IBM Spectrum Scale
– Use cases –

—

Tomer Perry, Ulf Troppens

SC19 Denver, CO hpc is now.

IBM

Outline

1. **What is IBM Spectrum Scale?**
    a. **Evolution**
    b. Key concepts
2. Primary Use Cases
    a. High performance computing (HPC)
    b. Data intensive application & workflows
    c. AI/ML/DL
3. Summary

The world is changing …

IBM Spectrum Scale

Luca Bruno/AP

2005

The world is changing ...

IBM Spectrum Scale

2005

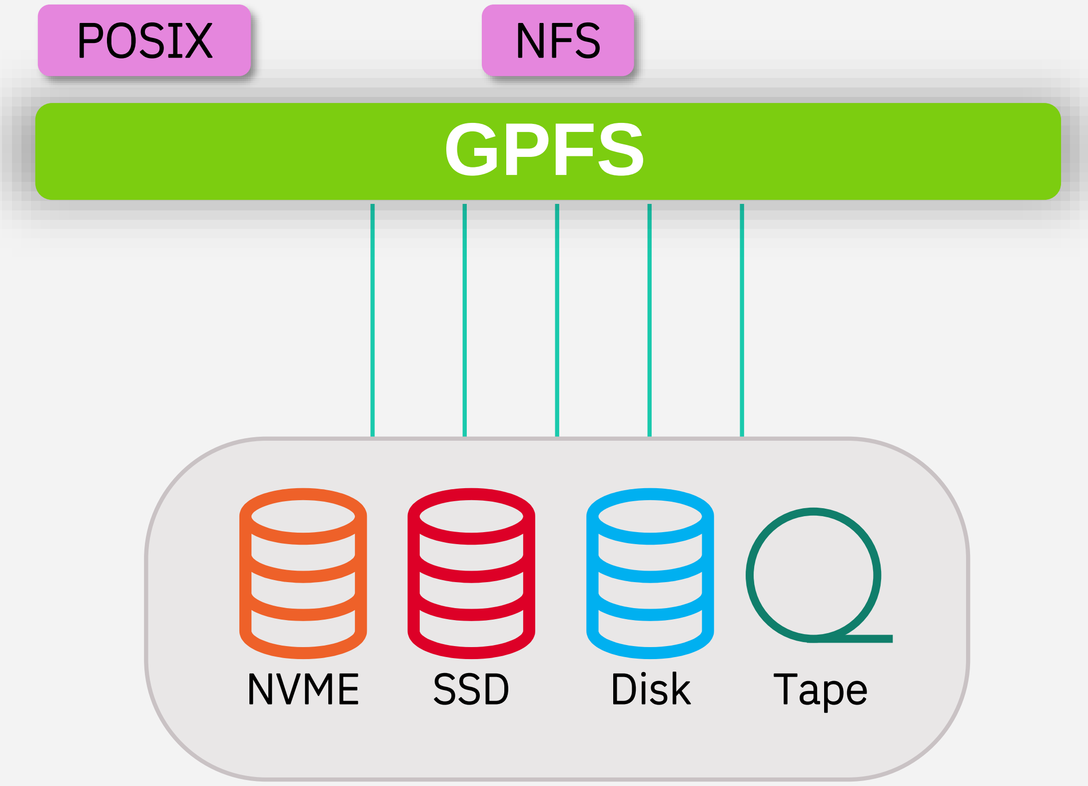
