TS7650G ProtecTIER Deduplication Gateway at:

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

## 4.5.3 IBM TS4500 and TS3500 tape libraries

Spectrum Archive Enterprise Edition, along with Spectrum Scale, allows you to connect tape libraries to the cloud. Tape has always been a cost effective solution. With the growth in data controlled by cloud environments, TS4500 and TS3500 can be used as an efficient tape tier.

Figure 4-33 shows how a TS4500 or TS3500 tape library can be configured as the tape tier in the storage cloud through Spectrum Scale and Spectrum Archive Enterprise Edition.

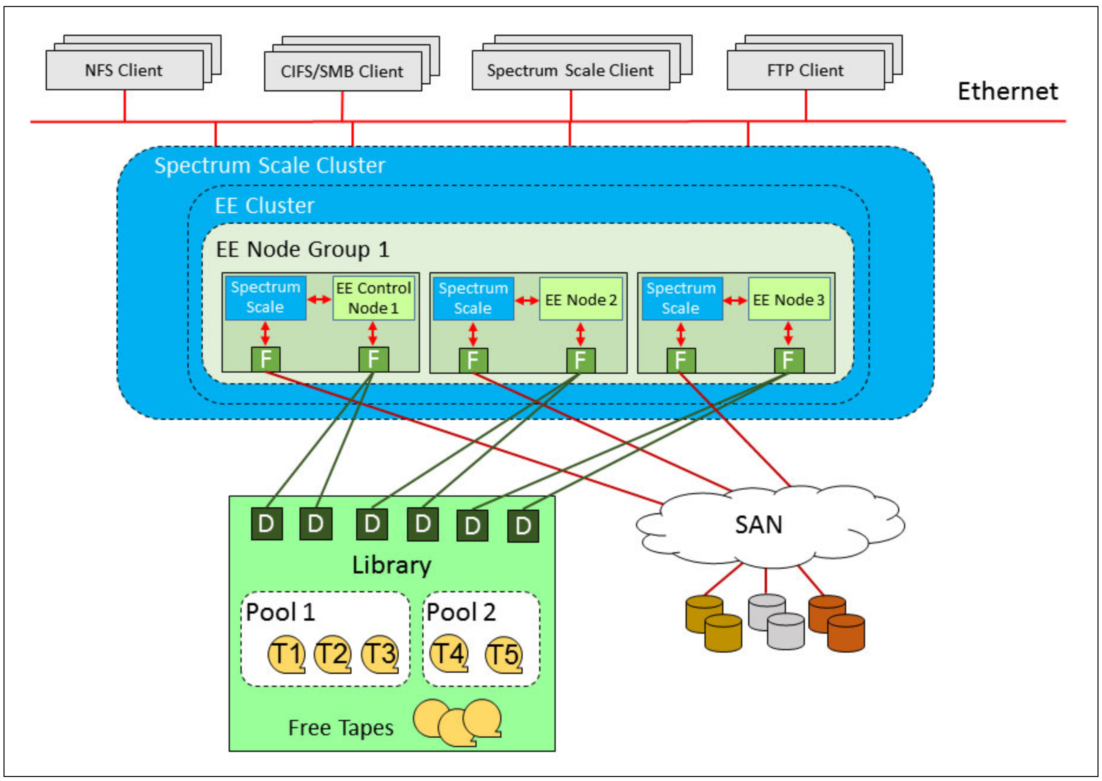



*Figure 4-33 TS4500/TS3500 tape library tape tier configuration for cold storage*

### TS4500 tape library

The IBM TS4500 tape library is a next-generation storage solution that is designed to help midsize and large enterprises respond to storage challenges. Among these challenges are high data volumes and the growth in data centers. These factors in turn increase the cost of data center storage footprints, the difficulty of migrating data across vendor platforms, and increased complexity of IT training and management as staff resources shrink.

In the TS4500, IBM delivers the density that today's and tomorrow's data growth requires, along with the cost efficiency and the manageability to grow with business data needs while preserving existing investments in IBM tape library products. You can now achieve both a low cost per terabyte (TB) and a high TB density per square foot. The TS4500 can store up to 5.5 PBs of uncompressed data in a single 10-square foot library frame and up to 175.5PBs of uncompressed data in a 17 frame library.

The TS4500 tap library has these additional highlights:

- Improve storage density with more than two times the expansion frame capacity and support for 33 percent more tape drives
- Proactively monitor archived data with policy-based automatic media verification
- Improve business continuity and disaster recovery with automatic control path and data path failover
- Help ensure security and regulatory compliance with tape-drive encryption and Write Once Read Many (WORM) media
- Support Linear Tape-Open (LTO) Ultrium 7, Ultrium 6, LTO Ultrium 5, and IBM TS1150 and TS1140 tape drives
- Increase mount performance and overall system availability with dual robotic accessors
- Provide a flexible upgrade path for users who want to expand their tape storage as their needs grow
- Reduce the storage footprint and simplify cabling with 10U of rack space on top of the library

For more information, see the *IBM TS4500 R3 Tape Library Guide*, SG24-8235.

### TS3500 tape library

TS3500 continues to lead the industry in tape drive integration with features such as persistent worldwide name, multipath architecture, drive/media exception reporting, remote drive/media management, and host-based path failover.

The IBM TS3500 tape library is designed to provide a highly scalable, automated tape library for mainframe and open-systems backup and archive. The library can scale from midsize to large enterprise environments. Here are the highlights of the tape library:

- Support highly scalable, automated data retention on tape by using LTO Ultrium and IBM 3592 tape drive families
- Deliver extreme scalability and capacity, growing from one to 16 frames per library and from one to 15 libraries per library complex
- Provide up to 2.25 exabytes (EB) of automated, low-cost storage under a single library image, improving floor space utilization and reducing storage cost per TB with IBM 3592 JD enterprise advanced data cartridges (10 TB native capacity)

Massive scalability (300,000+ LTO tape cartridges) can be achieved with the TS3500 Shuttle Complex as shown in Figure 4-34.

*Figure 4-34 TS3500 Shuttle Complex moves tapes cartridges between physical libraries*

For more information, see *IBM Tape Library Guide for Open Systems*, SG24-5946.

### 4.5.4 IBM DS8880

The IBM DS8000 series is the flagship disk storage system within the IBM System Storage portfolio.

The IBM DS8880 family now offers business-critical, all-flash, and hybrid data systems that span a wide range of price points:

- The IBM DS8888 is an all-flash offering. It scales to 192 gigabytes of raw Flash Card capacity in a two-rack footprint.
- The IBM DS8884 and IBM DS8886 are two high-performance hybrid models scaling to more than 4.6 petabytes (PB) of raw drive capacity. Nine types of drives can be managed in up to three different tiers (Flash cards and flash drives, SAS and Nearline SAS drives).

IBM enhanced the OpenStack Cinder driver with DS8880 support.

Integration of storage systems requires an OpenStack Block Storage driver on the OpenStack Cinder nodes. The driver is an IBM proprietary solution that supports OpenStack Block Storage on top of the OpenStack and Cinder open source technologies.

Figure 4-35 illustrates how Horizon (Dashboard) and Nova (Compute) interact with Cinder over the Ethernet control-path. It also illustrates the data path from Nova to the DS8880 storage system.

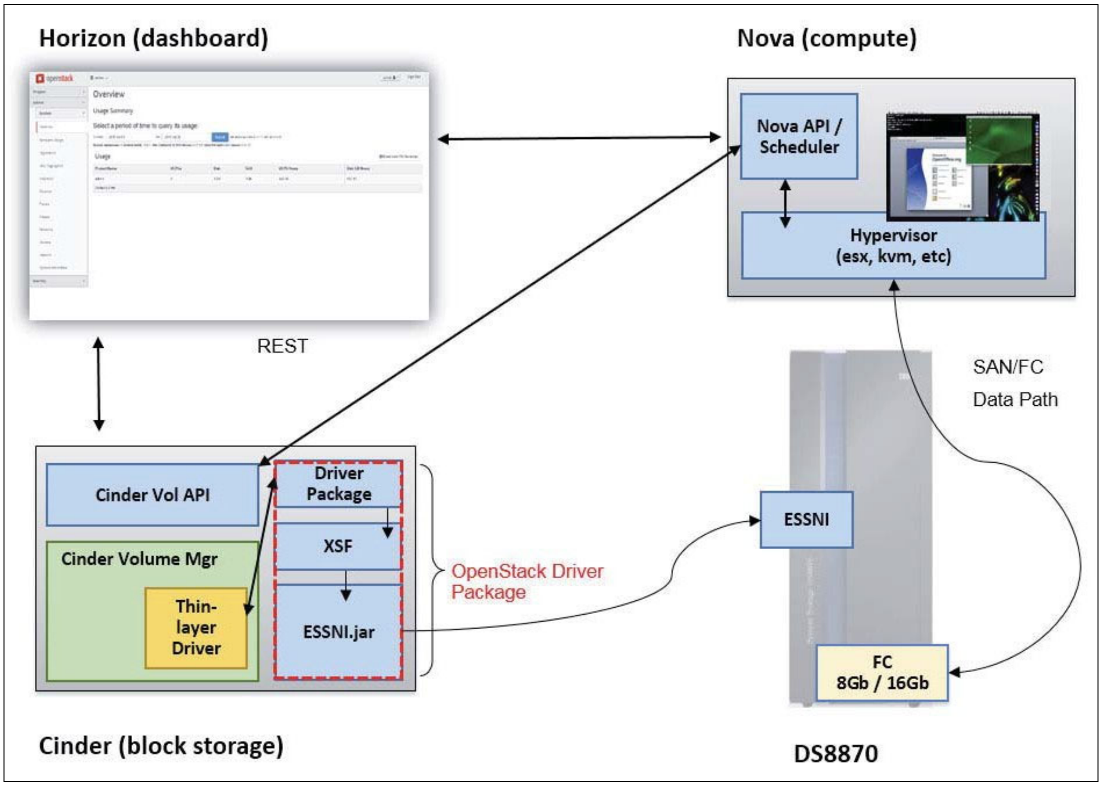



*Figure 4-35 OpenStack Cinder driver support for DS8880*

With the availability of the IBM Storage Driver for the OpenStack Cinder component, the IBM DS8880 storage system can now extend its benefits to the OpenStack cloud environment.

The IBM Storage Driver for OpenStack Cinder enables OpenStack clouds to access the DS8880 storage system. The IBM Storage Driver for OpenStack is fully supported by Cinder and provides block storage as a service through iSCSI and Fibre Channel to VMs. Cloud users can send requests for storage volumes from the OpenStack cloud. These requests are routed to, and transparently handled by the IBM Storage Driver. The IBM Storage Driver communicates with the DS8880 storage system and controls the storage volumes on it.

The last version of Cinder Driver Mikata (1.7.0) provides the following capabilities:

- Create/Delete Volume
- Volume Attach/Detach (through Nova Compute)
- Snapshots, Clones (FlashCopy with background copy)
- Backups (Copy Volume Images to Object Store)
- Swift, Ceph, and TSM Support
- Volume Types, Volume Retype
- Quotas
- Consistency Groups for FlashCopy

- Volume Retype: Ability to change the type of an existing Cinder volume to a new tier, add capabilities, and so on
- Volume Replication: Ability to do synchronous replication of Cinder volumes between two storage subsystems

IBM also provides Storage integration between VMware and DS8000. See Figure 4-36.

DS8880 supports VMware VAAI:

- Integration with vStorage APIs to improve performance
- **Full copy** (also known as XCOPY) primitive offloads work from production virtual servers to storage, which helps reduce host and SAN resource utilization
- **Hardware-assisted locking** (through Atomic Test & Set) primitive enables a finer grained level of locking (block-level instead of LUN-level) on VMFS metadata, which is more efficient and also scales better in larger VMware clusters
- **Write Same** (zero blocks) primitive offloads the process of zeroing the VMDK to the storage subsystem

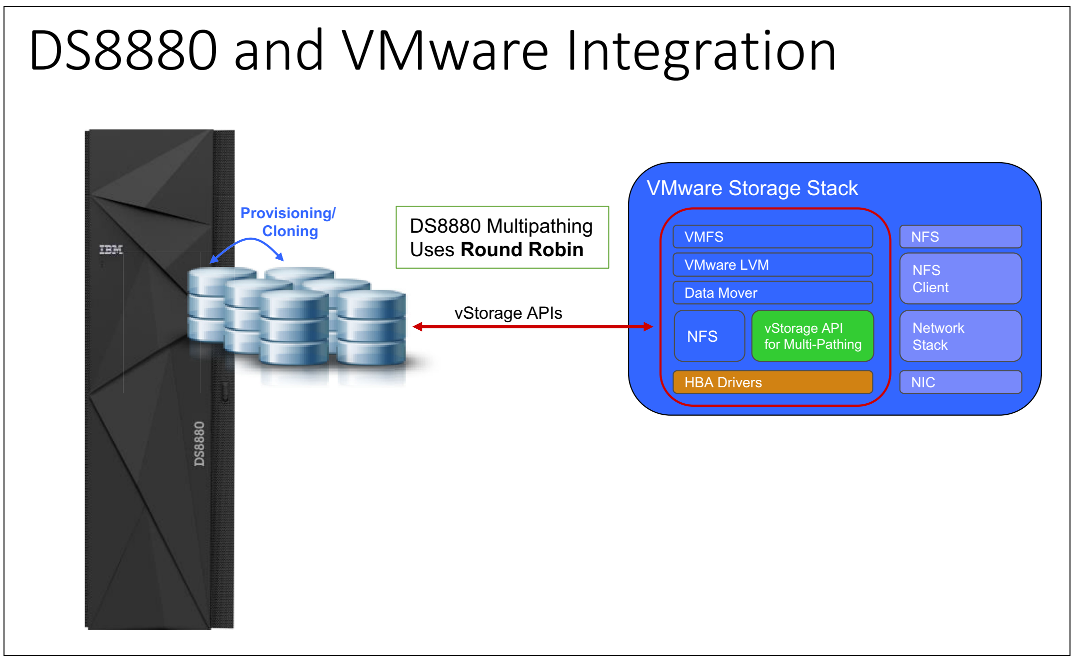



*Figure 4-36 Storage integration between VMware and DS8000*

Example tasks that can benefit from improved performance:

- VM creation/cloning/snapshots/deletion
- vMotion and storage vMotion
- Extending a VMFS Volume
- Extending the size of a VMDK file

DS8880 also supports these items:

- vCenter plug-in
- VMware vCenter Site Recovery Manager (SRM)
- VMware Web Client

For more information see the *DS8880 Product Guide (Release 8.2)*, REDP-5344.

For information about using the DS8000 in an OpenStack environment, see *Using IBM DS8870 in an OpenStack Environment*, REDP-5220.

## 4.5.5 IBM Storwize family

Designed for software-defined environments, the IBM Storwize family includes technologies that both complement and enhance virtual environments, as well as built-in functions such as Real-time Compression and Easy Tier technology that deliver extraordinary levels of efficiency. Available in a wide range of storage systems, the Storwize family delivers sophisticated capabilities that are easy to deploy, and help to control costs for growing businesses.

The IBM Storwize family consists of the IBM SAN Volume Controller, IBM Storwize V7000 and V7000 Unified, IBM Flex System® V7000 Storage Node, IBM Storwize V5000, IBM Storwize V3700, and the all flash memory systems including V7000F and V5000F. Benefits of the Storwize family include high-performance thin provisioning, Real-time Compression, IP replication, Easy Tier, an advanced GUI, and storage virtualization.

The Storwize family uses IBM Spectrum Virtualize software the same proven software as SAN Volume Controller, and provides the same interface and similar capabilities across the product line.

### Storwize V7000 and V7000 Unified

These are highly scalable virtualized, enterprise-class, flash-optimized storage systems that are designed to consolidate workloads into a single system for ease of management, reduced costs, superior performance, and high availability.

IBM Storwize V7000 Unified is a virtualized storage system that is designed to consolidate block and file workloads into a single storage system. This configuration provides simplicity of management, reduced cost, highly scalable capacity, performance, and high availability. IBM Storwize V7000 Unified storage also offers improved efficiency and flexibility through built-in SSD optimization, thin provisioning, and nondisruptive migration of data from existing storage. The system can virtualize and reuse existing disk systems, offering a greater potential return on investment (ROI). Integrated IBM Active Cloud Engine enables you to use all those features to build your storage cloud.

