# IBM Storage Scale Strategy

IBM Storage Scale User Group 2023 London, UK – June 27-28, 2023

Ted Hoover Product Manager, Storage for Data and AI

Wayne Sawdon CTO, IBM Storage Scale & Scale System



## Disclaimer



IBM's statements regarding its plans, directions, and intent are subject to change or withdrawal without notice at IBM's sole discretion. Information regarding potential future products is intended to outline our general product direction and it should not be relied on in making a purchasing decision. The information mentioned regarding potential future products is not a commitment, promise, or legal obligation to deliver any material, code, or functionality. The development, release, and timing of any future features or functionality described for our products remains at our sole discretion.

IBM reserves the right to change product specifications and offerings at any time without notice. This publication could include technical inaccuracies or typographical errors. References herein to IBM products and services do not imply that IBM intends to make them available in all countries.

	IBM Storage				
	Storage for Data and AI	Storage for Hybrid Cloud	Storage for Data Resilience		
Software	IBM Storage Scale IBM Storage Ceph IBM Cloud Object Storage	IBM Storage Fusion	IBM Storage Defender		
Hardware	IBM Storage Scale System IBM Cloud Object Storage	IBM Storage Fusion HCI System	IBM Storage FlashSystem   DS8000   Tape   Networking		

## A Global Data Platform for Unstructured Data

Unified File and Object Services coming Together to Provide an Industry Leading Data Services Offering



# Highlights of 2022 releases (1Q22-4Q22)

#### **Access Services**

#### Modernize protocol stack

- High Performance Object:
  - Containerized S3
  - Targeting AI/Analytics
  - Initial support ESS storage

#### AI / ML/ GPU acceleration

- Improved Pre fetch performance ~2X
- NVIDIA GDS over RoCE
- GDS Write acceleration
- GDS Perf enhancements

#### **Containerized Environments**

- OpenShift integration for upgrade resiliency
- FSgroup support for RWO
- VMWare 7 support for CNSA

#### **Caching Services**

#### **Data Caching**

- AFM policy-based tiering to object storage - AWS, Azure, Google
- AFM DR 60 min RPO
- AFM Containerization
- AFM to S3 download and upload performance
- AFM Recovery
  improvements for rename
  and remove operations

#### Core Services

- Multi-Rail Over TCP
- Improved Prefetch
  performance ~2x
- Fileset scaling
- IOPS and io500 improvements

#### **Management Services**

#### Visibility, Control and Automation

- Remote Fileset Access Control
- Independent filesets scalability
- Scale GUI HA, GUI HA in ESS Environment – Stretch Cluster (Active/Active)
- Online mmfsck
- RHEL 9 support

## Monitoring, Availability & Proactive Services

- Enhanced stretch cluster monitoring
- Monitor AFM memory queue alerts in mmhealth
- Detection of
  hung/unresponsive nodes

#### **Security Services**

#### Security

- Safeguarded Copy: Prevents modification or deletion of copies due to user error, malicious destruction, or ransomware attack
- Encryption -Using Hashicorp/ KMIP

#### Resiliency

Erasure Code Edition Enhancements:

- High density server (60 disks)
- Small ECE 3 nodes
- Trim Automated reclamation



Search less. Discover more.

## **IBM Storage Fusion Data Cataloging**

## Unified metadata management and insights for heterogeneous file and object storage, on-premises and in the cloud.

## Discover

Automatically ingest and index system metadata from multiple file and object storage systems, on-premises and in the cloud

## Classify

Automatically identify and classify data, including sensitive and personal identifiable information

## Label

Enrich data with system and custom metadata tags that increase the value of that data

## Find

Find data quickly and easily by searching catalogs of system and custom metadata

IBM Storage for Data and Al / © 2023 IBM Corporation

# IBM Storage Scale – Global Data Platform for Open Data Lakehouse



Capacity Storage (Object, File, HDFS)

A data lakehouse is a hybrid of data lakes and data warehouses combining the capacity and flexibility of a data lake with the governance and transactional update of a data warehouse on structured, semistructured and unstructured data.

IBM Storage Scale provides high performance storage for AI and Analytics including high performance shared scratch space for data shuffle.

Data may reside within Scale or be virtualized into Scale from any cloud, from any edge or from any legacy data silos, whether object, file or HDFS format

Data may be orchestrated to minimize the time to results.

# Put AI to work with watsonx

Scale and accelerate the impact of AI with trusted data.