Michael Sohn/AP

Luca Bruno/AP

2013

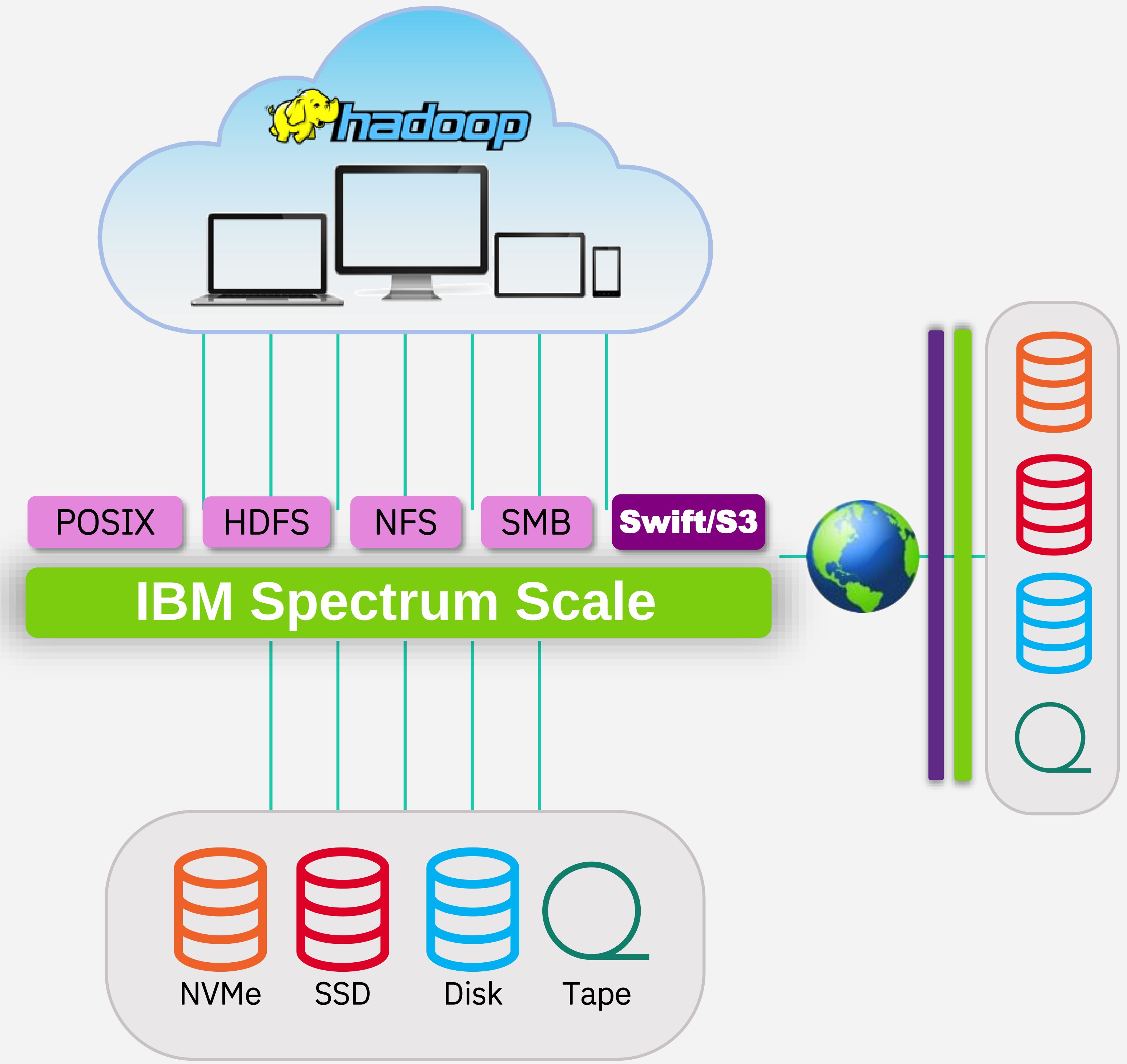
GPFS is changing …

- 1993: Started as “Tiger Shark” research project at IBM Research Almaden as high performance filesystem for accessing and processing multimedia data
- Next 20 years: Grew up as General Parallel File System (GPFS) to power the world’s largest supercomputers
- Since 2014: Transforming to IBM Spectrum Scale to support new workloads which need to process huge amounts of unstructured data



