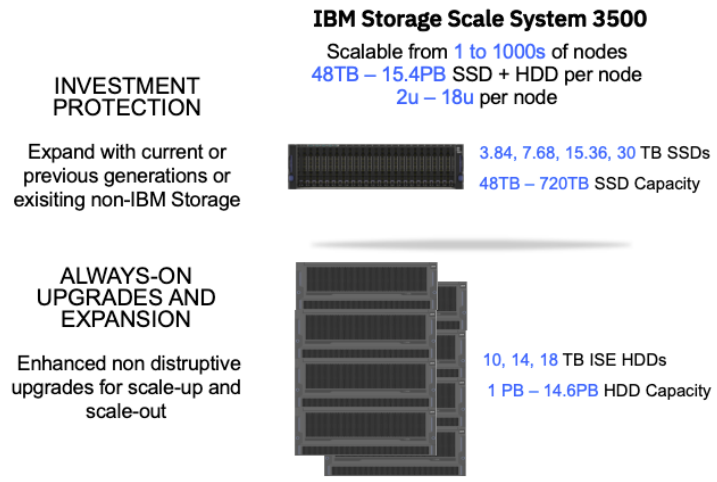
Solutions and Use Cases for IBM Storage Scale System

Solving data challenges with simple building blocks

Highlights

- Simple building blocks to create a global data platform
- Configure a system in minutes with simple wizzards and ansible configuration scripts
- Integrate to S3 object data to expand use cases and access global data for current applications
- Automate data lifecyle and application performance tuning with data management services
- Create a more sustainable infrastructure with a single node and a system that can turn off and on data when not used
- Combine with other IBM and non-IBM storage including the public cloud
- Access the same data from a file and an object application simutabously with data access services
- Create container native storage connecting to current applications

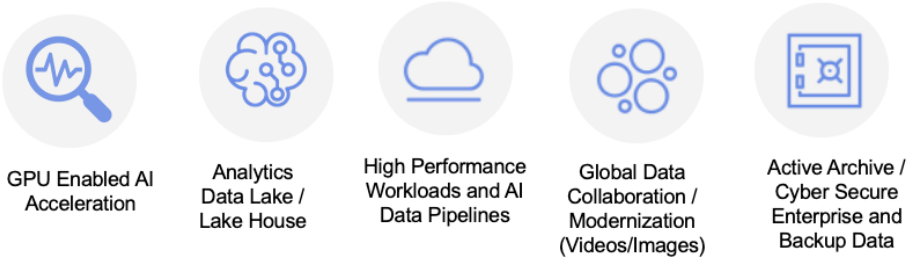


IBM Storage Scale is a high performance, cyber secured global data platform engineered to unlock customers data for maximum value. Our storage software is engineered to accelerate AI and data intensive applications. Our storage is engineered to unlock the value of data with data services that provide a global data strategy from edge to core to cloud. IBM Storage Scale is optimized for cost and performance, secured for faster recovery, protect from the unknown, and engineered to access data anywhere. IBM Storage Scale System is a plug and play appliance that provides market-leading performance, density and scalability that seamlessly integrate to a customer's file and object storage environment. These nodes are perfect for edge computing or core data center deployments and can consolidate data for hybrid cloud solutions that include the public cloud, other file and object storage and Red Hat OpenShift containerized nodes.

You might ask, “Why buy an integrated solution when I can buy just the software and build my own solution?” Of course, you can build your own if that is the route you wish to take. However, many customers opt for the integrated solution. Simple tasks can become more complex as many tasks are not integrated or may not be tested. Installation may include many parts to install, and performance tuning is difficult and hard to optimize. Finally, maintenance can include difficult-to-patch components such as OS, server firmware, switch firmware, software updates, etc.

IBM Storage Scale System comes as an integrated solution. Hardware is assembled and software is pre-installed and tested in the factory. The solution is ready to be mounted in a rack at the client site. All components make the deployment experience as smooth as possible, and a system can be configured in minutes ready to accept data. IBM Storage Scale System is a single vendor end-to-end solution, simplifies the deployment process as well as the support and maintenance tasks involved. The upgrade process is an all-encompassing online replacement of the existing solution stack with a newer version. This includes the OS, firmware, and software stack. This means that you get upgrades as painless as possible and can extend the life of your system for years. Each node can be separately upgraded so you can better plan your budgets and lifecycle management.

