Draft Document for Review September 8, 2025 7:59 pm REDP-5729-01

Redbooks
ibm.com/redbooks

IBM Storage Scale System Introduction Guide

Phillip Gerrard
Nilton Carlos Dos-Santos
Stieg Klein
Chris Maestas

Cloud

Storage

IBM

Redpaper

IBM Redbooks

IBM Storage Scale System Introduction Guide

August 2025

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

Second Edition (August 2025)

This edition applies to IBM Storage Scale System 6.2.3 and IBM Storage Scale 5.2.3

This document was created or updated on September 8, 2025.

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 https://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.

IBM®
IBM FlashCore®
IBM Services®
IBM Spectrum®
IBM Z®
Interconnect®
PartnerWorld®
Passport Advantage®
POWER®
Redbooks®
Redbooks (logo) ®
watsonx®
watsonx.ai®
watsonx.data®

The following terms are trademarks of other companies:

The registered trademark Linux® is used pursuant to a sublicense from the Linux Foundation, the exclusive licensee of Linus Torvalds, owner of the mark on a worldwide basis.

Ansible, OpenShift, Red Hat, are trademarks or registered trademarks of Red Hat, Inc. or its subsidiaries in the United States and other countries.

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

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

Preface

IBM Elastic Storage Systems (ESS) is now known as IBM Storage Scale System! For more information see the following IBM Storage Blog:
https://www.ibm.com/blog/evolving-the-ibm-storage-portfolio-brand-identity-and-strategy/

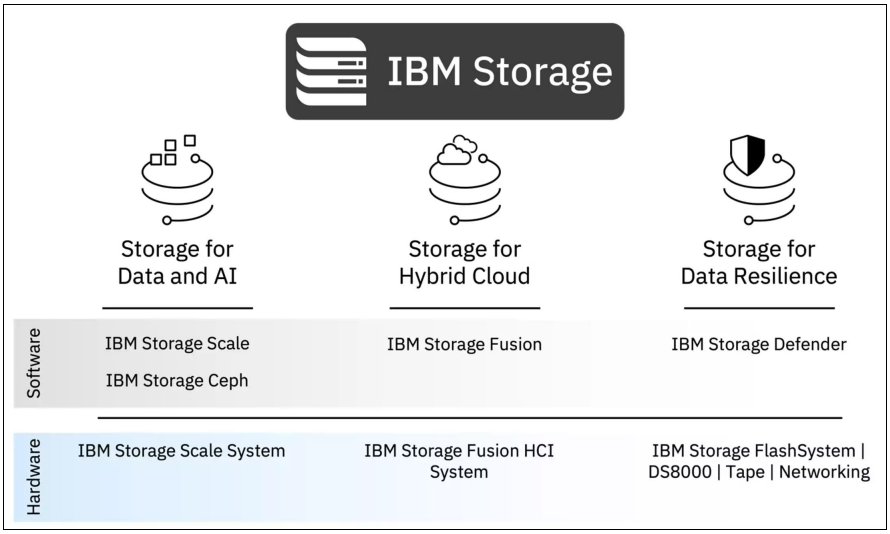

Figure 1 IBM Storage Family

This IBM Redpaper publication provides an overview of the IBM Storage Scale System (formally known as the IBM Elastic Storage System or Elastic Storage Server (ESS)). These scalable, high-performance data and file management solutions, are built on IBM Storage Scale technology. Providing reliability, performance, and scalability, IBM Storage Scale Systems can be implemented to support a wide range of diverse requirements.

The latest IBM Storage Scale System model is the 6000, which builds upon the innovations of the most recent Storage Scale System models. The 6000 is comprised of two 2U canisters for an over all 4U footprint. Each of the 6000's canisters provides eight Gen 5 x16 Peripheral Component Interconnect® Express (PCIe) slots. Target availability for the IBM FlashCore® Modules (FCM) in the 6000 is 1H2024.

This publication is designed to help you understand the solution and its architecture. It describes and provides an overview of the system components and tools used to order the best solution for your environment. This includes planning the installation and the integration of the solution into your environment, as well as correctly maintaining your solution.

The solution is created from the following combination of physical and logical components:

- Hardware
- Operating system
- Storage
- Network
- Applications

Knowledge of the IBM Storage Scale System components is key when planning an environment.

This paper is targeted toward technical professionals (consultants, technical support staff, IT Architects, and IT specialists) who are responsible for delivering cost-effective cloud services, as well as providing AI and big data ready solutions.

Authors

This paper was produced by a working with IBM Redbooks (ITSO), Tucson Center.

Phillip Gerrard is a Project Leader for the International Technical Support Organization working out of Beaverton, Oregon. As part of IBM® for over 17 years he has authored and contributed to hundreds of technical documents published to IBM.com and worked directly with IBM's largest customers to resolve critical situations. As a team lead and Subject Matter Expert for the IBM Spectrum® Protect support team, he is experienced in leading and growing international teams of talented IBMers, developing and implementing team processes, creating and delivering education. Phillip holds a degree in computer science and business administration from Oregon State University.

Nilton Carlos Dos-Santos is a Senior IT Specialist with more than 20 years of experience in several different roles. He joined IBM in 2008 and for the past 5 years he has been working in the IBM Storage space. He holds a bachelor’s degree in computer science and a post-graduate degree in Computer Network Systems. Nilton has also co-authored other IBM Redbooks® publications, including IBM Software Defined Infrastructure for Big Data Analytics Workloads, IBM Platform Computing Solutions for High Performance and Technical Computing Workloads, IBM Redpaper publication IBM Platform Computing Cloud Services and Getting Started with Docker Enterprise Edition on IBM Z®.

Stieg Klein is a Senior Storage Technical Specialist in IBM Advanced Technology Group (ATG). He joined IBM in 2012 and is focused on IBM software-defined storage (SDS), including IBM Storage Scale System (SSS) and IBM Storage Scale. He holds a Bachelor of Arts degree in Computer Science from the University of California, Berkeley.

Chris Maestas is the Chief Executive Architect for IBM File and Object Storage Solutions with over 25 years of experience deploying and designing IT systems for clients in various spaces. He has experience scaling performance and availability with various file systems technologies. He has developed benchmark frameworks to test out systems for reliability and to validate research performance data. He also has led global enablement sessions online and face to face, where he described how best to position mature technologies like IBM Storage Scale with emerging technologies in cloud, object, container, or AI spaces.

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

Larry Coyne
International Technical Support Organization

Rezaul Islam
IBM Systems

John Sing
IBM Global Sales

Thanks to the authors of prior editions:

Dino Quintero, Puneet Chaudhary, Lee Nee Helgeson, Brian Herr, Steven Normann, Marcelo Ramos, Richard Rosenthal, Robert Simon, Marian Tomescu, Richard Wale

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

- 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

Introducing the IBM Storage Scale System

This chapter introduces the IBM Storage Scale System family of products. In this chapter, we describe their characteristics, value, systems integration and where solutions fit in today's business environments.

The latest IBM Storage Scale System 6000 provides investment protection to expand or build a new Global Data Platform in the environment, using either new or currently in use storage devices. The Scale Storage System 6000 extends and expands on the IBM Storage Scale System 3500 models features and capabilities. IBM Storage Scale System can scale up or out with different storage mediums in the environment and supports NVIDIA's ConnectX-7 network interface card (NIC). With the appropriate transceivers and adapter cables the system is capable of 100, 200 or 400 Gb Ethernet as well as InfiniBand-HDR, InfiniBand-NDR connectivity, or a combination of both.

This chapter also describes some of the software and hardware characteristics of the IBM Storage Scale System, IBM Storage Scale RAID (Redundant Array of Independent Disks) software, and the storage building block concepts of the solution.

This chapter includes the following topics:

Note: In this IBM Redpaper publication, *IBM Storage Scale System* can refer to any of the following products:

- IBM Elastic Storage Server (the first and second generations of this solution).
- IBM Elastic Storage System (the third generation of this solution includes the IBM Elastic Storage System 5000 (IBM Storage Scale System 5000), IBM Elastic Storage System 3500 (IBM Storage Scale System 3500), IBM Elastic Storage System 3200 (IBM Storage Scale System 3200), and IBM Elastic Storage System 3000 (IBM Storage Scale System 3000), IBM Storage Scale System 6000.
- For more information about all the current IBM Storage Scale System products, see IBM Storage Scale System documentation.
- For more information, see the following IBM Redbooks publications:
  - *Implementation Guide for IBM Elastic Storage System 3000, SG24-8443*
  - *Implementation Guide for IBM Elastic Storage System 5000, SG24-8498*
  - *Implementation Guide for IBM Elastic Storage System 3200, SG24-8516*
  - *Implementation Guide for IBM Elastic Storage System 3500, SG24-8538*
  - *Implementation Guide for IBM Storage Scale System 6000, SG24-8566*

The general architecture/capabilities of these solutions are similar. Where suitable, we differentiate between the models and generations.

1.1 IBM Storage Scale System

The IBM Storage Scale System is an integrated hardware and software storage solution for high-performance, reliable file and object storage. Storage Scale System 6000 is designed for massive scalability, supporting thousands of nodes and yottabytes of capacity. It leverages IBM Storage Scale RAID erasure coding for enhanced data efficiency, hardware failure mitigation, and intelligent monitoring with dynamic data tuning. Installations and updates are delivered by means of containerized software that speeds and simplifies the maintenance process.

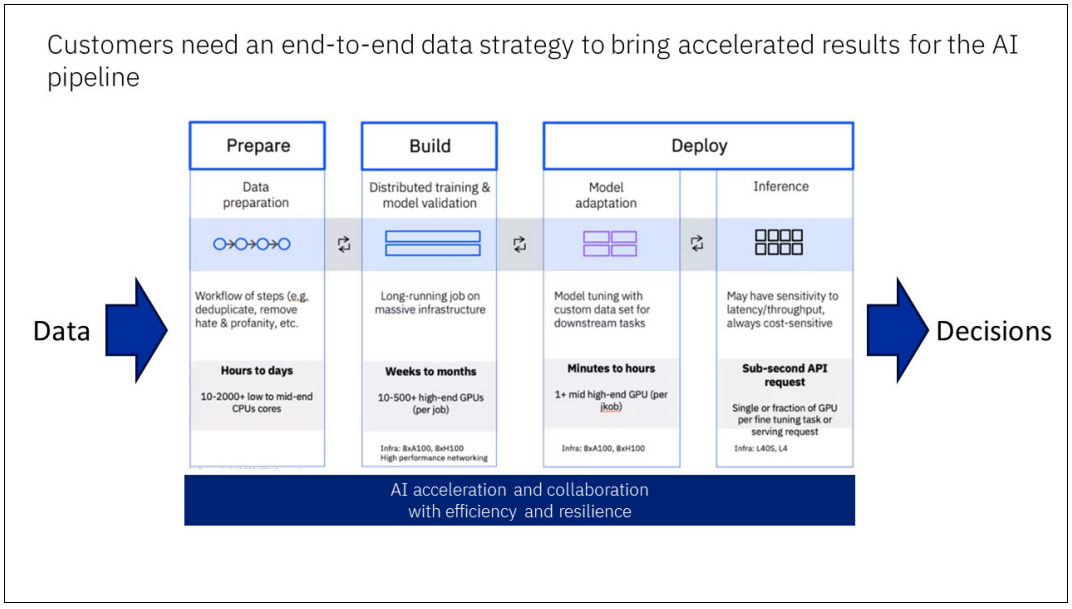

Figure 1-1 IBM Storage Scale end-to-end data strategy

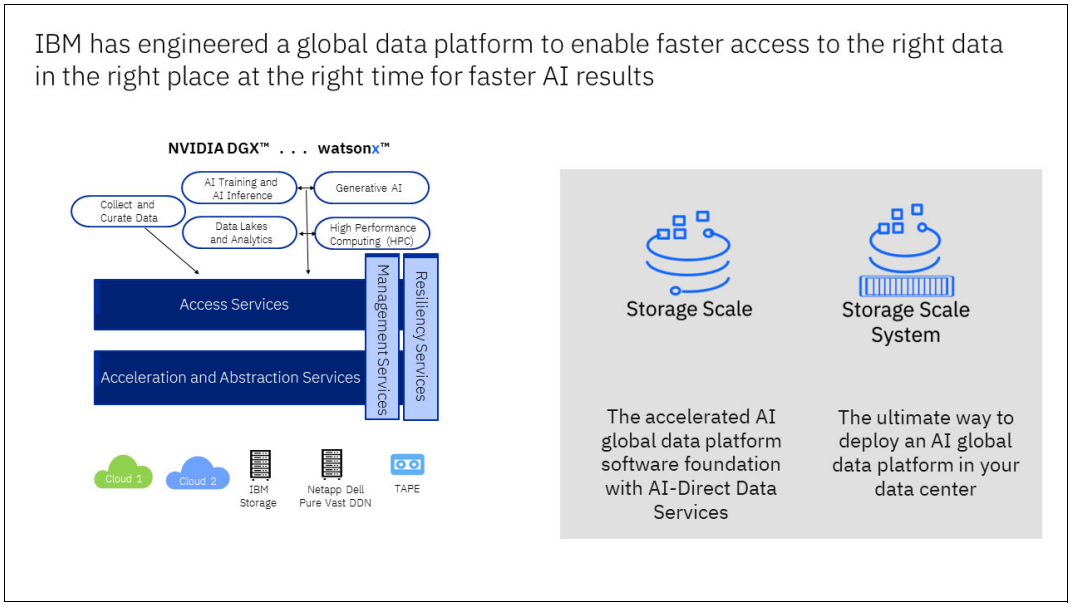

Figure 1-2 IBM Storage Scale with the Global Data Platform

Included in an IBM Storage Scale System is a pair of Storage Scale Data Servers, which are deployed as a fully integrated, tested storage building block solution (see Figure 1-3). Instead of hardware-based disk RAID controllers, IBM Storage Scale System uses software defined Scale Native RAID, providing declustered erasure code technology to deliver consistent high performance. This feature mitigates storage drive failures and RAID rebuild overhead and improves storage efficiency.

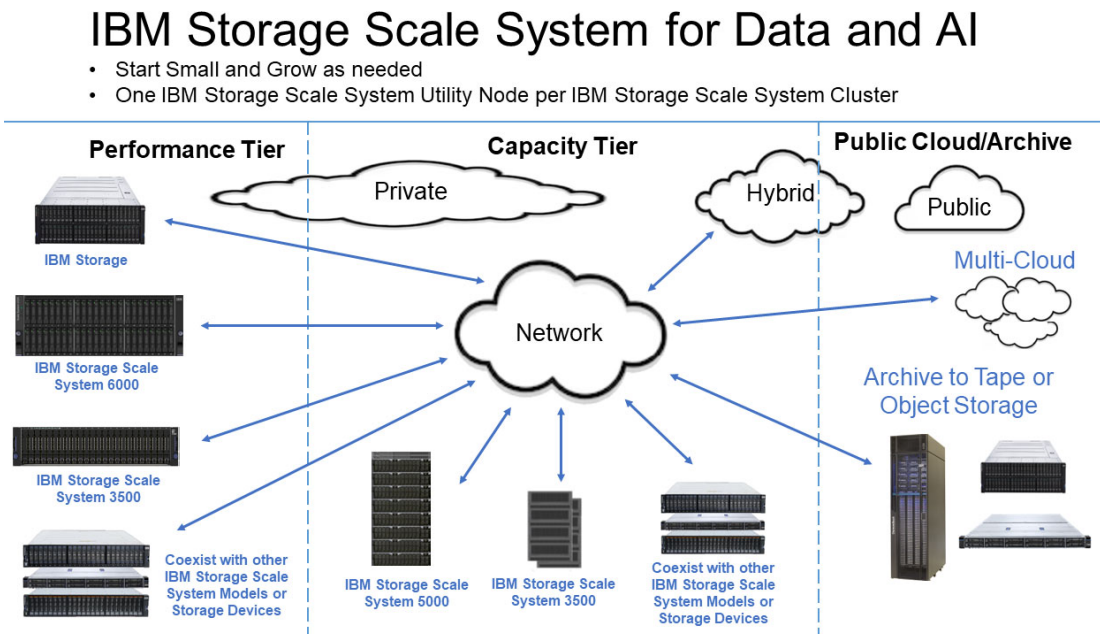

Figure 1-3 Third-generation IBM Storage Scale System integrated with IBM Storage Scale

IBM Storage Scale System with IBM Storage Scale RAID also provides superior data protection by reducing rebuild times to a fraction of the time required to recover using conventional hardware-based RAID controllers.

Data access to IBM Storage Scale System attached storage is available by using one of the following methods:

- IBM Storage Scale clients on user workstations, which provide a POSIX file system interface to data on IBM Storage Scale System.
- IBM Storage Scale clients on NVIDIA GPU systems, which provide GPU Direct Storage via cuda implementation, reducing latency by directly access storage from the GPU.
- IBM Storage Scale protocol nodes, which allow IBM Storage Scale System data access through Server Message Block (SMB), Network File System (NFS), or S3 object data access. In addition, Hadoop clusters can access IBM Storage Scale System data by using the IBM Storage Scale Hadoop Transparency Connector.
- IBM Storage Scale Container Native Storage Access (CNSA) or Container Storage interface (CSI) drivers that run on OpenShift or Kubernetes environments.

Multiple IBM Storage Scale System building blocks can also be integrated into an IBM Storage Scale cluster to suit the storage needs of the environment.

IBM Service supports IBM Storage Scale System as an integrated solution providing support for the full software and hardware stack.

1.2 IBM Storage Scale System and IBM Storage Scale

In this section, we describe how IBM Storage Scale System is part of an IBM Storage Scale cluster.

Within the IBM Storage Scale cluster, IBM Storage Scale System is an integrated IBM Storage Scale building block that reads and writes data for IBM Storage Scale users. IBM Storage Scale System combines storage hardware with IBM Storage Scale software to manage any data in storage.

In the following sections, we provide a brief overview of these topics:

- What is an IBM Storage Scale cluster
- What is an IBM Storage Scale Client and IBM Storage Scale Data Server
- How IBM Storage Scale System is an integrated IBM Storage Scale storage building block

Note: An IBM Storage Scale System must be installed and configured, including integration into an IBM Storage Scale cluster, before it can be used to hold user data. In isolation, an IBM Storage Scale System is not sufficient for an IBM Storage Scale cluster. IBM Storage Scale Server Management functionality is required as well.

In this section, we provide a high-level IBM Storage Scale and IBM Storage Scale System solution overview. For more detailed information, see Chapter 2, “IBM Storage Scale System architecture” on page 21.

1.2.1 IBM Storage Scale overview

IBM Storage Scale is IBM’s strategic high-performance parallel file system, a shared storage platform for end-to-end collaborative common enterprise, data platform, big data analytics, and AI workflows (see Figure 1-4 on page 9).

