

Smarter technology for all

Lenovo Ceph Storage Solutions

Lenovo

Leaders are Faced with Unprecedented Challenges

Changing Technologies

Traditional infrastructure limitations

Deploying infrastructure to support new and existing generative AI and AI workflows is a top initiative but AI technology is not performing as expected

Data Silos

Data that is difficult to access across silos

82% of enterprises are inhibited by data silos and 87% run multi-cloud environments, with 44% running applications siloed on different clouds

Sustainability & Costs

AI requirements drive the need for more resources

80% of enterprises are working with or planning to adopt AI and 86% have a sustainability strategy

Unknown Threats

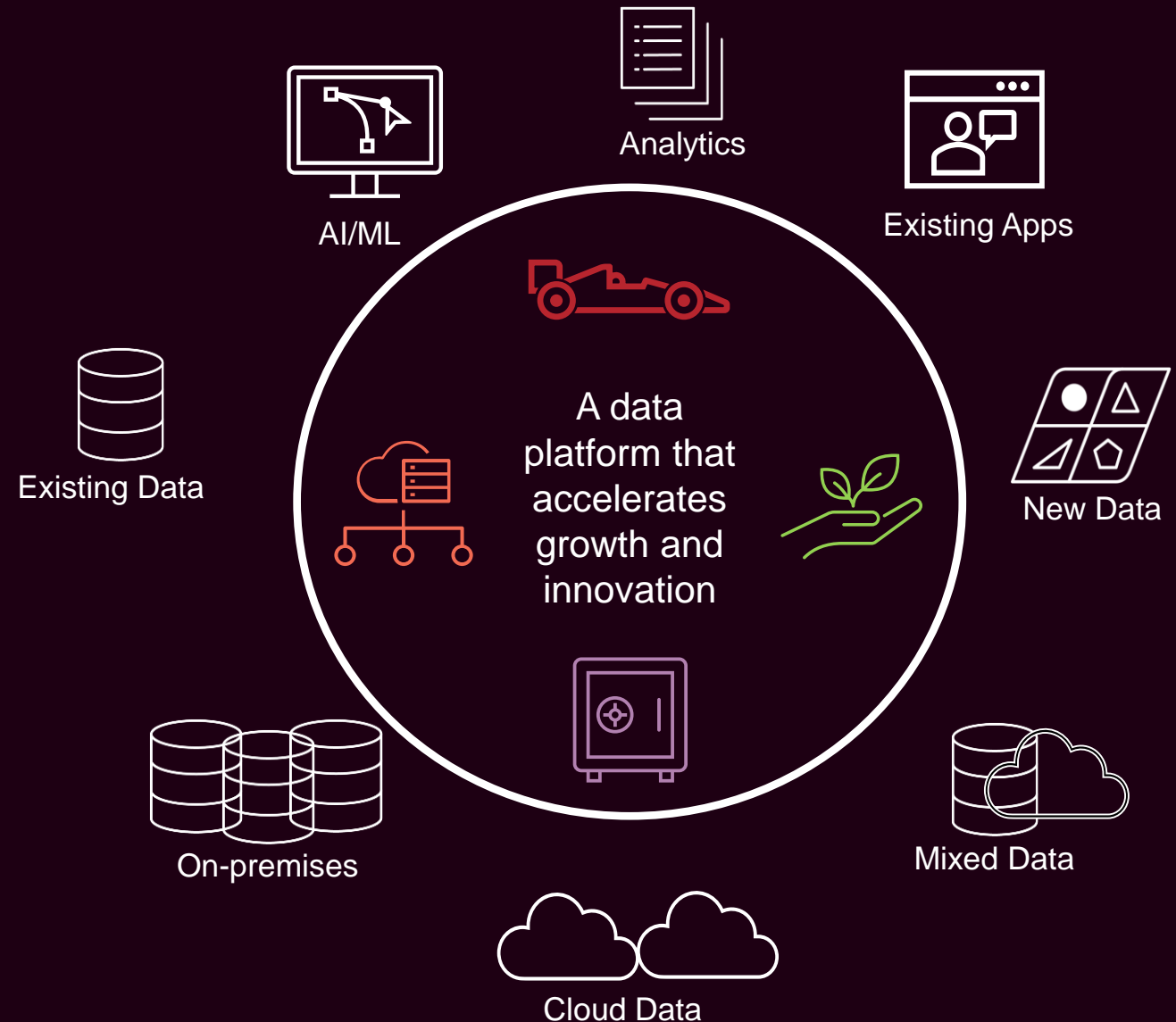
As AI becomes mission critical the data must be protected

As of 2023, over 72% of businesses worldwide were affected by ransomware attacks

... Leading to more cost and complexities with current infrastructure

What if you Could Break Free from Challenges and Enable...