Figure 4-37 describes how IBM Storwize V7000 Unified uses the best of IBM storage technologies.

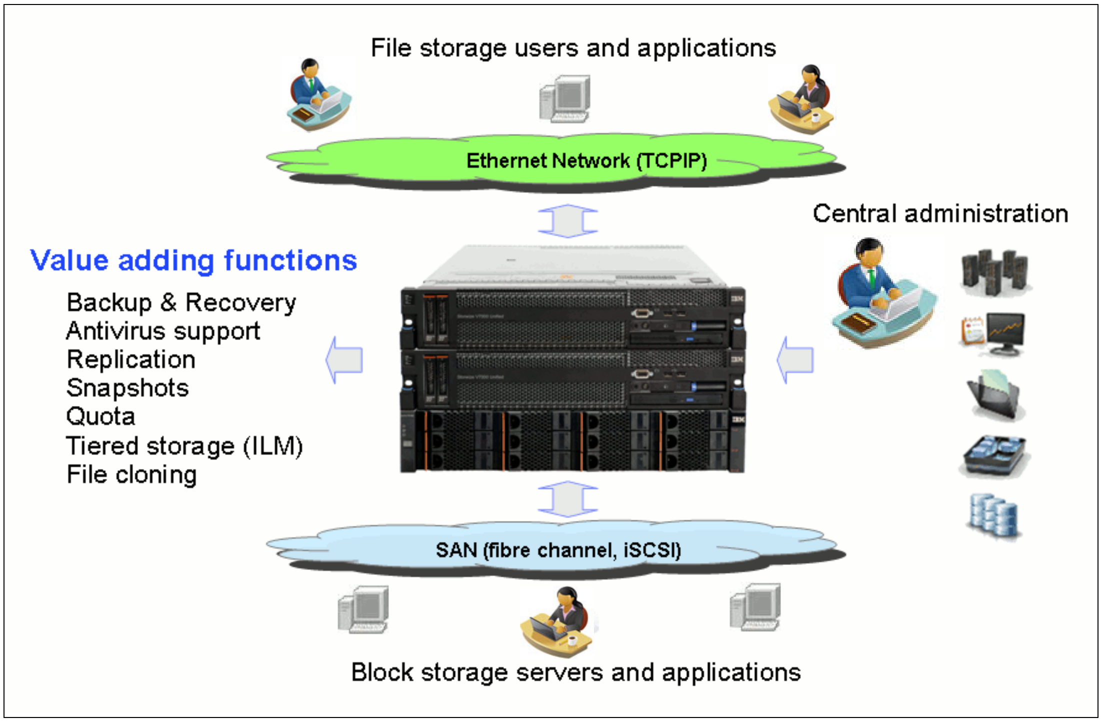



*Figure 4-37 IBM Storwize V7000 Unified*

### *IBM Storwize V7000 Unified highlights*

By using the proven technology of IBM Storwize V7000, and extending it with file capabilities and IBM Active Cloud Engine, the IBM Storwize V7000 Unified becomes an essential building block in storage cloud implementations.

IBM Storwize V7000 Unified provides the following features:

- Unified storage management
- Block capabilities
- File capabilities

### *Unified storage management*

IBM Storwize V7000 Unified provides a single GUI to manage both your block and file-level storage. This approach provides a streamlined user interface regardless of your workload. It also helps previous IBM Storwize V7000 users to easily provision file storage without needing to learn a new interface. Installation of IBM Storwize V7000 Unified is also simplified with the use of a Storwize USB key. This key enables you to quickly get the system running and start provisioning your storage, whether it is a block or file. The unified interface also provides coordinated monitoring and reporting facilities, which can be enhanced with IBM Spectrum Control.

For more information, see the IBM Spectrum Control web page:

https://www.ibm.com/developerworks/servicemanagement/sm/spectrum_control/index.html

### *Block capabilities*

IBM Storwize V7000 Unified provides storage efficiency by providing virtualization, thin provisioning, and Easy Tier functions to help you consolidate and optimize your storage. Data protection is provided with the standard and proven set of features that include FlashCopy, Metro and Global Mirror, and volume mirroring. IBM Storwize V7000 Unified also provides flexibility and investment protection, with the ability to virtualize your existing environment and perform online volume migrations without disrupting business continuity.

### *File capabilities*

File module software is based on IBM common network-attached storage (NAS) software and roadmap, which enables faster delivery of file functions and file ISV certifications across multiple products. IBM Active Cloud Engine is included to reduce costs through policy-based management of files and use of tiered storage, and to improve data governance. Information lifecycle is provided with automated movement of less frequently used files to lower tiers of storage, including tape in an IBM Spectrum Protect system. File replication, backup and recovery, and snapshot features are provided to extend your data protection to the file level. Data is also protected by providing antivirus capabilities with the Antivirus Connector, which integrates with the external antivirus scan nodes. Files can be provisioned to various host types by using open file protocols such as NFS, CIFS, FTP, HTTPS, and SCP.

## IBM SAN Volume Controller as an SDS appliance

Built with IBM Spectrum Virtualize software, IBM SAN Volume Controller (SVC) is a storage virtualization system that enables a single point of control for storage resources to help support improved business application availability and greater resource utilization. The objective is to manage storage resources in your IT infrastructure and to make sure that they are used to the advantage of your business. You also want to do it quickly, efficiently, and in real time, while avoiding increases in administrative costs.

IBM Spectrum Virtualize software in SVC helps make new and existing storage more effective. SVC includes many functions that are traditionally deployed separately in disk systems. By including these functions in a virtualization system, SVC standardizes functions across virtualized storage for greater flexibility and potentially lower costs.

SVC systems can handle the massive volumes of data from mobile and social applications, enable rapid and flexible cloud services deployments, and deliver the performance and scalability that is needed to gain insights from the latest analytics technologies.

### *IBM SVC highlights*

Improving efficiency and delivering a flexible, responsive IT infrastructure are essential requirements for any cloud deployment. Key technologies for delivering this infrastructure include virtualization, consolidation, and automation. SAN Volume Controller provides these technologies to help you build your storage cloud.

SAN Volume Controller’s enhanced storage capabilities with sophisticated virtualization, management, and functions have these benefits:

- Enhance storage functions, economics, and flexibility with sophisticated virtualization
- Employ hardware-accelerated data compression for efficiency and performance
- Use encryption to help improve security for data on existing storage systems
- Move data among virtualized storage systems without disruptions
- Optimize tiered storage, including flash storage, automatically with IBM Easy Tier
- Improve network utilization for remote mirroring and help reduce costs
- Implement multi-site configurations for high availability and data mobility

Figure 4-38 shows an overview of the IBM SAN Volume Controller architecture.

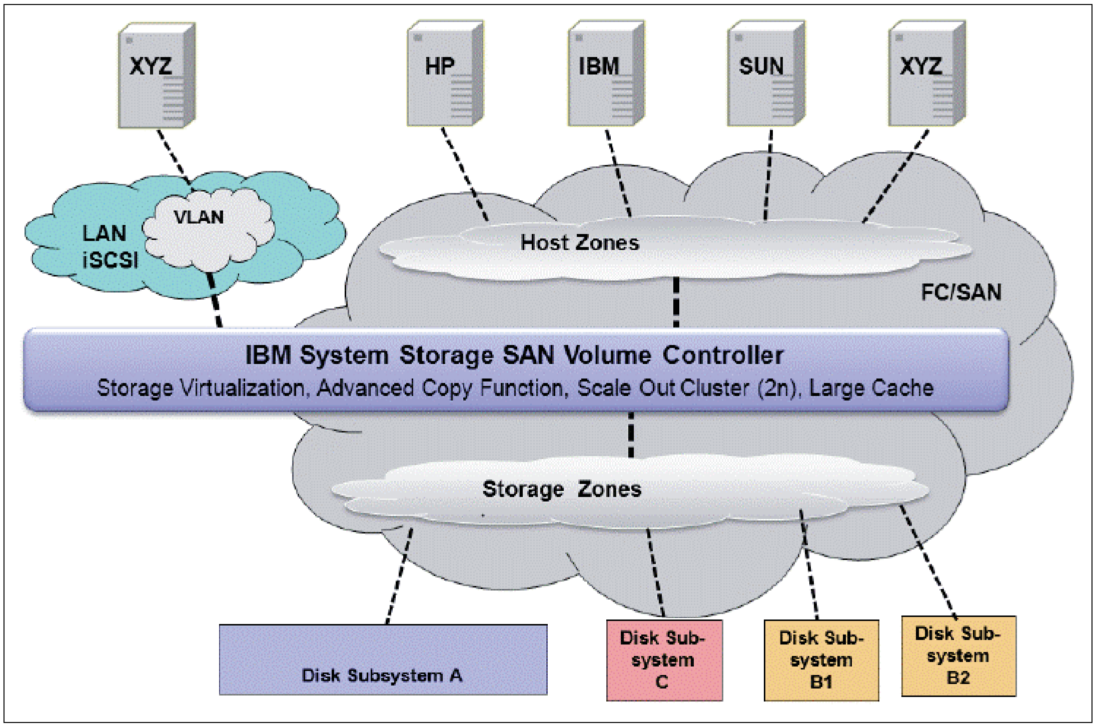



*Figure 4-38 SAN Volume Controller overview*

### Ongoing IBM contributions to OpenStack Cinder for the Storwize family

The Storwize family, which includes SAN Volume Controller, supports these items:

- Folsom, Grizzly, Havana, Icehouse, Juno
- iSCSI and Fibre Channel
- Advanced Storwize features such as Real-time Compression and Easy Tier
- Software-defined placement by using OpenStack filter scheduler
- Storage-assisted volume migration (Storwize family is the only storage to support this function in Havana)

With the OpenStack Havana release, a new administrator feature for migrating volumes between Cinder instances was added. Volumes can be migrated with Host Assisted Data Migration or by Storage Assisted Data Migration with the IBM Storwize family.

Figure 4-39 lists the common use cases for migrating volumes.

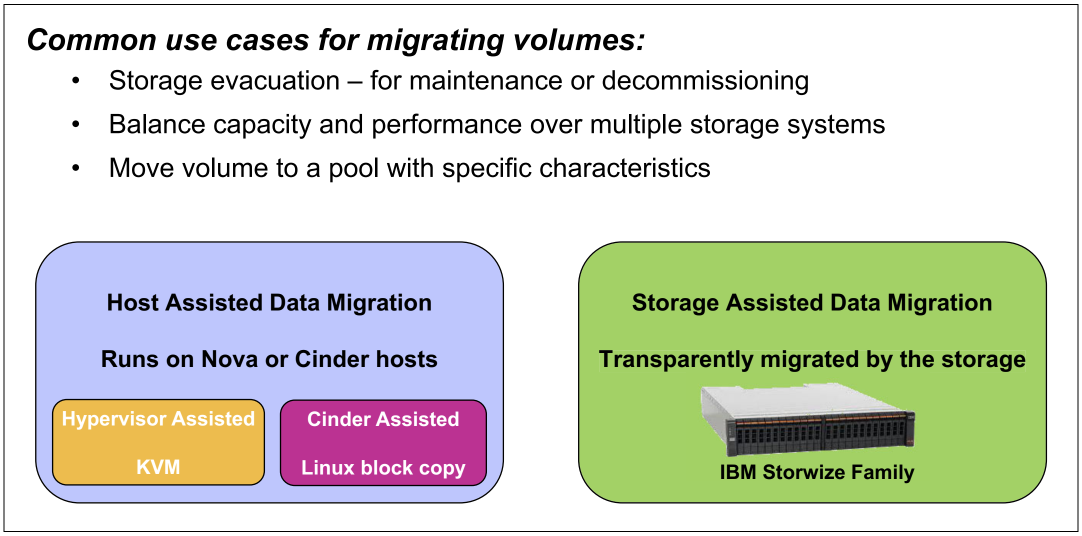



*Figure 4-39 Common use cases for volume migration in OpenStack environment*

IBM Storwize family is the only storage in the Havana release to support storage assisted migration. Volumes move between two storage pools that are managed by a Storwize family system.

These are the key benefits to using the Storwize family storage assisted migration:

- No interaction with the host
- No impact on VM and node
- Instantaneous
- No effect on VM operations or volume management

## 4.6 IBM cloud services

The previous chapters described how IBM can help you build a storage cloud on your storage cloud journey. IBM continues to rapidly expand our major commitment to cloud technologies and service offerings that allow our customers to enjoy the benefits of a storage cloud without having to build their own infrastructure. IBM provides these offerings as part of the IBM SmartCloud family, supporting public, hybrid, and private cloud delivery models.

Figure 4-40 shows IBM SoftLayer and IBM Cloud Managed Services offerings for public, private, and hybrid clouds.

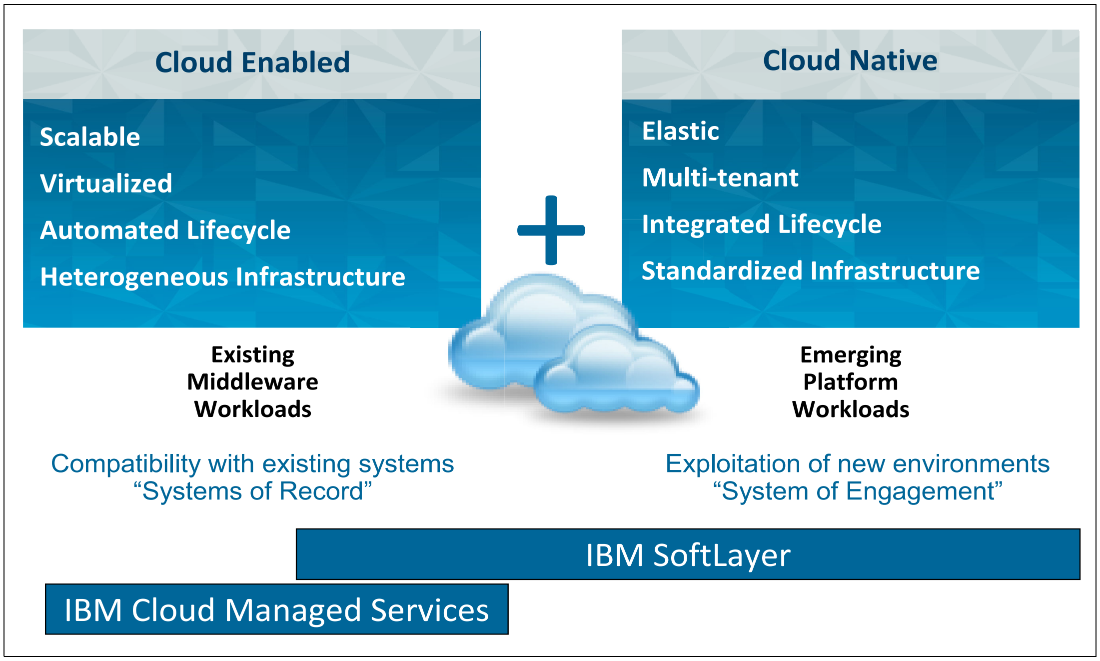



*Figure 4-40 IBM SoftLayer and IBM CLoud Managed Services Offerings*

Figure 4-40 shows the SoftLayer and Cloud Managed Services strengths to highlight the broad scope of the two offerings. The overlapping bars at the bottom of the figure illustrate that there are functional similarities between the offerings.

You can think of IBM Cloud Managed Services as ideal for these use cases:

- A highly built out and resilient, steady workload, transaction-oriented IT infrastructure
- Traditional three tier IT application models that consist of application, middleware, and database
- SAP, Oracle ERP solutions, and similar database management systems

You can think of IBM SoftLayer as ideal for these use cases:

- Cloud native and internet native applications
- Associated requirements for internet scale, widely varying and quickly changing workload levels
- Modern internet applications where resiliency is built into the application layer
- Web-centric, native cloud programming environments
- Bare metal server capability, allowing the client the option and flexibility to load their own software stack to meet their needs
- Full self-service API control of all aspects of the client’s SoftLayer infrastructure to allow the flexibility and control access to manage their SoftLayer infrastructure however the client wants

Within each of these two complementary IBM Cloud Services offerings are many individualized offerings, which include (but are not limited to) the following:

- Private Modular Cloud: Patterns of expertise that can be purchased and applied to your private, public, or hybrid cloud.
- IBM PureSystems® as a Service: New ability to purchase portions of the IBM PureSystems software stack running in off-premises, pay-per-use version in IBM SoftLayer.
- Software as a service: Off-premises, pay-per-use ability uses IBM software products inside SoftLayer that previously only existed in on-premises versions. Examples include IBM General Parallel File System.