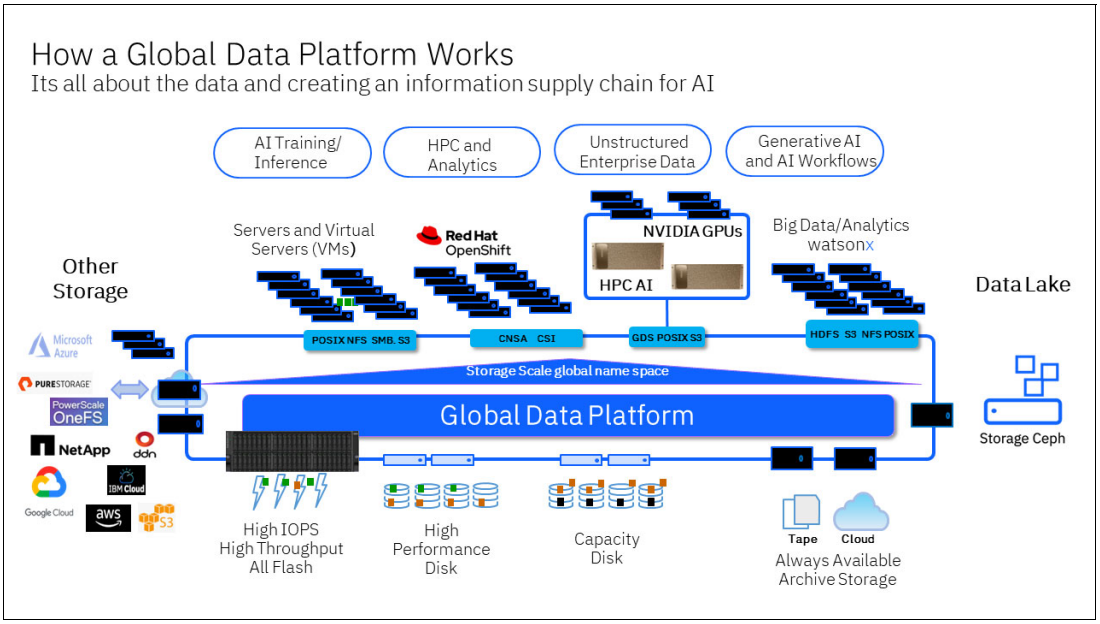

Figure 1-4 IBM Storage Global Data Platform with IBM Storage Scale overview

IBM Storage Scale

IBM Storage Scale is designed to provide the following major value propositions:

- Simplified data management by supporting enterprise workflows on a single common enterprise data platform via a variety of access methods to reduce data silos.
- A single global namespace that supports enterprise-level data over high-performance networks to accelerate your workloads to the max available bandwidth or POSIX transactions per second (File IOPS).
- Abstracts and allows for intelligent automatic tiering of data between storage pools, externally to tape, to object based and cloud resources. This delivers cost-effective storage economics by automatically managing and tiering data to different classes of storage including data targets on remote S3 or NFS storage that may already exist today.

Although multiple types of IBM Storage Scale cluster configurations are available, the configuration into which IBM Storage Scale System is commonly deployed is the IBM Storage Scale Network Shared Disk (NSD) configuration, as shown in Figure 1-5.

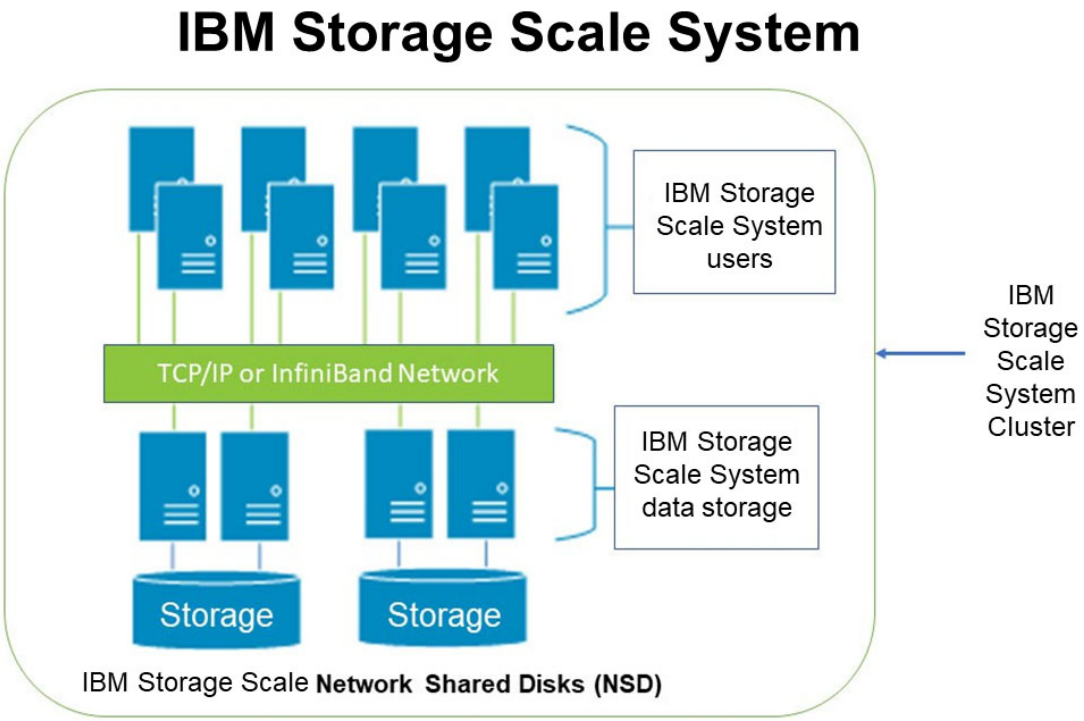

Figure 1-5 IBM Storage Scale cluster component overview

IBM Storage Scale System contains a pair of IBM Storage Scale NSD Data Servers, which are configured together as a tested, integrated, highly available (HA), and reliable IBM Storage Scale storage building block based solution.

As shown in this example, eight IBM Storage Scale nodes are workstations, servers, or users. Four nodes are IBM Storage Scale data server nodes. In this configuration the user workstations are running the IBM Storage Scale client in this configuration. The IBM Storage Scale client provides multi-threaded, highly parallel data access to the IBM Storage Scale Data Servers, as shown in Figure 1-6.

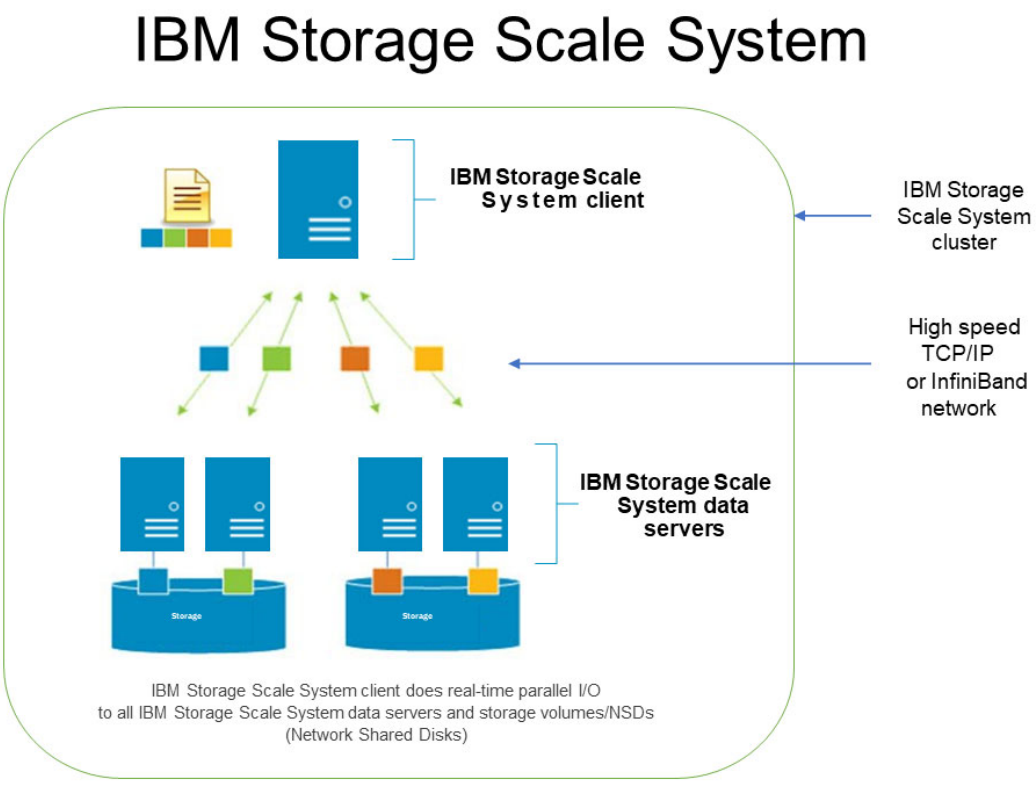

Figure 1-6 IBM Storage Scale clients and servers

The IBM Storage Scale client achieves high performance by performing simultaneous real-time parallel I/O to all IBM Storage Scale data servers, storage volumes and NSDs simultaneously.

IBM Storage Scale clusters can grow by adding nodes, whether they are IBM Storage Scale clients or IBM Storage Scale data servers, as shown in Figure 1-7.

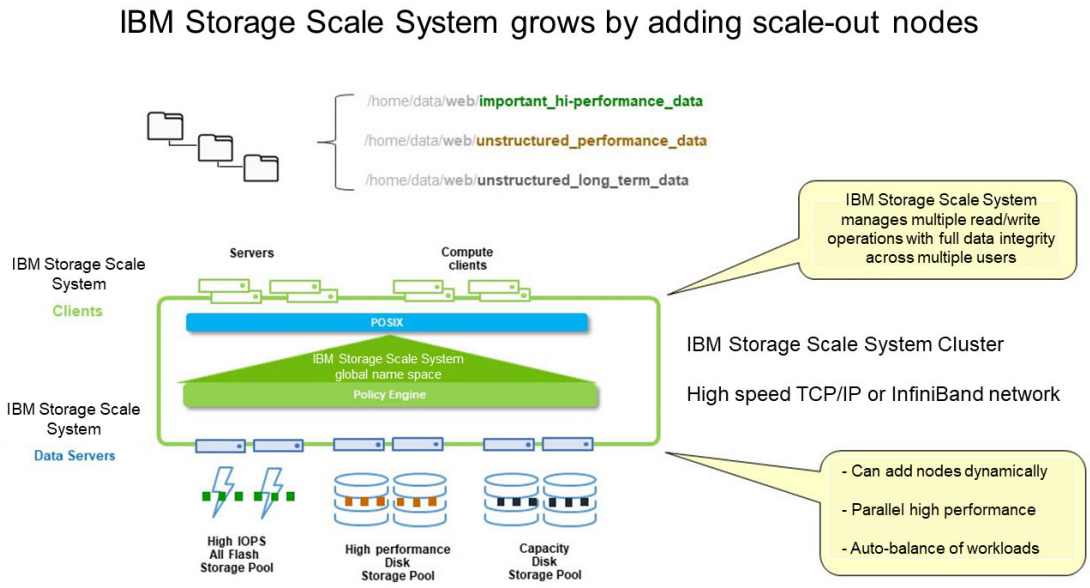

Figure 1-7 IBM Storage Scale grows by adding scale-out nodes

An IBM Storage Scale cluster can provide 1 - 256 logical POSIX file systems to users and workstations. The IBM Storage Scale client provides the appearance of a mountable POSIX file system to the applications and users on the workstation where the IBM Storage Scale client is installed.

IBM Storage Scale users are unaware of the physical distribution of data in the IBM Storage Scale data server physical storage pools. The automatically balanced data distribution is seamlessly determined by the IBM Storage Scale policy engine at the time that the data is imported. The policy engine can also transparently move data from one storage pool to another storage pool while the data is accessed and active.

The IBM Storage Scale parallel file system provides an enterprise with the capability for data management over large amounts of data, while also performing constant auto-balance of workload and storage by equally distributing I/O and data within a storage pool or among different storage pools.

The preferred method of accessing IBM Storage Scale data is to install the IBM Storage Scale client on every workstation or server that accesses IBM Storage Scale data. The IBM Storage Scale client provides multiple threads and communication with data servers to provide high-performance parallel throughput. While doing so, IBM Storage Scale also manages full read/write data integrity between users who are working with the data in the file system.

1.2.2 Protocol nodes overview

As an optional part of the IBM Storage Scale solution, protocol nodes allow access to IBM Storage Scale data by using NFS, SMB, or object protocols without installing IBM Storage Scale software on a node.

The primary benefit of protocol nodes is that applications, workstations, and users that do not have the IBM Storage Scale client can still access IBM Storage Scale data through one of these configured protocols. IBM Storage Scale is often used as an enterprise data lake or central enterprise data repository.

Protocol nodes are IBM Storage Scale nodes that are designed, configured, and set up to provide the following benefits:

- NFS, SMB, or object protocols on the customer-facing side
- A full IBM Storage Scale client that provides parallel access to IBM Storage Scale data on the storage-facing side

Figure 1-8 shows an example of protocol nodes that were added to an IBM Storage Scale cluster.

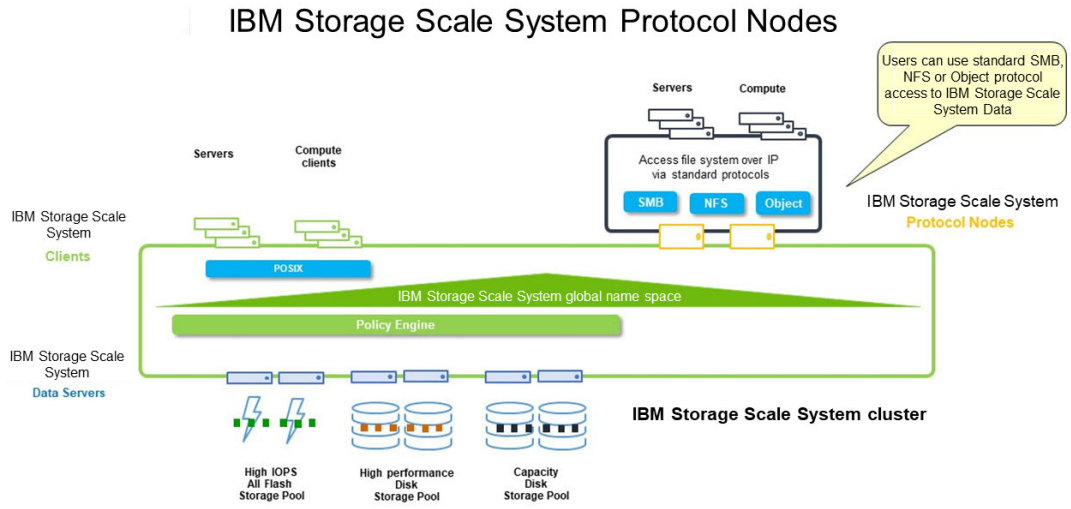

Figure 1-8 IBM Storage Scale System protocol nodes

When IBM Storage Scale data is accessed through protocol nodes, the limiting factor on performance is the single-threaded NFS, SMB, or object protocols. These protocols cannot deliver the same highly parallelized performance of the IBM Storage Scale client. However, the value of accessing the IBM Storage Scale data through NFS, SMB, or object protocols is a flexible means for allowing enterprise-wide user access to data stored in an IBM Storage Scale system.

1.2.3 IBM Storage Scale System storage building block solution for IBM Storage Scale

IBM Storage Scale System with IBM Storage Scale is a pair of IBM Storage Scale Data Servers that are cross-configured together into an integrated, highly available IBM Storage Scale building block. IBM Storage Scale System provides a fully tested, integrated, and supported IBM solution. Storage Scale allows you to deploy, manage, maintain and expand/grow the environment as needed, using a Storage Scale building block approach.

Two IBM Storage Scale System models are deployed as a cluster to provide most of the IBM Storage Scale storage, as shown in Figure 1-9.

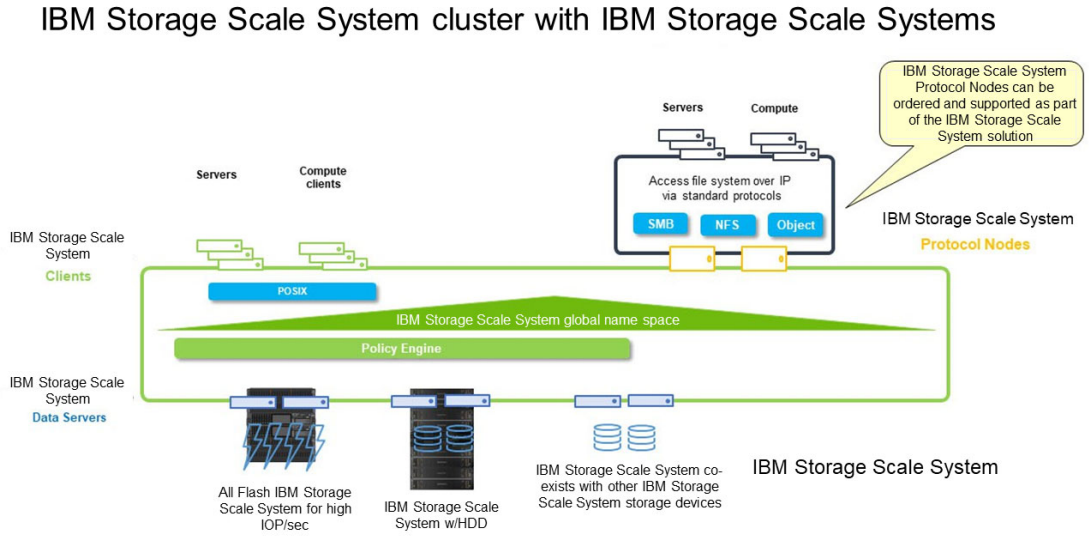

Figure 1-9 IBM Storage Scale Systems as IBM Storage Scale cluster building blocks

IBM Storage Scale System can coexist with other IBM Storage Scale data servers in the same IBM Storage Scale cluster. IBM Storage Scale System provides a powerful, flexible, fully integrated, and supported option for deploying IBM Storage Scale storage in the most efficient manner possible.

An IBM Storage Scale System storage building block is designed as a unit of storage expansion for an IBM Storage Scale cluster. If more storage is needed, an IBM Storage Scale System I/O enclosure or additional IBM Storage Scale System building blocks can be dynamically added to expand an IBM Storage Scale cluster.

To provide high reliability and consistent high performance, IBM Storage Scale System runs IBM Storage Scale native RAID (erasure coding), which is designed to maintain high performance even while transparently recovering from storage media or storage data server failures. IBM Storage Scale native RAID also provides checksum and disk hospital functions to detect and correct silent data corruption issues and maintain high availability (HA) of exabyte scale file systems.

As part of an IBM Storage Scale System order, you can also order IBM provided servers to use as IBM Storage Scale protocol nodes. These protocol nodes are managed by the IBM Storage Scale System solution software stack and the IBM Storage Scale System GUI. When including IBM Storage Scale System and IBM Storage Scale protocol nodes this provides a completely integrated solution for an IBM Storage Scale cluster.

For more information about planning for the usage of protocol nodes, see 2.2.5, “IBM Storage Scale System Utility Node Platform - Management Server/Protocol Node” on page 29.

1.3 IBM Storage Scale and Storage Scale System solution value

The following sections describe why the IBM Storage Scale and IBM Storage Scale System brings advantages to the business.

1.3.1 IBM Storage Scale

The unstructured and semi-structured data from AI workloads, advanced analytics, data lakes, and other data-intensive apps must be stored in distributed file and object systems to make it accessible to geographically dispersed applications, services, and devices.

IBM Storage Scale software is designed to address these requirements with global data abstraction services that provide connectivity from multiple data sources and multiple locations to bring together data wherever it lives, including non-IBM storage environments.

Storage Scale achieves this with active file management (AFM),which abstracts other storage systems, including customers' existing on-prem or public cloud storage, while providing a transparent caching layer that improves performance by unifying data into a single global namespace, accessible simultaneously via multiple protocols.