- Single view for large data sets with the ability to enhance data science productivity
- Which is easily scalable to manage growth
- With advanced data management for performance, global connectivity, sustainability and resiliency



Lenovo Ceph Storage Solutions Help Clients Manage Data

To solve these challenges, Lenovo has partnered with IBM to deliver a validated software-defined Ceph solution for large scale storage and data management built on industry leading x86 ThinkSystem Ready Nodes. Together, Lenovo and IBM, deliver a world-class open-source software-defined solution supporting large data sets with unparalleled support.

- Enterprise Support: Available 24x7x365 with Lenovo Premier support
- Scalability: Non-disruptively grow from one to hundreds of petabytes
- Cost-effective: At scale, implementing software defined storage reduces TCO
- Validated Design: Easy to configure and order with proven performance and manageability
- Security: Enterprise grade data protection and data governance

Lenovo Ceph Storage Solutions Help Clients Manage Data

Lenovo Ceph Storage Solutions validated design leverages ThinkSystem high performance reliable servers for secure & simplified operations delivering the benefits of open-source storage with enterprise-grade confidence and support for AI & large data repositories with 24x7 support.

Validated Design * Secure * Enterprise Support

- ✓ *Validated design = simple planning & purchasing*
- ✓ *24x7x365 enterprise support by Lenovo & IBM*
- ✓ *No single point of failure*
- ✓ *Self-healing and self-managing*
- ✓ *Automatically recover/copy data*
- ✓ *S3 fidelity for integration with modern applications*
- ✓ *Encrypt data at Rest, SSE-KMS, SSE-S3, SSE-C*
- ✓ *Certified WORM, Object lock capability to prevent data manipulation*
- ✓ *Scales to billions of objects, with deterministic and sustainable performance*
- ✓ *Analytics for critical data insights*

Storage Ceph Data Services

Data Efficiency

- Multi-protocol
- Hybrid/ Multi-cloud
- Replication
- Erasure coding
- Compression
- Performance-at-scale

Data Resilience

- Snapshots
- Clones
- Backup and recovery
- Migration options
- Business continuity
- Disaster recovery

Data Security

- Authentication and authorization
- Data policies
- At-rest encryption
- In-flight encryption
- Certified WORM
- Key management service integration

AI with Storage Ceph

STS Support

Role-based governance

Security

Store tables with S3 data encryption

D3N Caching

- D3N query acceleration
- Datacenter Data Delivery Network
- Natively implemented in IBM Storage Ceph RGW

Parquet Support

S3 Select pushdown

Bucket Notification

Notification of bucket events or changes. Can be used in context of serverless runtimes, feeds into messaging systems or similar.

Open-source Storage Software

For interaction with open-source data analytics applications. Also suitable for any other workload that requires file-, block or object storage.

Ceph Storage: Operational Benefits

- **Scalability:** Scale horizontally, allowing organizations to add storage capacity & performance by simply adding Ready Nodes.
- **Unified Storage:** Unified storage platform support file, object, and block storage requirements.
- **Data Protection:** Ceph's RADOS (Reliable Autonomic Distributed Object Store) provides data redundancy and resilience through replication or erasure coding, ensuring data availability and integrity.
- **Performance Optimization:** Includes features like caching mechanisms, tiering & QoS (Quality of Service) to enhance performance.
- **Management & Monitoring:** Centralized management interfaces allow administrators to configure, monitor, and manage the storage infrastructure efficiently. This includes features for automation and integration with cloud management platforms.
- **Multi-Tenancy & Security:** Capabilities for secure multi-tenancy, access controls, and encryption to protect data at rest and in transit.
- **Compliance & Data Governance:** Features to comply with regulatory requirements and implement data governance policies effectively.
- **Cloud Integration:** Ability to integrate with public cloud and hybrid cloud architectures, providing seamless data mobility & workload portability.
- **High Availability:** Built-in mechanisms ensuring high availability, including fault tolerance, automatic failover, & recovery capabilities.
- **Analytics & Insights:** Advanced analytics & reporting capabilities to gain insights into storage performance, capacity utilization, & trends.
- **Container & Kubernetes Support:** Integration with container orchestration platforms like Kubernetes, enabling deployment of persistent storage for containerized applications.
- **APIs & Integration:** Extensive APIs and integration capabilities to facilitate automation, orchestration, and integration with third-party applications and tools.
- **24x7x365 Support:** Lenovo Premier support for both hardware and software reducing risk and complexity.