Leverage foundation models to automate data search, discovery, and linking in watsonx.data watsonx.ai watsonx.data watsonx.governance Scale AI workloads, for Train, validate, tune Enable responsible, transparent and all your data, anywhere explainable data and AI workflows and deploy AI models Leverage governed enterprise data in watsonx.data to seamlessly train or fine-tune foundation models

Enable fine-tuned models to be managed through market leading governance and lifecycle management capabilities

# Storage Scale CloudKit



## Advantages

- Cloud agnostic (initial support for AWS)
- Easy to use, guided interface
- Rapid deployment and configuration of Storage Scale on the Cloud (in minutes)
- Flexible deployment (supports multiple deployment models)
- Ability to provision multiple clusters from a single CloudKit instance

## What is Storage Scale Cloudkit?

- Command Line Interface tool to create Storage Scale clusters on the cloud
- Provides end to end automation to create and bring up a Storage Scale cluster on public clouds
  - Automates infrastructure provisioning on the cloud
  - Automates the deployment of Storage Scale on the cloud
  - Applies Storage Scale best practises for deploying on the cloud

#### **CloudKit can Provision**

- Single cluster with Storage Scale (Storage) NSD servers and Compute nodes
- Separate Storage Scale Storage Cluster (with all NSD servers) and Storage Scale Compute Cluster
- Separate Storage Scale Storage Cluster (with all NSD servers)

# **Modernization of Scale**

# Security Improvements

Removal of SSH dependency

Removal of root requirement for control plane



Remote Administration

Fine-Grained Role Based Access Control Declarative policy rules based on Open Policy Agent

# Control Plane Designed For Applications / Operators

Retain CLI for human management





MCOT / MROT

- With this feature, no bonding/teaming or dynamic routing protocols are needed for Spectrum Scale to use multiple physical interfaces concurrently when communicating with a single destination

Two issues which have been seen with bonding, which MROT does a great job of addressing:

- Bouncing/flapping link detection. 1. Bonding is good at removing a link from the configuration if it fails completely but can have issues with intermittent packet loss.
- 2. Achieving even/equal use of all available interfaces. The hash used to map physical interfaces to virtual interfaces may sometimes not be well balanced.



Multiple connection over TCP, introduced with 5.1.1

Multi-Rail over TCP with 5.1.5

8882B	458B:	66B	4282B:	100k	7954B:	8502B	262B	i			
66B	114B:	130B	66B:	4478B	8438B:	66B	186B^C				
[root@ess5kio1 ~]# dstat -n -N enp1s0f1,enp1s0f0,enP51p1s											
net/enP51p1-net/enP51p1-net/enp1s0f-net/enp1s0f											
recv	send:	recv	send:	recv	send:	recv	send				
1639k	4804M:	438k	2244M:	2012k	7122M:	559k	1513M				
<b>11</b> 88k	4401M:	408k	2083M:	2181k	6980M:	552k	1801M				
1651k	5585M:	569k	2557M:	2107k	7033M:	634k	1762M				
1349k	4162M:	566k	2740M:	2404k	7756M:	537k	2257M				
1090k	3513M:	593k	2174M:	1687k	7469M:	565k	1913M				
1090k	3396M:	376k	1729M:	1586k	7027M:	639k	1972M				
618k	3362M:	240k	1202M:	877k	5539M:	317k	2064M				
1004k	3213M:	373k	1054M:	1006k	5089M:	361k	1739M				
806k	3111M:	127k	1037M:	896k	5070M:	215k	1432M				
500k	1908M:	165k	1131M:	662k	3424M:	151k	829M				
67k	328M:	248k	1283M:	137k	1427M:	76k	658M				
63k	206M:	172k	1186M:	199k	1166M:	62k	477M				
53k	190M:	113k	693M:	200k	1132M:	80k	709M				
35k	236M:	77k	526M:	179k	1180M:	138k	1169M				
<b>11</b> k	102M:	301k	1312M:	23k	374M:	269k	1316M				
1268B	14k:	73k	298M:	34k	13k:	40k	288M				
60B		60B		60B		60B	0 ^C				

[root@ess5kio1 ~]# 🗌

# IBM Storage Scale System (IBM Elastic Storage System 3500)

The simplest and fastest way to deploy a global data platform for AI and Hybrid Cloud workloads

#### Dual redundant side-by-side control canisters in 2U

Dual path to all drives Reed-Solomon 8+2P or 8+3P encoding Redundant Power Supplies, Fans, Network All FRUs are Hot Swap with Status LEDs

12 or 24 High performance PCI Gen4 NVMe drives 3.84 TB, 7.68 TB,15.36 TB or 30 TB drives 46 TB to 720 TB raw capacity

12 or 24 FCM 3.0 NVMe 38 TB drives Inline compression up to 116 TB each drive 24 x 116 TB = 2784 TB raw capacity in 2U

Four x16 PCIe Gen4 PCI adapter slots per canister 2x IB NDR / RoCE per canister

Up to 8x 4U x102 SAS HDD enclosures

Support for Self Encrypting Drives

Scale-up to PB and scale-out to YB for GB/s+ performance and capacity to manage your entire data ecosystem with lower cost and enterprise security and resiliency your business requires



Scale from 1 to 1000s of nodes up to 91 GB/s per node up to 30M IOPS per rack 48 TB to 921 TB raw Flash capacity per node 510 TB to 16.3 PB raw HDD capacity per node



# IBM Storage Scale System (ESS 3500)

Flash

Capacity

HDDs<sup>1</sup>

Launched in 2022

48TB-920TB Flash Grow to 16PB of (2.8PB HDD Capacity/2U Compressed) BB<sup>1</sup> 4U102 SAS JBOD Up to 125 GB/s per 2U More than 16M IOPs / 2U Up to 8 JBODs

Scale-up to PB and scale-out to YB for GB/s+ performance and capacity to manage your entire data ecosystem with lower cost and enterprise security and resiliency your business requires

"I don't have a full-time person who looks after Spectrum Scale on my team...For the most part, it looks after itself."