Built on a massively parallel file system,Storage Scale supports deployment across x86, IBM Power, IBM Z, ARM-based POSIX clients, virtual machines, and Kubernetes. It delivers high-performance access to unstructured data across any protocol, location, or format, helping customers accelerate AI pipelines and unlock the value of their data estate.

1.3.2 IBM Storage Scale System

Today's AI and big data applications require large, high-performance, manageable, and flexible storage to be successful. As more and more data is generated, datasets grow and new language models are created, it is also important to be able to allow that storage to grow with the enterprises needs.

The rapid rise of artificial intelligence (AI), high-performance computing (HPC), analytics, and hybrid cloud is transforming the business world. Unlike traditional applications that rely on structured databases, these modern workloads and data lakes process vast amounts of unstructured data, including documents, audio, images, and videos.

To stay competitive, organizations must rethink their storage strategies and leverage AI to unlock the value of their data—wherever it resides. IT leaders face key challenges, including:

- Accessing and analyzing data and workloads scattered across the globe.
- Managing the growing AI infrastructure and ensuring scalability for evolving workloads.
- The increasing time needed by AI training and inferencing workloads.

Addressing these challenges requires specialized software and hardware:

- IBM Storage Scale is software-defined file and object storage for both structured and unstructured data.
- IBM Storage Scale System 6000 is a hardware implementation of Storage Scale software and is optimized for the most demanding AI, HPC, analytics, and hybrid cloud workloads.

IBM Storage Scale System 3500 is for customers requiring an enterprise-ready entry-level or mid-level system.

The IBM Storage Scale System is designed to provide an integrated and tested IBM Storage Scale storage building block to meet these growing storage needs. Incorporating IBM Storage Scale System storage servers flexibly adds to the overall IBM Storage Scale capacity, bandwidth, and performance, all within a single global namespace.

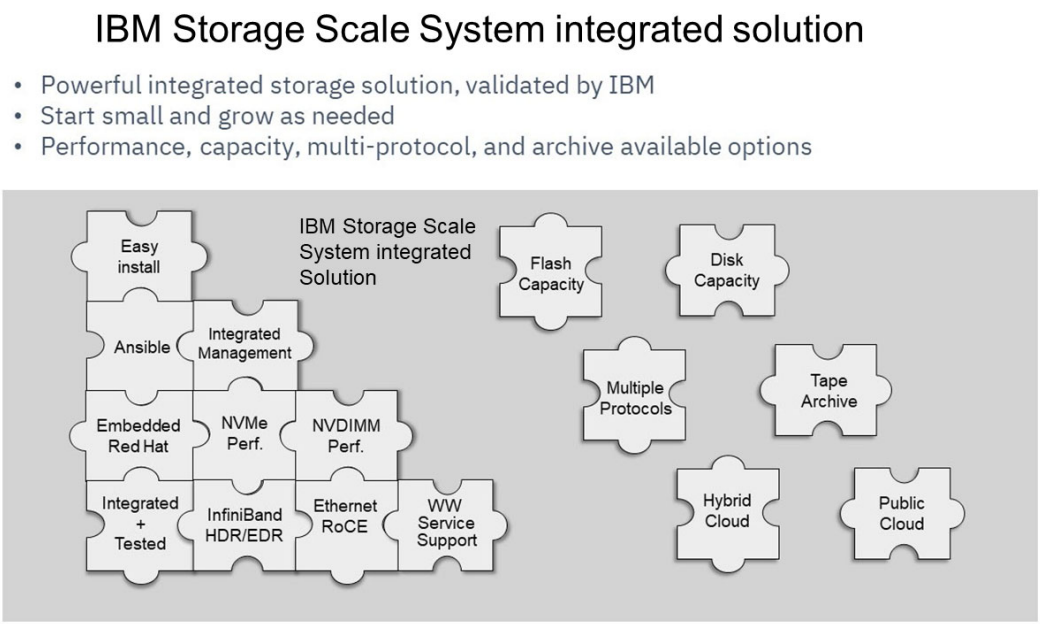

Figure 1-10 IBM Storage Scale System integrated solution

As shown in Figure 1-10, IBM Storage Scale System is an integrated IBM Storage Scale solution that significantly accelerates the time-to-value of deploying IBM Storage Scale environments. IBM Storage Scale System provides the following benefits:

- An integrated and tested IBM Storage Scale storage building block
- Predictable, consistent high performance at any level of scale
- Tested and integrated with the functions of IBM Storage Scale
- Supported as an integrated solution worldwide by IBM Service and Support

For more information about the IBM Storage Scale System, see the following web pages:

- IBM Storage Scale System
- IBM Elastic Storage Server documentation

1.3.3 Integrated and reliable storage building block

Every part of the IBM Storage Scale System solution is engineered to provide high performance enterprise wide data availability and reliability, from an entry level system up to massive scale multi node configurations.

IBM engineering and testing teams work cohesively on designing, building, testing, and delivering an end-to-end IBM Storage Scale System solution which is integrated and tested for reliability. The complete IBM Storage Scale System solution is again verified during manufacturing.

In the deployment stage, IBM Storage Scale System-specific tools and scripts are used for deployment. IBM Systems Lab Services are available and recommended to provide installation of the IBM Storage Scale System solution and, to assist you in integrating IBM Storage Scale System into your client environment.

IBM Storage Scale System provides optimized storage configurations that include the following benefits:

- Optimum adapters that are strategically placed on the servers

- HA and redundancy for disks, drawers, and adapters
- Optimal cabling performance
- Tested and integrated firmware and software versions
- Policy-managed optimized placement of data
- HA access to data
- Automated storage management

IBM Storage Scale System is capable of 100, 200 or 400 GB Ethernet or InfiniBand-HDR-200, InfiniBand-NDR-200/400 connectivity, or a combination of both.

1.3.4 Predictable and consistent high-performance

IBM Storage Scale System uses IBM Storage Scale RAID, which is a declustered RAID erasure code technology that recovers from multiple disk failures in minutes, versus hours or days with older technology RAIDs. This reduced recovery time gives the IBM Storage Scale System solution predictable performance and data protection including 8+2 and 8+3 RAID protection, and platter-to-client data protection.

A high-performance parallel file system environment, such as IBM Storage Scale, can run only as fast as its slowest component. IBM Storage Scale System is designed to use IBM Storage Scale RAID to provide consistent high performance by masking and mitigating performance effects of storage hardware failures, even if multiple drive failures occur. In this way, IBM Storage Scale System assures that the larger IBM Storage Scale parallel file system always runs at optimum efficiency and consistently delivers expected high performance, even if storage hardware or media failures occur.

1.3.5 Simplified IBM Storage Scale System installation

IBM Storage Scale System technology deployment methods use Red Hat Ansible driven container orchestration, which results in faster installation, upgrades, with less effort and skills required for deployment and maintenance (see Figure 1-11).

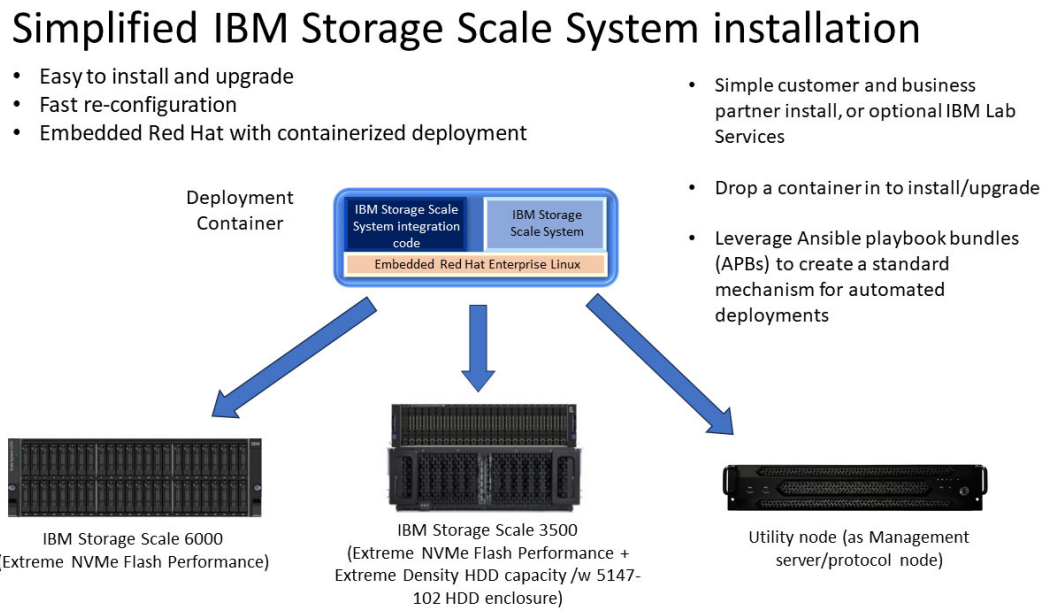

Figure 1-11 Simplified IBM Storage Scale System installation

As shown in Figure 1-11, IBM Storage Scale System provides a fully tested and integrated solution that is composed of Red Hat Enterprise Linux (RHEL), IBM Storage Scale, and IBM Storage Scale System integration code. Installation modules are done using Ansible-driven container orchestration and Ansible playbook bundles (APBs). This creates a standard, fast, and manageable mechanism for faster automated deployment of IBM Storage Scale System.

1.3.6 Supported as a solution

The IBM Storage Scale System is supported as a solution. Clients who engage IBM Service and Support for assistance on IBM Storage Scale System receive support from a dedicated IBM team that is cross-trained on all IBM Storage Scale System (hardware and software) components for the solution.

This specialized IBM Storage Scale System Service and Support team can debug and advise on all aspects of IBM Storage Scale System.

If more detailed support is needed for an IBM Storage Scale System component, IBM Storage Scale System service and support records show and document the problem resolution progress for the overall IBM Service and Support team. The resulting fast time to resolution contributes to the added value of purchasing an integrated, tested IBM Storage Scale solution from IBM.

For more information about IBM service and support of the IBM Storage Scale System, see IBM Storage Scale System (SSS) Support Reference Guide (IBMid log in required).

1.4 IBM Storage Scale System models

IBM Storage Scale System is an IBM Storage Scale storage building block that is available in various models. These models offer a flexible portfolio of IBM Storage Scale storage building block options with NVMe or hard disk drive (HDD) storage.

The general physical arrangement for an IBM Storage Scale System is an IBM rack that features the following components:

- A pair of IBM Storage Scale I/O Data Servers
- Storage enclosures
- Network switches

An example of the various models of IBM Storage Scale System over time is shown in Figure 1-12.

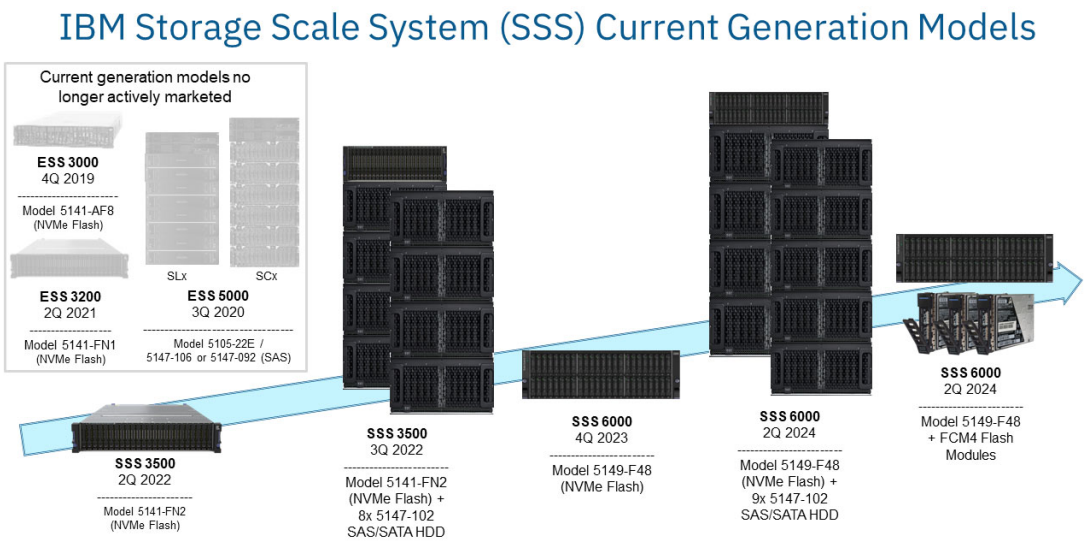

Figure 1-12 Current generation IBM Storage Scale System models

2

IBM Storage Scale System architecture

This chapter describes the architecture, hardware, and software of the IBM Storage Scale System, including the following components:

- IBM Storage Scale Management Server (IBM EMS)
- IBM Storage Scale Utility Node (Services for Management, Protocol, Active file Management (AFM), Protect and in the future Data Cataloging)
- I/O /data servers
- Storage enclosures

This chapter also describes the different building block models and software components and features, such as:

- Operating system
- Ansible playbooks and container-based software upgrades
- IBM Storage Scale
- IBM Storage Scale RAID

Note: IBM Storage Scale System is an integrated packaged solution. Users should *not* install different kernel levels or drivers into the IBM Storage Scale System. The user also should *not* run any non IBM Storage Scale System (client) application or non IBM Storage Scale System workloads on the I/O nodes or IBM Storage Scale System Utility Node Servers.

This chapter includes the following topics:

2.1 Architecture overview

In this section, we describe the overall hardware and software architecture of the IBM Storage Scale System solution.

An IBM Storage Scale System is defined as a combine set of hardware and software that are tightly coupled together and tested as a single unit. The major components of an IBM Storage Scale System solution release include the following items:

- Server hardware
- Storage hardware
- IBM Storage Scale software
- Embedded Red Hat Enterprise Linux (RHEL) operating system

The minor components, which are integration-tested with the major components, include (but not limited to) the following items:

- Server firmware
- Server host bus adapters (HBAs) and related firmware
- NVIDIA network adapter, drivers, firmware, and network switches
- Storage enclosure and drive firmware
- IBM racks, power distribution units (PDUs), and cabling management

In addition, many of the hardware and software components of the IBM Storage Scale System solution communicate with each other over an IP network. IBM Storage Scale System requires all networking connections to be in place and provisioned to install the system. This IP network can use NVIDIA network switches, which are tested with IBM Storage Scale System by IBM, or the network switches can be provided by the client. These IP network components consist of the following switches:

- Low speed 1/10 Gb Ethernet (GbE) network switches for management and service networks
- High-speed network switches for data read and write over a high-speed data network

2.2 IBM Storage Scale System Hardware components

In this section, we describe the following hardware components:

- Solution models
  - 6000
  - 3500
- Server hardware
  - Current generation IBM Storage Scale System 6000 and 3500 use x86 based servers for storage.
- Storage enclosures
- Network interface cards (NICs) and network switches
- Rack

IBM Storage Scale Systems are a rack-mounted storage solution. The initial storage building block in an IBM Storage Scale System cluster includes the following components:

- One IBM Storage Scale Management Server (one is required for every IBM Storage Scale cluster).
- One or more IBM Storage Scale System, each of which include redundant I/O data servers

Current generation IBM Storage Scale System models include various models that provide NVMe, serial-attached SCSI (SAS), or Serial Advanced Technology Attachment (SATA) hard disk drives (HDDs) in selected sizes.

The IBM Storage Scale System 6000, which is the latest third-generation IBM Storage Scale System model, was announced and delivered in 4Q2023. The IBM Storage Scale System 6000 design builds on the IBM Storage Scale System 3500, with more powerful processors and additional PCIe Gen 5 x 16 lanes that are both faster and wider than those supported by the 3500. Additionally, several design updates over the 3500 improve serviceability and provide additional system configuration flexibility, e.g. optional campus connection.

The IBM Storage Scale System 3500, was announced and delivered in 2Q2022. The IBM Storage Scale System 3500 design is based on prior IBM Storage Scale System models. The IBM Storage Scale System 3500C/H models support up to eight external storage enclosures.

2.2.1 IBM Storage Scale System actively marketed models

This section describes the current IBM Storage Scale System models that are available to order from IBM:

- IBM Storage Scale System 6000: IBM machine type 5149-F48 for the 4U24 from factor IBM Storage Scale System with embedded NVMe offer options of 3.84TB, 7.68TB, 15.36TB or 30.75TB drives or IBM FlashCore Modules with 19.2TB and 38.TB drives of flash storage, can optionally be seamlessly extended with expansion enclosures, Each unit supports up to 91 SAS HDDs with a choice of 12 TB, 16 TB and 20 22 TB Self Encrypting Drives (SED), scaling-up to 18PB per rack over 24Gb SAS, ensuring security and performance.
- IBM Storage Scale System 3500: IBM machine type 5141-FN2 for the 2U24 form factor IBM Storage Scale System with embedded NVMe flash storage, and optionally 1 - 8 of the 4U102 form factor IBM Storage Scale System Storage Enclosures model 5147-102

For more information about previous IBM Storage Scale System models, see the IBM Redbooks publication “IBM Elastic Storage System Introduction Guide”.

2.2.2 IBM Storage Scale System 6000

IBM Storage Scale System 6000 is an All-flash, integrated, end to end NVMe storage solution, providing industry leading performance, scalability and unrivaled data protection and availability. The 6000 provides 2x the performance and storage capacity relative to the previous generation systems and is the most sustainable storage solution to date.

IBM Storage Scale System 6000 is a high performance, accelerated storage offering optimized for data intensive workloads for advanced AI, analytics and high performance computing. The 6000 delivers significant advancements to the data services provided by the global data platform.

The Storage Scale System 6000 leverages the power of Storage Scale combined with NVMe/flash technology in a x86 PCIe Gen5 4U48 form factor to deliver the ultimate high-performance storage for AI, data analytics and any high-performance file and object use cases. With support for IBM's Flash Core Module technology, the 6000 can support up to

1.8PB of raw flash capacity and 5.5PB of compressed flash storage in a single 4U building block.

The Storage Scale System 6000 can provide over 250 gigabytes per second (GBps) of throughput performance with low latency to optimize all types of I/O patterns and data layouts. This delivers data to applications faster and helps ensure maximum GPU resource utilization. The modular and scalable architecture supports 24 or 48 drives and is expandable to 1.8PB of flash capacity.

The 6000 supports up to 16 ports of 100Gb RDMA over Converged Ethernet (RoCE), InfiniBand-HDR-200, InfiniBand-NDR-200/400 or a combination of both. The two redundant high-performance controllers in a 2U platform support high availability and reliability. Embedded baseboard management controller (BMC) and remote control capability enable system administrators to manage from a distance.

IBM Storage Scale System 6000 is an ideal storage platform for: Compute intensive AI model training and inferencing for GPU accelerated workloads requiring low latency, and high OPS. Accelerated AI, delivering performance, scalability, data virtualization, and data resiliency for IBM's watsonx®.data and watsonx.ai® Industry leading throughput for data intensive high performance computing solutions.

IBM Storage Scale System is designed to enable you to break data access barriers and create new storage dynamics for NVIDIA AI solutions. With IBM Storage for data and AI, you can create a global data platform with top-rated performance to accelerate data access greater than 125GB/s per node with parallel data access for GPU-accelerated computing with NVIDIA DGXtm systems.

IBM Storage Scale System is an accelerated storage offering optimized for data intensive workloads for advanced AI, analytics, and high performance computing.