Lenovo Ceph Storage: Use Case Examples

Generative AI, Data lakes and Data analytics

- Client analysis/support
- Healthcare Treatment/diagnosis
- Client care analysis
- Engineering design
- Research/business analysis
- Large pattern analysis
- High throughput database
- Computer aided engineering (CAE)
- Public safety/smart cities
- Hybrid Cloud, S3 fidelity

Backup, Archive, Data collaboration

- DR backup data
- Archival of less used data
- Video/Voice/Image applications
- Remote data collaboration
- Design collaboration
- Video surveillance/production

Storage for Data & AI: Case Study

Preserve Software Code for Future Generations



Business Challenges

Client needed a rock solid and flexible storage system to store tens of billions of files efficiently anywhere in the world and scale to large capacity as requirements continue to grow.

Organizational Outcomes

- 10+ Billion Files now stored in the Ceph Cluster
- Used by millions of projects around the world
- Easy to scale without downtime as requirements grow

Technology Choices

- Storage Ceph

Storage for Data & AI: Case Study

Support millions of users for smart city application



Business Challenges

Client wanted to build an infrastructure of the future that was flexible and could easily store, manage and connect disparate data on various city applications, such as traffic conditions and waste management, together through a mobile app accessible by citizens.

Organizational Outcomes

- Support millions of concurrent users accessing data
- Easy to scale without downtime as requirements grow
- Used by multiple applications

Technology Choices

- Storage Ceph

Storage for Data & AI: Case Study

Large retail client used the Cloudera Hadoop Analytics Ecosystem to provide insights that are used by the different business domains



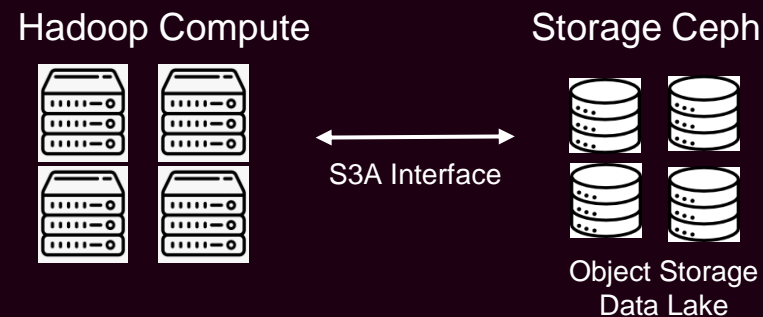
Business Challenges

- Client wanted to replace their current Data Lake
- Facing huge yearly maintenance costs
- Complicated platform lifecycle forced to work with old versions
- Not able to run multiple versions of Hadoop on the same dataset
- Current HDFS platform not performant for streaming data

Organizational Outcomes

- Reduce Costs
 - Breaking compute from storage
 - Reduced raw storage with EC
 - RGW Multisite out-of-the box DR
- Enhanced Functionality
 - Ability to ingest data with batch or streams
 - Access same datasets with different HDP versions
 - Ceph fulfills the lifecycle and scale needs of client Data teams

Solution: Ceph Storage



Lenovo Ceph Storage Solutions: Operational Benefits

Scalability

On-demand horizontal scaling and autonomic re-balancing. Scaling positively impacts both capacity and performance

Reliability

Zero data loss by having resiliency, self-healing capabilities and autonomic failure domains

CapEx Efficiency

Storage Ceph runs on validated industry leading Lenovo x86 ThinkSystem Ready Nodes

Data Security

Storage Ceph offers encryption, Object lock, certified WORM functionality for regulatory compliance

Flexibility of Unified Storage

With common storage pool operations and unified storage services with Object, File, and Block access protocols

Hybrid Cloud Compatibility

Seamless connectivity and access to data on-prem, compatible with public cloud providers, with a high fidelity of similarity

Lenovo Ready Nodes + Ceph: A Complete Solution

Ready Nodes offer a pre-configured HW platform designed to simplify & accelerate deployment of software-defined storage solutions. Validated hardware configurations ensuring compatibility and performance optimization for the intended workload.