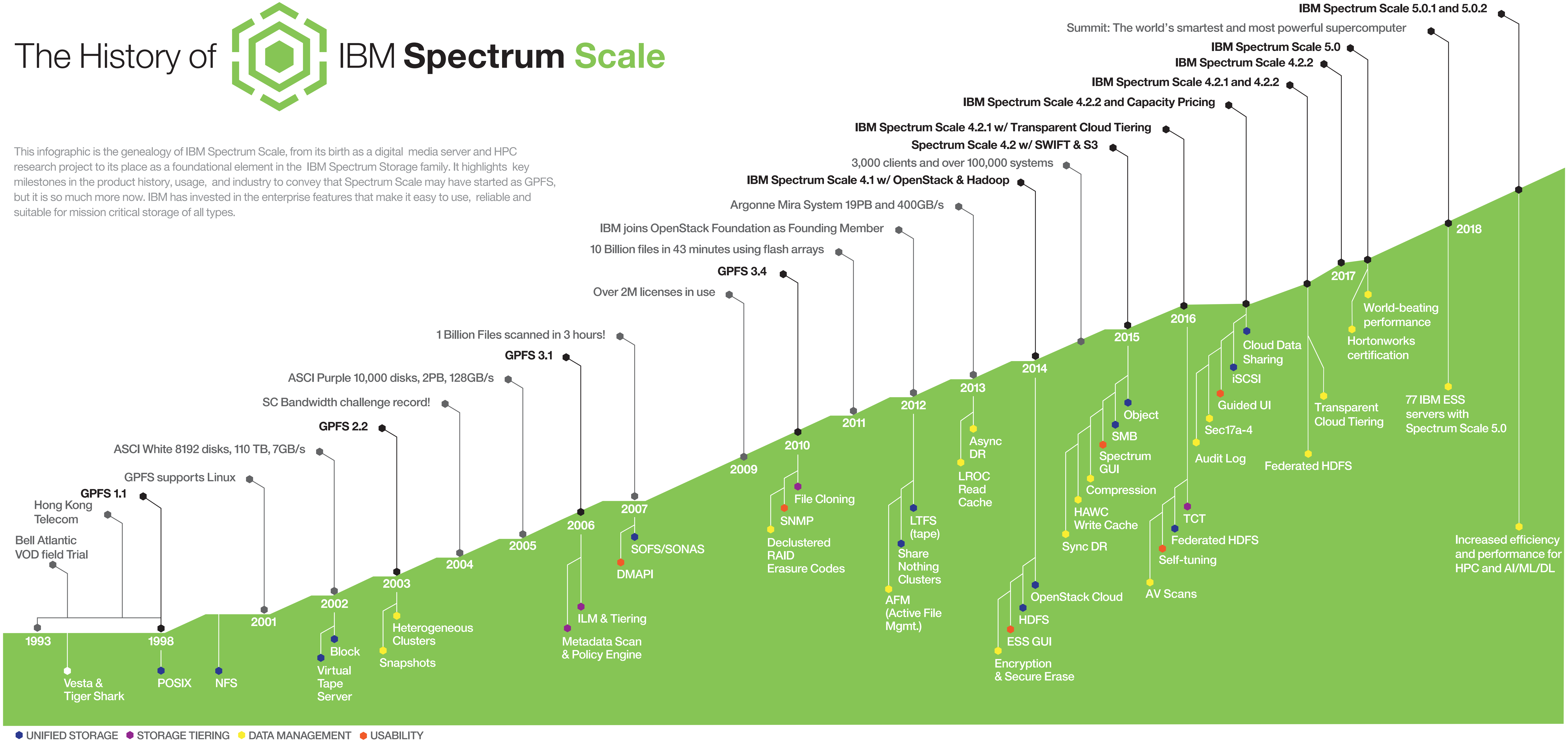
IBM Spectrum Scale

- Based on GPFS, a robust, fast and mature parallel file system
- BUT: If you still just think GPFS, you miss:
  - Support for workflows which for example inject data via object, analyze results via Hadoop/Spark and view results via POSIX
  - Storing and accessing large and small objects (S3 and Swift) with low latency
  - Automatic destaging of cold data to on premise or off premise object storage
  - Exchange of data between Spectrum Scale clusters via object storage in the cloud
  - Storing and starting OpenStack VMs without copying them from object storage to local file system
  - GUI , REST API, Grafana Bridge
  - And many, many more