Regarding positioning IBM Cloud Managed Services and IBM SoftLayer, see Figure 4-41 for a short description of the high-level characteristics of each.

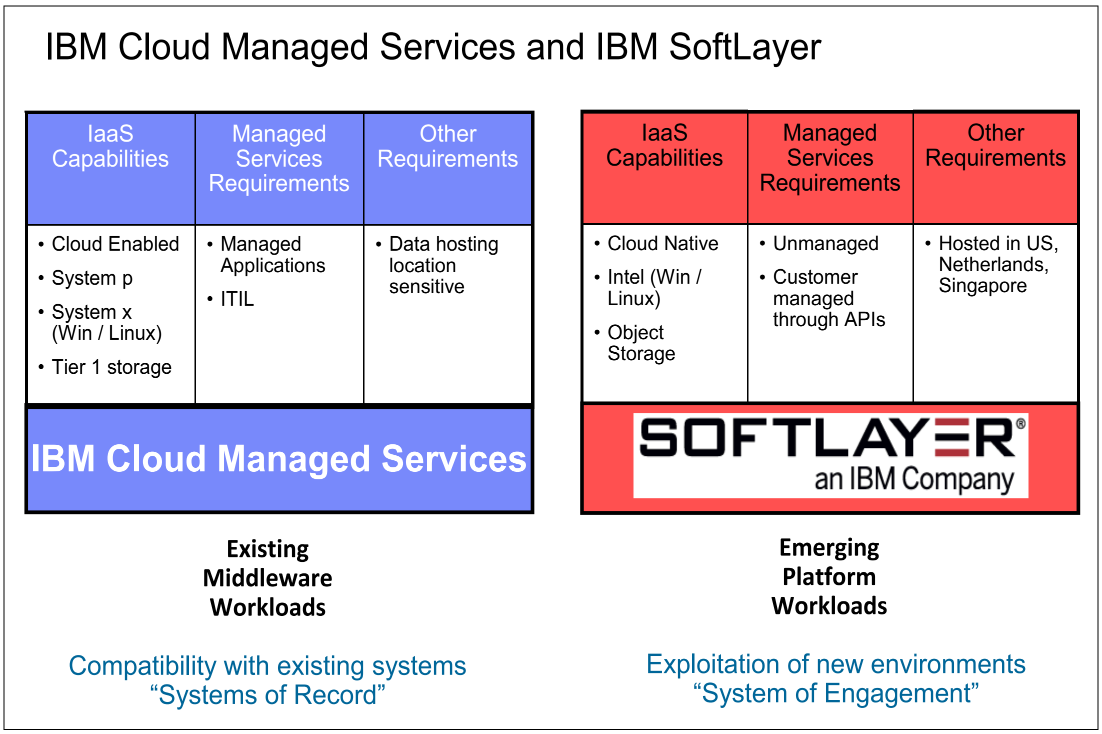



*Figure 4-41 IBM Cloud Managed Services and IBM SoftLayer characteristics*

Figure 4-41 highlights the major characteristics of IBM Cloud Managed Services and SoftLayer to show the breadth of the offerings together. For example, the Managed Services Requirements column shows the focus areas of Cloud Managed Services and SoftLayer. Unmanaged is displayed for SoftLayer because they also have options for managed services.

### 4.6.1 IBM SoftLayer

SoftLayer offers storage that is attached to compute servers, and also stand-alone SaaS. SoftLayer provides a complete object storage solution with OpenStack Swift that includes powerful tagging, search, and indexing capabilities. This solution allows you to assign rich metadata tags for ease of finding and serve objects when requested. IBM is establishing SoftLayer as the foundation of the IBM cloud portfolio. It is the scalable, secure base for the global delivery of cloud services that span the IBM middleware and SaaS solutions. SoftLayer's flexibility and global network also facilitate faster development, deployment, and delivery of mobile, analytic, and social solutions as you adopt cloud as a delivery platform for IT operations and managing your business.

SoftLayer offers the following features:

- Complete self-service capability to acquire, spin up, allocate, and de-allocate IT infrastructure for public, private, and hybrid clouds
- Wide choice and flexibility in options in the specific infrastructure to be provisioned, including bare metal server capability
- APIs are provided to help IT users and administrators manage all aspects of the SoftLayer provisioned infrastructure

For more information about IBM SoftLayer, see the following website:

http://www.softlayer.com/

SoftLayer is designed to support an automated cloud environment, from private dedicated servers (including bare metal) to shared (public) multi-tenant model, and provides pay-as-you-go capabilities.

### 4.6.2 IBM Cloud Managed Services

IBM Cloud Managed Services is a multi-tenant, IBM hosted, IaaS cloud delivery offering that provides fully managed, highly secure cloud compute environment that is optimized for production applications. Cloud Managed Services provides operating system management, patching, backup, and security with selectable service level agreements (SLAs) on a per virtual server basis up to 99.95%.

Cloud Managed Services offers three storage options that include Flash Storage (IBM FlashSystem 900), High Performance (IBM XiV), and Base Storage (IBM V7000) to meet even the most demanding storage requirements of cloud workloads. Individual virtual servers in Cloud Managed Services can support up to 96 TBs of storage on AIX, and 48 TBs on Microsoft Windows and Linux. Other storage features include availability zones, local and remote mirroring for high availability and disaster recovery implementations, FlashCopy, shared disks, and robust backup and restore capabilities.

IBM Cloud Managed Services provides these features and benefits:

- Managed services including OS systems administration integrated into every VM
- High availability and continuity services including alternative site disaster recovery
- Best-of-class security services isolate data and servers, which provides protection from outside threats
- Enhanced security services to protect Payment Card Industry (PCI) or Health Insurance Portability and Accountability Act (HIPAA) regulated environments
- Ability to deploy applications globally in any of IBM's 13 locations

- Option for a dedicated, centrally managed cloud at an IBM, business partner, or client data center
- Solutions for SAP, SAP HANA, Oracle applications, and other ERP/CRM workloads
- Easily create hybrid cloud environments with SoftLayer and Cloud Managed Services by using SoftLayer Interconnect Network

Learn more about IBM Cloud Managed Services:

http://www.ibm.com/services/us/en/it-services/cloud-services/cloud-managed-services/index.html
http://www.ibm.com/marketplace/cloud/managed-cloud/us/en-us

For information about other storage-related IBM cloud service offerings, see the following websites:

- IBM Cloud Managed Backup

  http://www.ibm.com/services/us/en/it-services/business-continuity/cloud-managed-backup
- IBM Data Vault

  http://www.ibm.com/services/us/en/it-services/managed-data-vault.html
- IBM Storage Services

  http://www.ibm.com/services/us/en/it-services/smart-business-storage-cloud.html
- IBM Federal Community Cloud

  https://www.ibm.com/shop/americas/content/home/en_US/government-contracts.html

### IBM Cloud Automated Modular Management

IBM Cloud Automated Modular Management allows you to manage your SoftLayer cloud environment with automated management tools.

With IBM Cloud Automated Modular Management, you can transfer the complexity and challenges of monitoring and managing an increasingly complex environment to a trusted service provider. IBM offers discreet IT management components that are available as self-managed or fully managed services that use policy-based automation to help you significantly reduce management costs:

- Reduced IT administration costs with automation, improved visibility, and efficient monitoring features
- Improved utilization and quality of service by using reports on health and usage of resources and applications
- Simplified business operations with fulfillment, billing, and management of this service with your SoftLayer account

For more information, see:

https://www.ibm.com/support/knowledgecenter/SSVV6P/ammintro.html

### IBM Cloud Transformation Services

Cloud transformation services from IBM helps you adopt cloud the correct way. The services can align your cloud and business strategies, applying a more realistic framework and creating a detailed roadmap for implementation. These services help you optimize your potential for transformation while also reducing complexities and risks.

IBM seasoned cloud and strategy specialists help you with a five-step approach to cloud transformation:

- A cloud envisioning workshop
- Opportunity prioritization
- Roadmap creation
- Cloud readiness and governance assessment
- Adoption plan creation

These steps cover all quadrants of the IBM Cloud Strategy Framework. For more information, see:

http://www.ibm.com/services/us/en/it-services/cloud-services/cloud-transformation-services/index.html

## 4.6.3 IBM Bluemix

IBM Bluemix is the IBM open cloud platform that provides mobile and web developers access to IBM software for integration, security, transaction, and other key functions. It also includes software from IBM Business Partners.

Built on Cloud Foundry open source technology, Bluemix makes application development easier with platform as a service (PaaS). Bluemix also provides pre-built mobile backend as a service (MBaaS) capabilities. It simplifies the delivery of an application by providing services that are ready for immediate use and hosting capabilities to enable internet scale development.

With Bluemix, you can access the public Bluemix platform, set up a dedicated Bluemix platform, or both.

At its core, Bluemix is an environment for you to develop apps and use services that provide ready-to-use functions. Bluemix also provides an environment to host application artifacts that run on an application server such as Liberty. By using SoftLayer, Bluemix deploys virtual containers that host each deployed app. In this environment, the app can use pre-built services (including third-party services) to make app assembly easy.

As a developer, you can interact with the Bluemix infrastructure by using a browser-based user interface. You can also use a Cloud Foundry command-line interface, called **`cf`**, to deploy web apps.

Clients, which can be mobile apps, apps that run externally, apps that are built on Bluemix, or developers that are using browsers, interact with the Bluemix-hosted apps. Clients use REST or HTTP APIs to route requests through Bluemix to one of the app instances or the composite services.

Figure 4-42 shows the high-level Bluemix architecture.

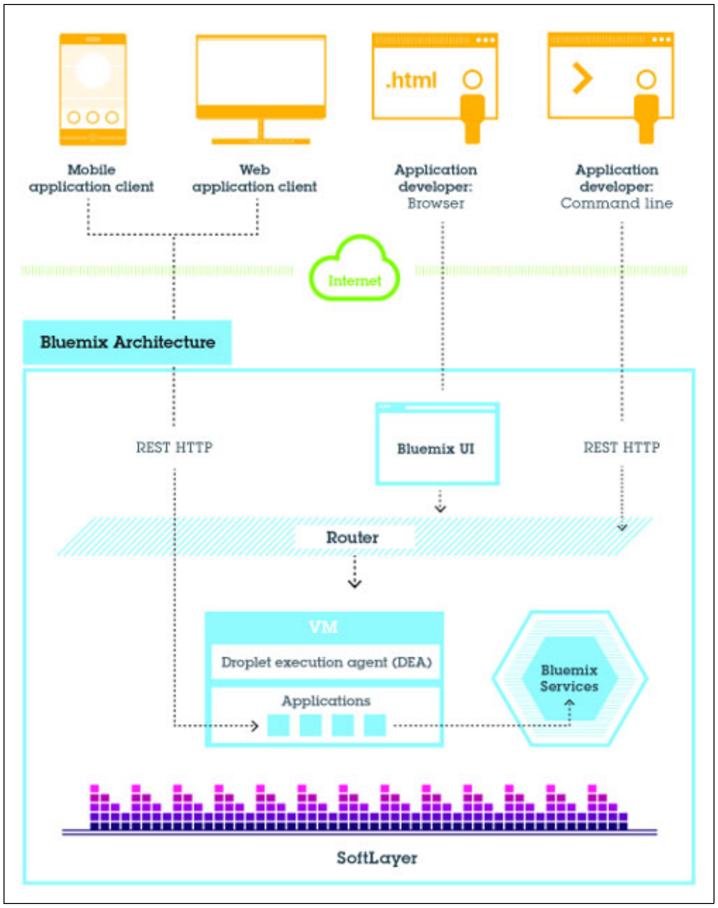



*Figure 4-42 Bluemix architecture*

Learn more about Bluemix at these web pages:

- IBM Bluemix:

  https://www.ibm.com/cloud-computing/bluemix

- Build and run your app on the Bluemix cloud platform at:

  https://ibm.biz/BdXkEe

- What is IBM Bluemix? in IBM developerWorks:

  http://www.ibm.com/developerworks/cloud/library/cl-bluemixfoundry

Learn more about Cloud Foundry at:

https://www.cloudfoundry.org

5

# What are others doing in the journey to storage cloud

Storage cloud is a reality within numerous infrastructures worldwide. IBM clients have successfully implemented storage hybrid cloud solutions across industries in large enterprises and in small and medium businesses to improve IT agility and support for business process requirements, while controlling costs. Proven technology with skilled implementation assistance and services make IBM a leader in smart storage cloud deployment.

This chapter describes hybrid storage cloud solutions across various industries. Each description covers the client's needs, proposed solution, and results.

This chapter includes the following sections:

- Large enterprise OpenStack cloud supporting XIV
- IT storage cloud orchestration
- National library public cloud
- Video surveillance solution for public safety
- Telecommunication company public cloud
- Life science healthcare hybrid cloud
- University disaster recovery on public cloud
- Media and entertainment company hybrid cloud

## 5.1 Large enterprise OpenStack cloud supporting XIV

A large internet commerce company provides online marketplace services to online buyers and sellers. One of the core environments provides cloud-based e-commerce capabilities to commercial-level users of the marketplace. These capabilities include hosting, website building, transaction and order fulfillment, and customer relationship management tools.

### 5.1.1 Business needs

To remain competitive in the very fast pace change of the worldwide internet commerce marketplace, this customer needs to continuously improve their functional services provided to their users. During this process, they must maintain absolute reliability and consistency of service regardless of Internet-scale peaks and spikes in workloads, all while operating at competitive Internet-scale costs of operation.

### 5.1.2 Proposed solution

To provide these capabilities, this company chooses OpenStack software as the cloud computing platform on which to provide cloud services to marketplace commercial customers. OpenStack cloud software is implemented on a large and growing commercial enterprise class Intel server farm, as shown in Figure 5-1.

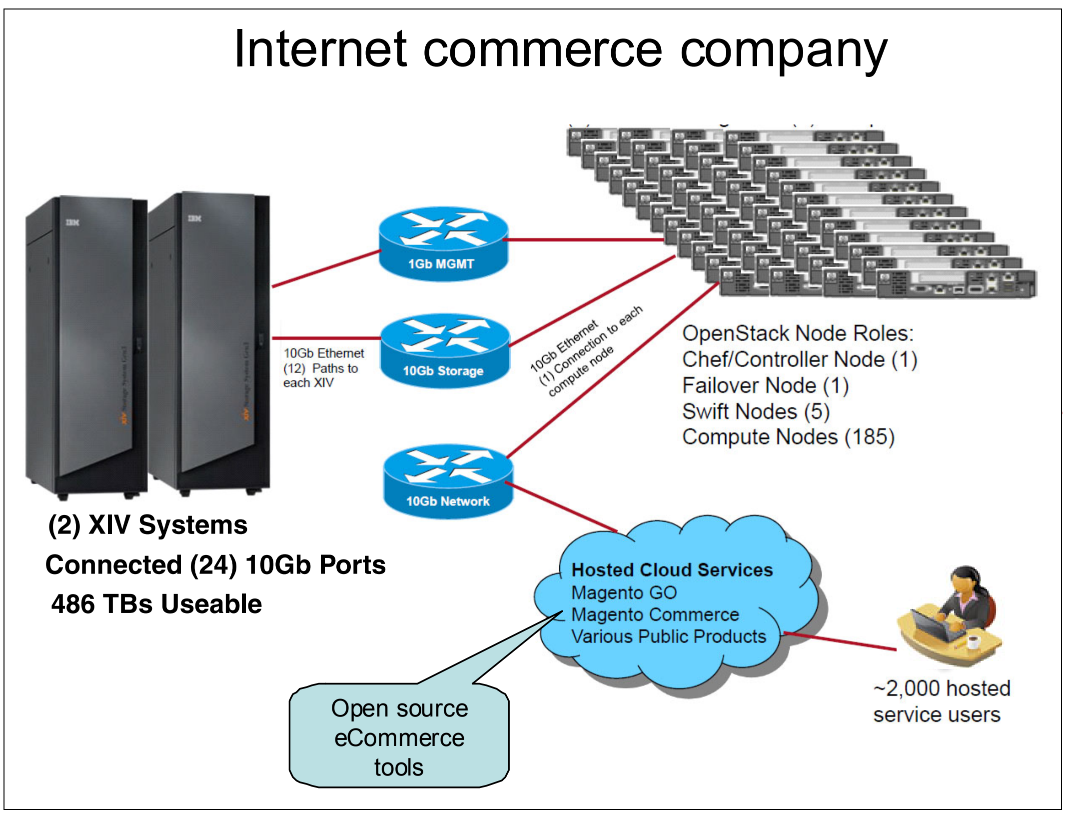



*Figure 5-1 Internet commerce company using OpenStack cloud software with XIV storage*

What is notable is that the OpenStack environment is implemented entirely on commodity enterprise class components, except for storage, which was chosen to be XIV storage.