IBM Storage Ceph Software License Types

- Storage Ceph is a Subscription license model
 - 1 license = 1TB of raw storage

Options

- IBM Storage Ceph Premium Edition, Object only
- IBM Storage Ceph Premium Edition, Block, File & Object
- Terms: 1, 3, or 5-year terms

- Both options include
 - IBM Storage Insights, one license for every 1TB managed (management analytics)
 - Red Hat Enterprise Linux, one license for every 40TB managed (OS to run the Ceph cluster)

Lenovo Storage ThinkSystem Ready Nodes

Lenovo servers achieved the highest levels of 'five nines' or 99.999% reliability among all mainstream server hardware platforms.

— ITIC Global Server Hardware. Server OS Reliability Survey

Ready Nodes

Lenovo Ready Nodes are a simple, cost-effective way for clients to deploy Ceph software on industry leading hardware

Validated Building Blocks

A simple building block approach to software defined storage and infrastructure, validated and supported by Lenovo

Incremental Growth

Start with a cluster configuration that meets current needs and then simply scale with additional nodes as needed (Minimum cluster is four nodes)

Lenovo Product Integration

Investment protection while Lenovo Storage Ready Nodes can be integrated into existing Storage software configurations

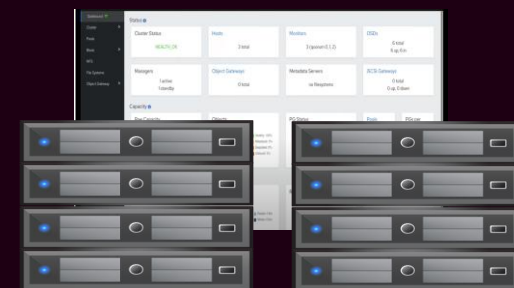
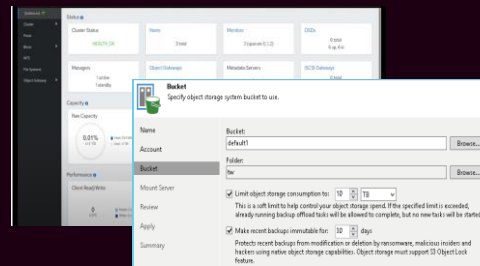
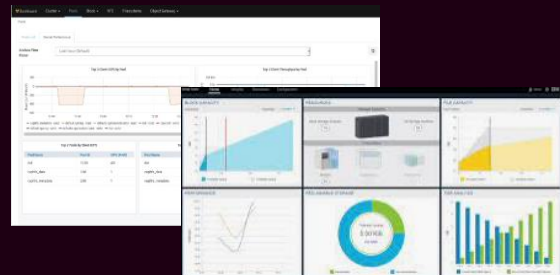
Use Cases

- Large Data repository, Data Lakehouse
- Backup/archive target
- Cyber secure backup
- S3 Hybrid Cloud
- Containers, persistent storage

Enterprise Support

Lenovo support 24x7x365 with Premier support

Lenovo Ceph Storage Solutions: Process



1

Install Lenovo ThinkSystem Ready Nodes with Red Hat Enterprise Linux.

Start with four server nodes.

2

Install Storage Ceph with Storage Insights.

Storage Insights is included for use with Lenovo Ceph Storage.

3

Configure object storage buckets and access keys and parameters.

Configure your applications with relevant bucket access information.

4

Add additional resources and capacity according to needs.

Perform Day-2 operations, like in example, setup remote replica cluster locations.

Product Definition

Lenovo Ceph Storage Solutions Hardware: Lenovo Ready Nodes (HS350X V3 & SR650 V3)

Pre-configured server platform designed to simplify & accelerate deployment of software-defined storage solutions - validated hardware configurations, including CPU, memory, storage, and networking components, ensuring compatibility and performance optimization for the intended workload.