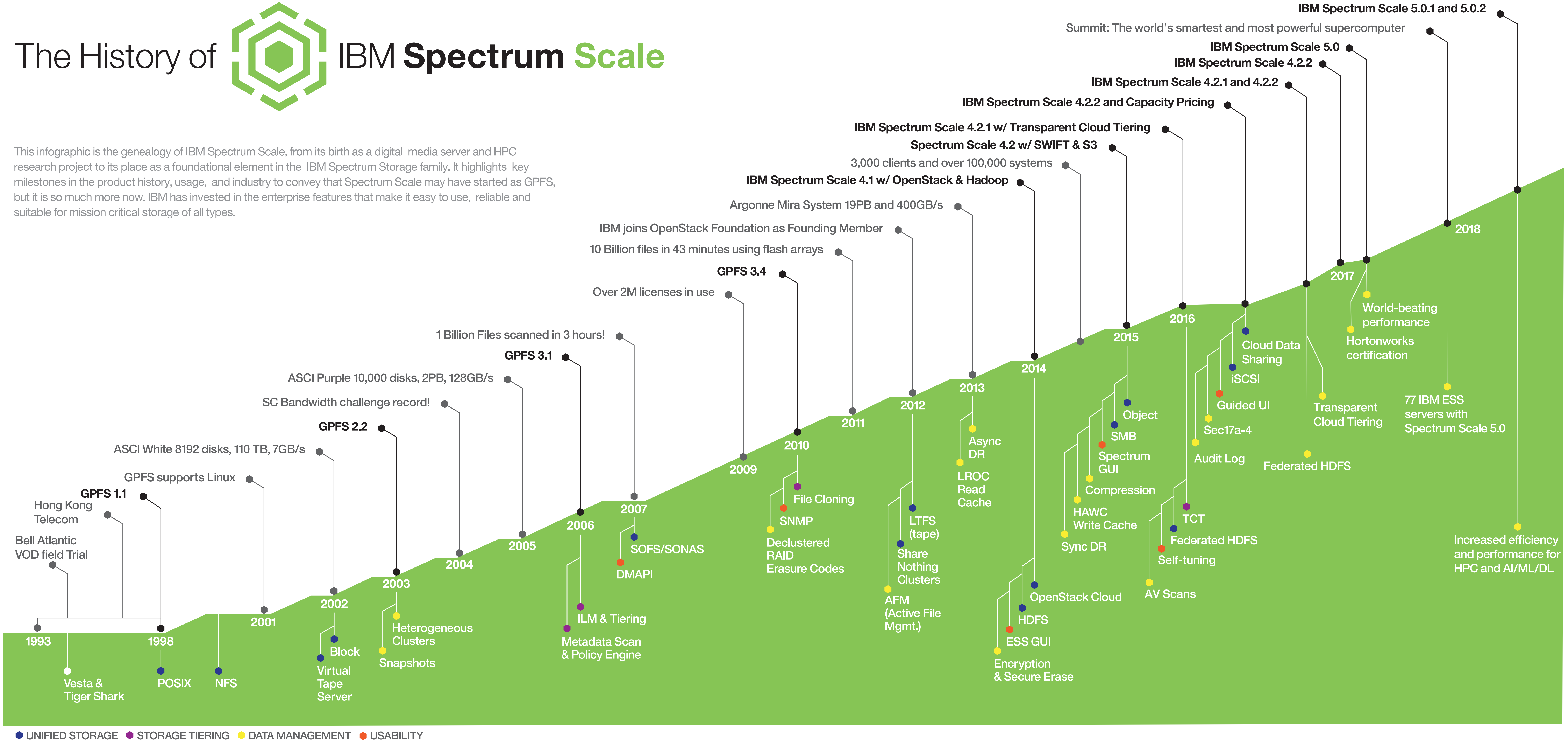
The History of IBM Spectrum Scale

This infographic is the genealogy of IBM Spectrum Scale, from its birth as a digital media server and HPC research project to its place as a foundational element in the IBM Spectrum Storage family. It highlights key milestones in the product history, usage, and industry to convey that Spectrum Scale may have started as GPFS, but it is so much more now. IBM has invested in the enterprise features that make it easy to use, reliable and suitable for mission critical storage of all types.