In the client's worldwide customer-centric, internet-based cloud environment, extreme variability in workloads is the norm. The ability of XIV to deliver a ensured SLA for the cloud storage service is uniquely valuable. The consistent, auto-tuning performance of XIV results in greater predictability of I/O response time.

### 5.1.3 Benefits of the solution

With XIV and its support of the OpenStack Cinder component, the client is able to meet a demanding SLA. The client is able to import a lot more concurrent users on the same compute and storage platform, and can scale out to multiple compute platforms. They can do it while delivering a predictable performance curve.

The consistent, auto-tuning performance, and predictability of I/O response time are the characteristics that led this client to choose XIV storage for their cloud OpenStack environment to enforce their crucial SLAs.

## 5.2 IT storage cloud orchestration

This example involves a large bank servicing with over 45,000 users. The average number of requests for storage resources was increasing, and so the pressure to manage the growing IT infrastructure was also increasing.

### 5.2.1 Business needs

The bank faced these issues related to the growth of the business and the manual efforts that made it difficult to support that growth and meet the SLAs:

- A 100% capacity growth on storage resources is requested every year.
- The number of service requests handled by the storage services team doubles every year.
- The storage services team had to manage services that included installing storage devices, provisioning resources, and troubleshooting. With only a limited number of people, the service request management was getting out of hand.
- More than 50% of requests per year are for storage allocation.
- The existing service request management system did not include approval work flow management. The approval management was manual and depended on email.
- The reclaiming resources and expiration of storage resources involved more manual processes. These processes depended on the resource user.
- Metering and charge back were non-existent because of a lack of adequate tools.
- Predefined service levels were set for each team or department with standard approval authority.

#### Client background

The client had the following background and IT infrastructure already in place:

- Category: Banking Industry.
- Storage Infrastructure: 10 petabytes of storage that included block storage devices like XIV, Spectrum Virtualize (under Spectrum Virtualize control: HP 3PAR and PureStorage), Spectrum Scale, and Spectrum Control.
- Number of users: More than 45,000.

- The following information is also relevant:
  - IBM Security Directory Server was used as the authentication and user management system.
  - The IT environment includes many IBM Power, Oracle SPARC, HP x86, and VMware ESX servers.
  - The client has two data centers: One for production and one for disaster recovery (DR).
  - The storage resources were already monitored through Spectrum Control.
  - The storage services provided included file and block services, and were supported by a small team of people.

### 5.2.2 Proposed solution

IBM proposed a solution with IBM Cloud Orchestrator to meet the requirements. Table 5-1 shows the requirement versus the proposed solution matrix.

*Table 5-1 IT Services requirement versus solution matrix*

| **Requested function** | **Matching IBM Cloud Orchestrator feature** |
|---|---|
| Request process simplification | Self-service portal |
| Approval process management | Approval work flow management |
| Department-wide resource allocation and workflow | Departmental model |
| Automated resource allocation and workload reduction | Automated storage creation on request approval |
| Central resource reservation at customized service levels | Resource pools through storage environments and service levels |
| Metering | IBM Cloud Orchestrator-based metering (IBM Cloud Orchestrator Enterprise Edition includes SmartCloud Cost Management) |
| Alignment to existing policies | IBM Cloud Orchestrator policy management |
| Driving resource expiration with warnings | IBM Cloud Orchestrator expiration settings during the request |
| Reduction in resource allocation turn-around times | IBM Cloud Orchestrator resource allocation automation integrated with IBM Spectrum Control |

## Solution description

The IBM Cloud Orchestrator was the only missing piece to enable the solution. The client had all the other prerequisites already in place for an IBM Cloud Orchestrator solution implementation. Figure 5-2 shows a diagram of the proposed solution.

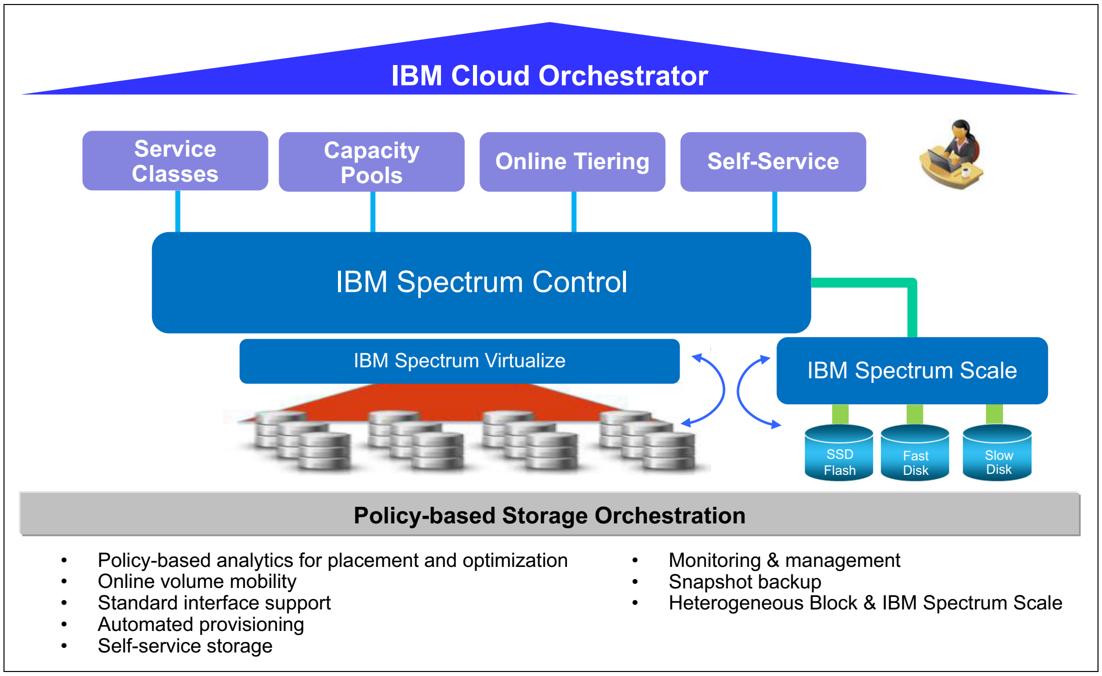



*Figure 5-2 Proposed solution diagram*

## 5.2.3 Solution benefits

Deploying Cloud Orchestrator based provisioning of the storage cloud provided significant benefits by overcoming the pitfalls of the resource request system that they had in place:

The following benefits were realized by the solution:

- Automation and administrator effort reduction: with a Cloud Orchestrator, Spectrum Control, Spectrum Virtualize, and Spectrum Scale based solution in place, the rapid growth in the storage capacity can be managed effectively by the storage administrators.
- Automated metering: The IBM Cloud Orchestrator-provided reports are helpful for the storage department because they remove the need for manual report collection, estimation, and validation.
- Turn around time reduction for resource allocations: The turn around times for resource provisioning was reduced from a few days to a few hours.
- Expiration and reclaiming the resources: Cloud Orchestrator solution has an automated expiration that is based on policy selection.
- Process simplification: The overall resource provisioning process was simplified. The users could complete tasks with a few clicks in the IBM Cloud Orchestrator self-service portal instead of entering more information into the call service management system. The approval and resource creation policy was fully automated and removed the delays caused by manual intervention in the earlier call service management process.
- Effective monitoring on utilization of resources: Because the resource expiration is automatic and supported by an able metering and charge back system, the efficiency in reclaiming resources and monitoring the resource utilization improved dramatically.

- Other uses: The customer understood that IBM Cloud Orchestrator can also be used for provisioning of virtual machines, networks, applications, and advanced patterns. They also began researching how to use the capabilities of IBM Cloud Orchestrator.

## 5.3 National library public cloud

The client is a national library whose mission is to collect and archive documents and publications about the country from the early 1900s. The organization helps develop international standards and works with other entities on national and international matters. The central location develops and manages the central database for the national library, and produces and distributes bibliography services across the entire country.

### 5.3.1 Business needs

The national library needed a dynamic infrastructure consisting of a storage solution that could start with a few terabytes of storage and eventually grow into several petabytes over the next few years. In addition, the client was looking for a storage cloud solution to be able to synchronize information in various sites across the country. When the organization received a mandate by the federal government to digitize the national cultural assets, it sought a new storage solution to meet the requirements.

### 5.3.2 Proposed solution

After the proof of concept (POC) was presented, the client engaged IBM to implement a Spectrum Scale solution. In the POC, client data was used to show how its base configuration with two interface nodes and two storage devices could be used to store digitized cultural assets, such as scanned books and documents. With Active File Management, the replication and the hierarchical storage lifecycle of the information between the two main sites was transparently managed. The second step of the solution proposed was to implement in the small sites across the country with smaller Storwize V7000 Unified systems to ensure fast access to the information from the local library.

Figure 5-3 shows the high-level diagram of Spectrum Scale protocol architecture.

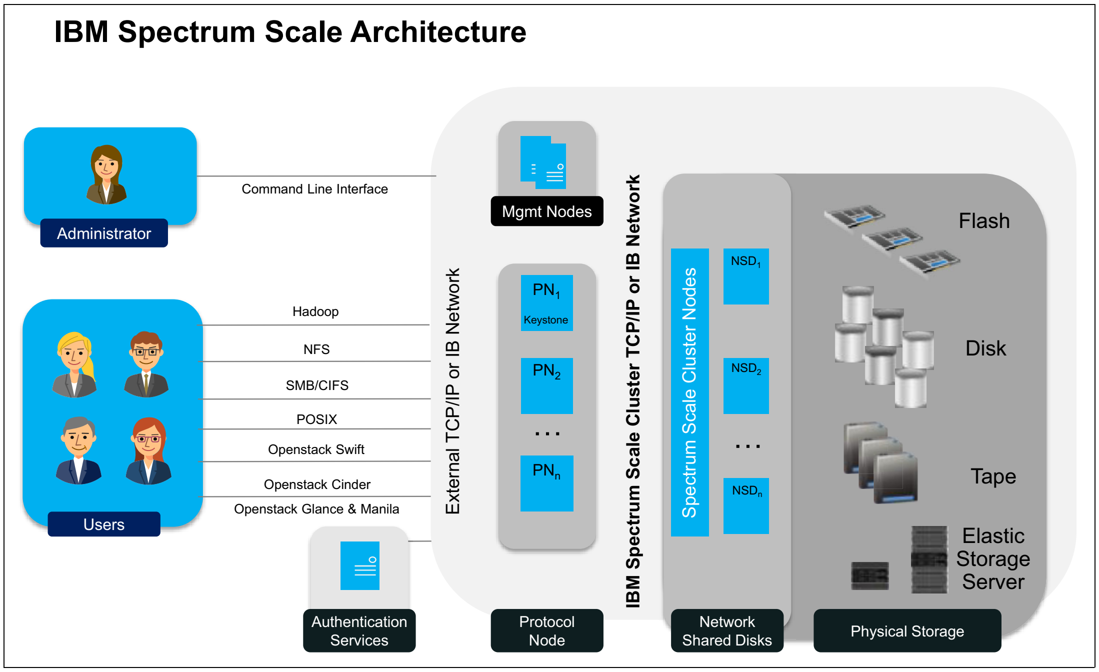



*Figure 5-3 High-level diagram of Spectrum Scale protocol architecture*

## 5.3.3 Benefits of the solution

By engaging IBM, the client gained the ability to manage multiple petabytes of storage and up to a billion files in a single file system. It also achieved operational efficiency with automated policy-driven tiered storage. The client reduced its total cost of ownership (TCO) with automated lifecycle management and migration to tape.

The solution provides the capability to the national library to distribute bibliography services across the country through the public intranet.

The customer is looking to the possibility to modify the Spectrum Scale infrastructure, including IBM cloud object storage as new pool using the evolution of the cloud infrastructure from Private to Hybrid. The proposed architecture is represented in Figure 5-4.

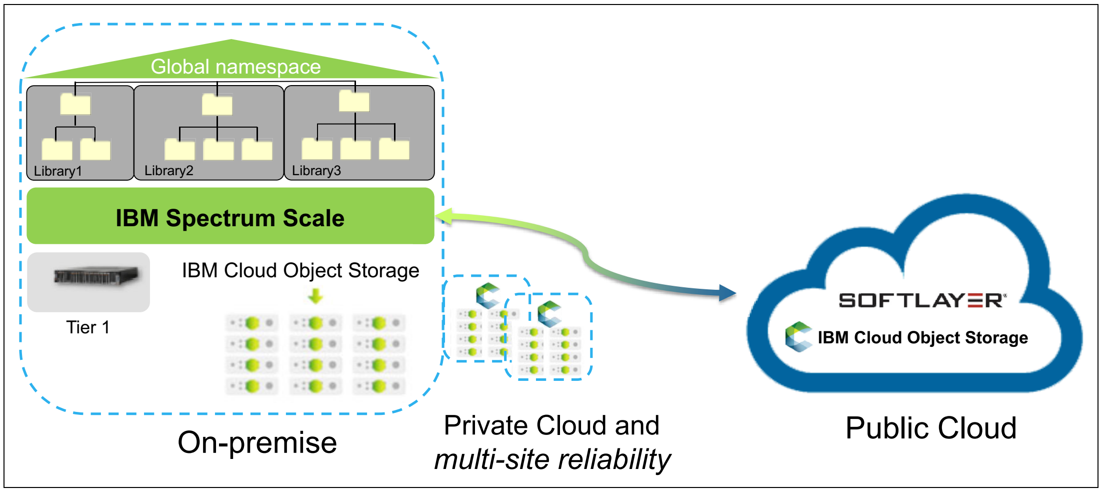



*Figure 5-4 High-level diagram of Spectrum Scale evolution to Hybrid storage cloud infrastructure*

## 5.4 Video surveillance solution for public safety

The client is a government organization that took an initiative to improve public safety in the areas under its responsibility. The scope of the organization is to be able to monitor critical public locations for identification and rapid response to accidents, crimes, and so on.

### 5.4.1 Business needs

The requirement was to store at least three days of recorded content, providing access to the police and justice departments when necessary. The client was looking for a storage infrastructure that could support the simultaneous storage of many thousands of video streams. The solution needed to allow hundreds of simultaneous reads and, for particular public locations, the images needed to be archived for 10 years for security reasons. The solution needed to be implemented in various cities across the country.

### 5.4.2 Proposed solution

In addition to the deployment of an end-to-end video surveillance solution, IBM provided a Spectrum Scale solution to implement a scalable global file system that is geographically distributed in the country. To address the archive requirements, the solution includes Spectrum Archive. Tape libraries were part of the solution to provide enhanced, cost-effective scalability to meet business requirements to preserve video images for 10 years.

Spectrum Scale was paired with a specialized video software provider, delivering a broadband cable network solution and video network displays.

Figure 5-5 shows the general view of the project, with examples of the various video collectors and their integration with Spectrum Scale Active File Management replication.