IBM Storage Scale System 6000 is an ideal storage platform for:

- Compute intensive AI model training and inferencing for GPU accelerated workloads requiring low latency, and high OPS.
- Accelerated AI, delivering performance, scalability, data virtualization, and data resiliency for IBM's watsonx.data® and watsonx.ai
- Industry leading throughput for data intensive high performance computing solutions

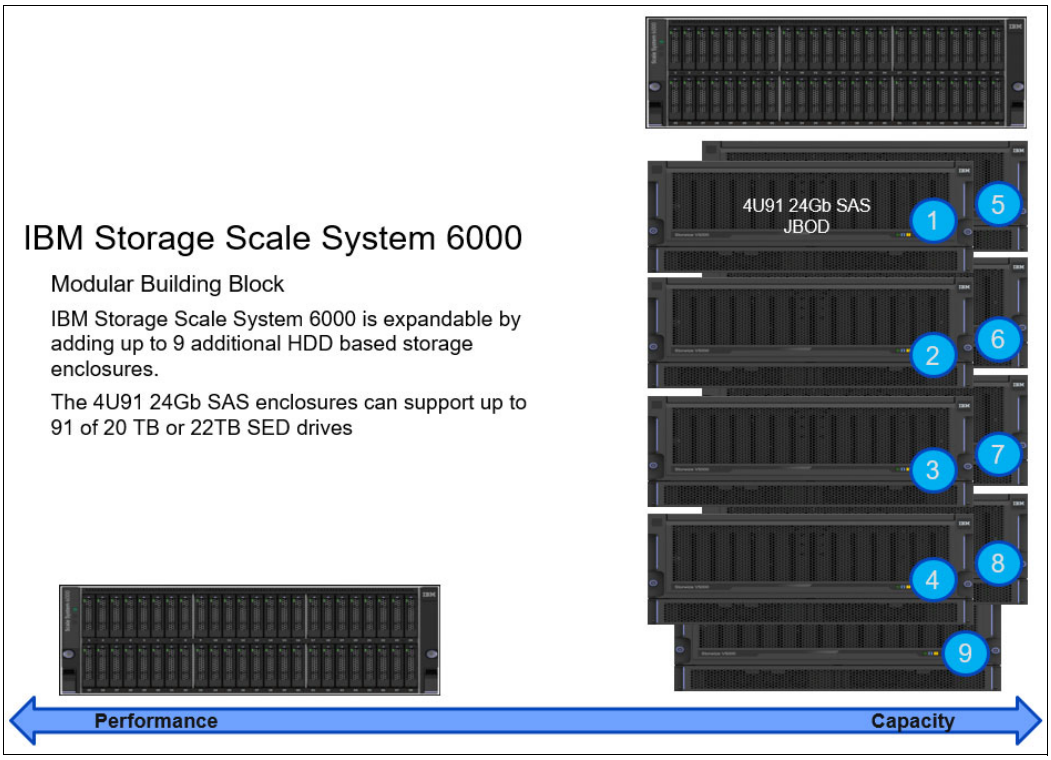

Figure 2-1 IBM Storage Scale System 6000 expansion components example

Figure 2-2 IBM Storage Scale System 6000 front view

Figure 2-3 IBM Storage Scale System 6000 rear view

Table 2-1 Overview of the IBM Storage Scale System 6000

Product	Number of DIMMs Per Server Canister	Total Memory Per Server Canister	Total Raw Capacity Per ISS6k unit	Server Canister Features
IBM Storage Scale System 6000	24 (32 GiB DIMM)	768 GiB	Up to 1440 TB NVMe	Dual socket AMD EPYC 9454 SP5, 48 core processors.
	24 (64 GiB DIMM)	1536 GiB		Dual 960 GB NVMe boot drives.

2.2.3 IBM Storage Scale System 3500

IBM Storage Scale System 3500 is a hybrid model offering high-performance local NVMe storage and capacity storage through SAS or SATA connected storage enclosures.

IBM Storage Scale System 3500 can include 12 or 24 drives of 3.84, 7.68, 15.36, or 30.74 TB NVMe drives for user data. The model 3500C capacity model includes, four NVMe drives that are used exclusively for log tip and other internal scale use. For more information, see IBM Storage Scale System 3500.

IBM Storage Scale System 3500 can include 1 - 8 storage shelves (MTM- 5147-102) of SAS or SATA drives. The first storage shelves can contain 52 or 102 drives. Extra shelves contain 102 drives. The drives are 10, 14, 18, 20 or 22 TB. For more information, see IBM Storage Scale System 5147-102 Storage Enclosure.

IBM Storage Scale System 3500 supports non-disruptive capacity upgrades including the addition of up to a maximum of 8 storage shelves. All HDDs must be the equivalent size and performance.

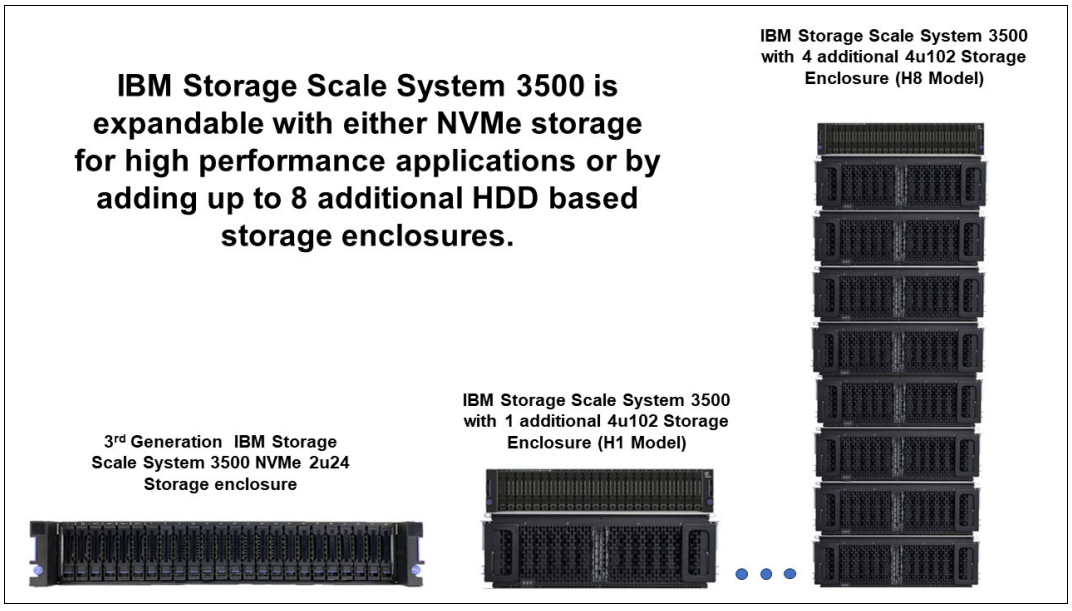

Figure 2-4 IBM Storage Scale System models and expansion example

2.2.4 Positioning various IBM Storage Scale System models

The different models of IBM Storage Scale System storage building blocks provide varying levels of performance and capacity according to the needs of the IBM Storage Scale storage pool in which they are deployed. The various models of third-generation IBM Storage Scale System can provide NVMe, SAS or SATA storage in selected disk sizes.

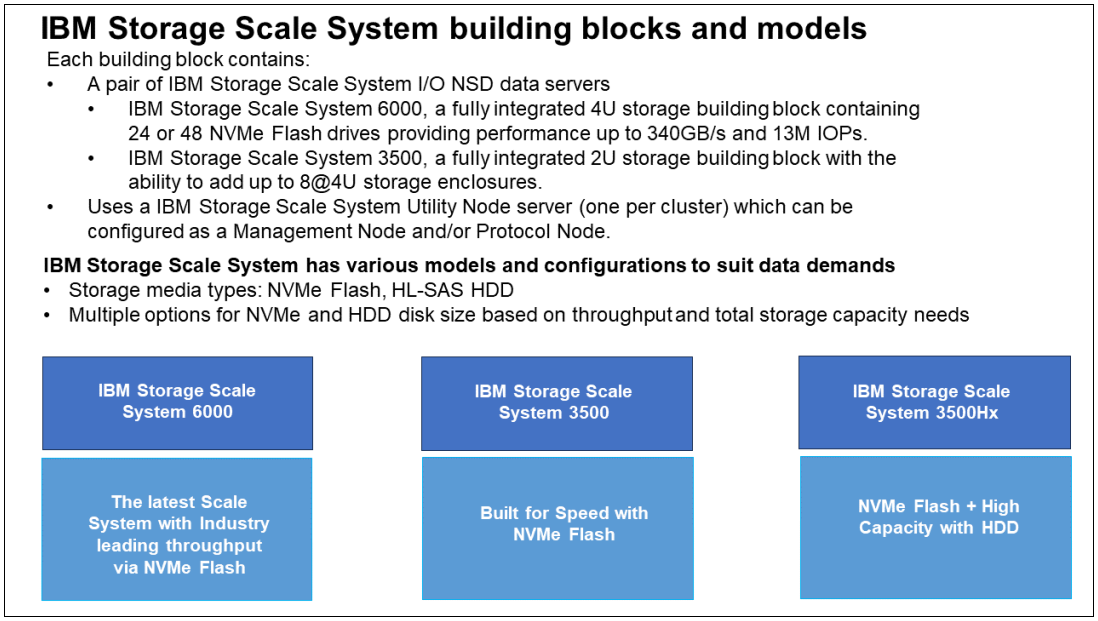

Figure 2-5 General positioning of available IBM Storage Scale System models

IBM Storage Scale System models are available in different categories. The “x” in the models (see Figure 2-6 on page 28) denotes the number of storage enclosures in that model.

Consider the following points:

- IBM Storage Scale System 6000 models are built on NVMe flash storage for high performance. Each IBM Storage Scale System 6000 canister supports up to eight PCIe5 cards. PCI cards are able to be used for either networking or HBA attachment for external storage expansion. A single NIC supports up to two high-speed networking interfaces. A single HBA supports up to two high-capacity external storage enclosures.
- IBM Storage Scale System 3500 models are built on NVMe flash storage for high performance, and optionally high capacity 4U102 HDD storage enclosures. Each IBM Storage Scale System 3500 canister supports up to four PCIe4 cards. A single card either supports up to two high-speed networking interfaces or up to two high-capacity storage enclosures.
- IBM Storage Scale System components also require the following:
  - An IBM Storage Scale System Management Server (see 2.2.5, “IBM Storage Scale System Utility Node Platform - Management Server/Protocol Node” on page 29).
  - An IBM Storage Scale System protocol node (see 2.2.5, “IBM Storage Scale System Utility Node Platform - Management Server/Protocol Node” on page 29).

IBM Scale System Models at a glance						
Models	6000 NVMe / FCM		3500 / 3500Hx NVMe	6000Hx 6000Cx	3500Hx 3500Cx first external enclosure	3500Hx 3500Cx additional external enclosure
	Half or full 2U48 24 or 48 drives		Half or full 2U24 12 or 24 drives	Full only @4U91 storage enclosure 91 drives	Half or full @4U102 storage enclosure 52 or 102 drives	Full only 2 to 8 @4U102 storage enclosure 102 drives
Drive sizes	NVMe 3.84TB 7.68TB 15.36TB 30.72TB	FCM 19.2TB 38.4TB	NVMe 3.84TB 7.68TB 15.36TB 30.72TB	HDD 12TB SED 16TB SED 20TB SED 22TB SED	HDD 10TB SED 14TB SED 18TB SED 20TB SED 22TB SED	
* all Flash media is SED capable				6000Cx and 3500Cx has 4 NVMe		

Figure 2-6 IBM Storage Scale System models

IBM Storage Scale System usable capacity is a percentage of the raw capacity. The usable capacity varies, depending on the IBM Storage Scale RAID parity that is selected for use. Generally, approximately 73% of the raw capacity is the usable capacity when the default 8+2P parity is used, which is commonly selected parity for most scenarios.

IBM or IBM Business Partners can provide the exact usage capacities of the various models on request by using the IBM Storage Modeler (StorM) to calculate the exact usable capacity. See your IBM or IBM Business Partner representative for assistance in determining the usable capacity that your IBM Storage Scale System model and implementation provides.

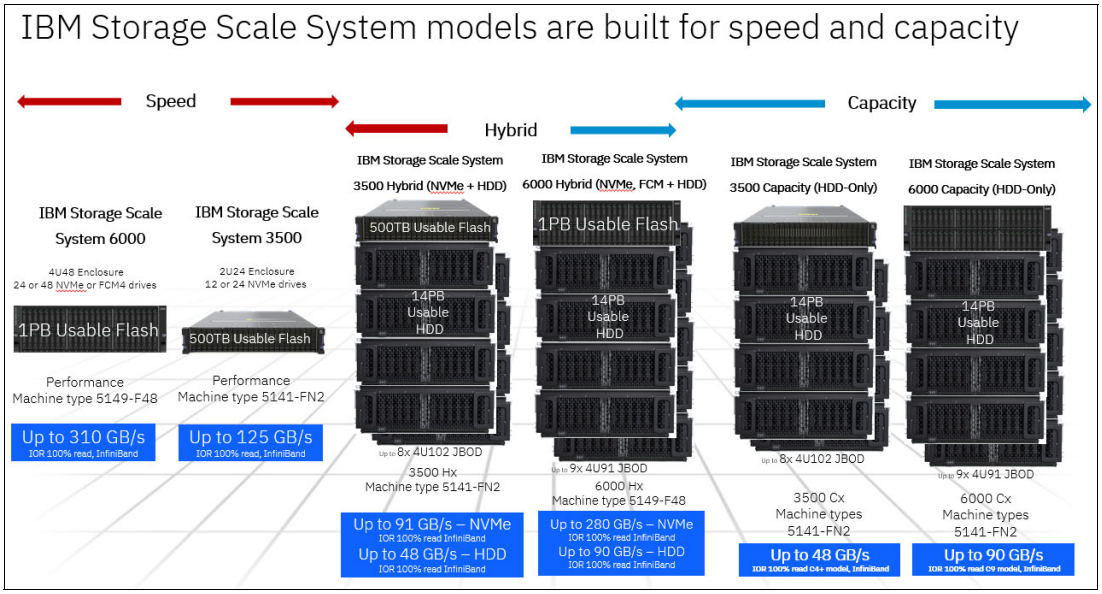

Figure 2-7 IBM Storage Scale System configuration options

For more information about the various IBM Storage Scale System models specifications, see Appendix A, “IBM Storage Scale System models” on page 59.

2.2.5 IBM Storage Scale System Utility Node Platform - Management Server/Protocol Node

Released as part of IBM Storage Scale 6.1.8.1 in August 2023 the IBM Storage Scale System utility node for the 3500 model (MTM 5149-23E) is able to be configured for use as a management server, and/or protocol node. IBM Storage Scale 6.1.9.1 extends this functionality to the 6000 model as well. Figure 2-8 and Figure 2-9 on page 29 show the front and back of the Utility Node.

Figure 2-8 IBM Storage Scale Utility node front

Figure 2-9 IBM Storage Scale Utility node front

A dual processor version was released in December 2023 with increased memory as a protocol node and first deployable in IBM Storage Scale System 6.1.9.

At release there is a restriction when using IBM Storage Scale 6.1.8.x, the utility node cannot be ordered in combination with any POWER® node (including P9 EMS or P9 Protocol node). The utility node supports the Storage Scale Server 6000 and Storage Scale Server 3500.

2.2.6 Storage enclosures

A storage enclosure is a specialized storage drawer which holds and powers flash or HDD storage, while providing the connections to allow communication to one or more separate firmware service modules that run the entire enclosure.

Note: Software updates for Generation 2 system types continues through IBM Storage Scale System v6.1.9.x.

At the time this publication was created, no end of support date for IBM Storage Scale System v6.1.x has been announced.

IBM Storage Scale System v6.2.x supports the 6000 and 3500 models. The v6.2.x release also support the systems 3000, 3200 and 5000 however models have been withdrawn from marketing and are outside the direct scope of this document.

The Storage Scale 3500 external IBM 5147-102 Storage Enclosure is shown in Figure 2-10.

Figure 2-10 IBM 5147-102 and 5149-091 Storage enclosures

2.2.7 NVIDIA DGX SuperPod

IBM Storage Scale System is an approved storage solution for NVIDIA DGX SuperPOD

The NVIDIA DGX SuperPOD with NVIDIA DGX H200 , or NVIDIA DGX B200 systems is an artificial intelligence (AI) supercomputing infrastructure, which provides the computational power necessary to train today's state-of-the-art deep learning (DL) models and to fuel future innovation. The DGX SuperPOD delivers groundbreaking performance, deploys as a fully integrated system, and is designed to solve the world's most challenging computational problems.

This DGX SuperPOD storage reference architecture (RA) is the result of collaboration between DL scientists, application performance engineers, and system architects to build a system capable of supporting the widest range of DL workloads. The performance delivered by the DGX SuperPOD with DGX systems enables the rapid training of DL models at great scale. The integrated approach of provisioning, management, compute, networking, and fast storage enables a diverse, multi-tenant system that can span data analytics, model development, and AI inference.

IBM Storage Scale is a parallel file system, and its optimal performance is achieved when multiple processes or workloads are running in parallel. To maximise the results of a high-performance computing (HPC) and AI cluster solution with IBM Storage Scale 6000 and NVIDIA DGX SuperPOD architecture refer to IBM Storage Scale System 6000 with NVIDIA DGX SuperPOD Deployment Guide and NVIDIA DGX SuperPOD with IBM Storage Scale and IBM Storage Scale System publications

2.2.8 Networking switches

The IBM Storage Scale System solution requires IP or InfiniBand networking switches to interconnect the IBM Storage Scale System hardware components.

As a best practice, purchase and use NVIDIA networking switches. These network switches are tested by IBM to work with IBM Storage Scale System.

Clients also can provide and support their own Ethernet or InfiniBand networking switches.

Note: Specific IBM Storage Scale System models support NICs that provide Virtual Protocol Interconnect (VPI). VPI allows individual ports of the dual-port network card to be configured as ethernet or InfiniBand. If you use the VPI feature to reconfigure your NIC, configure or add suitable NICs to your IBM Storage Scale System Management Server so that the device can attach to that network.

Types of networks that are used by IBM Storage Scale System

The IBM Storage Scale System features the following types of networks that are used to interconnect the IBM Storage Scale System components:

- Management and service networks, which must have separate 1/10 GbE networks for management and service traffic
- Data network, which is a high-speed network for the data that is written into and read out of the IBM Storage Scale IBM Storage Scale System file systems

Regardless of the source and type of network switches between the IBM Storage Scale System, the IBM Storage Scale cluster and the users, the networking must be solid, robust, reliable, and able to provide consistent low-latency responses. Non-blocking, high-speed switches are highly recommended.

Experience shows that IBM Storage Scale and IBM Storage Scale System can easily generate throughput rates that can overwhelm an over-subscribed network. IBM Storage Scale provides a network load generation testing tool (**`nsdperf`**), which is available to anyone as an open source repository on GitHub. This tool should be used to test network performance, network latency, network topology, and network readiness level before IBM Storage Scale System is installed.

For more information about the **`nsdperf`** tool, see this GitHub repository.

Management network switches

IBM Storage Scale System components require a low speed Ethernet connection for scale management and separate baseboard management controller (BMC) traffic. As part of an IBM Storage Scale System order, IBM can supply the low speed management network switches for IBM MTM 5149-S54.

As a best practice, order and use this 1/10 GbE network management switch because it is integrated into the IBM Storage Scale System solution at the manufacturing stage. This integration allows the entire IBM Storage Scale System to be installed and initially tested without external network dependencies.

Alternatively, the client can provide a 1/10 GbE network switch management infrastructure. In this case, the client is responsible for providing the suitable management IP networking port counts and network switch configuration and definitions, including VLANs.