The History of IBM Spectrum Scale

This infographic is the genealogy of IBM Spectrum Scale, from its birth as a digital media server and HPC research project to its place as a foundational element in the IBM Spectrum Storage family. It highlights key milestones in the product history, usage, and industry to convey that Spectrum Scale may have started as GPFS, but it is so much more now. IBM has invested in the enterprise features that make it easy to use, reliable and suitable for mission critical storage of all types.



High Performance Computing

Data Intensive Applications

Data Intensive Workflows

Outline

1. **What is IBM Spectrum Scale?**
    a. Evolution
    b. Key concepts
2. Primary Use Cases
    a. High performance computing (HPC)
    b. Data intensive application & workflows
    c. AI/ML/DL
3. Summary

IBM Software-Defined Storage portfolio

IBM Spectrum Scale

IBM Spectrum Storage

IBM's comprehensive set of award-winning storage software delivered across appliance, converged and cloud.

Management	Monitoring & Control: IBM Spectrum Control	Cloud-based Storage Management and Support: IBM Storage Insights	Metadata-Driven Data Insight: IBM Spectrum Discover		Container & VM APIs: IBM Spectrum Connect
Storage Services	Archive: IBM Spectrum Archive	Backup: IBM Spectrum Protect	VM Data Availability: IBM Spectrum Protect Plus		Copy Data Management: IBM Spectrum CDM
Infrastructure Software / Software Defined	Scale-Out NAS: IBM Spectrum NAS	Virtualized Block: IBM Spectrum Virtualize	Scale-Out File: IBM Spectrum Scale	Scale-Out Object: IBM Cloud Object Storage	Scale-Out Block: IBM Spectrum Accelerate
Infrastructure Platforms	High IOPS All Flash	Disk	Cloud		Servers

Spectrum Scale value proposition

IBM Spectrum Scale

Highly scalable high-performance unified storage software
for files and objects with integrated analytics