Two options: Easy to configure, quote, and manage



ThinkSystem HS350X V3

- Intel Xeon Gold 32C 2.1GHz Processor
- 24 x 3.5 HDD
- 8,12,16,22TB SATA
- 10/25GbE NIC

Capacity / Backup / Archive / Media



ThinkSystem SR650 V3 (TLC or QLC)

- Intel Xeon Gold 32C 2.1GHz Processor
- 28 x NVMe
- TLC = 3.8, 7.7 & 15.4TB
- QLC = 15.4, 30.7TB (& 61TB future statement)
- 10/25GbE NIC

Performance / Data & AI / Lakehouse

Lenovo Ceph Storage Solutions Protocol Support

Object Storage

- AWS S3 Supports all common AWS S3 bucket and object API calls
- Advanced S3 features such as object versioning, object lock, S3-select with table format support (Apache Parquet, CSV and JSON)
- OpenStack Swift Compatible with the OpenStack Swift object storage API

Block Storage

- Supports NVMe/TCP block protocol, VMware ESXi 7.0U3 and 8.0+
- Supports latency-sensitive applications running on operating systems that support NVMe/TCP
- Ceph RBD, Block storage access through the native Ceph RBD client part of RHEL or OpenStack. Used for KVM/QEMU virtual machine workloads and native Linux applications requiring TCP/IP-accessible block storage

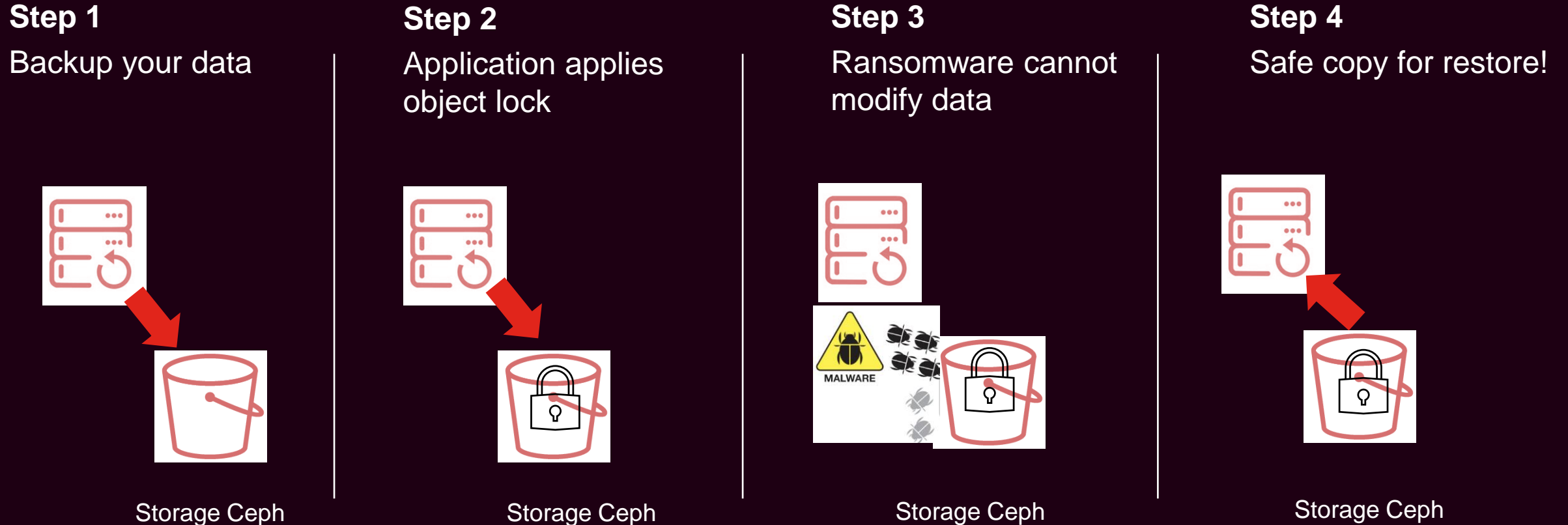
File Storage

- NFSv3 and NFSv4, Export of CephFS volumes and sub-volumes via NFSv3 and NFS v4.1
- NFSv4 gateway to object storage, support for data ingest and export of object storage
- data via NFS shares
- CephFS, export of CephFS volumes and sub-volumes via native Linux kernel or FUSE client running on RHEL or other Linux distribution