High-speed data network switches

IBM Storage Scale System components also require a high-speed data network IP or InfiniBand connection for the reading and writing of user data. As part of IBM Storage Scale System testing, IBM tests with NVIDIA data network switches.

As a best practice, order and use the qualified data network switches where possible. These switches are tested and integrated with the IBM Storage Scale System solution.

NVIDIA high-speed network switch drivers and firmware are tested and delivered integrated within the IBM Storage Scale System software solution stack.

Alternatively, the client can provide their own high-speed data network switch infrastructure. In this case, the client is responsible for provisioning networking port counts, firmware, network configuration, and definitions. The client also is responsible for maintaining and troubleshooting any networking issues.

Note: Consider the following points:

- IBM Storage Scale System 6000 use a PCIe5, which fully supports both ports of the 200 GbE or InfiniBand NICs or single port (IB only) x16.
- IBM Storage Scale System 3500 uses a PCIe Generation 4 bus, which fully supports both ports of the 200 GbE or InfiniBand NICs to run at full rated speed.

  Experience shows that throughput on IBM Storage Scale System is often limited to the amount of network NIC bandwidth that is available. The best performance can be achieved over InfiniBand.
- Suitable network cables and connectors must be specified and ordered for your specific network requirements. If you use the VPI feature to reconfigure your high-speed network ports, make sure that your IBM Storage Scale System Management Server also can access the newly configured network.
- RDMA over Converged Ethernet (RoCE) might be available on third-generation IBM Storage Scale System models for specific configurations, networks, and environments only. If you have this requirement, submit a Request for Price Quotation (RPQ) to IBM by contacting your IBM or IBM Business Partner representative.
- For more information about estimating and sizing your performance, or for specifying the suitable network cables and connectors, contact your IBM or IBM Business Partner representative.

2.2.9 Racking

The IBM Storage Scale System solution can be ordered with or without the building block being integrated into an IBM rack.

An IBM Storage Scale System solution that is ordered with feature codes that specify integration into the rack in IBM manufacturing result in a fully integrated, tested, and pre-cabled solution.

The IBM 5149-R42 is the default for most models of IBM Storage Scale Systems. The rack provides 42 EIA (42U) of usable space for IBM Storage Scale System components and network switches. A choice of various electrical PDUs is available to be specified for inclusion in the rack side areas.

It is advantageous to order IBM Storage Scale System as a racked system because IBM Storage Scale System Manufacturing delivers an IBM Storage Scale System with the rack

mostly pre-cabled and pre-tested. This configuration provides faster time to install with higher quality.

If installing the IBM Storage Scale System in a customer-supplied rack, it is the customer's responsibility to correctly specify all physical planning for electrical power, connectors, cables, and racking locations.

Clients also must plan the physical installation and cabling of the IBM Storage Scale System components in the client-supplied rack. An IBM TSS Services contract is available and recommended to perform physical planning, specification, and physical installation of the IBM Storage Scale System components in a client rack on behalf of the client.

For the rackless IBM Storage Scale System order, the solution must be assembled and integrated into the rack at the client site.

For the purposes of cable management and serviceability, each IBM Storage Scale System building block is integrated and deployed within the same rack.

If you require to split the IBM Storage Scale System components between adjacent racks, ask your IBM or IBM Business Partner representative to submit an RPQ special bid request. IBM examines the request and responds regarding whether IBM can grant and support the specific requested physical racking.

2.3 Software components

This section describes the software components that are used on the IBM Storage Scale System solution.

2.3.1 Software solution stack overview

The IBM Storage Scale System solution provides an integrated and tested stack of these operating system, adapter drivers, firmware, IBM Storage Scale software, management software, and installation scripts into a full IBM Storage Scale System software stack. This software solution stack is supported as an integrated IBM Storage Scale storage building block solution by IBM Service and Support.

In the following sections, we briefly describe the following IBM Storage Scale System solution stack software components:

- RHEL operating system
- IBM Storage Scale high-performance parallel file system
- IBM Storage Scale RAID software-defined storage (SDS)
- IBM Storage Scale System solution installation and management software, which includes (but not limited to) the following components:
  - IBM Storage Scale System-specific documentation for installation and upgrade scripts
  - Other tools for the IBM System Services Representative (IBM SSR) to use while installing IBM Storage Scale System, such as **`essinstallcheck`**, **`essnetverify`**, and **`essutils`**.
- Third-generation IBM Storage Scale System systems deploy a container-oriented management software stack in the IBM Storage Scale System Management Server that includes Red Hat Ansible playbooks for enhanced ease of use during installation and

orchestration. Installation is faster, and skill requirements are significantly lower than previous-generation IBM Storage Scale System models.

The IBM Storage Scale System solution stack levels are released as a version, release, and modification level.

The IBM Storage Scale System solution components are periodically upgraded to newer release levels, tested as an integrated solution, and released as a new level of the IBM Storage Scale System solution software. Additionally, generalized recommendations are made available to help clients implement a code update strategy. It is a full-field perspective, and as such, a customized recommendation that accounts for specifics such as business upgrade windows, length of time since last update, and decommission plans. You might require assistance from local support teams. For more information, see IBM Storage Scale Software Version Recommendation Preventive Service Planning.

For more information about the release levels of the IBM Storage Scale System software solution and the levels of the software components for that IBM Storage Scale System release level, see question 2.2 in the
IBM Storage Scale RAID FAQ for IBM Storage Scale System in IBM Documentation.

2.3.2 Operating system

The IBM Storage Scale System solution runs RHEL as the operating system on the IBM Storage Scale Data Servers.

Each IBM Storage Scale System solution release level integrates and tests a suitable current level of RHEL, including any necessary RHEL fixes and errata that are required for the successful operation of the IBM Storage Scale System solution stack. IBM periodically provides new IBM Storage Scale System solution release levels that incorporate newer levels of RHEL. These releases are provided often enough to assure that a current level of RHEL is always available.

Embedded RHEL licensing on IBM Storage Scale Systems

IBM uses an embedded RHEL license. This license is included with every order.

2.3.3 IBM Storage Scale

In this section, we describe IBM Storage Scale, the high-performance parallel clustered file system software that is used in IBM Storage Scale System. IBM Storage Scale software is released several months before the equivalent IBM Storage Scale System release. Intentionally, the release number of the Storage Scale software differ in the first digit from the Storage Scale System software. At the time of publication the current software version is 5.2.3, The current Storage Scale System version is 6.2.3.

IBM Storage Scale is highly distributed, clustered file system software that provides high-speed concurrent data access to applications that run on multiple nodes and clusters. In addition to providing parallel high-performance file storage capabilities at exabyte scale, IBM Storage Scale provides tools for tiering, management, administration, and archiving of that enterprise-level data. IBM Storage Scale is the IBM strategic SDS for enterprise big data, analytics, and AI applications.

In the most common IBM Storage Scale deployment architecture, IBM Storage Scale data is accessed by IBM Storage Scale clients and users over a LAN network, accessing disk volumes that are known as Network Shared Disk *(NSDs)* that are attached to IBM Storage

Scale nodes that are known as *NSD Data Servers*. In this IBM Redpaper publication, these nodes also are referred to as *Data Servers*.

2.3.4 IBM Storage Scale RAID

IBM Storage Scale RAID is an SDS controller that performs all of the storage controller functions that are normally associated with hardware storage controllers. IBM Storage Scale RAID integrates all HA and features of an advanced storage server into IBM Storage Scale SDS.

IBM Storage Scale RAID runs on the IBM Storage Scale System NSD Data Servers. IBM Storage Scale RAID provides sophisticated data placement and error correction algorithms to deliver high levels of storage reliability, availability, and serviceability, and performance.

IBM Storage Scale RAID implements a declustered erasure code parity schema, distributing data, redundancy information, and spare space across all disks of the IBM Storage Scale System enclosures. With this approach, a significant improvement is realized on the application performance. Also, storage rebuild time overhead is reduced (disk failure recovery process) compared to conventional RAID controllers.

IBM Storage Scale RAID implements large cache for performance by using memory on the IBM Storage Scale System NSD Data Servers. The large cache intelligently improves read and write performance, particularly for small block I/O operations.

IBM Storage Scale RAID mitigates performance effects of storage rebuilds

If storage failures occur, IBM Storage Scale RAID reconstructs lost or erased stripes for I/O operations dynamically. By using the highly distributed erasure coding, IBM Storage Scale RAID mitigates the performance effect of storage media failures.

IBM Storage Scale RAID end-to-end checksums

IBM Storage Scale RAID includes integrated end-to-end checksums that detect data corruption that might otherwise go undetected by a conventional storage controller. Unlike conventional storage controllers, IBM Storage Scale RAID is integrated with the IBM Storage Scale file system and performs end-to-end checksum comparison all the way out to the IBM Storage Scale client code on the workstations. This feature ensures data integrity at a file system level, detecting and automatically correcting data corruption errors that might occur in conventional storage environments.

In an environment where a customer experienced excessive file system checks and suffered downtime to repair file systems, the use of IBM Storage Scale RAID end-to-end checksums mitigates file system check problems. This feature assures availability of data and removes application outages that are caused by file system checks.

IBM Storage Scale RAID disk hospital

One of the key features of IBM Storage Scale RAID is the disk hospital. This powerful function asynchronously diagnoses errors and faults in the IBM Storage Scale System storage media, down to the level of the individual drive and the individual performance of each drive. IBM Storage Scale RAID is fully aware of and tracks the performance of each individual drives because all drives do not perform equally. IBM Storage Scale RAID uses the individual performance history of each drive to make intelligent data allocation and data retrieval decisions.

Extensive health metrics down to the level of the individual drive are maintained by the disk hospital. Performance variation is continually monitored. If or when a disk metric exceeds a threshold, the storage media is marked for replacement according to the disk maintenance replacement policy for the declustered array.

As an example, disk hospital features the following metrics:

- relativePerformance, which characterizes response times. Values are compared to the average speed. If the metric falls below a specific threshold, the hospital adds “slow” to the pdisk state, and the disk is prepared for replacement.
- dataBadness, which characterizes media errors (hard errors) and checksum errors.

For more information about IBM Storage Scale RAID implementation and best practices, see the Administering IBM Storage Scale RAID manual for your particular IBM Storage Scale System model. at IBM Documentation.

2.3.5 IBM Storage Scale System solution installation and management scripts

In this section, we provide an overview of the components of the IBM Storage Scale System solution installation and management scripts. This overview includes the following information:

- IBM Storage Scale System-specific documentation for installation and upgrade scripts
- Ansible playbooks
- IBM Storage Scale System specific tools and utilities for the IBM SSR and administrators to use while installing or maintaining IBM Storage Scale System

For more information about these IBM Storage Scale System solution components, see the IBM Storage Scale System solution release-specific level information at IBM Documentation.

Installation, upgrade, and administering guides

IBM provides manuals and documentation for deploying and administering IBM Storage Scale System, including the following publications:

- Quick Deployment Guide (QDG), which documents IBM Storage Scale System-specific scripts for installing, deploying, and upgrading IBM Storage Scale System for experienced users.
- IBM Storage Scale RAID Administration, which focuses on administering IBM Storage Scale RAID on IBM Storage Scale System.
- Problem determination guide, which provides more information about monitoring, troubleshooting, and maintenance procedures.

The IBM Storage Scale System product documentation is available at IBM Documentation.

Ansible playbooks and container-based software installation

Third-generation IBM Storage Scale System models include a software update process and methodology that are based on Ansible playbooks.

On this IBM Storage Scale System Management Server node, Ansible is configured to store the container and cluster configuration and definitions for all the managed IBM Storage Scale System Data Servers. Network services (such as Dynamic Host Configuration Protocol (DHCP) and SSH) are enabled to allow IBM Storage Scale System Management Server to deploy IBM Storage Scale System solution stack components to all IBM Storage Scale System nodes in the cluster.

essutils are IBM Storage Scale System installation and deployment toolkits that facilitate IBM SSR hardware setup, installation, deployment, and upgrade tasks. Any authorized IBM Storage Scale System administrator also can use these tools.

For more information about **essutils** and **gssutils**, see the Quick Deployment Guide, IBM Storage Scale System Hardware Planning and Installation Guide for the system of your choice; and IBM Storage Scale System software deployment preparation, which are available at IBM Storage Scale System documentation. PDF versions are also available, e.g. IBM Storage Scale System 6000 PDFs.

3

IBM Storage Scale System planning and integration

This chapter provides guidelines and considerations for the correct planning, installation, and configuration of the IBM Storage Scale System.

This chapter also describes configurations and integration considerations for a smooth IBM Storage Scale System deployment into an existing or a new IT environment. In this chapter, we describe the following elements:

- Roles and responsibilities
- Planning
- Networking
- Installation
- Maintaining hardware
- Maintaining software
- Upgrades

This chapter includes the following topics:

3.1 IBM Storage Scale System planning overview

In this section, we make recommendations and highlight best practices for various roles and responsibilities when an IBM Storage Scale System is deployed.

3.1.1 Roles and responsibilities

In this section, we discuss best practices were learned from successful IBM Storage Scale System installations regarding who is responsible for each of the areas of expertise.

Overall planning

An experienced project leader should be appointed to coordinate and manage the IBM Storage Scale System installation project. Deploying an IBM Storage Scale System requires coordination across different IT departments and involves various team members (IBM and client).

The various client and IBM teams that the project manager must coordinate with include (but are not limited to):

- Physical planning for the data center, including the networking team
- Networking setup, switch installation, cabling, and verification
- Hardware setup, installation, and verification
- Software planning and integration of the file system into the environment
- Acceptance testing and criteria for moving into production

Networking

IBM Storage Scale System components are connected by using IP or InfiniBand networking. IBM Storage Scale System data users access IBM Storage Scale System by way of high-speed IP or InfiniBand networking.

The networking planning tasks that are required include (but are not limited to) include the following examples:

- Suitable networking switches must be ordered or allocated, including cable planning, ordering, racking, and proper connectors.
- Network addressing must be planned for IBM Storage Scale System and the integration of that network into existing client IP or InfiniBand network.
- After the network is designed and approved, IP addresses must be implemented and tested.
- Multiple VLANs must be designed, configured, and tested on the management network switches.
- Suitable networking authentication and firewall rules must be configured and tested to assure that they comply with client network audit and security requirements.
- Testing of the network before IBM Storage Scale System is installed (by using tools, such as **`nsdperf`** or Network Readiness Tools that are provided by IBM Storage Scale as no-charge, open-source tools) must be done to identify and verify the network's ability to provide
  IBM Storage Scale System with the required network bandwidth, latency, and consistent performance.

These tasks require the client IBM Storage Scale System infrastructure implementation team to complete the following tasks:

- Identify all IP or InfiniBand networking-related requirements.
- Engage early with the infrastructure networking team.
- Document and communicate all required IP network provisioning, implementation, and network implementations, including all network stress testing and acceptance criteria.
- Monitor these tasks closely to assure that the networking is in place and load-tested before IBM Storage Scale System is installed. Experience shows that networking addressing mismatches, VLAN issues, configuration issues, and a failure to remedy these issues are the major cause of unforeseen delays in IBM Storage Scale System installation progress.

The client IP or InfiniBand networking team is responsible for the following tasks:

- Receive the specified networking requirements.
- Consult and collaborate with the IBM Storage Scale System implementation team to design, optimize, and agree upon a suitable IP or InfiniBand networking design and implementation plan.
- Implement the agreed-upon networking changes.
- Run requested network load tests by using tools that are provided by IBM Storage Scale before IBM Storage Scale System is installed, including agreeing to and implementing any necessary quality assurance and acceptance criteria.
- Adjust and tune the network based on pre-installation tests or during the IBM Storage Scale System implementation.

The following best practices from successful IBM Storage Scale System installations assure success:

- Engage early with the client networking team.
- Obtain collaborative buy-in and well-defined ownership of networking-related task completions.
- Operatively running network performance tests by using IBM Storage Scale tools far enough in advance of the IBM Storage Scale System installation allows time for necessary tuning adjustments to be made.
- Ongoing IBM Storage Scale System implementation team and client networking team dialog and interaction during the IBM Storage Scale System installation process quickly remedies any networking-related installation issues.

Installation

In this phase, the following IBM and client teams must be involved and coordinated:

- Physical site administration and engineering
- Networking team and network engineering
- Hardware implementation team, including client IT administrators, IBM System Services Representative (IBM SSR) for IBM Storage Scale System physical installation and setup of physical hardware
- Software implementation team, potentially including IBM System Lab Services consultants, and client operating system and software administrators

These teams collaborate to implement the following phases:

- Physical site planning

  This planning includes the receipt and staging of IBM Storage Scale System components after shipment, physical floor space planning and racking; electrical power planning and provisioning, including specifying suitable electrical power, cables, connectors, and power distribution units (PDUs); and assuring proper airflow and air conditioning.

- Software planning and integration

  This planning includes defining the IBM Storage Scale file system implementation parameters, requirements, integration testing, and acceptance with the client applications and procedures.

- Hardware physical setup, installation, and verification

  This planning and scheduling includes the IBM Storage Scale System hardware setup with the IBM SSR and the IT physical installation teams.

- Acceptance testing

  Running well-defined criteria for acceptance testing and moving the IBM Storage Scale System system into production are featured in this phase.

Maintaining hardware

After the IBM Storage Scale System is installed and running in production, many parties are responsible for various IBM Storage Scale System hardware tasks, including the following examples:

- Client IBM Storage Scale System administrators:
  - Monitoring IBM Storage Scale System physical hardware by using provided tools and the GUI
  - Basic hardware problem determination
  - Optional basic maintenance, such as replacing a failed solid-state drive (SSD) or hard disk drive (HDD)
- IBM SSR: Responsible for physical IBM Storage Scale System hardware maintenance and repair.

Maintaining software

After the IBM Storage Scale System is installed and running in production, many parties are responsible for various IBM Storage Scale System software tasks, including the following examples:

- Client IBM Storage Scale administrators:
  - Monitoring and tuning IBM Storage Scale performance
  - Data management, which includes monitoring data and storage usage patterns, and writing and running policies that manage data, backups, and snapshots

Learning Services and Classes

IBM Learning Services or IBM Business Partners can provide education courses about the customer topics for maintaining IBM Storage Scale and IBM Storage Scale System hardware. More worldwide resources might be available.

Classes that are available for IBM Storage Scale and IBM Storage Scale System include (but are not limited to) the following offerings:

- IBM Storage Scale Basic Administration
- IBM Storage Scale Advanced Administration
- Monitoring the IBM Storage Scale System

For more information, contact your IBM representative or IBM Business Partner.

Upgrades

The IBM Storage Scale System provides an integrated software solution stack, which is composed of specific software release levels and the following fix levels:

- RHEL kernel levels and components
- Firmware (System/Chassis/Drive)
- Network interface card (NIC) drivers
- IBM Storage Scale and IBM Storage Scale RAID software

Note: It is a client responsibility to plan and upgrade their IBM Storage Scale System solution software.

IBM recommends upgrading the IBM Storage Scale System software solution stack at least once a year and to keep the IBM Storage Scale System current on levels of RHEL, firmware, and IBM Storage Scale.

IBM System Lab Services is available and it is recommended to engage Lab Services to assist and perform these software upgrade actions on behalf of the client, if required.

Technical and Delivery Assessment

An IBM Technical and Delivery Assessment (TDA) is a technical inspection of a completed solution design. Technical subject matter experts (SMEs) who were not involved in the solution participate to ensure the client's satisfaction.