- IT Manager, Univ. of Birmingham

Up to 125 GB/s per 2U BB 48TB to 920TB flash capacity per 2U BB 48TB to18PB total capacity per 2U BB More than 16M IOPs per 2U BB Scale 1 to 1000s of BBs Global Data Platform Built-in policy optimization engine Enterprise resiliency and security Container-native OpenShift access

# IBM ESS 3500 SED Support Self-encrypting drive



ESS/GNR integrated key management (GKLM) MEK (Master Encryption Key)

- AES (256 bit) secure key
- FIPS 140-2 issued by NIST

#### DEK (Data Encryption Key)

- 1. Internal to drive.
- 2. No external access.
- Maintained and generated by drive AES (256 bit) secure key FIPS 140-2 issued by NIST

## SED Supported Drives

- 3.84 TB, 7.68 TB, 15.36 TB, 30.74 TB NVMe drives
- 20 TB SAS HDD
- 10TB SAS HDD (2023)

SED Requirements

- ESS 6.1.5+
- GKLM 4.1.0.1+
- ISKLM V2.6+
- FIPS 140-2 AES

## **Management and Security**

- Leverages ESS commands to enable/manage SED
- Automatic SED enabled of replacement drives
- Automatic Crypto / Secure erase of drives
- No performance degradation

# IBM ESS 3500 FCM Support



ESS/GNR compression support using FCM drives

Flash Core Module (FCM)

- 2.5" U.2 high-performance flash drives (38.4TB)
- Hardware self-encryption
- Real-time inline hardware data compression (3:1)
- Leverages NVMe protocol (Gen 4 PCIe)
- Industry leading QLC endurance (15K program/erase cycles)

#### **ESS** Compression Features

- Disk hospital integration
- RAS / health monitoring
- Call home integration
- Automatic Trim support

## ESS Compression Requirements

- New ESS deployments
- ESS 6.1.6+ (1Q23)
- 12 or 24 FCM 3.1 38.4 TB drives

## Future Enhancements

- Compression migration support
- Additional FCM capacities

## **Key Highlights**

- Support for up to 3:1 data compression
- Up to 2.7PB of compressed flash capacity
- High throughput and IOPs
- Cost-competitive
- Transparent to user/application

# IBM FlashCore<sup>™</sup> Module 3 Capacity and Performance

2.5" dual ported U.2 NVMe Gen 4 PCle Industry leading density at 38.4 TB per drive Inline hardware FIPS 140-3 encryption Inline hardware 3:1 compression = 116 TB!

Internally tiered storage -> MRAM -> SLC -> 3D QLC Performance comparable to TLC

Industry leading QLC endurance 15K Program/Erase cycles Compared to 1500 for enterprise QLC

IBM Unique QLC management (100+ patents) read calibration, heat binning, health binning, error correcting codes, optimized voltage

Continuous health monitoring keeps wear across all cells within 5%





# Integrated NVMe-OF Extreme performance Tier



Elastic Storage System 3500

## Measured over 16 M IOPs and 110 GB/s

#### Use Case

Data analytics (AI/ML) needing very high rand IOPS with high throughput

High performance Scratch / Shuffle space

## **System Config**

3.84 TB, 7.68 TB, 15.36 TB or 30.74 TB

4x CX6-VPI Adapters / canister

#### **Performance and Features**

- Integrated extreme high IOPs storage Pool
- Dedicated performance pool (12x drives)
- Easy configuration and setup
- Automatic data migration between pools
- Integrated RAS support

# **ESS Fabric Hospital**

The ESS Fabric Hospital is designed to identify and isolate problems that are impacting I/O availability.



- Collects new metric gpfs\_fabhospital\_errorIOCount per Canister/Adapter/Port/Enclosure/Slot
- Leverages *mmhealth custom thresholds* to group I/O error information based on hardware topology and allow users to configure warning/error threshold levels.
  - SASPortErrorThreshold → all enclosures/slots attached to this port see errors
  - SASEnclosureErrorThreshold  $\rightarrow$  all slots in an enclosure see errors

ESS Fabric Hospital is only supported on the ESS 3500 platform and later models and is currently limited to SAS-based error monitoring.

# IBM **TechXchange** Conference 2023

A technical learning event designed with technologists for technologists who use IBM products and solutions.

September 11–14, 2023 MGM Grand, Las Vegas