Application Workloads for the Storage Scale System

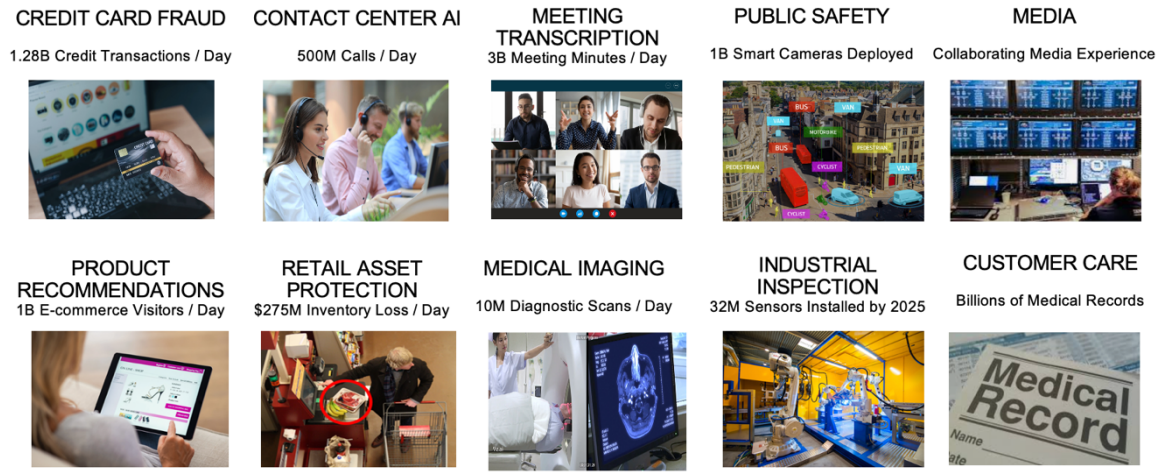


The five primary use cases for IBM Storage Scale and IBM Storage Scale System are:

- GPU Enabled AI Acceleration
- Analytics, data lake and data lake house workloads
- High performance computing and AL/ML data pipelines
- Global data collaboration and data modernization especially with videos and images
- Active archive, cyber security protection and enterprise data backup

IBM Storage Scale System provides an extreme high-performance tier of Storage Scale file and object storage with up to 91GB/s of throughput and up to 16M IOPs[1], for a broad variety of applications. The Storage Scale System is designed to keep GPUs active to solve AI problems faster and running at peak performance. Like all previous generations, IBM Storage Scale System runs the proven IBM Storage Scale RAID erasure coding, which provides superior

consistent high-performance, mitigation of storage hardware failures, and intelligent monitoring, management, dynamic tuning of IBM Storage Scale System and IBM Storage Scale data.



These are some of the other ways customers are using the IBM global data platform to help transform organizations with data. The Global Data Platform helps build the AI foundation for multiple multiple workloads that changing the way data is used. From Financial Services to Services providers, the benefits of the Global Data Platform are helping customers around the world with new modern applications that require data and access to data in new and unique ways.

The opportunity for AI to transform every industry is huge. From credit card fraud detection on billions of transactions, to public safety with over 1 Billion smart city cameras, to product recommendations for 300M e-commerce visitors / day, the demand for fast, easy inference deployment is greater than ever with the AI economy upon us.

Contact your IBM Sales representative or value added business partner for details about pricing and your specific configuration.

GPU Enabled AI Acceleration

One of the fastest growing use cases for AI is using GPU acceleration for optimizing the AI data pipeline. NVIDIA is the recognized leader for GPU acceleration and IBM is a partner with

NVIDIA bringing more data faster to improve Data Science productivity. IBM provides a Scalable, software-defined infrastructure powered by the Global Data Platform for AI and NVIDIA DGX systems. IBM Storage for scalable data and AI with NVIDIA DGX is the perfect engine for your

data pipeline on which companies can build their shared data services. Storing data is important for GPU acceleration and NVIDIA use cases because GPUs (graphics processing units) are designed to handle large amounts of data in parallel. By storing data on fast and efficient storage systems, such as IBM Storage Scale System, GPUs can access the data quickly, enabling faster processing and analysis. IBM Storage Scale System is engineered to specifically accelerate GPU enabled and NVIDIA workloads by enhancing:

- Data transfer speed: GPUs require a significant amount of data to operate efficiently, and storing data on high-speed storage devices allows the GPU to access that data quickly, reducing the time needed to transfer data between storage and processing units. By providing up to 91GB/s and up to 16M IOPs per node and GPU direct storage IBM Storage Scale optimizes data transfer speed.
- Data pre-processing: In many cases, data must be pre-processed before it can be used in GPU-accelerated applications. By storing the data in a pre-processed format on IBM Storage Scale System, GPUs can access and use the data immediately, without the need for additional pre-processing.
- Large data sets: NVIDIA GPUs are often used to process large data sets, such as those generated by scientific simulations or machine learning algorithms. Storing these data sets on IBM Storage Scale System enables global data to be transparently accessed, reducing the time needed for analysis and improving overall performance.
- Real-time data processing: In applications that require real-time data processing, such as autonomous driving or video surveillance, data must be stored on IBM Storage Scale System to ensure that the GPU can access the data quickly enough to process it in real-time.

Analytics, data lake and data lake house workloads

Storing data is crucial for analytics, data lakes and data lake houses because it allows organizations to collect, organize, and analyze large volumes of data from various sources.

IBM Storage Scale System helps solves two difficult challenges for Analytics. For the data users, IBM Storage increases productivity and provides faster results. Users no longer have to worry about moving data or accessing the data located in remote silos as data is transparently

connected. Users can also use the global data catalog that is included to help curate and organize the data for much easier and faster analysis. All data is visible no mater where its located, including data in the cloud. For the Data Admin, they experience lower cost and increased service-level agreements (SLAs) by ensuring optimized resources and a single copy of data that is protected but not duplicated in multiple locations. The system is much easier to manage as the same cluster spans multiple environments. There are several other reasons why data storage is important for Anlytics:

- Data preservation: Storing data on IBM Storage Scale System helps to ensure that it is available for analysis when needed. IBM Storage Scale System with tranparent archiving to tape or cloud ensure that data is efficeintly stored and readily available. This is especially important for historical data that may be required for trend analysis, forecasting, or compliance purposes.
- Data integration: Data from various sources can be integrated into a single repository, which makes it easier to access and analyze. This can include data from IBM storage, non-IBM storage and cloud storage.
- Scalability: Storing data in a data lake using IBM Storage Scale System with up to 633YB of maximum capacity allows organizations to store large volumes of data and scale their infrastructure as needed to accommodate growing data volumes.
- Data governance: Storing data in a centralized location with IBM Storage Scale System allows organizations to enforce data governance policies and ensure that data is properly secured, managed, and used in accordance with regulatory requirements.
- Analysis and insights: Storing data in IBM Storaeg Scale System enables organizations to perform advanced analytics and gain insights into business performance, customer behavior, and operational efficiencies.

High performance computing and AL/ML data pipelines

Storage is a critical component of high-performance computing (HPC) and AI/ML data pipelines because these systems require fast and reliable access to large volumes of data. To deliver the benefits of HPC and AI/ML data pipelines, it is important to have a good storage foundation.

Here are some reasons how IBM Storage Scale Systems breaks barriers for HPC and AI/ML data pipelines and provides the best storage foundation:

- Data accessibility: HPC and AI/ML requires fast and reliable access to data from multiple sources. IBM Storage Scale System ensures that data is easily accessible and can be quickly retrieved when needed. The IBM Storage system is capable of handling large data sets (PBs to YBs) and provides fast read/write speeds to support the needs of the workloads.
- Data integration: HPC and AI/ML workloads need to integrate data from various sources, including IBM and non-IBM storage and even the public cloud. IBM Storage Scale System provides a unified storage layer that can handle multiple data types and formats, making it easier to integrate data from diverse sources.
- Data security: IBM Storage Scale System secures storage for sensitive data and provides secure storage solutions, such as end-to-end and data at rest encryption with user-controlled keys, access controls with multi-layer autehntication, and fast cyber secure recoveries for PBs of data. The system is designed to ensure that data is protected from unauthorized access and cyber-attacks.
- Data scalability: IBM Storage Scale System supports large-scale data processing and storage. IBM Storage provides the necessary storage infrastructure to handle the growth of data, ensuring that workloads can scale as needed and use new or exiting resources all on-line without downtime to the users.
- Data performance: HPC and AI/ML workloads requires fast data processing and access to deliver insights in real-time. IBM Storage Scale provides high-performance storage that can meet the demands of demanding workloads like AI with up to 91GB/s throughtput with ms latency and up to 16M IOPs per node that can scale to 1000s of nodes.