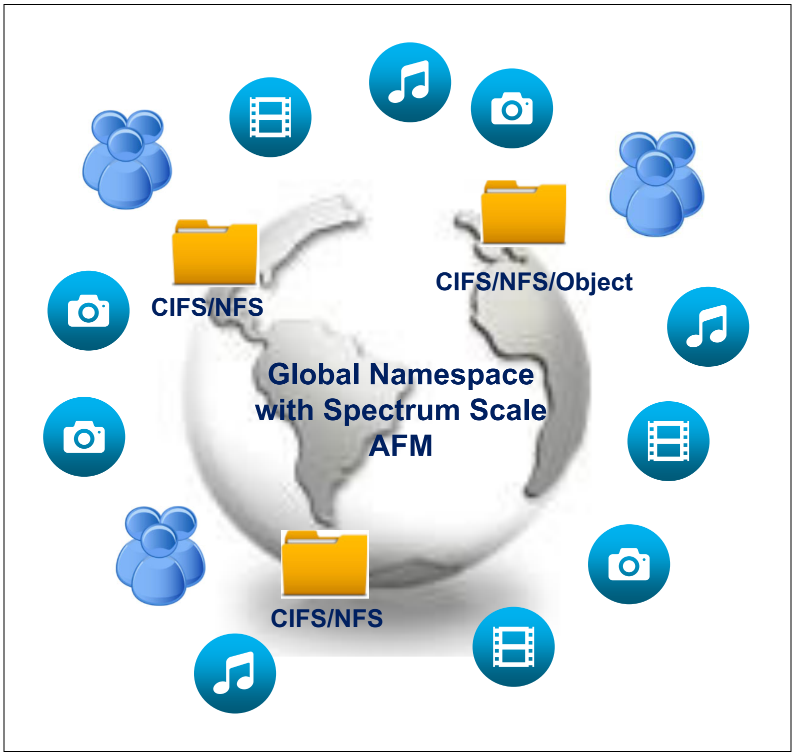



*Figure 5-5 Video collector with Spectrum Scale Active File Management*

Figure 5-6 summarize the implemented architecture.

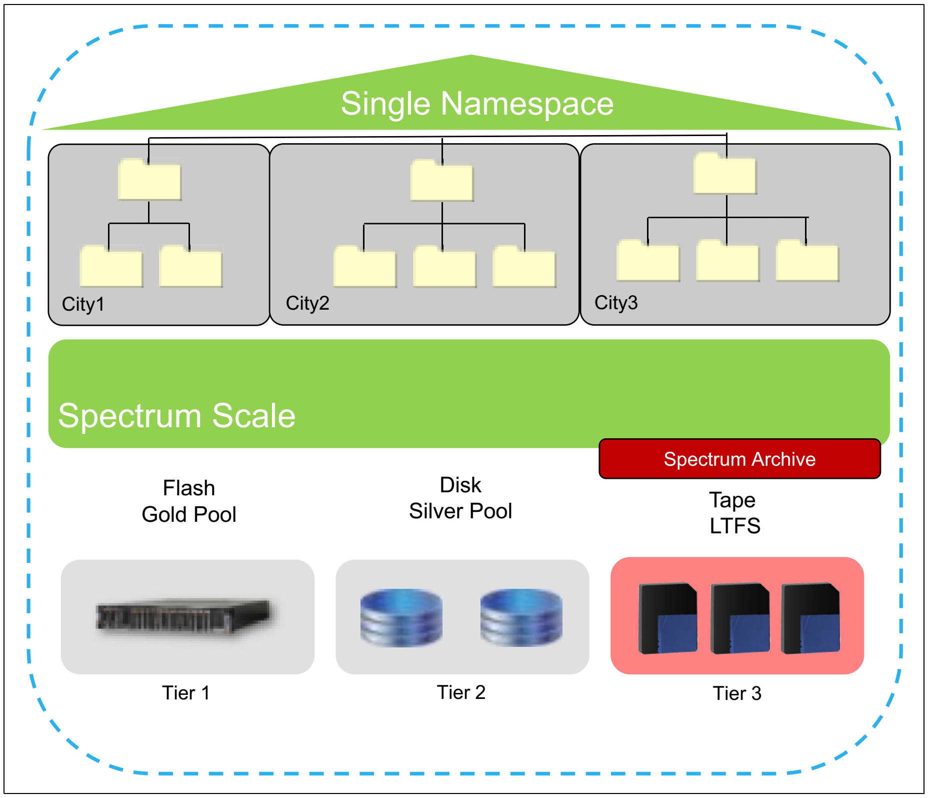



*Figure 5-6 Spectrum Scale and IBM Spectrum Archive architecture*

### 5.4.3 Benefits of the solution

The high-performance storage cloud solution was able to ensure collection of all the information that was created by the various video collectors. The implementation of data replication and collaboration was delivered by Active File Management. In addition, by using Spectrum Archive, the objective to provide a cost-effective solution for the archiving requirements was addressed. Information managed by Active File Management is available from all police stations and the courts of the country in case of accidents, crimes, and civil and criminal cases.

## 5.5 Telecommunication company public cloud

The client is a telecommunications company, based in Europe that offers technological infrastructures and platforms in which voice and data are converted to advanced telecommunications services. One of the missions of the company is to sell cloud IT services with the main focus on these offerings:

- Infrastructure as a service (IaaS): Offers CPU, storage space, and other IT components
- Platform as a service (PaaS): Addresses application deployment

The client's cloud computing environment is managed by a proprietary application.

## 5.5.1 Business needs

The client wants to add disaster recovery and high availability of the information, and backup for all its customers to its IaaS cloud offering. This emerging model of delivering disaster recovery as a service (DRaaS) offered by service providers is gaining popularity, mainly because of the pay-as-you-go pricing model, and infrastructure resource sharing that can lower costs. This model can offer several options and use various technologies depending on the SLAs:

- Both primary production and disaster recovery instances that are provisioned in the cloud are managed by a managed service provider.
- Primary production on-premises and disaster recovery instances are provisioned in the cloud.

The company has many data centers that are distributed across the country, and the storage infrastructure is based on different storage systems from various vendors, including IBM XIV and IBM DS8800. The goal of this project was to find a solution that could reduce the complexity in the management activities to save time and expedite the delivery of the service.

## 5.5.2 Proposed solution

The solution took a phased implementation approach because of the complexity of the existing environment and the need to ensure access to the information for the customers of the IaaS cloud that is hosted by the client.

### Phase 1

IBM proposed implementation of SVC in a split-cluster configuration between the two production sites (150 km (93 mi) apart) to provide an abstraction layer between the physical storage resources and the applications. This configuration hides the complexity of the environment, makes management easier, and increases utilization. The Spectrum Virtualize software in SVC represents a key enabler at the infrastructure level to implement an efficient data replication mechanism for high availability and disaster recovery as a cloud service. SVC creates a single point of control and management. For disaster recovery, Spectrum Virtualize provides consistent synchronous and asynchronous mirroring across sites (campus and long distance).

Figure 5-7 shows the high-level description of Phase 1 of the solution with SAN Volume Controller in split cluster configuration and Global Mirror for disaster recovery.

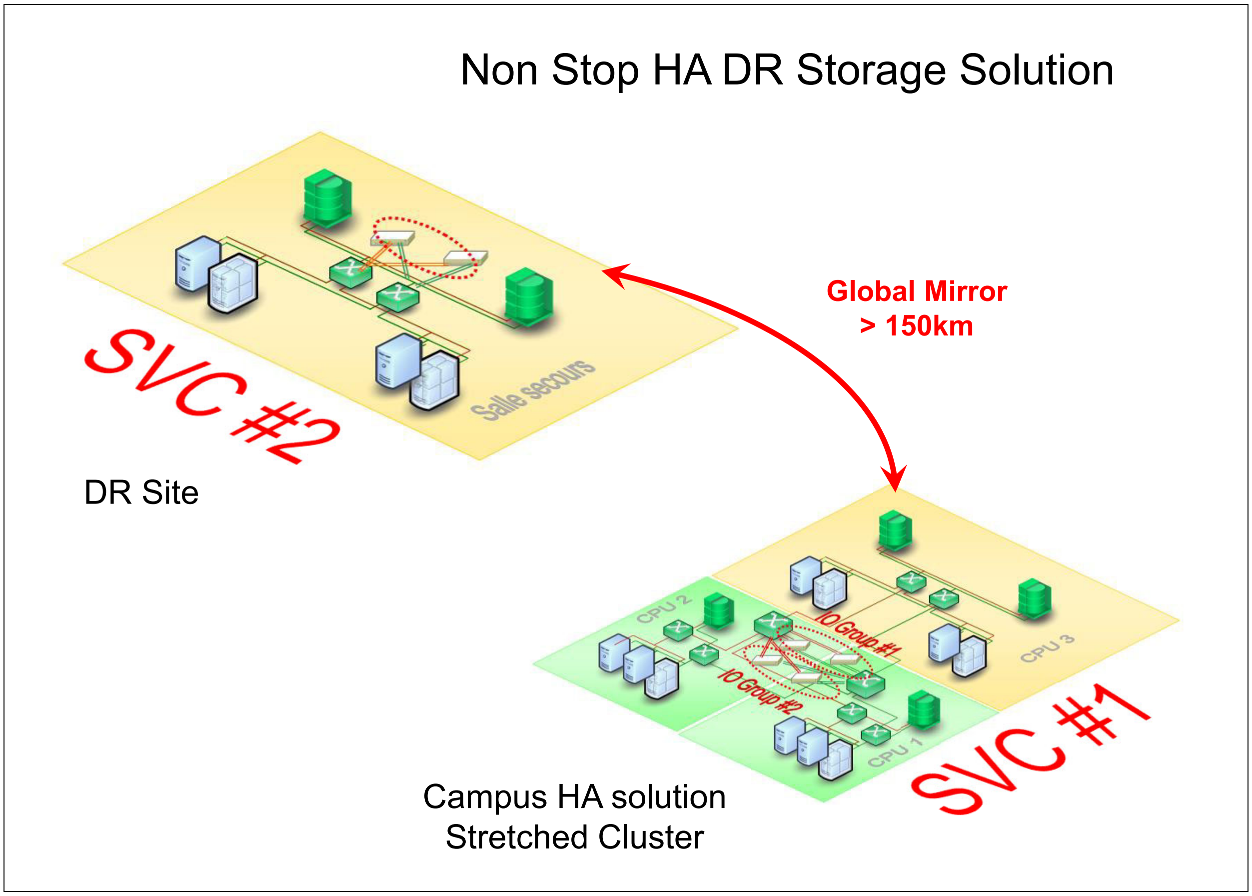



*Figure 5-7 SAN Volume Controller architecture*

### Phase 2

As the IT environment grew and the needs to also address unstructured data requirements, the client had a vision to move from an OpenStack cloud service management to a proprietary cloud service management. This change was intended to establish a standard support matrix through a proprietary vendor.

A solution to address the business requirement was formulated and proposed to the client with the following things in mind:

- Must co-exist with the existing IT environment
- Must reuse the existing solution components
- Must align to the future vision of the client
- Must reduce capital investments

The solution included a proposal to use OpenStack drivers that can work with the existing environment. The proposal also introduced Spectrum Scale. The customer had a OpenStack IceHouse version that included the drivers for Spectrum Scale.

Spectrum Scale through IBM Elastic Storage Server was an automatic choice because it could address the needs of very large amounts of data that needs stringent lifecycle management. In addition, Spectrum Scale can be used as a back end for OpenStack object storage service.

Spectrum Control was used to simplify and monitor the entire environment.

For the entire cloud client base, the solution is able to deliver replication and collaboration on the access of the information. It does so by using both Spectrum Virtualize replication functions and Active File Management with Spectrum Scale. Figure 5-8 represents the final architecture.

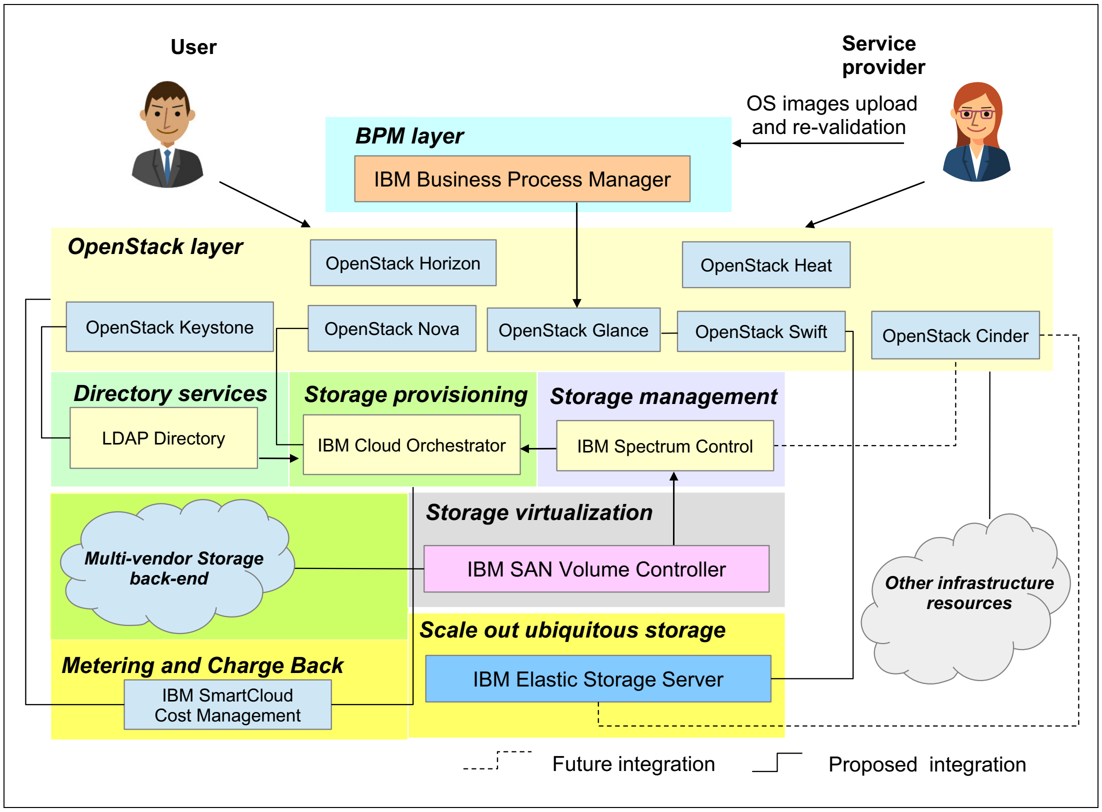



*Figure 5-8 Final architecture for public cloud*

## 5.5.3 Benefits of the solution

Spectrum Virtualize provides flexibility in a multivendor environment. It supports multiple hypervisors and operating systems (VMware, Power, and others) and a wide variety of storage devices from IBM and others (EMC, HP, HDS, Oracle, Dell, NetApp, Fujitsu, NEC, and Bull). The mirroring that is integrated with recovery automation tools provides management of disaster recovery SLAs as requested by the client. All the storage operations are managed from a central point with Spectrum Control and IBM Spectrum Protect Snapshot.

Spectrum Scale Active File Management provides these advantages:

- Allow multiple sites to collaborate on information exchange
- Lower storage costs by moving files transparently to the most appropriate tier of storage
- Control storage growth by moving older files to tape and deleting unwanted or expired files

## 5.6 Life science healthcare hybrid cloud

The client is a health sciences company whose mission is DNA sequencing and analysis research. The client creates and delivers excellence in biomedical research to better understand chronic human diseases and aging, as influenced by metabolism, genetics, and the environment.

### 5.6.1 Business needs

The client's intention was to build a large-scale computing infrastructure to allow the storage of massive genomic data sets to perform complex processing of that data. The solution needed to provide high-performance computing (HPC) scalability and performance in an environment with a progressive growth of DNA sequencing data. The HPC application requires collecting data directly on a global file system. The company has one main research center. In a second phase of the project, the client needs extra capacity for just a few days per month for especially large analysis jobs. The client expected 2 PB of data growth each year. For this quantity of data, the customer needed a solution to integrate hierarchical storage management (HSM) to make sure that the data is stored on the most cost effective medium possible.

### 5.6.2 Proposed solution

The proposed solution provides IBM Power8 as the ideal server platform for HPC DNA analysis, and Spectrum Scale (Elastic Storage Server) as a storage appliance. All the integrated information lifecycle management (ILM) functions are provided by IBM Spectrum Scale Storage to move files from fast disks to economical disks. HSM with the help of Spectrum Archive archives files to tape while still keeping them accessible. IBM Spectrum LSF® software is used for management and control of the HPC/Analysis workloads. For the extra capacity requirement, IBM SoftLayer provides Spectrum LSF bursting capability together with Spectrum Scale. Spectrum Scale uses the Active File Manager functions to keep the compute data file sets in sync with the compute jobs both onsite and offsite.