Container Storage

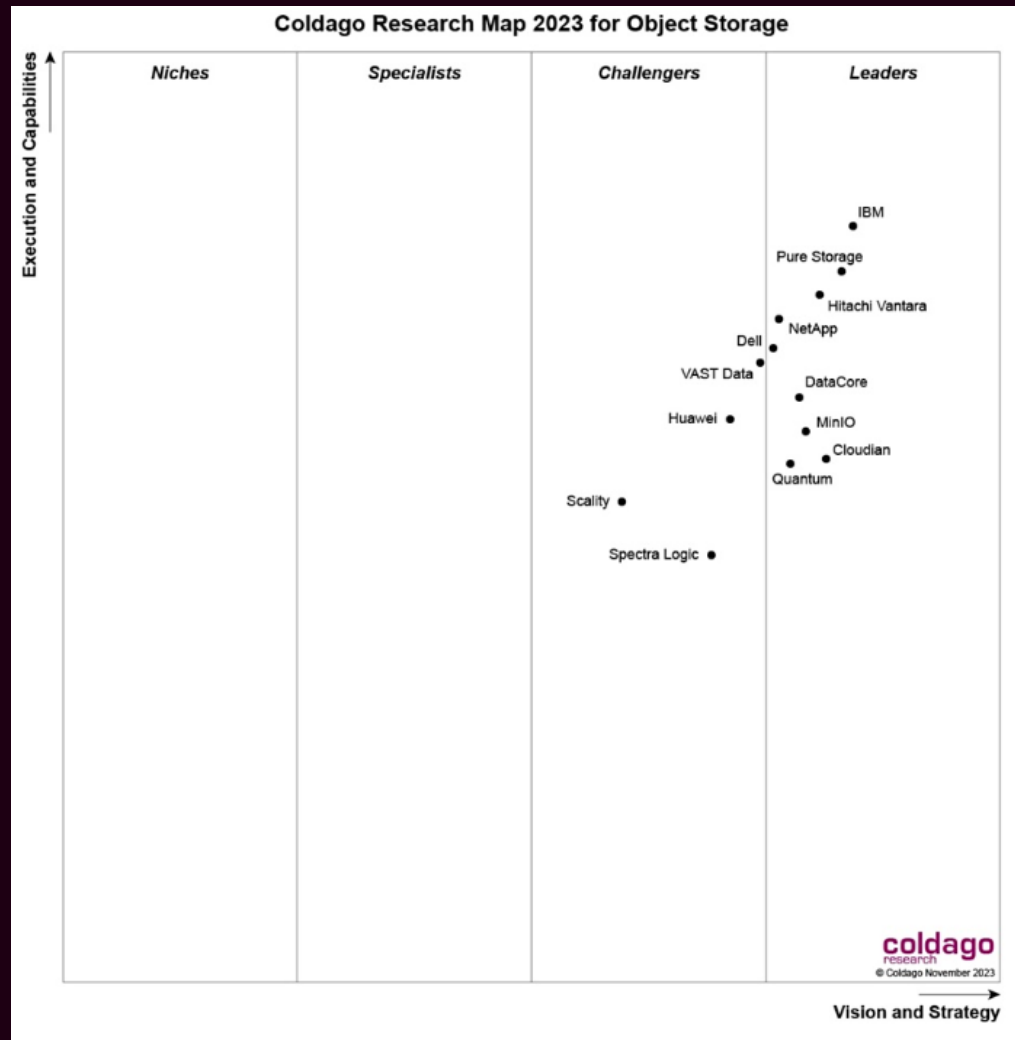
- Kubernetes CSI-drivers, supports CephFS and Ceph RBD CSI-drivers to provide persistent storage for containers

Safely Store Backups and Protect Data from Ransomware



IBM Ceph Storage is a Leader in Object Storage

Gartner Magic Quadrant and Coldago Research Map



Smarter
technology
for all

Lenovo

thanks.