Global data collaboration and data modernization

Many organizations are moving their storage to cloud-based solutions such as Amazon Web Services (AWS), Microsoft Azure, IBM Cloud and Google Cloud to modernize and centralize data for better collaboration. This works if you want to move everything to the cloud and lift and shift all your current applications. What if you have data that is in the cloud and in the data center and located in different locations and on multiple platforms. That is the problem that IBM Storage Scale System solves with this use case. IBM Storage Scale System is designed to

modernize data for new and existing workloads and at the same time increase collaboration of users and applications by connecting data with the following capabilities:

- Centralized storage: IBM Storage Scale System stores data in a centralized location ensuring that all team members have access to the same information. This is particularly important for data collaboration, as it allows teams to work together on the same dataset, regardless of their location.
- Scalability: As data sets grow and application data requirements expand, IBM Storage Scale System ensures that the data grows with them. IBM Storage Scale System can handle increasing amounts of data without sacrificing performance or reliability with our distributed parallel architecture and the capabilities in our global data platform.
- Data integrity: IBM Storage Scale System ensures that data is stored securely and is protected from loss or corruption. This is particularly important for modernization efforts, which often involve integrating data from legacy systems to new platforms.
- Accessibility: IBM Storage Scale System storage solutions enable data to be accessed from anywhere, at any time. This is essential for teams that are distributed across different locations and time zones.
- Integration: IBM Storage Scale System integrates with modern data processing and analysis tools, such as machine learning and artificial intelligence, which can help organizations unlock the full value of their data.

Active archive, cyber security protection and enterprise data backup

Protecting data is critical to an organization and there are many different reasons why companies need a comprehensive resiliency plan. Ransomware attacks have become a significant and growing threat to organizations in recent years. Ransomware is a type of malicious software that encrypts files on a victim's computer or network, making them inaccessible until a ransom is paid to the attacker. These attacks can be incredibly disruptive, causing significant downtime, data loss, and financial damage.

In addition, protecting file and object data is important for compliance purposes. Many organizations are required to store data for regulatory or legal reasons, and losing or compromising that data can have serious consequences.

File and object data may also contain sensitive information, such as personal or financial data, intellectual property, or confidential business information. If this data is lost or becomes

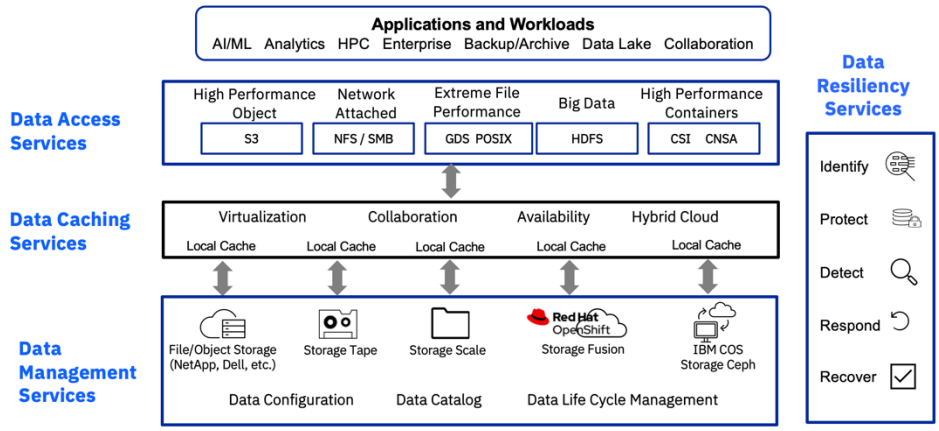
unavailable for longs periods of time which could be as short as a few hours or days it put the business at risk for short term or even long term problems. If this data falls into the wrong hands, it can be used for malicious purposes, including identity theft, fraud, or corporate espionage.