In a pre-sales TDA, which is performed before the IBM Storage Scale System solution is ordered, the IBM IMPACT tool and process is used to examine and approve the IBM Storage Scale System solution design to determine the following information:

- Will the IBM Storage Scale System solution work as expected?
- Is the configuration and implementation plan sound?
- Will the IBM Storage Scale System solution meet client requirements and expectations?

In a pre-installation TDA, SMEs also evaluate the client's environment for readiness to install, implement, and support the proposed solution.

IBM and IBM Business Partner teams are responsible to conduct these pre-sale and pre-installation processes. Contact your IBM representative or IBM Business Partner to schedule and perform these TDA and StorM processes.

For more information about the IMPACT and pre-installation TDA tools, which help your IBM account team, client technical specialist, and IBM Business Partner select the best IBM Storage Scale System and IBM Storage Scale solutions to match your business requirements and needs, see the following web pages:

- Preinstall TDA
- TDA and StorM
- IMPACT Tool

These resources are available to IBMers or Business Partners after logging in by using their IBMid or IBM PartnerWorld® ID.

3.2 IBM Storage Scale System hardware planning

In this section, we describe the necessary planning and implementation steps that must be followed as part of installing the IBM Storage Scale System hardware.

3.2.1 Planning for network switches and networking

Suitable network planning and configuration are one of the most important success factors for an IBM Storage Scale System installation. This planning is crucial because the IBM Storage Scale and IBM Storage Scale System cluster components are connected by way of an extensive set of IP addresses and associated networking switches and cabling.

The following networks are required to install an IBM Storage Scale System:

- Service network

  This network connects the Baseboard Management Controller (BMC)/Flexible Service Processor (FSP) on the management server and I/O server nodes, depending on the IBM Storage Scale System generation).

- Management and provisioning network

  This network connects the management server to the I/O data server nodes. The management server runs the Dynamic Host Configuration Protocol (DHCP) on the management and provisioning network. If an IBM Storage Scale System Management Server is not included in the solution order, it is assumed that an IBM Storage Scale System Management Server exists in the same IBM Storage Scale cluster that can be used.

- Clustering network

  This high-speed data network is used for IBM Storage Scale System data transfer to and from the rest of the IBM Storage Scale cluster and client node access.

  Note: This network must provide consistent, congestion-free network packet delivery with consistent low latency. IBM can supply tools before the installation to test the network.

  For more information about for estimating and sizing suitable network performance, contact IBM or your IBM Business Partner.

- External and campus management network

  This public network is used for external and campus management of the IBM Storage Scale System Management Server, and optionally to other nodes in the scale cluster.

Figure 3-1 shows an example network topology for IBM Storage Scale System

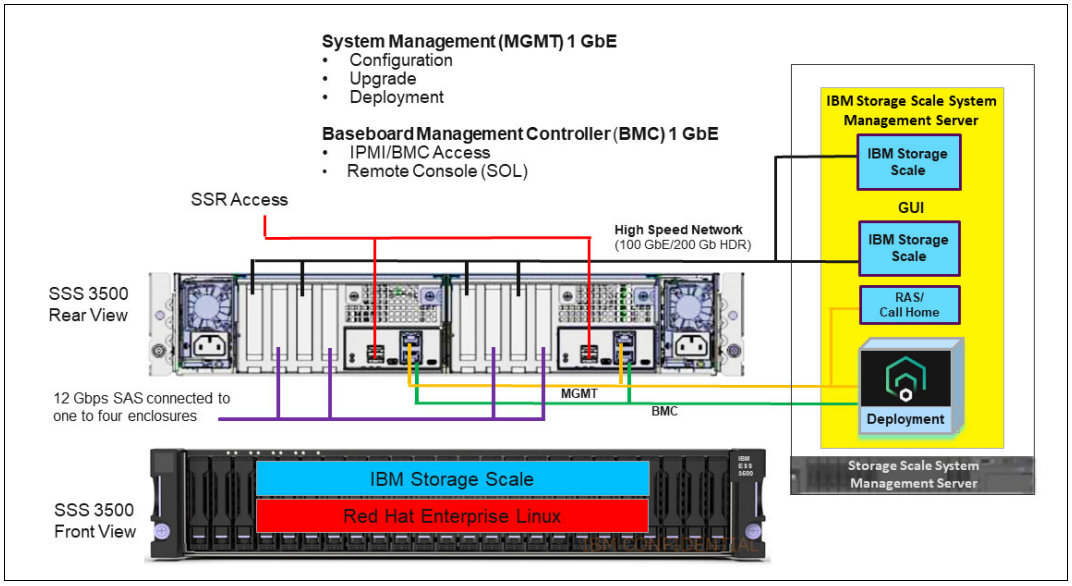

Figure 3-1 IBM Storage Scale System network topology example

An overview of the many networking steps that are required in the installation of an IBM Storage Scale System are described next. This example is representative (not all inclusive) of the necessary network planning and implementation steps that must be completed before installing the IBM Storage Scale System hardware. For more information about network planning, see the IBM Storage Scale System Quick Deployment Guide that is available at IBM Documentation.

Example: Network IP addressing, planning, and implementation for IBM Storage Scale System

In this section, we present an example of typical networking-related items for installing an IBM Storage Scale System. The following template can be used for the installation process:

1. Confirm Private IP range for DHCP server.
2. Confirm Private Service network with defined IP addresses and private management network with defined IP addresses. Separate the networks by using switches or VLAN.
3. Confirm Public network connections for IBM Storage Scale System Management Server. Multiple IP addresses are needed.
4. Select **Host** → **IP mappings** to confirm the mappings for the following items:
   - IBM Storage Scale System Management Server
   - IBM Storage Scale System Data Servers: I/O server 1, I/O server 2, I/O server 3, I/O server 4, and so on
   - 200 Gb Ethernet (GbE), 100 GbE, 40 GbE, 25 GbE, 10 GbE, or InfiniBand switches
5. Set up domain names for management private network.
6. Set up domain names for high-speed interconnect.
7. Set up Data I/O partition and partition profile names.
8. Confirm Server names.
9. Confirm that the ethernet or InfiniBand switches are in place and cabled.
10. Set up Bonding if it is used.

11. Set up Public network in place and cabled to IBM Storage Scale System Management Server.
12. Confirm that all building block components are correctly cabled in the rack (Data I/O servers, IBM Storage Scale System Management Server, and switches).
13. Set up and confirm dual feed power to racked IBM Storage Scale System components.
14. Set up the IBM Storage Scale System Management Server console or terminal, including the network connections.
15. Prepare for and install RHEL errata.
16. Define and specify IBM Storage Scale installation parameters and requirements. For example, how many file systems, what are the block sizes, how do we allocate and split metadata, and what replication factor is used.
17. Confirm all physical NVMe, SSD, or HDD disks are in place. IBM Storage Scale System installation checks with scripts.
18. Confirm that all physical cabling in place. IBM Storage Scale System installations are checked by scripts.
19. Confirm that Wi-Fi access exists in lab to set up Webex (for access to and by support teams).

This template shows the processes that IBM Systems Lab Services can perform as part of a contract to install an IBM Storage Scale System. As part of that contract, IBM Systems Lab Services provides and works with the customer on a detailed IP networking planning document.

3.2.2 Planning for servers that are used in IBM Storage Scale System

In this section, we review planning information for the physical servers that are used in an IBM Storage Scale System. We also describe specific recommendations for memory, NICs, and cabling and connectors.

IBM Storage Scale System Management Server

Each IBM Storage Scale cluster that uses an IBM Storage Scale System requires at least one IBM Storage Scale System Management Server. The IBM Storage Scale System Management Server acts as the central management focal point for all IBM Storage Scale System systems that are in the cluster; runs specific functions, such as the IBM Storage Scale System GUI; and stores and distributes the master boot image for all the IBM Storage Scale System systems that are in the cluster by using Red Hat Ansible playbooks.

IBM Storage Scale System Data Servers

Every IBM Storage Scale System is an integrated IBM Storage Scale storage building block that consists of two IBM Storage Scale Data Servers, and each data server is cross-connected to all the storage for high availability (HA). In normal operation, half of the IBM Storage Scale System storage is running through the first data server, and the other half of the IBM Storage Scale System storage is running through the second data server. If a planned or unplanned outage occurs on one of the data servers, the storage is failed over to

the remaining data server in the IBM Storage Scale System, which provides HA and resiliency.

Note: Both data servers must have the same amount of memory in each IBM Storage Scale System building block. When possible, as a best practice, all IBM Storage Scale System systems in a specific IBM Storage Scale storage pool should have the same size and speed. If this configuration is not possible, keep the IBM Storage Scale System similar in size and speed. IBM Storage Scale parallelism means that in a specific storage pool, the speed of data reads/writes is gated by the slowest storage building block.

Network interface cards

The IBM Storage Scale System features the following types of networks that are used to interconnect the IBM Storage Scale System components:

- Management and service

  Mandatory, separate 1/10 GbE network for management and service traffic.

- Data

  A high-speed network for the user data being written into and read out of the IBM Storage Scale and IBM Storage Scale System file system.

Standard NVIDIA server NICs are used to provide high-speed network connectivity in an IBM Storage Scale System. The network cards that are available vary depending on the IBM Storage Scale System generation.

It is important to ensure that the network bandwidth requirements are clearly studied, documented, and communicated. Specify a suitable number of NICs to meet expected IBM Storage Scale System performance requirements. Plan and specify the correct cables and connector types.

Note: IBM Power IBM Storage Scale System Data Servers do not support every possible IBM Power PCI adapter.

Cabling and connectors

In addition to specifying the correct number of NICs to provide bandwidth requirements, make sure that physical planning for cable lengths, adapter connector types, and switch connector types, are clearly studied, documented, and communicated.

Physical machine room placement of IBM Storage Scale System components relative to the physical location of the IP network switches must be clearly understood and documented. Physical cable orders, including wanted colors, connector types, and required cable routing must be studied, specified, and verified with network teams, server teams, physical site teams, and infrastructure teams.

Many network connector types cable types are available. Requirements must be identified for the required networking components, including the following examples:

- Transceivers
- Optical versus copper
- Connector form factor
- Cable lengths

3.2.3 Planning for storage configuration

IBM Storage Scale System storage configuration for ordering purposes is relatively straightforward.

First, determine your IBM Storage Scale storage requirements. Request that your IBM representative or IBM Business Partner use the StorM tool to help identify usable capacity requirements and to estimate wanted performance.

After storage capacity and performance requirements are known, identify which models of IBM Storage Scale System feature the suitable type and quantity of storage media available. Within each model, you choose the suitable drive size.

Each IBM Storage Scale System model includes predefined storage configurations and requirements, which are pre-tested and supported as a solution by IBM. Cable lengths, installation parameters, adapters, drive preparation, and location, all are pre-specified within the IBM Storage Scale System architecture.

3.2.4 Planning for IBM Storage Scale protocol nodes

First announced in 2014, IBM Storage Scale protocol nodes provide external Server Message Block (SMB), Network File System (NFS), object, and Hadoop access to IBM Storage Scale data.

IBM Storage Scale requires that all protocol nodes in an IBM Storage Scale clusters are alike in processor architecture, performance and memory.

For more information about protocol nodes, search *Quick Deployment Guide*, in IBM Documentation located here.

3.2.5 Planning for physical, electrical, and cooling installation

Preparation for installation of your IBM Storage Scale System is as important as the physical installation of the system. To have a smooth and successful installation, preparation and planning processes are needed.

StorM and pre-installation meetings with IBM TSS and IBM Service personnel must be completed. Also, IBM Storage Scale Server (IBM Storage Scale System) environmental specifications provide detailed information about your system, including dimensions, electrical, power, temperature, environmental requirements, and noise emissions. By following the preparation process, IBM service personnel can efficiently install the physical hardware and also deployment of the software.

The specifications for IBM Storage Scale System can be found in IBM Documentation.

3.3 IBM Storage Scale System software planning

In this section, we highlight key items to be aware of for planning the installation of the IBM Storage Scale System software stack; in particular, IBM Storage Scale and IBM Storage Scale RAID.

3.3.1 Planning for IBM Storage Scale file system

The IBM Storage Scale file systems that are placed into the IBM Storage Scale System must be thoroughly planned, documented, and communicated. Make sure that user application, authentication, firewall, and backup and restore requirements are understood and included.

Before the arrival and installation, the IBM Documentation manual for IBM Storage Scale and IBM Storage Scale RAID Administration should be consulted. These manuals contain detailed instructions and best practices recommendations for implementing the requirements of the client's environment.

3.3.2 Planning for IBM Storage Scale RAID

The IBM Storage Scale RAID Administration manual provides best practices and recommendations for setting up IBM Storage Scale RAID. Consult this manual and follow the recommendations.

The IBM Storage Scale System and IBM Storage Scale Server implementation of IBM Storage Scale RAID uses just a bunch of disks (JBOD) arrays, provides the required redundancy protection, and delivers usable disk capacity, required spare capacity, and maintenance strategy.

In IBM Storage Scale System, the IBM Storage Scale RAID implementation best practices are enforced as de facto standards by the IBM Storage Scale System installation scripts, which provide configuration parameters that can achieve the best storage performance.

Each JBOD array is connected to each of the two Network Shared Disk (NSD) Data Servers in the IBM Storage Scale System to protect against server failure. Each NSD Data Server has two independent paths to each physical disk to protect against path failure and provide higher throughput to the individual disks.

3.4 Ordering IBM Storage Scale System

IBM Storage Scale System is always part of a larger IBM Storage Scale system. Because IBM Storage Scale and IBM Storage Scale System are sophisticated technologies to assure high client satisfaction, IBM requires that a TDA and the StorM tools are used to elicit requirements, assure that designs are reviewed, assessed, and approved.

The IBM Storage Scale System solution process recommends a pre-sale and a pre-delivery TDA. The purpose of these two assessments is to match the client's requirements and needs. It also evaluates the client's current IT environment to create the best proposed IBM Storage Scale System solution that matches the client's needs.

Before confirming the IBM Storage Scale System order, the IBM team or the Business Partner uses the IBM StorM tool. StorM is a question and answer tool that elicits and documents the various technical requirements for a solution that uses IBM Storage Scale

System. Contact your IBM representative or Business Partner to request that a pre-sales IBM Storage Scale System TDA evaluation is performed.

IBM eConfig is used by the IBM or Business Partner team to configure the IBM Storage Scale System for ordering. eConfig is a tool that helps the IBM or Business Partner team to configure the correct combination of IBM Storage Enclosures, servers, NICs, and associated feature codes to make up the wanted IBM Storage Scale System.

Note: An approved pre-installation TDA *should* be performed by the IBM account team or IBM Business Partner to have an IBM Storage Scale System order to be entered and shipped.

As part of the order, various decisions regarding hardware must be made. Be prepared to supply requirement information to the IBM Business Partner. The following requirements often are included in the information that is sent:

- Storage requirements (total capacity, drive type, and performance)
- Server memory and NIC requirements
- Network cables, cable lengths, connector types, and network switches
- Physical electrical power requirements (phase and voltage)
- Any other more requirements (Your IBM or Business Partner representative works with you to determine these requirements.)

IBM or the Business Partner then configures the IBM Storage Scale System. IBM Storage Scale System components can be installed in an IBM rack (preferred) or installed in customer-supplied racks.

Figure 3-2 on page 51 shows on the left, the size of various IBM Storage Scale System components in rack units. Shown on the right of Figure 3-2 on page 51 are examples two IBM Storage Scale System single rack configurations. These configurations are examples only. Configurations are tailored to your requirements however, when planning the rack layout, it is expected that components are racked from the bottom up beginning with the heaviest and ending with the lightest components.

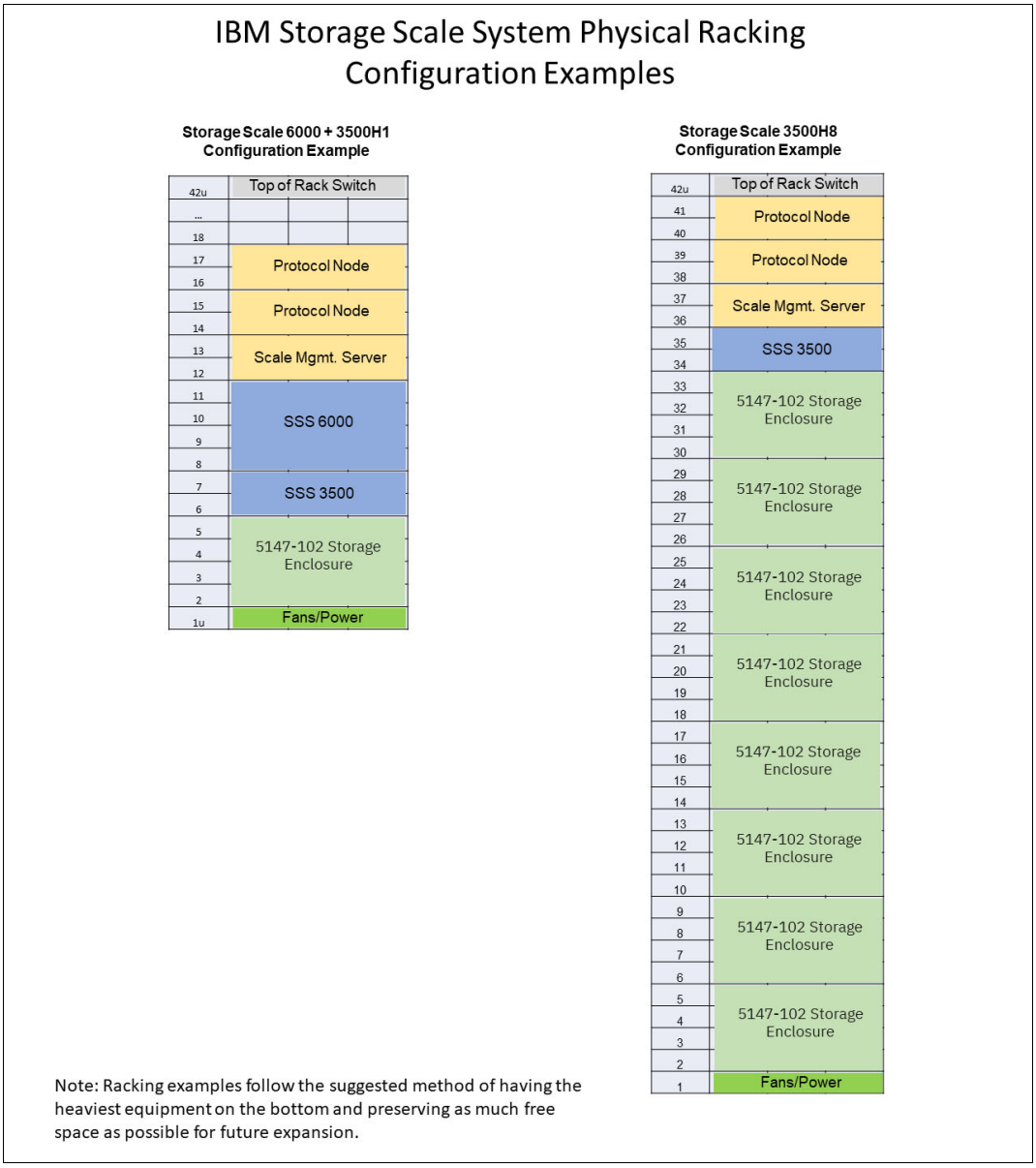

Figure 3-2 IBM Storage Scale Server physical racking examples

3.5 IBM Storage Scale System installation

In this section, we describe the process that is used to physically install the IBM Storage Scale System and deploy it into the client's IBM Storage Scale environment.

