High Performance Object Storage for HPC, AI and Data Analytics

Presenter:
Madhu Thorat, IBM

1. The world is changing …

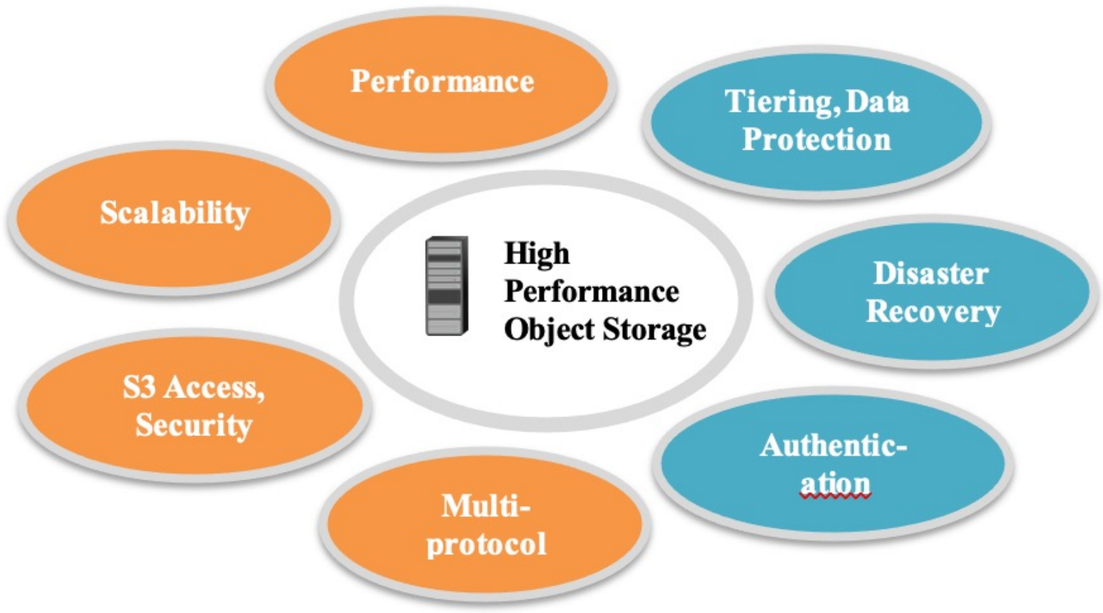
2. Need of High Performance Object Storage

More than 80% of enterprise unstructured data will be stored in scale-out file systems and object storage systems in enterprise and cloud data centers, an increase from 30% today.