This use case is one that overlaps other use cases but can also be a major requirement for choosing a file and object storage solution that we break it into its own specific solution or use case for IBM Storage Scale System. The capabilities of IBM Storage Scale System are key differentiators when it comes to active archive, cyber security, and enterprise data backup. IBM Storage Scale System provides a rich information lifecycle policy management (ILM) engine that offers intelligent automated tiering with extreme scalability. IBM Storage Scale System also supports the ability to have separate storage pools with different storage characteristics such as flash, SSD, disk, and IBM Storage tape or S3 cloud storage. The policy engine provides seamless migration and file and object placement between storage pools under a single global namespace.

For cyber security IBM Storage Scale provides an air-gapping approach with IBM Safeguarded Copy. The cyber secure capability creates a logical air-gap relationship between the source files and objects and the secure recovery environment. These restore points can be used to quickly bring data back online following a cyberattack if a crash-consistent restore operation is required. We refer to this as a secure operational air-gap as the data set being analyzed for restore operations is separated from the source data, but often in the same datacenter.
IBM Safeguarded Copy for cloud scale file and object storage is a new protection mechanism for data on IBM Storage Scale and Storage Scale System. Safeguarded Copy sessions secure data to prevent it from being compromised, either accidentally or deliberately. Safeguarded Copies can be used to take many frequent copies of a production environment (for example, hourly copies maintained for a number of days), while Spectrum Scale snapshots continues to be used to take a small number of less frequent copies (such as weekly copies maintained for 1-2 weeks).

IBM Storage Scale System provides capabilities to not only securely backup data within the IBM Storage Scale System with the IBM Storage Scale Backup option, the system can also provide an excellent target for global enterprise data because of the resiliency, ease of scalability and the integrated life cycle management of data that can transparently move data to tape or cloud optimizing both accessibility and cost of storing long term backups. For more information on data protection and disaster recovery on IBM Storage Scale you can reference additional information here.

The IBM Global Data Platform



Our storage solutions are designed to optimize, secure and unlock customers data. We accomplish these goals with two software products and 4 differentiated data services. Our data services include:

Data Access Services with Multi-protocol Performance that connects directly to your applications
Data Caching or Core Services that provide Global Connectivity from multiple data sources and multiple locations to bring together data from IBM and non-IBM storage environments.
Data Management Services with policy automation that transparently helps you mange the flow of data and take much of the complexity out of data to day data management
Data Resiliancy Services that provide cyber secure automation to ensure your data is protected and safe and quickly recoverable when needed

IBM has been about inventing things since its inception in 1911. With the Global Data Platform, IBM has engineered a platform that adds values for multiple file and object concurrent workloads or use cases. Since IBM Storage is all about data, it follows that IBM Storage innovation is about achieving value from that data and we summarize our focus on three main business needs for customers which are, unlocking data, de risking data and optimizing data. These three business needs form the three key differentiators for our use cases.

- Unlock data with ultimate parallel performance optimized for large and small files required for AI workloads and connect existing global resources that can be accessed in place anywhere by anyone
- De Risk data that can be accessed in multiple locations concurrently and always-on with fast cyber secure recovery for PBs of data
- Optimizing data using less resources, lower cost/energy and up to 380% ROI with automated data reduction, archiving and using existing storage and cloud resources

Why IBM?

Data matters. When planning high performance infrastructure for new or existing applications it's easy to focus on compute resources and applications without proper planning for the data that will drive the results for the applications. As customers strive to unlock IT, de risk IT and enable green IT, IBM storage is all about solving those difficult challenges. Our products are all about solving hard problems faster with data so that custoers can react faster to competition and to market dynamics by accelerating insights and providing seamless access to data that is not limited by borders or data silos. Our story is not just about another storage product but is about innovation and a storage portfolio that is powered by our global data platform.

For Further Information

For further information on IBM Storage file and object products please visit
https://www.ibm.com/storage/artificial-intelligence

Next steps

Contact your IBM or IBM Business Partner Representive
(https:/ /www.ibm.com/partnerworld/bpdirectory/)