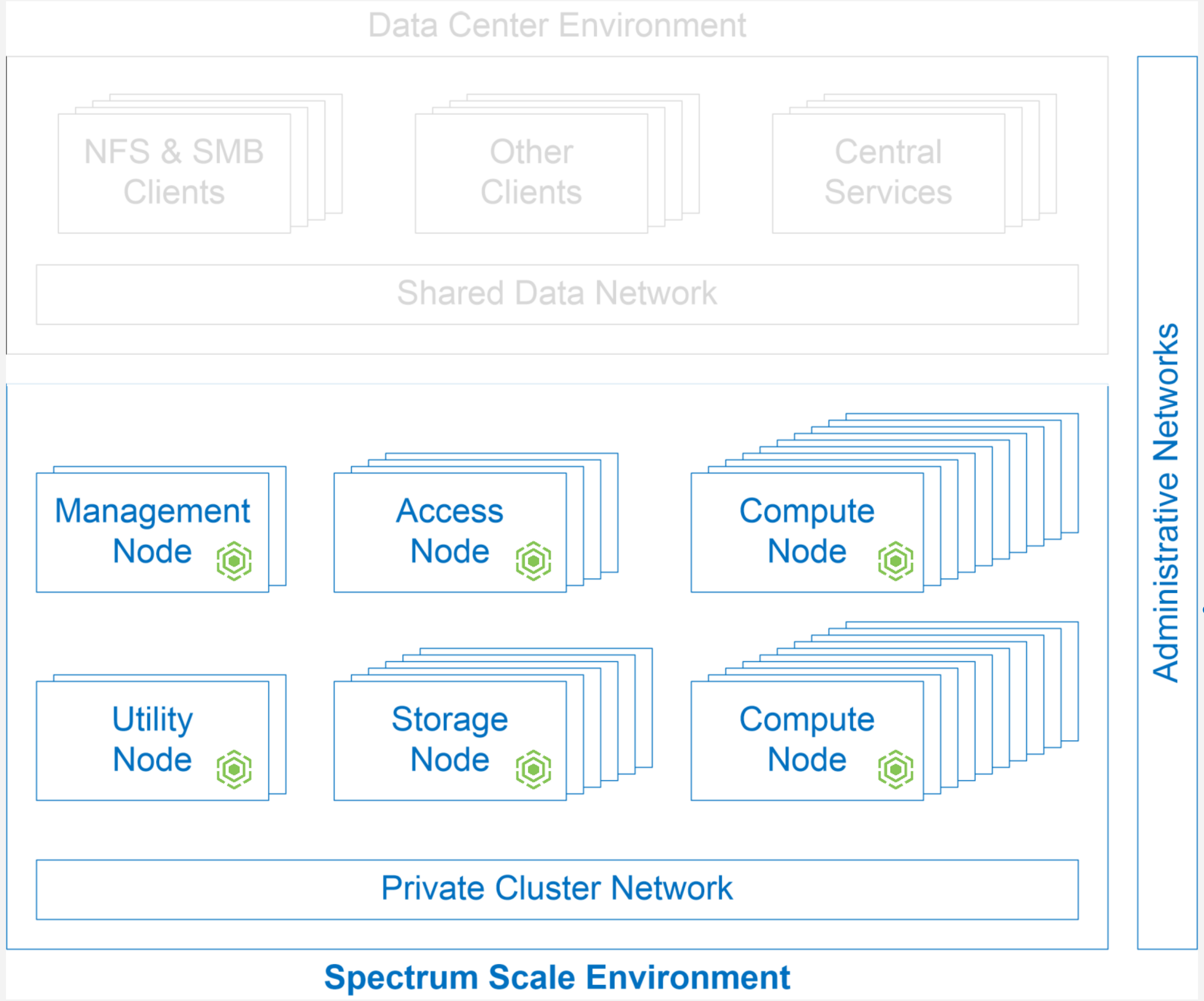
Remove data-related bottlenecks
2.5TB/s demonstrated throughput for a 250PB filesystem

Enable global collaboration
HDFS, files and object across sites

Optimize cost and performance
Automated data placement, movement and compression

Ensure data availability, integrity and security
End-to-end checksum, Spectrum Scale RAID, NIST/FIPS certification

Spectrum Scale environment



➔ The Shared Data Network provides remote access to the Spectrum Scale environment.
➔ The Private Cluster Network connects all components of the Spectrum Scale environment.

Compute Nodes (NSD Clients)

- Run applications to access and analyze data stored in one or more Spectrum Scale filesystems
- Most nodes of a Spectrum Scale environment are Compute Nodes.

Storage Nodes (NSD Server)

- Provide the storage capacity for the Spectrum Scale filesystems

Data Access Nodes (Remote & Local Access)

- Access to Spectrum Scale filesystems using protocols like NFS, SMB, HDFS and Object