Figure 5-9 highlights the high-level overview of the solution.

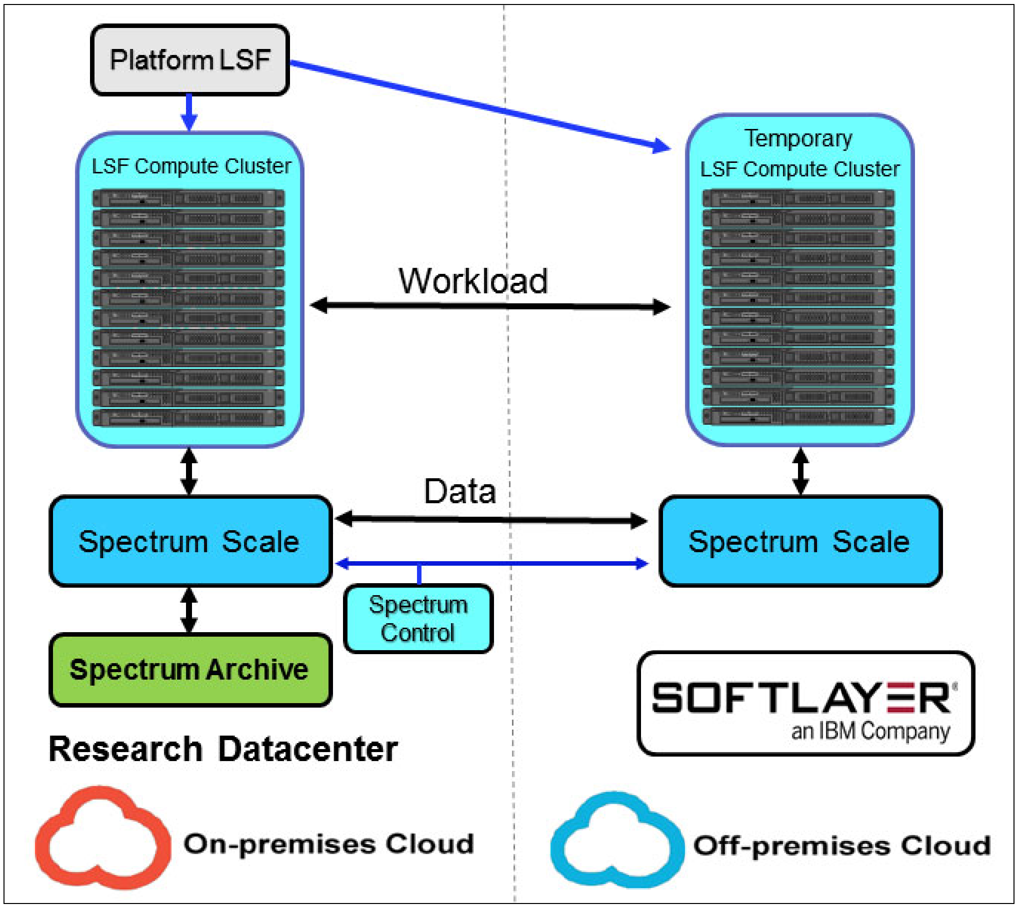



*Figure 5-9 Spectrum Scale, Platform LSF providing Workload Scheduling, and Active File Management providing file synchronization services between onsite and offsite cloud*

Figure 5-10 highlights the interaction between Spectrum Scale and Spectrum Archive.

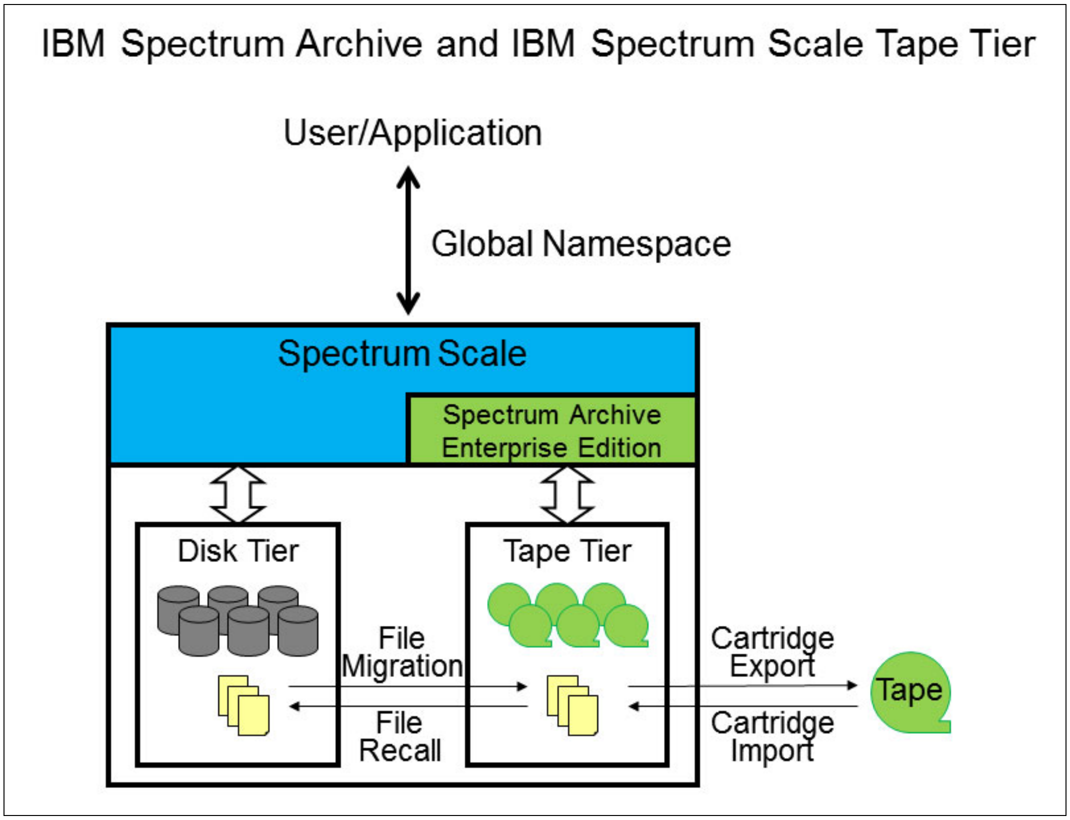



*Figure 5-10 Spectrum Scale and Spectrum Archive interaction to archive to or recall from tape*

## 5.6.3 Benefits of the solution

Genomics and cancer research companies, and university research generate gene sequencing and other vital information in laboratories. This information is received from various sequencers and other research instruments worldwide into a central repository for analysis. DNA-based risk assessments and other high-value-added research is performed. Results can be shared or sold to other genomic organizations. Using the Spectrum Scale storage management capabilities, as data ages, the results can automatically be stored to lower-cost storage, and, with the help of Spectrum Archive, to tape, the lowest cost storage type. This process reduces the TCO while retaining the ability to retrieve the information relatively quickly (just minutes versus a known cloud competitor's SLA of 3-5 hours) and good bandwidth (160 MBps versus internet speed).

The stub information that is needed to recall the data on tape remains in the researcher's folder. In the future, data can be recalled from tape and compared to newer results. Using the Public SoftLayer Cloud reduces the investment in hardware, software, and facilities that are only needed during a couple of days per month. The money that is saved can be spent on areas more central to the research company.

## 5.7 University disaster recovery on public cloud

The client, a university, is an institution that is dedicated to higher learning and research.

### 5.7.1 Business needs

The university wanted to use the Public Cloud as a DR solution to save costs. They had outgrown their existing DR data center and it would cost a lot of money to build a new data center. Funding for that was not going to be available for many years to come. In fact, the university had more important research areas that needed money instead of IT. The university's existing environment was based on IBM XIV and x86 Servers with a mix of physical and VMware Virtual Servers. The university was happy with their XIV Systems with their easy-to-use GUI and "no tuning required" performance.

### 5.7.2 Proposed solution

The proposed storage solution provides Spectrum Accelerate on SoftLayer Bare Metal Servers as the solution to replicate on-premises XIV systems to public cloud for DR purposes. Spectrum Control Base was included to integrate with VMware and Storage Control to manage and monitor the whole environment.

Figure 5-11 shows the use of Spectrum Accelerate.

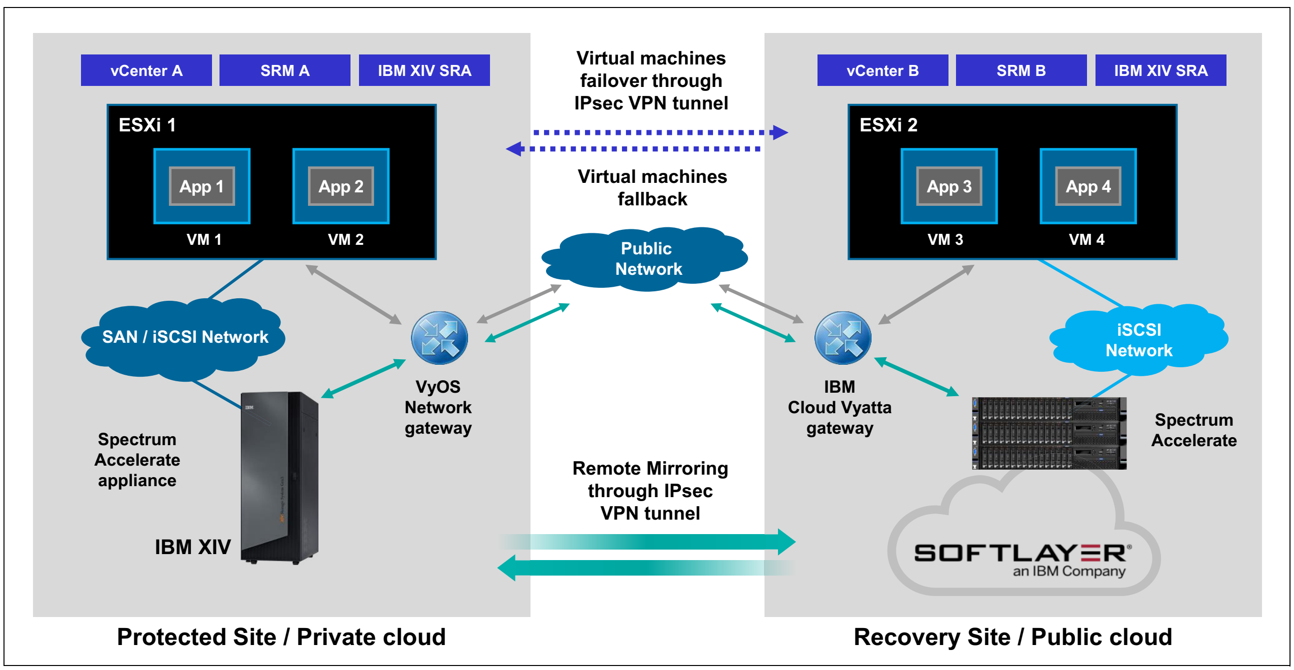



*Figure 5-11 Spectrum Accelerate as DR solution for XIV on premises*

### 5.7.3 Benefits of the solution

Spectrum Accelerate provides XIV as software so that you can provision XIV on bare metal servers in the IBM SoftLayer Public cloud. This configuration saves money on data center space, power, and personnel. It has the same GUI as XIV, so the university's existing storage skills can be reused without major retraining costs. Spectrum Control Base provides a rich integration layer for VMware to make XIV a full member of a VMware Private Cloud. Spectrum Control provides advanced management, monitoring, and capacity and performance management for proactive systems management.

## 5.8 Media and entertainment company hybrid cloud

This example involves an international multimedia publishing group that operates daily newspapers, magazines, books, radio broadcasting, news media, and digital and satellite TV.

### 5.8.1 Business needs

The company defined a new strategy that is based on high-quality editorial production, and rethinking products and offers. The client redesigned the business model, mainly regarding new organization of work to develop multimedia and a digital business model. The media company needs a dynamic storage solution that was able to provide up to 20 PB of data, distributed on separate tiers that include a tape library to ensure a cost-effective solution. The client was looking for a storage cloud solution capable of replicating data in separate sites, where the editors and journalists are based. They needed a system that offered zero downtime while delivering predictable performance in all the digital media information lifecycle phases:

- Create
- Manage
- Distribute
- Archive

The storage cloud solution needs to deploy a pay-per-use model for its customers to allow them to purchase access to old and recent TV programs.

### 5.8.2 Solution

The proposed solution was a Spectrum Scale Storage system with a multitier storage pool. It included Active File Manager tape management by IBM Spectrum Protect to move information to tape for archiving requests. The regional sites of the media company used Storwize V7000 Unified with Active File Manager (formerly Active Cloud Engine) to share the information managed by the local site and vice versa. Spectrum Control Storage Insights (storage as a service (SaaS) running on IBM SoftLayer) is used for collecting and managing storage usage information. IBM SmartCloud Cost Management is used for reporting for chargeback of the customers of the web services. IBM Aspera® On Demand Managed File Transfer (SaaS running on IBM SoftLayer) is used for file transfers to and from external parties for the fastest and the most cost-effective file-transfer.

Figure 5-12 shows the design based on Spectrum Scale.

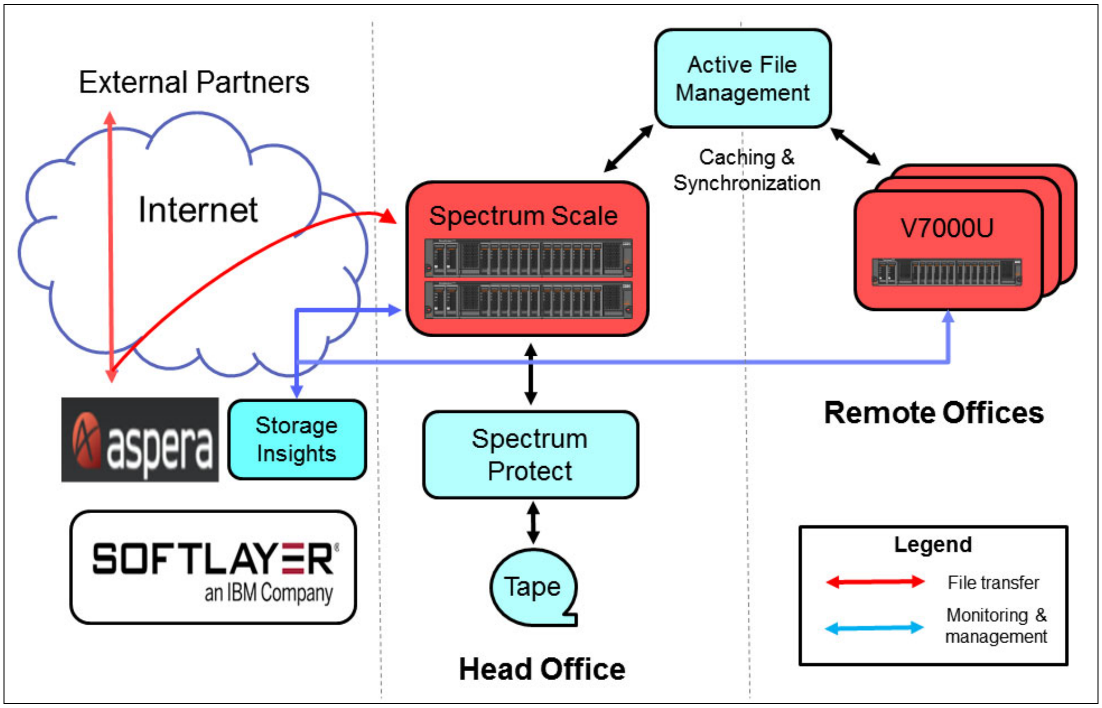



*Figure 5-12 Solution based on Spectrum Scale with multitier storage pools*

## 5.8.3 Benefits of the solution

With Spectrum Scale, the client is able to manage more than 100 million files with tens of thousands of users. The solution provides concurrent profile logons, many of which have over a thousand small files. The embedded function of replication that is supplied by Active File Management, combined with IBM Spectrum Protect for backup and the archiving policies of HSM, ensure the continuous availability of the data. By using the ILM approach, the client is able to move information to the correct storage tier, including tape, to obtain a cost-effective solution.