The installation process starts with the client preparing their environment to be ready for the IBM Storage Scale System. After the client's environment is prepared, the IBM expert service personnel install the hardware. Depending on the solution, the deployment of IBM Storage Scale System also is performed by IBM specialized service personnel.

3.5.1 Physical installation, racking, and cabling

IBM Technical Support Services (TSS) should be engaged to help you plan for the physical installation and racking of the IBM Storage Scale System components.

IBM TSS works with you to help you plan for the many items to consider for racking, installation, and servicing of the IBM Storage Scale System. Your TSS representative can help you plan for many factors, including the following examples:

- Electrical power, power connectors, PDU, and amperage
- Network adapter and connectors
- Electrical power cables, network cables, and cable lengths
- Shipping location of the system
- Lift tools
- Ladders
- IBM SSR physical installation instructions (WCII) and IBM SSR scheduling and confirmation

If you have an IBM Storage Scale System HDD model, allow time for the physical installation of the many HDDs into the storage enclosures. HDD IBM Storage Scale System models are shipped with the storage enclosures empty to mitigate shipping damage. Each HDD is shipped in a sealed protected package. In a large IBM Storage Scale System, many HDDs are used and it takes time for the IBM SSR to install each HDD into its slot in the storage enclosures.

When planning the physical installation, remember to consider future expansion needs. Some factors include the following examples:

- When purchasing new Storage Scale System hardware from IBM, is a new rack also needed?
- Is there sufficient space in the current or desired rack location which can be reserved to allow for expansion in the future?
- Is the location of the rack in the data center suitable for future cabling and electrical power supply expansion?
- Is the location of network switches and the IBM Storage Scale System Management Server in a location that has suitable cable lengths and access for current or future networking needs?

3.5.2 Integrating into client environment

Integrating the IBM Storage Scale System into the clients' environment consists of many actions. In this section, we provide an overview of key best practices, practical advice, and experiences and lessons learned for a successful IBM Storage Scale System deployment.

Integrating into customer networking environment

In any IBM Storage Scale or IBM Storage Scale System solution, the network interconnect always plays a central role. The goal of this section is to guide you in lessons that are learned from an IBM Storage Scale or IBM Storage Scale System network readiness discussion.

Networking is a vast expertise area, and you should not expect to need to become a networking expert. The following questions guide you through initial discussions you must have with your networking infrastructure team to assure that your IP network is ready for your

IBM Storage Scale or IBM Storage Scale System solution to integrate IBM Storage Scale System into your environment.

Will a new or existing network be used

Document if you are installing IBM Storage Scale or IBM Storage Scale System on an existing network or if this installation is on a new network that is designed specifically for IBM Storage Scale or IBM Storage Scale System.

IBM Storage Scale or IBM Storage Scale System is a high-performance, high-data rate big data solution. Best practices include providing the IBM Storage Scale or IBM Storage Scale System cluster with a dedicated network infrastructure that is co-designed with the solution.

However, especially in a commercial environment, many (if not most) circumstances are such that you cannot implement a dedicated network for IBM Storage Scale or IBM Storage Scale System. Instead, the IBM Storage Scale or IBM Storage Scale System traffic must coexist on an IP network. Especially in this case, consider the following points to validate if the network can provide the high bandwidth and consistent low latency IP networking that IBM Storage Scale or IBM Storage Scale System solution requires:

- IP Network vendors and switch models.

  Document and provide a list of the IP network switch vendors and switch models that are to be used to provide the network infrastructure for this IBM Storage Scale or IBM Storage Scale System cluster. IBM Storage Scale or IBM Storage Scale System depends upon IP networking for all of the IBM Storage Scale nodes and users to connect and communicate with each other.

  NVIDIA is the preferred high-performance IP networking switch vendor for IBM.

- Provide a network topology diagram.

  Provide a copy of the network topology and document where the IBM Storage Scale System will be placed and how they will be connected.

- Total number of nodes in IBM Storage Scale or IBM Storage Scale System cluster.

  Document the number and type of IBM Storage Scale and IBM Storage Scale System nodes that are planned for your solution. Provide the number of protocol nodes and the number of IBM Storage Scale NSD Data Servers.

Are there sufficient port counts and port speeds that are available for the network switches

Document the number of ports and port speeds that are required by your IBM Storage Scale System and assure that number of ports and speeds are available within the networking infrastructure.

What are the interconnects between switches

Document the inter-switch links, where they are located, and what traffic traverses them.

Can existing networking links and switches be used

IBM Storage Scale or IBM Storage Scale System can generate unusually high workload on a network. If IBM Storage Scale or IBM Storage Scale System is to be placed on an existing network, document how heavily used the network is, and what the anticipated usage will be after IBM Storage Scale System is added.

Best practice is to validate whether the client's network can sustain the intended IBM Storage Scale or IBM Storage Scale System workload, use the IBM network performance tool **nsdperf** (written by IBM Storage Scale Development). The **nsdperf** tool

mimics IBM Storage Scale workloads and provides an open-source, no-charge method to test the network before installation.

For more information, see this GitHub repository.

What other network traffic types and amounts must coexist on these network links

Document what other traffic is to compete for bandwidth on the IP network. IBM Storage Scale depends on low latency, consistent delivery of many small metadata messages to work properly. If any congestion exists in the IP network, IBM Storage Scale can encounter problems. Therefore, the best practice is to have a dedicated network infrastructure for the IBM Storage Scale or IBM Storage Scale System, if possible.

Communicate this need for low latency and consistent, congestion-free delivery of IBM Storage Scale metadata messages to the network infrastructure team.

What are the basic workload types

Document what other traffic IBM Storage Scale is to compete with for network bandwidth. Identify usage cases that are to be used, looking to check the workload mix (small block IOPs versus large block sequential, for example). Engage a qualified IBM Storage Scale architect with networking skills to project how the IBM Storage Scale or IBM Storage Scale System network infrastructure performs in your specific environment.

Sustained bandwidth required in GBps

Determine performance expectations. In addition to documenting GBps sequential streaming workload requirements, what other specific workloads are to be running? What tools or criteria are to be used to quantify that performance? What specific criteria determine “good” or “not so good” performance?

Is validation required to confirm that the existing network can provide necessary GBps and IOPS

If it is deemed necessary to examine the network further and validate that it is ready for an IBM Storage Scale or IBM Storage Scale System solution after reviewing these questions and answers, use the **`nsdperf`** tool. For more information about this tool, see this GitHub repository.

Integrating into a customer software environment

IBM Storage Scale and IBM Storage Scale System provide various file or object interfaces to the user applications. Determine, inventory, and document the user applications and their required interfaces that the applications require. From that information, design the file systems, directory structures, and data architecture that must be supported by the IBM Storage Scale or IBM Storage Scale System solution.

Design the IBM Storage Scale or IBM Storage Scale System cluster to deploy the wanted POSIX, Linux, or UNIX file systems. If necessary, design and deploy the IBM Storage Scale protocol nodes and their associated shares, mount points, and IP addresses that the software applications require.

Integrating into a customer security and authentication environment

IBM Storage Scale System is a part of an IBM Storage Scale cluster. IBM Storage Scale System and IBM Storage Scale support security and authentication by using various interfaces, including LDAP and Active Directory. Determine, document, and inventory the required security, authentication, and access control requirements. Then, use that information to design the IBM Storage Scale or IBM Storage Scale System solution.

For more information about IBM Storage Scale authentication and security, see *IBM Storage Scale Security*, REDP-5426.

3.5.3 Reliability, availability, and serviceability

IBM Storage Scale System uses servers and design that have a rich tradition of reliability, availability, and serviceability (RAS).

IBM Storage Scale System's reliability (and the availability of the applications it supports) is a function of much more than just the reliability of the processors, or even of the entire system hardware. A full description of a system design for RAS must include all of the hardware, firmware, operating system, applications, operating environment, duty cycle, and so on.

Reliability in hardware is all about how often a hardware fault requires a system to be serviced; the less frequent the failures, the greater the reliability.

Availability is how infrequently such a failure affects the operation of the system or application. For high levels of availability, correct system operation must not be adversely affected by hardware faults. That is, a highly available (HA) system design ensures that most hardware failures do not result in an application outage.

Serviceability is concerned with identifying what fails and ensuring an efficient repair (of that component, firmware, or software).

IBM Call Home: IBM Storage Scale System hardware and software

The IBM Storage Scale System Management Server and I/O server hardware includes IBM Call Home capability.

When a serviceable event occurs on one of the monitored components (such as servers or storage), the Call Home generates a Call Home event.

IBM Storage Scale software also has its own separate Call Home capability for purposes of automatically collecting and securely uploading data that is related to IBM Storage Scale to IBM Service and Support.

3.6 Monitoring IBM Storage Scale System

Monitoring the IBM Storage Scale System system includes system health, performance, and capacity monitoring. You can monitor the system through IBM Storage Scale System GUI, restful API or with the help of CLI.

For more information, search for the topic "Monitoring IBM Storage Scale RAID" at IBM Documentation.

3.6.1 GUI

The IBM Storage Scale System GUI provides robust facilities to monitor and manage the system.
IBM Documentation provides more information about how you can configure the IBM Storage Scale System GUI to monitor the performance of the following functional areas in the system:

- Network
- System resources

- IBM Storage Scale RAID
- NSD server
- IBM Storage Scale client
- NFS
- SMB
- Object
- Waiters
- Active File Management (AFM)

For more information, see *Monitoring and Managing the IBM Elastic Storage Server Using the GUI*, REDP-5471.

3.7 Maintaining and repairing IBM Storage Scale System

For IBM Storage Scale System features, a full library of documentation for maintaining the hardware and software of the IBM Storage Scale System solution is available at IBM Documentation. At the website, select the IBM Storage Scale System solution release level that you are using.

3.7.1 IBM Storage Scale System IBM Solution service and support

IBM provides service and support for IBM Storage Scale System as a solution. IBM Storage Scale System-specific keywords and procedures are in place so that you, as a client, when calling or electronically requesting IBM Storage Scale System support, are routed to an IBM Storage Scale System cross-component solution support team that understands the interaction of all of the IBM Storage Scale System components. They also can diagnose and respond to your query in a holistic fashion, and where necessary, direct your query to suitable Level 2 or Level 3 resources.

For more information, see the *IBM Storage Scale System (ESS) Support Reference Guide*.

3.8 Upgrading IBM Storage Scale System

Maintaining and upgrading your IBM Storage Scale System software solution stack is a customer's responsibility.

You also can choose to engage IBM System Lab Services or an IBM Business Partner to perform the upgrades for you.

The IBM Storage Scale System solution is an integrated package of hardware, firmware, and software. All components of an IBM Storage Scale System are integrated, tested, and supported by IBM as a solution. IBM periodically provides a refreshed or up-leveled IBM Storage Scale System software solution stack.

Existing IBM Storage Scale System customers are entitled to download the newer IBM Storage Scale System solution stack from IBM Fix Central. IBM recommends that you plan to upgrade your IBM Storage Scale System solution stack at least once a year.

For hardware upgrades, you should contact and work with your IBM or Business Partner specialist to plan and run the IBM Storage Scale System upgrades when necessary.

3.8.1 IBM Storage Scale System hardware and firmware

IBM Storage Scale System hardware and firmware updates are included in the IBM Storage Scale System solution stack. Upgrading to a new level of IBM Storage Scale System solution stack also upgrades the server and storage firmware levels, if necessary.

Typically, the IBM Storage Scale System upgrades are done non-disruptively. During a planned maintenance window, fail over the storage to one of the I/O data servers, upgrade the I/O data server, and then, fail over the storage to this newly upgraded server and upgrade the other server.

Storage enclosure firmware can typically be applied non-disruptively.

IBM Storage Scale System also supports nondisruptive addition of more storage enclosures, which allows an IBM Storage Scale System to be non-disruptively upgraded to a larger capacity.

The client is responsible for keeping the IBM Storage Scale System operating system, software, firmware, and GUI current with the recommended levels of software. For more information about upgrade instructions, see IBM Documentation.

Note: IBM and IBM Business Partners might offer IBM Services® to assist or perform these upgrades on behalf of the client.

For all hardware upgrades, clients work with the IBM or their Business Partner to match their needs with the IBM Storage Scale System upgrades and services offerings that are available.

3.8.2 IBM Storage Scale System software

IBM Storage Scale System provides a fully integrated and tested solution stack of software that consists of the following components at a specific release and fix level (including all prerequisites) that were fully integrated and tested:

- RHEL
- IBM Storage Scale
- Cluster Export Services code (CES) (SMB, NFS, and S3 Object Storage)
- IBM x86 server firmware
- RHEL kernel at specific level
- RHEL system errata at specific level
- RHEL Network Manager at specific level
- RHEL device drivers
- NVIDIA OpenFabrics Enterprise Distribution (OFED) network adapter drivers
- Electronic Service Agent (ESA)

Keeping the IBM Storage Scale System software solution stack current is a customer responsibility. IBM recommends that clients plan to upgrade their IBM Storage Scale System software solution stack at least once a year. Typically, the IBM Storage Scale System hardware and software solution stack upgrades are performed at the same time.

For more information about the levels of the software components for each IBM Storage Scale System software solution release level, see IBM Documentation.

Upgrading the IBM Storage Scale System solution software is done by a full replacement of the software solution stack with a newer level.

New levels of the IBM Storage Scale System solution software stack are available at IBM Fix Central. For more information about these newer levels of IBM Storage Scale System software, log on to IBM Fix Central by using your IBMid. At the website, in the Product Selector field, enter `IBM Storage Scale System`. Then, select your IBM Storage Scale System version or release level.

A list of available IBM Storage Scale System solution software fix packs for download is shown. When you start the download, IBM FixCentral checks your IBMid authentication profile to assure that you set up your IBMid with the suitable authentication and client customer number information. This information authenticates that you are authorized to download IBM Storage Scale System solution software.

For more information about applying the IBM Storage Scale System Fix Pack, see the IBM Storage Scale System Quick Deployment Guide.

IBM System Lab Services is available to perform the IBM Storage Scale System software upgrade for you. As a best practice, take advantage of these services. For more information, contact your IBM representative.

A

IBM Storage Scale System models

In this appendix, we describe information and differences between the three generations of IBM Storage Scale Systems (IBM Storage Scale System) that are used today. Understanding the nomenclature and the differences between these generations is helpful when you are designing your IBM Storage Scale System environment, especially if different generations of IBM Storage Scale System coexist in the same IBM Storage Scale cluster.

This appendix includes the following topics:

IBM Storage Scale System model specifications

IBM Storage Scale System 6000 specifications

The specifications of the IBM Storage Scale System 6000 model are listed in Table A-1.

Table A-1 IBM Storage Scale System NVMe flash building block model

IBM Storage Scale System Model	*Enclosure U/Number drives IBM machine type - model*	*Drive Count*	*Drive size*
IBM Storage Scale System 6000	*4U48 Machine type 5149-F48*	*24 or 48*	*3.84 TB 2.5 Inch PCIe Gen4 NVMe Flash Drive* *7.68 TB 2.5 Inch PCIe Gen4 NVMe Flash Drive* *15.36 TB 2.5 Inch PCIe Gen4 NVMe Flash Drive* *30.72 TB 2.5 Inch PCIe Gen4 NVMe Flash Drive* *19.2TBu NVMe FlashCore Module* *38.4TBu NVMe FlashCore Module* *3.84TB U.3 PCIe 5.0 NVMe SSD* *7.68TB U.3 PCIe 5.0 NVMe SSD* *15.36TB U.3 PCIe 5.0 NVMe SSD* *30.72TB U.3 PCIe 5.0 NVMe SSD* *30.72TB 2.5 Inch QLC NVMe SSD* *61.44TB 2.5 Inch QLC NVMe SSD*

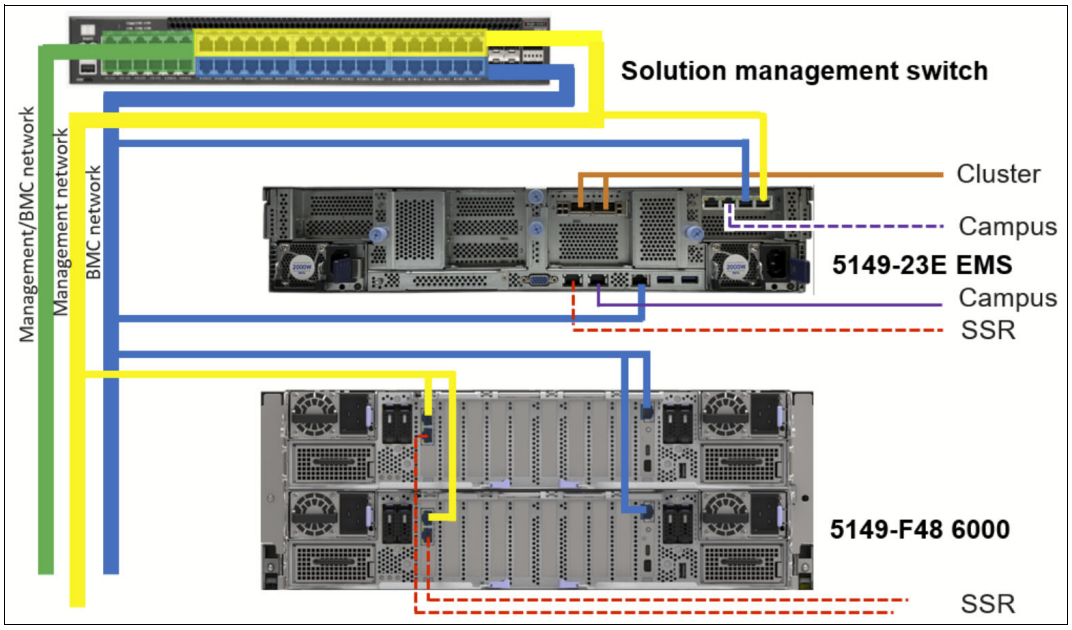

Figure A-1 IBM Storage Scale System 6000 back panel view and connections

IBM Storage Scale System 3500 specifications

The IBM Storage Scale System 3500 is one of the third-generation IBM Storage Scale System offerings.

The IBM Storage Scale System 3500 model includes NVMe flash storage.

The specifications of the IBM Storage Scale System 3500 model are listed in Table A-2.

Table A-2 IBM Storage Scale System NVMe flash building block model

IBM Storage Scale System model	Enclosure U/Number drives IBM Machine Type - Model	NVMe drives	Drive size
IBM Storage Scale System 3500 or **IBM Storage Scale System 3500 H0**	2U24 Machine type 5141-FN2	NVMe 12 or 24	3.84 TB 7.68 TB 15.36 TB

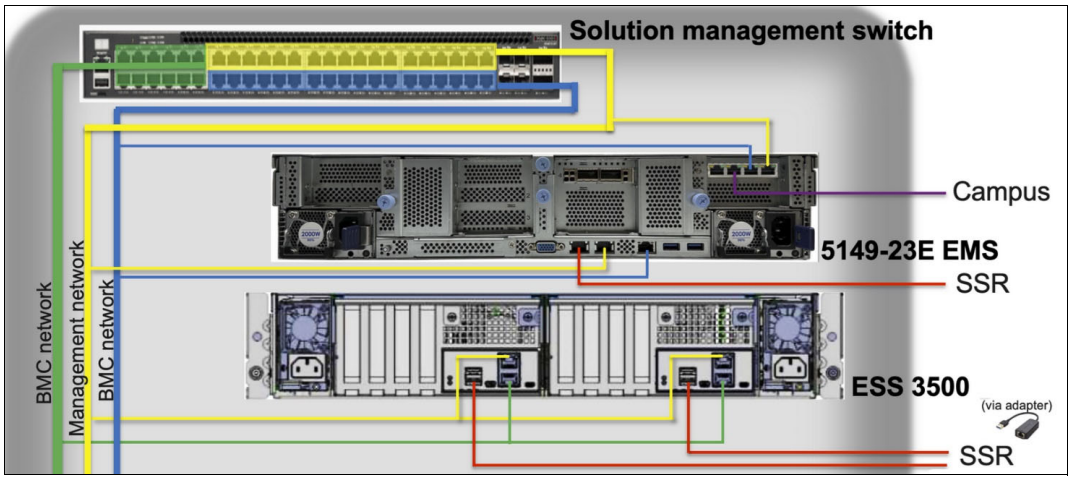

Figure A-2 IBM Storage Scale System 3500 rear panel view and connections

The IBM Storage Scale System 3500Hx, shown in Table A-3, includes machine type 5141-FN2 and 1-4 model 5147-102.