2.1. Major Benefits of High Performance Object Storage

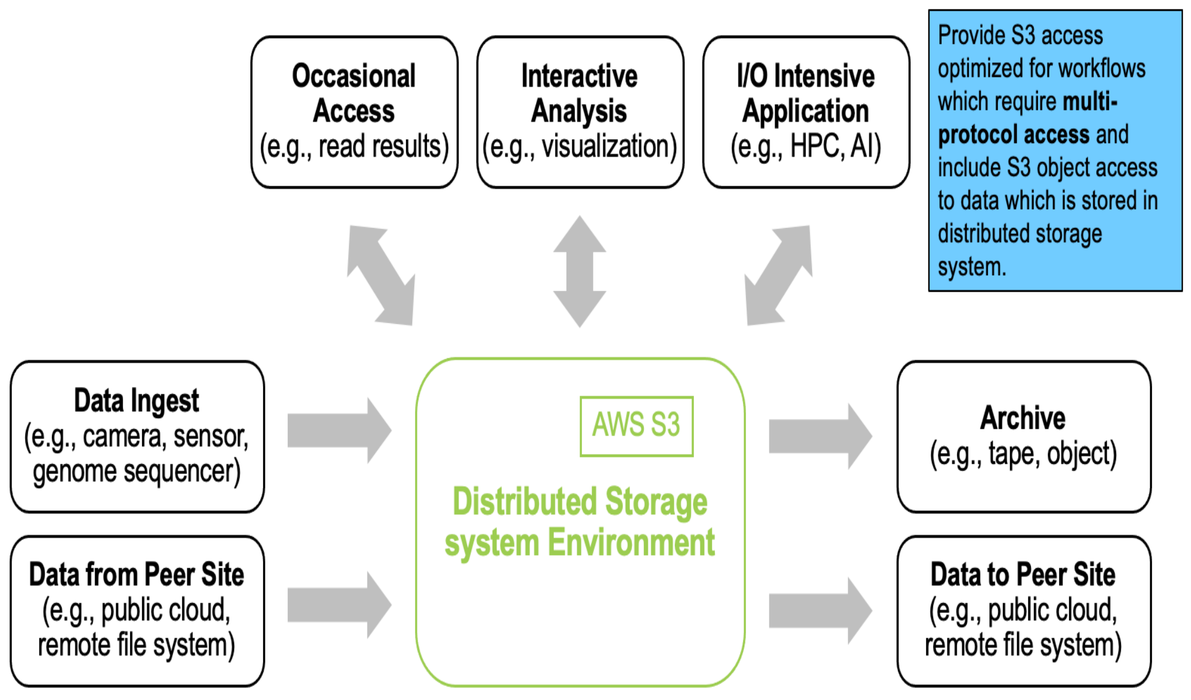
Suitable for Unstructured Data
Meets the demands for high speed, low latency, and scale from storage
Cost-effective as it provides low-cost storage for large capacity of data
AWS S3 protocol for object storage access is now de-facto standard for object access, and used by many vendors