The IBM cloud storage solution, which is based on Spectrum Scale and its Active File Manager Technology, ensures that the remote locations have excellent access-response time to the media content.

The Spectrum Scale capability to manage multiple file systems with multiuser file sharing, managed by the HSM policies, provided a secure and cost-effective solution for the requirements. Spectrum Control Storage Insights provides intuitive analytics-based monitoring of storage usage. The information that is collected by IBM SmartCloud Cost Manager provides the customer with a solution to start pay-per-use services. IBM Aspera on Demand High-speed Managed File transfer sends and receives files from and to external partners in the quickest possible time over the internet using as little bandwidth as possible.

The collaborative benefits of the solution for every phase of the digital-media data process are shown in Figure 5-13, Figure 5-14, and Figure 5-15 on page 159. Figure 5-13 shows the value for Broadcast.

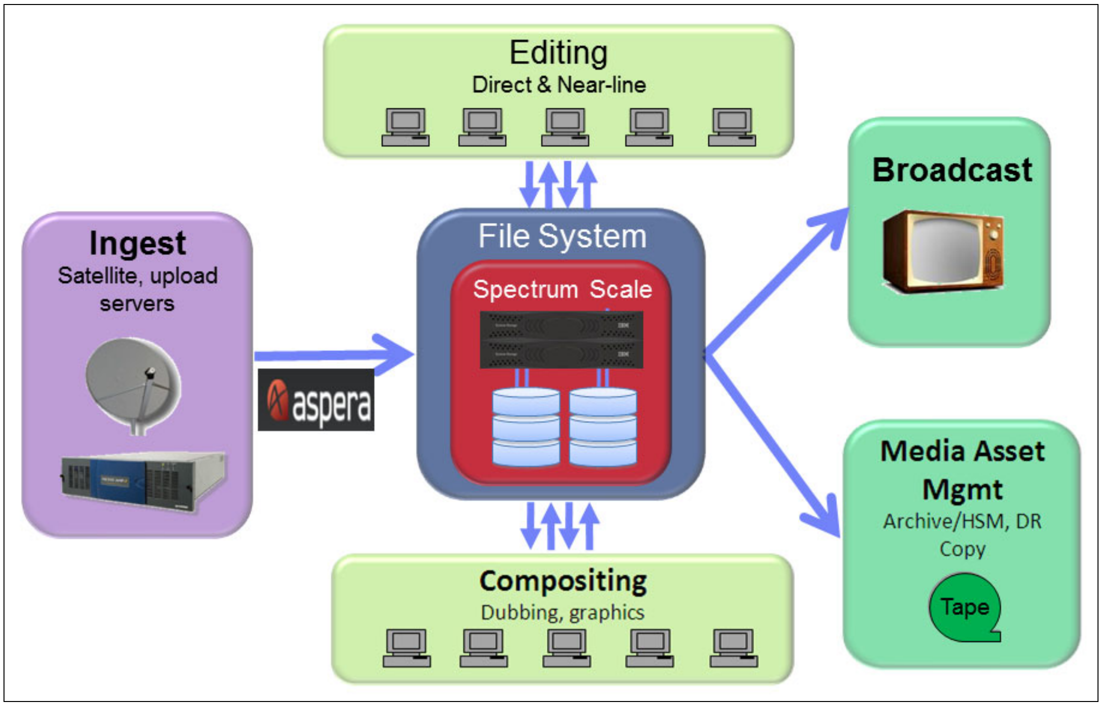



*Figure 5-13 Cloud storage value for Broadcast*

Figure 5-14 shows the value for Post-Production.

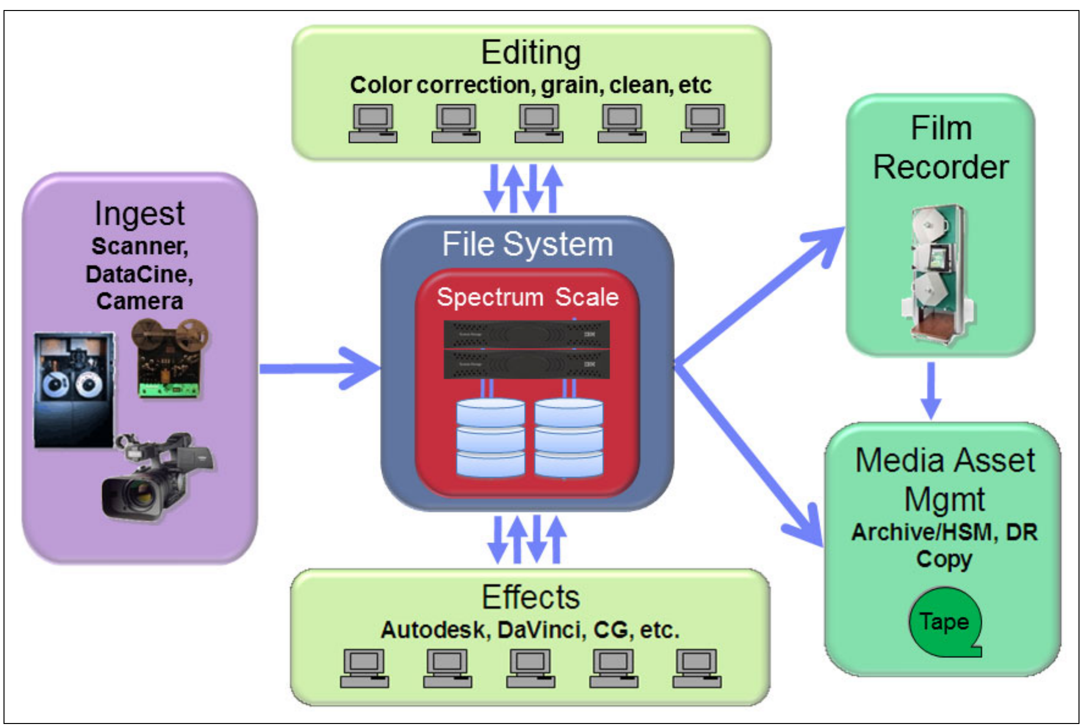



*Figure 5-14 Cloud storage value for Post-Production*

Figure 5-15 shows the value for Content Archiving and Distribution.

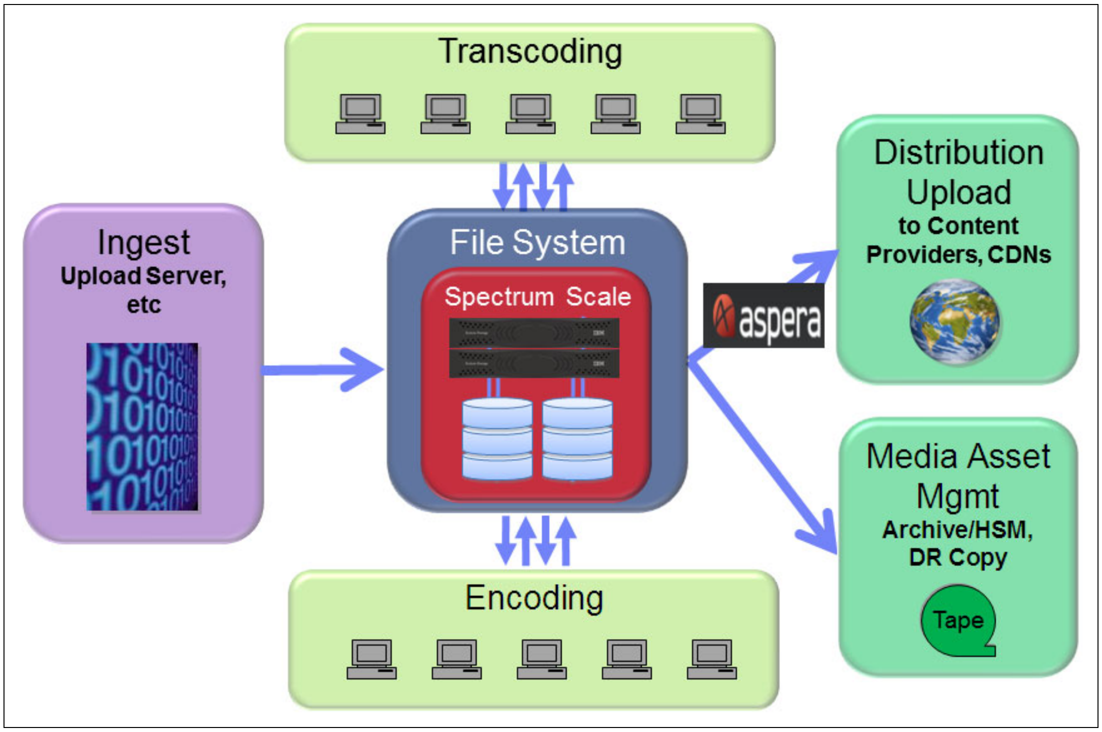



*Figure 5-15 Cloud storage value for Content Archiving and Distribution*

# 6

# Your next steps

Getting started on the journey to a smart storage cloud implementation can be relatively straightforward or fairly complex, depending on the scope of the project under consideration. It is important to understand your current organizational capabilities and challenges, and identify the specific business objectives to be achieved by deploying a smart storage cloud solution in your enterprise. You should ask the following questions:

- What strategy should the organization follow to build a “cloud-ready” infrastructure?
- Should your storage infrastructure be cloud-based?
- Does IBM have storage cloud offerings that meet the organization’s needs (SoftLayer and Cloud Managed Services)?

IBM personnel can assist you in your journey to smart storage cloud by developing a high-level architecture and implementation plan with a supporting business case to justify investment. This plan will be based on a compelling return on investment (ROI), and on improved service levels and lowered costs.

This chapter helps you review your storage strategy, identify where you are on the journey to storage cloud, and your next steps.

This chapter includes the following sections:

- Review your storage strategy
- Identify where you are in the journey
- Take the next step

## 6.1 Review your storage strategy

Before embarking on any journey, it is important to understand where you are currently, and where your chosen destination is. Developing your own cloud storage strategy should reflect these important considerations, which help you to define the path of your journey. Take the time that is needed to ensure that you understand how cloud storage can help your business. Justify your move by using ROI, total cost of ownership (TCO), and other business measures that are relevant to your organization. Be sure that you consider technical or compliance concerns, and develop risk-mitigation plans.

Remember that although storage cloud can be a key component of an overall cloud computing approach, you should determine how a storage cloud strategy fits within your broader cloud computing architectural plans. Overall integration of these system parameters is essential to successful implementations:

- Performance
- Availability and resiliency
- Data management
- Scalability and elasticity
- Operations
- Security
- Compliance

Consider your security needs and how a storage cloud is affected by the confidentiality of the data that you need to store. Data that is highly sensitive, or subject to security-compliance regulations, might not be able to be stored on a public network. Therefore, your storage cloud might need to be located behind an enterprise firewall, indicating a private cloud solution requirement. The same might be true for instances where users need to easily access, share, and collaborate, without compromising data security, integrity, availability, and control of the data.

Your storage strategy must consider the requirements of the various business units within your company, along with customer expectations of your IT organization. Competitive pressures might dictate that a storage cloud is the only way to meet the quick service provisioning, elastic resourcing, and pay-as-you-go charging model that your customers are looking for.

A framework for aligning a cloud implementation, optimized to business requirements, is shown in Figure 6-1. The figure focuses on key practice areas across IT architecture, process, and organizational structure.

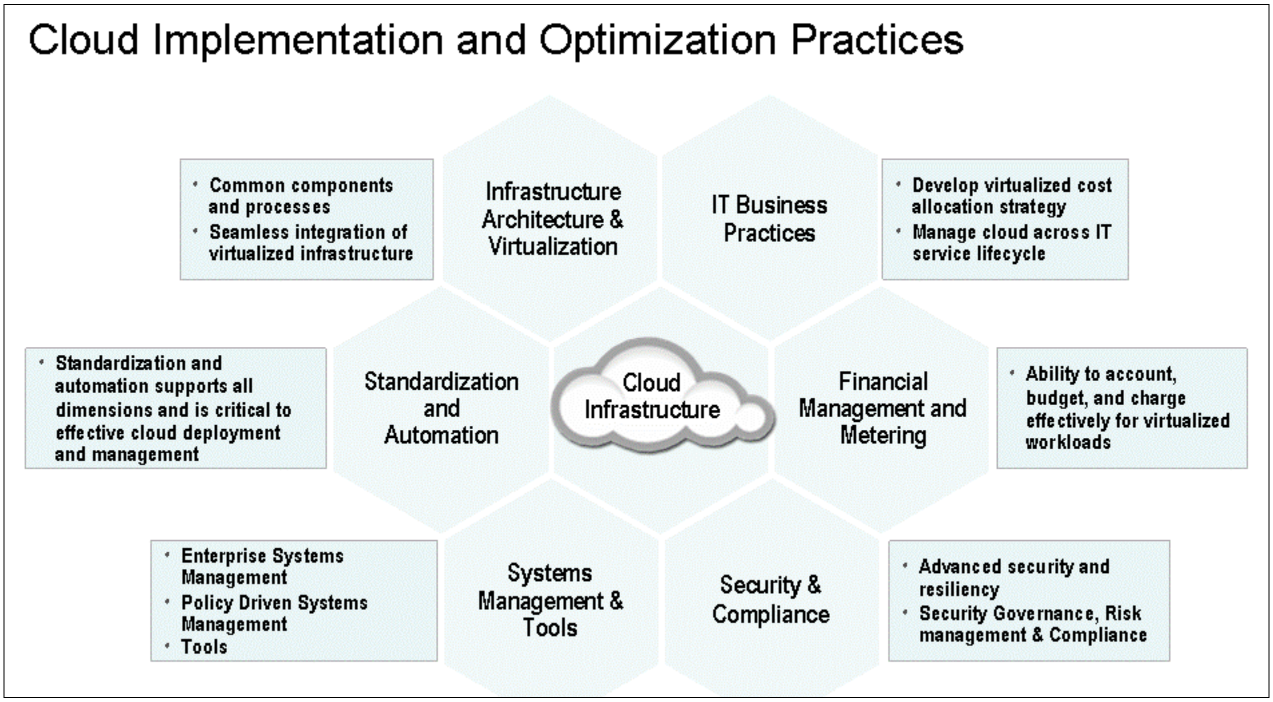



*Figure 6-1 Framework for cloud infrastructure implementation and optimization practices*

When considering the use of any new technology, a common mistake is to focus on the technology itself, rather than on the business requirements to be addressed by a technical solution. To stay on track with your storage strategy, identify several significant use cases in your organization where technology can be helpful. Start by analyzing a use case and its importance to your business, and then determine how the introduction of a storage cloud will affect your business operations and costs.

With the use-case approach, you can gain an understanding that a private cloud is not only a storage infrastructure, but rather a system of cloud storage clients, backup, and archive solutions, special purpose data movers, management, and support. When these components are combined with cloud storage infrastructure, a complete solution for storage is achievable.

## 6.2 Identify where you are in the journey

As described in Chapter 2, "What is a storage cloud" on page 29, the journey to delivering a storage cloud depends on where you are in the storage journey, and how far you are from a traditional storage infrastructure.

From internal experiences and from hundreds of cloud engagements with clients worldwide, IBM has identified key steps in the deployment of a storage cloud. These steps can overlap, so there is no need to necessarily complete one step before moving to the next. Rather, the steps represent a progression. For example, in some organizations, the consolidation step might require major effort because the infrastructure might be highly heterogeneous and distributed. For others, consolidation might be more evolutionary, and performed simultaneously with other steps. Although there is no single approach to completing these steps, they are all important considerations in the journey to a storage cloud. Figure 6-2 illustrates a high-level approach to the development of an optimized storage cloud strategy.