Table A-3 IBM Storage Scale System 3500Hx enclosure specifications

IBM Storage Scale System Model	Enclosure U/Number drives IBM machine type - model	NVMe	HDD enclosure	HDD drives	Drive size
IBM Storage Scale System 3500H1	2U24 Machine type 5141-FN2 + 1@4U external storage Model 5147-102	Quantity: 12 or 24 Size: 3.84 TB 7.68 TB 15.36 TB	1	52 or 102	HDD: 10 TB, 14 TB, 18 TB, 20 TB, 22 TB
3500H2	2U 5141-FN2 + 2@4U 5147-102		2	204	
3500H3	2U 5141-FN2 + 3@4U 5147-102		3	306	
3500H4	2U 5141-FN2 + 4@4U 5147-102		4	408	
3500H5	2U 5141-FN2 + 5@4U 5147-102		5	510	
3500H6	2U 5141-FN2 + 6@4U 5147-102		6	612	
3500H7	2U 5141-FN2 + 7@4U 5147-102		7	714	
3500H8	2U 5141-FN2 + 8@4U 5147-102		8	816	

The IBM Storage Scale System 3500Cx, shown in Table A-4, includes machine type 5141-FN2 logtip only NVMe drives and 1 - 8 model 5147-102.

Table A-4 IBM Storage Scale System 3500Cx enclosure specifications

IBM Storage Scale System Model	Enclosure U/Number drives IBM machine type - model	NVMe	HDD enclosure	HDD drives	Drive size
IBM Storage Scale System 3500C1	2U24 Machine type 5141-FN2 + 1@4U external storage Model 5147-102	Quantity: 4 (used for logtip)	1	52 or 102	HDD: 10 TB, 14 TB, 18 TB, 20 TB, 22 TB
3500C2	2U 5141-FN2 + 2@4U 5147-102		2	204	
3500C3	2U 5141-FN2 + 3@4U 5147-102		3	306	
3500C4	2U 5141-FN2 + 4@4U 5147-102		4	408	
3500C5	2U 5141-FN2 + 5@4U 5147-102		5	510	
3500C6	2U 5141-FN2 + 6@4U 5147-102		6	612	
3500C7	2U 5141-FN2 + 7@4U 5147-102		7	714	
3500C8	2U 5141-FN2 + 8@4U 5147-102		8	816	

Note: IBM Storage Scale for IBM Storage Scale System (current IBM Program IDs are 5765-DAT for Data Access Edition and 5765-DMT for Data Management Edition) is a specific IBM Program ID for use on IBM Storage Scale System. This license model is based on “per raw” TB capacity of the drives. The ”per raw TB” capacity of a drive refers to the total storage capacity of the drive without considering any overhead or formatting losses. For example, In the Storage Scale System 6000, drive capacities include 15.36TB for flash and 22TB for HDD – these are the raw drive capacities used for licensing purposes.

It uses a different metric than the IBM Storage Scale software-only capacity license (5641-DAx or 5641-DMx), which is licensed by TiB or PiB as specified by the mmlslicense command or the GUI. (The “x” indicates 1, 3, or 5 years of IBM software service and support.)

Also, unlike IBM Storage Scale software (which IBM clients can manage through IBM Passport Advantage®), IBM Storage Scale for IBM Storage Scale System does not have a Passport Advantage part number. Renewals with 5765-DAT and 5765-DMT can be managed with by working with your IBM Technical Support Services (TSS) representative.

For more information about how to use this licensing model on your IBM Storage Scale System, see IBM Storage Scale Licensing question 2.2 and *IBM Storage Scale: IBM Storage Scale System Licensing Information*.

IBM Storage Scale RAID technical overview

The following section provides a high-level, technical overview of the IBM Storage Scale RAID that is used in all IBM Storage Scale System models.

The IBM Storage Scale RAID software that is used in the IBM Storage Scale System solution runs on SAS disks in just a bunch of disks (JBOD) arrays. IBM Storage Scale RAID on IBM Storage Scale System provides JBOD cost reduction while simultaneously providing enterprise class reliability.

Different IBM Storage Scale System models provide solid-state drives (SSDs or NVMe) when more performance is needed. The IBM Storage Scale System solution does not require or use any kind of external RAID controller or acceleration.

IBM Storage Scale RAID supports multiple RAID codes and distributes client data, redundancy information, and spare space across the disks in such a way that if a physical disk loss or even a group of physical disk loss occurs, it does not affect data availability.

Instead of relying on the disks or conventional RAID controller only, IBM Storage Scale RAID implements erasure coding, which is embedded within the IBM Storage Scale file system to provide reliability and high performance. This configuration detects and reports storage media faults and read or write errors, and resolves other data integrity problems.

IBM Storage Scale RAID also implements an end-to-end checksum from the storage media all the way out to the IBM Storage Scale client.

IBM Storage Scale RAID erasure codes

IBM Storage Scale RAID in the IBM Storage Scale System supports different data protection algorithms and can detect and correct up to two or three concurrent storage media faults per erasure code array.

The options for RAID configuration are eight stripes of data plus two or three parity stripes that use Reed-Solomon codes or one stripe of data plus two or three replica stripes. The data plus parity or replica stripes, which are known as *tracks*, are shown in Figure A-3.

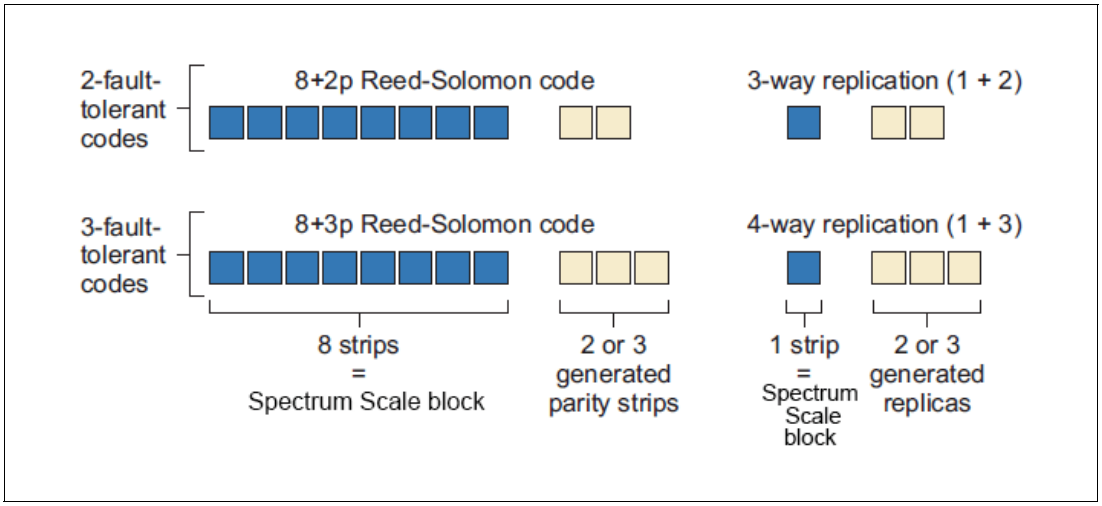

Figure A-3 IBM Storage Scale RAID tracks

End-to-end checksum

If you use the IBM Storage Scale software client to access data on the IBM Storage Scale System, the
IBM Storage Scale client is aware if the IBM Storage Scale file system is based on IBM Storage Scale RAID Network Shared Disks. If this configuration is used, an 8-byte checksum is calculated during a write operation, which is appended to the data, and sent over the network to the IBM Storage Scale RAID server. The checksum is verified and then IBM Storage Scale RAID writes the data along with its checksum on the disks and logs the version number on its metadata.

When a read operation is requested, IBM Storage Scale RAID verifies checksum and version on its metadata. If it is OK, it sends the data to the client. If it is not OK, the data is rebuilt based on parity or replication and then sent to the client along with newly generated checksum.

The end-to-end checksum feature provides a robust means to prevent and correct silent disk errors or missing disk writes. This feature provides an essential level of reliability to ensure data integrity and prevent file system checks, which ensure reliability and uptime for ever larger petabyte-scale file systems.

Declustered RAID arrays

IBM Storage Scale RAID implements its own data and spare disk layout scheme that reduces overhead and mitigates the performance impact to users when recovering from disk failures. IBM Storage Scale RAID spreads or declusters user data, redundancy information, and spare space across all the disks of the array instead of leaving all spare space in a single disk. A conventional 1+1 RAID layout is compared to a declustered array in Figure A-4.

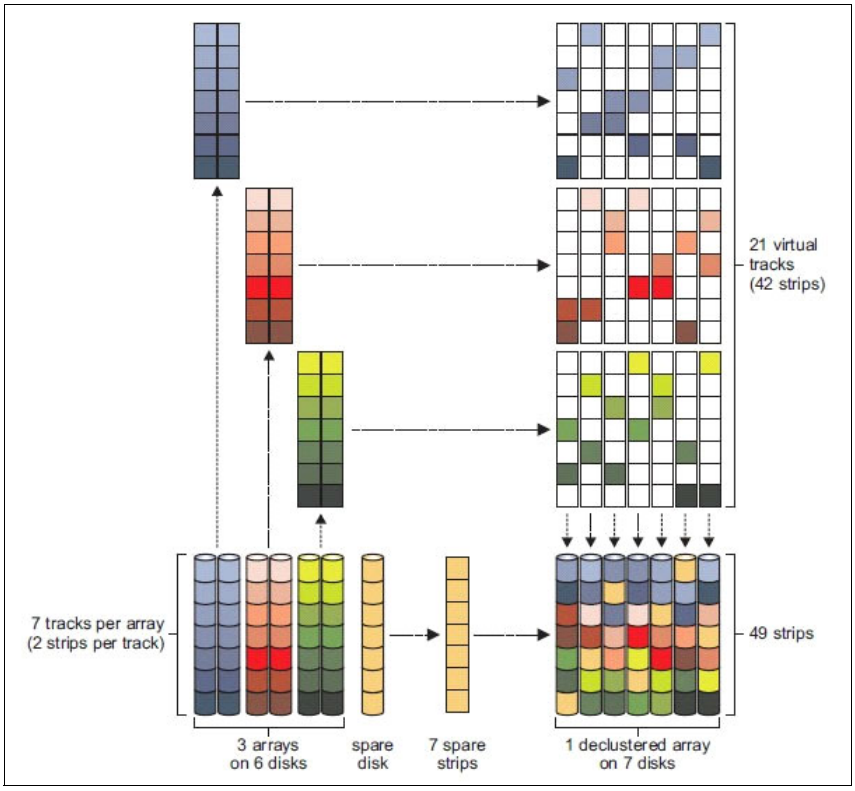

Figure A-4 Declustered array versus a 1+1 array

Consider an example where seven stripes of data are on each disk. Figure A-4 shows the left three arrays of two disks in a replicated 1+1 configuration and a spare. On the left, you can see the data stripes spreading all over the seven disks of the declustered array.

In a failure of one disk, all data from the remaining disks of the array must be replicated to the spare disk on the traditional 1+1 array. On the declustered array, the replication occurs on spare space of all the remaining disks, which can decrease the rebuild impact from three to four times.

The IBM Storage Scale RAID overhead reduction is shown in Figure A-5.

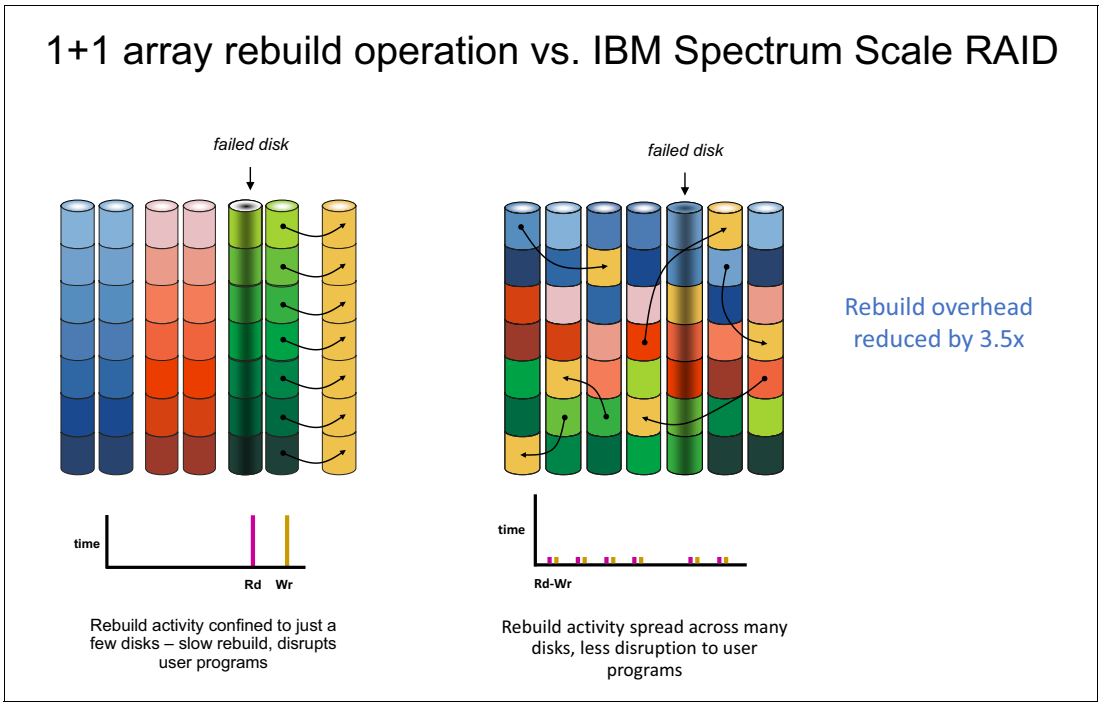

Figure A-5 Array rebuild operation

On the IBM Storage Scale System, using RAID 8+2 or 2-way replication (1+2), if one disk loss occurs, the rebuild operation starts with low priority with even lower impact for the clients. With this array configuration, the rebuild is considered critical and run on high priority only if two concurrent disk losses occur.

By using 8+3 RAIDs or 3-way replication (1+3), the rebuild operation becomes critical only if three concurrent disk losses occur in the same declustered array.

Abbreviations and acronyms

AFM	Active File Management
AI	Artificial Intelligence
APB	Ansible playbook bundles
ATG	Advanced Technology Group
BMC	Baseboard Management Controller
CES	Cluster Export Services
CIFS	Common Internet File System
CLI	Command line Interface
CTDB	Clustered Trivial Data Base
DHCP	Dynamic Host Configuration Protocol
EDR	Extended Data Rate
EMS	IBM ESS Management Server
EoS	End of Service
ESA	Electronic Service Agent
ESS	IBM Elastic Storage Server or IBM Elastic Storage System
EUS	Extended Update Support
FSP	Flexible Service Processor
GB	Gigabyte
GbE	Gb Ethernet
GPFS	General Parallel File System
HA	High Availability or Highly Available
HBA	Host Bus Adapter
HDD	Hard Disk Drive
HDR	High Data Rate
HMC	Hardware Management Console
HPT	High Performance Tier

IBM	International Business Machines Corporation
ITSO	International Technical Support Organization
JBOD	Just a Bunch Of Disks
MES	Miscellaneous Equipment Specification
MTM	Machine Type and Model
NFS	Network File System
NIC	Network Interface Card
NSD	Network Shared Disk
NVMe	Non-Volatile Memory Express
OFED	OpenFabrics Enterprise Distribution
PDU	Power Distribution Unit
POSIX	Portable Operating System Interface
QDG	Quick Deployment Guide
RAID	Redundant Array of Inexpensive Disks
RAS	Reliability, Availability, Serviceability
RDMA	Remote Data Memory Access
RHEL	Red Hat Enterprise Linux
RoCE	RDMA over Converged Ethernet
RPQ	Request for Price Quotation
SAS	Serial-Attached SCSI
SATA	Serial Advanced Technology Attachment
SCSI	Small Computer System Interface
SDS	Software-Defined Storage
SMB	Server Message Block
SME	Subject Matter Expert
SSD	Solid-State Drive
SSH	Secure Shell
SSS	Storage Scale System
SSR	System Services Representative
StorM	IBM Storage Modeler
TB	Terabyte
TDA	Technical and Delivery Assessment
TSS	IBM Technical Support Services
VLAN	Virtual Local Area Network
VPI	Virtual Protocol Interconnect

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 Redbook

Many IBM Redbooks publications and IBM Redpapers are available about IBM Storage Scale System and IBM ESS. For more information, see the following search results on the IBM Redbooks publications website:

- IBM Storage Scale Server
- IBM Storage Scale System
- IBM Storage Scale

You can search for, view, download, or order documents and other Redbooks, Redpapers, web docs, drafts, and additional materials at the following website:

ibm.com/redbooks

Online resources

The following websites are also relevant as further information sources:

- IBM Documentation Frequently Asked Questions for IBM Storage Scale Server:

  https://www.ibm.com/docs/en/ess-p8?topic=SSYSP8/gnrfaq.html
- IBM Documentation Frequently Asked Questions for IBM Storage Scale:

  https://www.ibm.com/docs/en/STXKQY/gpfsclustersfaq.html
- IBM Documentation for IBM Storage Scale Server:

  https://www.ibm.com/docs/en/storage-scale-system
- IBM Documentation for IBM Storage Scale Server Utility Node:

  https://www.ibm.com/docs/en/sssun

- IBM Documentation for IBM Storage Scale:

  https://www.ibm.com/docs/en/Storage-scale

- IBM Education on IBM Storage Scale Server or on IBM Storage Scale:

  https://www.ibm.com/training

  At this web page, search for:

  - Storage Scale System
  - IBM Storage Scale

- IBM ESS classes from IBM Learning Services: *Introduction to IBM Storage Scale Server* (no-charge course):

  https://www.ibm.com/services/learning/search?query=DL08017G

- IBM ESS IBM Storage Scale RAID Quick Deployment Guide:

  https://www.ibm.com/docs/en/ess-p8/6.1.4?topic=ess-deployment-guide

- IBM Network Switches sourced from NVIDIA:

  - https://www.ibm.com/docs/en/power9?topic=networking-switches
  - https://www.ibm.com/docs/en/power8?topic=POWER8/p8hdx/p8hdx_network_switches.htm

- IBM Rack:
  https://www.ibm.com/docs/en/power9/9080-M9S?topic=hss-rack-installation-specifications-racks-that-are-not-purchased-from

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

REDP-5729-01

ISBN DocISBN

Printed in U.S.A.