Utility Nodes

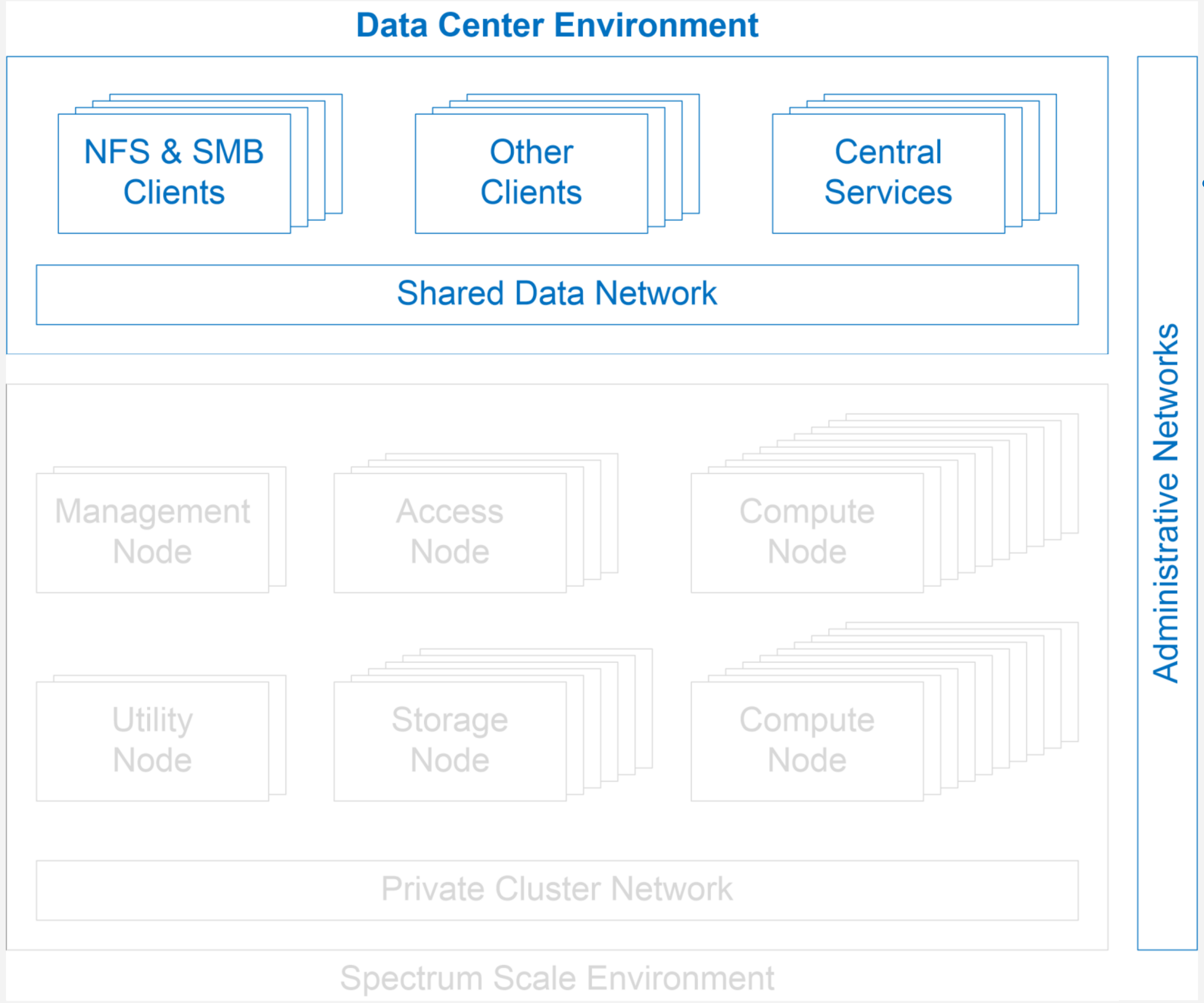
- Dedicated nodes for selected data management tasks such as backup, external tiering and hybrid cloud workflows.

Management Nodes

- Provides administration services (e.g., Spectrum Scale GUI, Performance Monitoring).

Data Center environment

IBM Spectrum Scale



➔ The Shared Data Network provides remote access to the Spectrum Scale environment.
➔ The Private Cluster Network connects all components of the Spectrum Scale environment.

NFS&SMB Clients

- Users and applications accessing data stored on a Spectrum Scale filesystem using NFS and/or SMB

Other Clients

- User and applications accessing data stored on a Spectrum Scale filesystem (e.g., Swift/S3, HDFS, Aspera, rsync, scp, etc.)
- Administrative workstations (e.g. GUI client, REST API client, SSH client, etc.)

Central Services

- External infrastructure services required for the whole solution such as
  - Authentication and ID mapping (e.g. AD, LDAP),
  - Time synchronization (e.g., NTP),
  - Name resolution (e.g., DNS), etc.