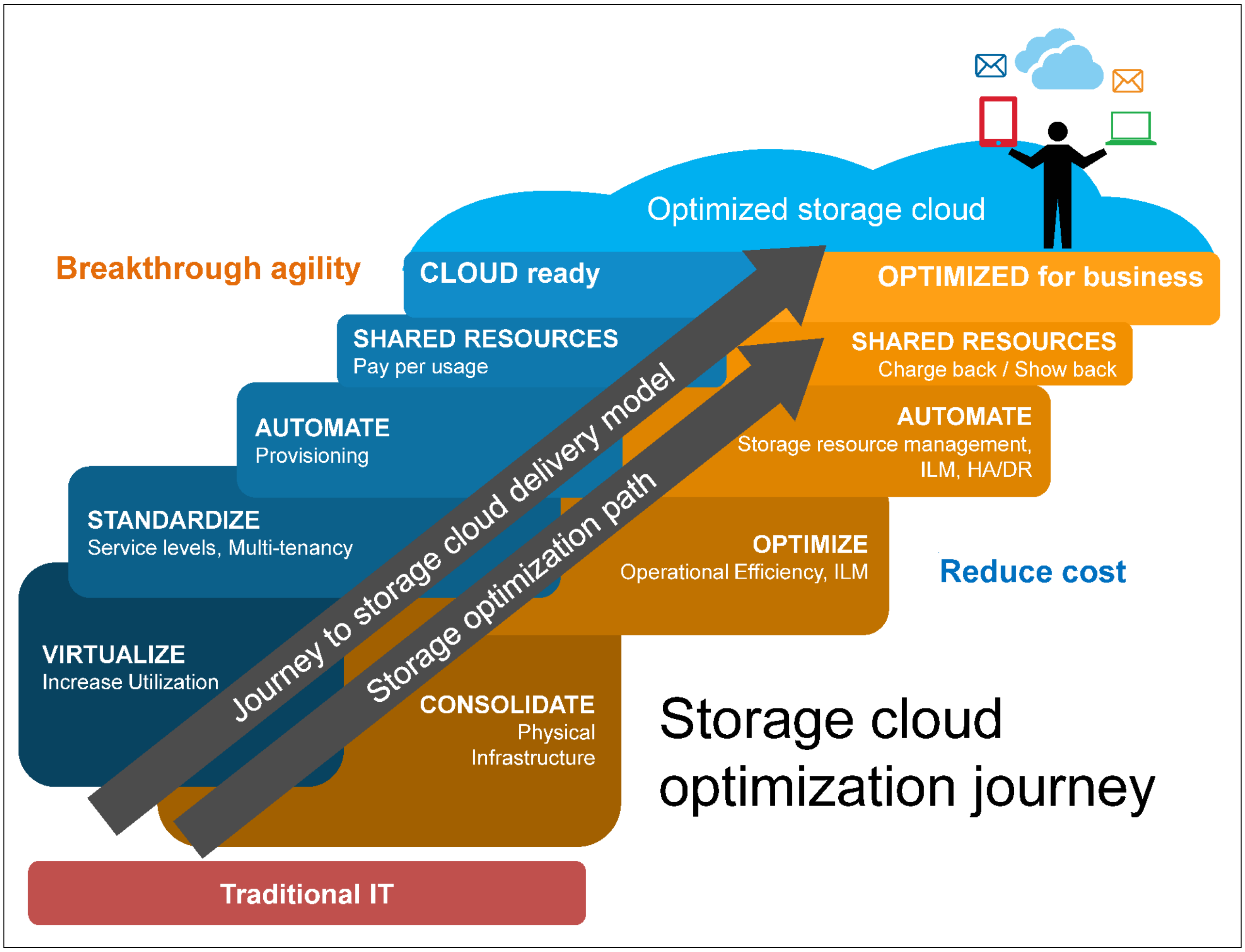



*Figure 6-2 The overall cloud journey from traditional IT to storage cloud*

## 6.2.1 Consolidate physical infrastructure

Consolidating assets reduces infrastructure complexity, increases economies of scale, and enables more efficient IT management focused on fewer aspects, which can all lower operational costs.

## 6.2.2 Virtualize: Increase utilization

Storage virtualization complements consolidation by making better use of existing resources. The virtualization step is about pooling storage resources so that the available storage appears to be a single storage system, whereas in reality it might be distributed across many storage devices. Another important aspect of storage virtualization that demonstrates increased utilization is the implementation of features, which, to the storage user, appear to be more storage than the user was granted. Thin provisioning is one such feature, which is described in more detail in 3.2.4, “Thin provisioning” on page 51.

### 6.2.3 Optimize operational efficiency

Optimizing operational efficiency is achieved in part by consolidation and virtualization, but also by implementing advanced storage features such as the following items:

- Data tiering: Categorizing data storage performance and data performance requirements, and matching those automatically
- Data deduplication: Removing duplicate data
- Compression: Reducing the physical space that is taken by data
- Self-tuning: Disk failure management, array rebuilding, call home, and proactive test on performance degradation

Each of these items is discussed in further detail in 3.2, “Storage efficiency” on page 50.

### 6.2.4 Automate

Automated processes result in significant cost reductions within the storage management discipline. Opportunities to automate can include the following items:

- A service catalog with self-provisioning capability
- Monitoring user activity and integrating this data with a chargeback system, which enables pay-per-use
- Policy-driven automation for data movement (replication, tier management, backup)

### 6.2.5 A different approach

So what really differs about provisioning cloud storage services by using these steps? The traditional IT approach to storage tends to pull together resources and deploy them in support of a business function workload in silos. The resources are dedicated to the workload and are unable or ill-suited to support other workloads when they might be needed.

By contrast, cloud storage uses a pool of optimized shared resources in an environment that uses virtualization of the physical assets to support multiple workloads. To achieve efficient delivery of the storage services, self-service and self-management are required. These features in turn rely on standardization of the assets and automation to enable a responsive user experience.

By following these steps to storage cloud, your infrastructure should be able to provide resources in support of any storage cloud delivery model (public, private, hybrid, or community), and will finally be cloud-ready.

When you have a roadmap for your journey to storage cloud, you can take the next step.

## 6.3 Take the next step

Now that you have identified where you are in your storage cloud journey, where you want to be, and what you need to do to get there, you are ready to take the next step. Do you have sufficient resources to take the next step on your own? Do you have sufficient skills to evaluate the options? Will a technology partner make a cost-effective contribution?

Cloud storage, as with any other emerging technology, is experiencing growing pains. Some facets are immature, fragmented, and lack standardization. Vendors are promoting their own particular technology as the emerging standard. Although standards are lacking, IBM believes that a set of web services API-based capabilities, accessed through non-persistent connections on public or private networks, provides the fundamental frame of reference for accessing storage cloud services. This definition allows for both public service offerings and private use, and provides a basis for expansion of solutions and offerings.

As a leader in cloud computing, IBM has the resources and experience to help businesses implement and use cloud services, including storage cloud. IBM offers hardware and software technologies and key services to help you take advantage of cloud computing. IBM can assist you in planning, designing, building, deploying, and even managing and maintaining a storage cloud environment.

Whether on your premises or someone else's, IBM can make the journey move more quickly, and in many cases deliver value to your business much more rapidly, ultimately saving you money.

Clients that have implemented an IBM Smart Business Storage Cloud solution are projecting savings as follows:

- A large client with 1.5 PB of usable unstructured file system capacity projects savings of over $7.1 million (USD) over the course of five years in hardware acquisition and maintenance, and environmental and administration costs.
- A medium client with 400 TB of usable unstructured file system capacity projects savings of over $2.2 million in hardware acquisition and maintenance, and environmental and administration costs.
- A small client with 200 TB of usable unstructured file system capacity projects savings of over $460,000 in hardware acquisition and maintenance, and environmental and administration costs.

The latest information related to IBM cloud offerings is available at the following website:

`http://www.ibm.com/cloud-computing/us/en/index.html`

IBM personnel can assist you by developing a high-level architecture and implementation plan with a supporting business case to justify investment based on a compelling return on investment, with improved service levels and lowered costs for your cloud infrastructure. IBM consultants use a unique cloud adoption framework, the Cloud Computing Reference Architecture (CCRA), and the IBM Cloud Workload Analysis Tool to help you analyze your existing environment and determine which cloud computing model is best suited for your business. They help you identify the business areas and workloads that, when changed to a cloud computing model, can enable you to reduce costs and improve service delivery that is in line with your business priorities.

Figure 6-1 on page 163 illustrates the comprehensive structured approach that IBM brings to a cloud implementation engagement. This approach helps IBM to perform a rigorous analysis of your IT and application infrastructure, and provides recommendations and project planning for streamlining your infrastructure and processes. The IBM methodology incorporates key practices that were learned from engagements with leading businesses around the globe, and partnering with them on their storage cloud journey.

# Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this paper.

## IBM Redbooks

The following IBM Redbooks publications provide additional information about the topic in this document. Some publications referenced in this list might be available in softcopy only.


- *Active Archive Implementation Guide with IBM Spectrum Scale Object and IBM Spectrum Archive*, REDP-5237
- *IT Modernization using Catalogic ECX Copy Data Management and IBM Spectrum Storage*, SG24-8341
- *Cloud Computing Patterns of Expertise*, REDP-5040
- *A Deployment Guide for IBM Spectrum Scale Object*, REDP-5113
- *Enabling Hybrid Cloud Storage for IBM Spectrum Scale Using Transparent Cloud Tiering*, REDP-5411
- *Harnessing the Power of ProtecTIER and Tivoli Storage Manager*, SG24-8209
- *IBM DS8880 Architecture and Implementation (Release 8.1)*, SG24-8323
- *DS8880 Product Guide (Release 8.2)*, REDP-5344
- *IBM Private, Public, and Hybrid Cloud Storage Solutions*, REDP-4873
- *IBM SmartCloud: Building a Cloud Enabled Data Center*, REDP-4893
- *IBM Spectrum Accelerate Deployment, Usage, and Maintenance*, SG24-8267
- *IBM Spectrum Accelerate Reference Architecture*, REDP-5260
- *IBM Spectrum Control Base: Enabling VMware Virtual Volumes with IBM XIV Storage System*, REDP-5183
- *IBM Spectrum Scale (formerly GPFS)*, SG24-8254
- *IBM Spectrum Scale in an OpenStack Environment*, REDP-5331
- *IBM Spectrum Scale and ECM FileNet Content Manager Are a Winning Combination: Deployment Variations and Value-added Features*, REDP-5239
- *IBM Spectrum Scale Security*, REDP-5426
- *IBM Spectrum Virtualize and IBM Spectrum Scale in an Enhanced Stretched Cluster Implementation*, REDP-5224
- *IBM System Storage SAN Volume Controller and Storwize V7000 Best Practices and Performance Guidelines*, SG24-7521
- *IBM XIV Storage System Architecture and Implementation*, SG24-7659
- *Implementing IBM FlashSystem 900*, SG24-8271
- *Implementing the IBM SAN Volume Controller and FlashSystem 820*, SG24-8172
- *Implementing IBM Storage Data Deduplication Solutions*, SG24-7888
- *Implementing the IBM Storwize V7000 and IBM Spectrum Virtualize V7.6*, SG24-7938

- *Implementing the IBM Storwize V7000 Gen2*, SG24-8244
- *Implementing the IBM Storwize V7000 Unified Disk System*, SG24-8010
- *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V7.6*, SG24-7933
- *Introducing and Implementing IBM FlashSystem V9000*, SG24-8273
- *Regain Control of your Environment with IBM Storage Insights*, REDP-5231
- *Understanding IBM Spectrum Scale for Linux on z Systems (Express Edition)*, TIPS1211
- *Using IBM DS8870 in an OpenStack Environment*, REDP-5220
- *Using XIV in OpenStack Environments*, SG24-4971
- *VersaStack Solution by Cisco and IBM with SQL, Spectrum Control, and Spectrum Protect*, SG24-8301


You can search for, view, download or order these documents and other Redbooks, Redpapers, Web Docs, draft and additional materials, at the following website:

**ibm.com**/redbooks

# Online resources

These websites are also relevant as further information sources:

- IBM Client Demonstration Center

  https://www.ibm.com/systems/clientcenterdemonstrations

  **Note:** The IBM Client Demonstration Center (for Business Partners, IBMers, and anyone with an IBMid) provides a catalog of remote demonstrations (video or live connection) which consist of self contained material for customer demonstrations of IBM solutions. Most of the demonstrations are provided with predefined scenarios and some also allow for the development of new scenarios. Demonstrations can also be considered as 'ready to use' material for enablement or training.

- IBM System Storage Tumblr provides product videos, customer reference videos, case studies, white papers, infographics, and subject matter expert videos around IBM Storage solutions.

  http://www.ibmstorageexperience.tumblr.com

- IBM storage news, hints and technical discussions by EMEA storage experts

  https://www.ibm.com/developerworks/community/blogs/storageneers

- The Storage Community sponsored by IBM

  http://storagecommunity.org/

- IBM Cloud Computing Reference Architecture wiki

  https://www.ibm.com/developerworks/community/wikis/home?lang=en#!/wiki/Wf3cce8ff09b3_49d2_8ee7_4e49c1ef5d22

- IBM Cloud Computing Redbooks

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

- *IBM SmartCloud Virtual Storage Center Solution*, IBM Redbooks Solution Guide

  http://www.redbooks.ibm.com/abstracts/tips0991.html

- IBM System Storage for Cloud

  http://www.ibm.com/systems/storage/solutions/cloud

- IBM Cloud Offerings

  http://www.ibm.com/cloud-computing/us/en/index.html

- IBM Cloud business continuity and resiliency services:

  – IBM Cloud Managed Backup

    http://www.ibm.com/services/us/en/it-services/business-continuity/cloud-managed-backup

  – IBM Managed Data Vault

    http://www.ibm.com/services/us/en/it-services/managed-data-vault.html

  – -IBM Storage Services

    http://www.ibm.com/services/us/en/it-services/smart-business-storage-cloud.html

  – IBM Federal and State Contracts

    http://www.ibm.com/shop/americas/content/home/en_US/government-contracts.html

  – IBM Resiliency Services

    http://www.ibm.com/services/us/en/it-services/business-continuity/index.html

- IBM SAN Volume Controller Knowledge Center

  http://www.ibm.com/support/knowledgecenter/STPVGU/welcome

- IBM Spectrum Archive Enterprise Edition Knowledge Center

  http://www.ibm.com/support/knowledgecenter/ST9MBR_1.2.0/ltfs_ee_intro.html

- IBM Spectrum Accelerate Knowledge Center

  http://www.ibm.com/support/knowledgecenter/STZSWD/landing/IBM_Spectrum_Accelerate_welcome_page.html

- IBM Spectrum Control Knowledge Center

  http://www.ibm.com/support/knowledgecenter/SS5R93

- IBM Spectrum Scale Knowledge Center

  http://www.ibm.com/support/knowledgecenter/STXKQY/ibmspectrumscale_welcome.html

- IBM Tivoli Storage Manager Knowledge Center

  http://www.ibm.com/support/knowledgecenter/SSGSG7/landing/welcome_ssgsg7.html

- IBM Spectrum Virtualize Knowledge Center

  http://www.ibm.com/support/knowledgecenter/STVLF4

- Blueprints for IBM Spectrum Protect

  https://www.ibm.com/developerworks/community/wikis/home?lang=en#!/wiki/Tivoli%20Storage%20Manager/page/IBM%20Spectrum%20Protect%20Blueprints

- Demonstration of the Operations Center GUI for IBM Spectrum Protect

  https://www.ibmserviceengage.com/on-premises-solutions

- IBM Spectrum Scale

  http://www.ibm.com/systems/storage/spectrum/scale/index.html

- IBM Spectrum Scale Wiki

  https://ibm.biz/BdFPR2

- IBM Elastic Storage Server

  http://www.ibm.com/systems/storage/spectrum/ess

- IBM Storwize V7000 Unified Knowledge Center

  http://www.ibm.com/support/knowledgecenter/ST5Q4U/landing/v7000_unified_welcome.htm

- IBM XIV Storage System Knowledge Center

  https://www.ibm.com/support/knowledgecenter/STJTAG/com.ibm.help.xivgen3.doc/xiv_kcwelcomepage.html

- Thoughts on Cloud: Cloud computing conversations led by IBMers

  http://thoughtsoncloud.com

# Help from IBM

IBM Support and downloads

**ibm.com**/support

IBM Global Services

**ibm.com**/services

REDP-4873-03

ISBN 0738455911

Printed in U.S.A.