3. High Performance Object Storage Features for HPC, AI and Analytics

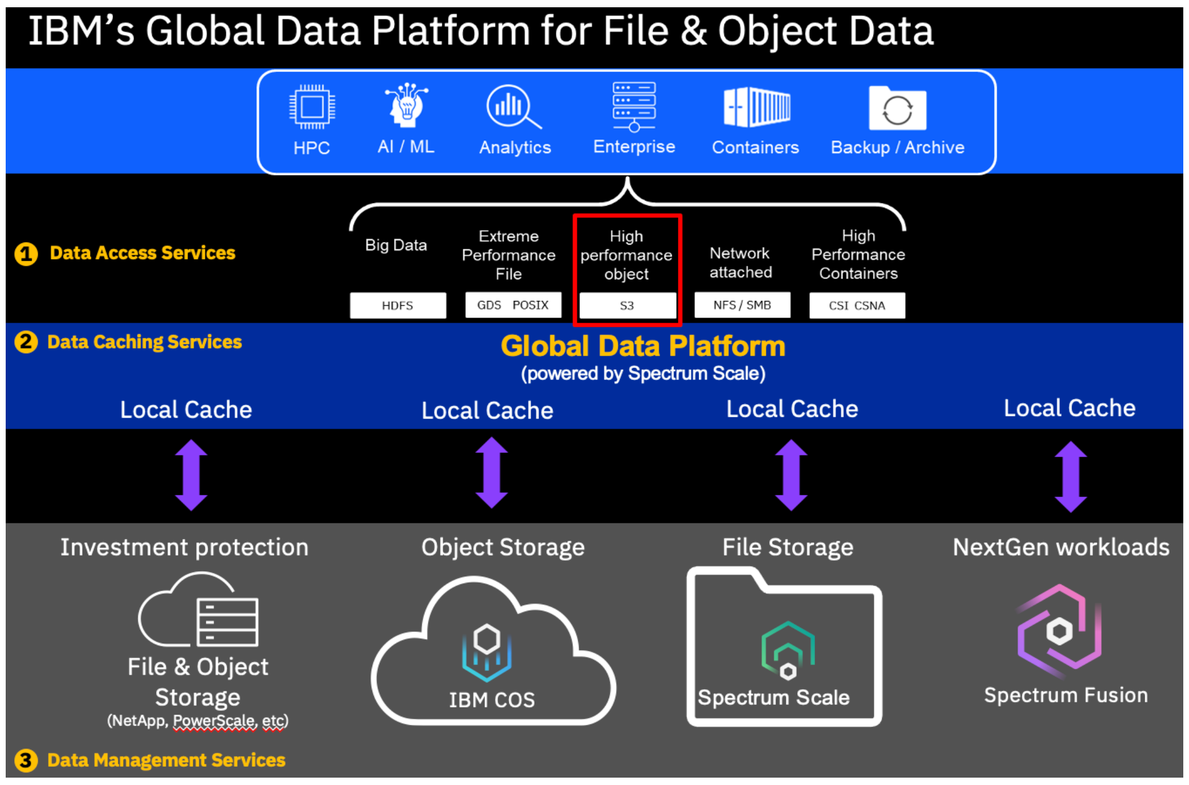


Performance	Provide faster object storage access
Scalability	Scalable storage supported by scalable compute nodes
AWS S3 Access	Makes it suitable for cloud-native applications
Mutli-protocol Access	Unified data access with multiple protocols like S3, NFS, Samba etc
Tiering, Data Protection	Efficient storage usage by assigning data to various categories of storage and safeguarding data.

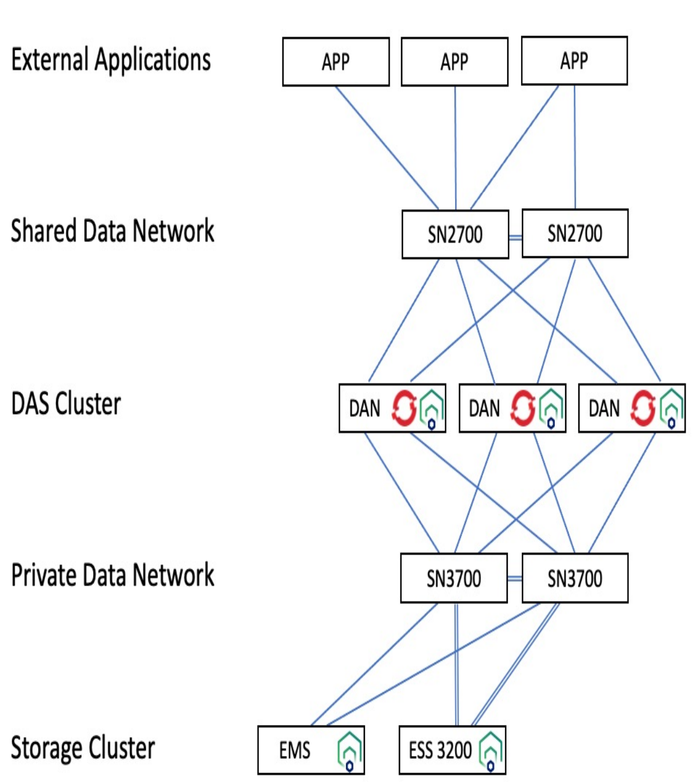
4. Use Case: Data Scientists accessing AI application data at high speed with High Performance Object Storage



5. IBM Spectrum Scale DAS (Data Access Services) S3 in IBM Global Data Platform



6. IBM Spectrum Scale DAS S3 cluster – Reference Architecture and Performance Numbers



61 GB/s read for 1GB objects
23 GB/s write for 1GB objects
Performance was limited by single ESS 3200 using 24 NVMe drives

Performance is measured for DAS Cluster (aggregated throughput)

Blog post with details:
https://community.ibm.com/community/user/storage/blogs/silvana-de-gyves-avila1/2022/05/20/ibm-data-access-services-performance-evaluation

References:

1. IBM Spectrum Scale DAS: https://www.ibm.com/docs/en/scalecontainernative?topic=516-spectrum-scale-data-access-services
2. Red hat OpenShift Data Foundation: https://www.redhat.com/en/technologies/cloud-computing/openshift-data-foundation