Spectrum Scale key capabilities

Scaleable performance

- Billions of files and hundreds of petabytes
- Demonstrated 2.5TB/s aggregated throughput
- Extend storage cache to compute for faster reads and writes

Automated data management

- Integration of NVMe, SSD, disk, tape and object in single filesystem
- Policy-based data placement, data movement and compression to optimize costs
- Integrated replication and scalable backup and restore for data protection
- Audit logging, immutability, encryption and checksums for compliance

Unified data access

- Proprietary NSD protocol for very high performance
- Built-in NFS, SMB, HDFS and object for application integration and end-user access
- Support for containers
- Custom access nodes for integration of 3rd-party applications such as IBM Aspera, 0MQ, scp, etc.

Flexible deployment options

- On-premise vs. cloud vs. hybrid
- Single site vs. multi site
- Reference Architectures vs. custom solutions
- IBM Elastic Storage Server vs. many other IBM or 3rd-party storage systems

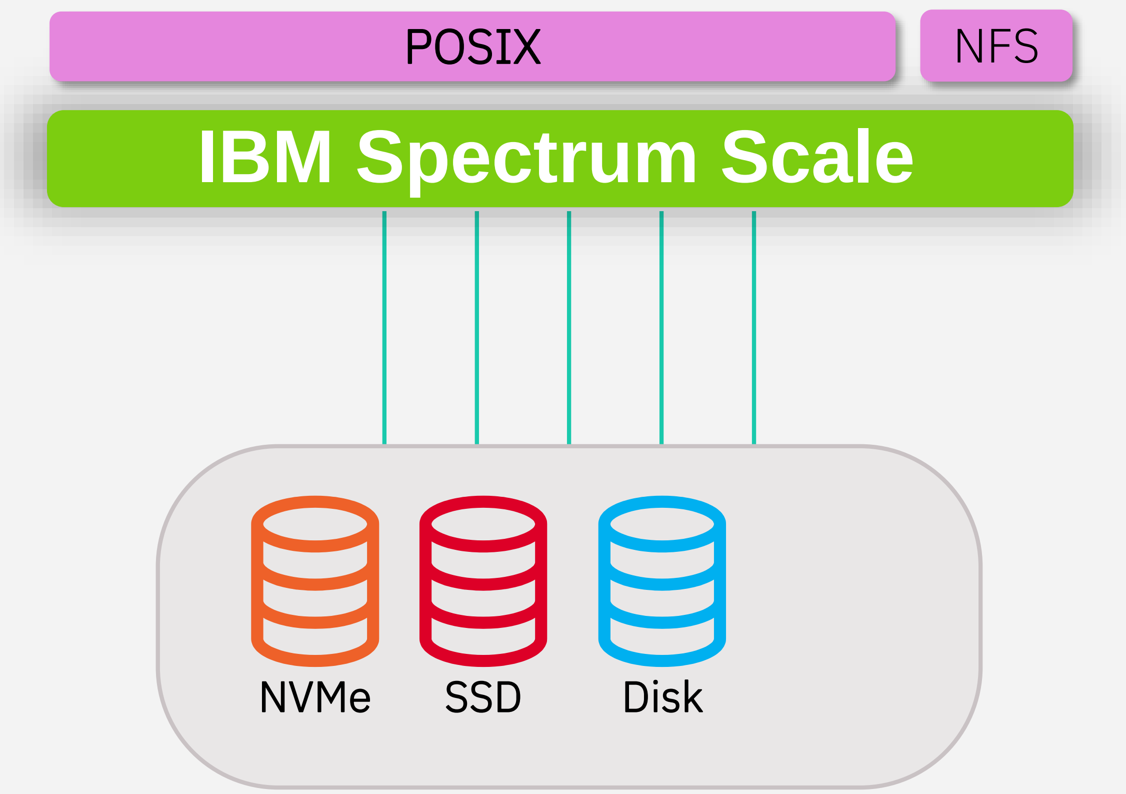
Outline

1. What is IBM Spectrum Scale?
    a. Evolution
    b. Key concepts
2. **Primary Use Cases**
    a. **High performance computing (HPC)**
    b. Data intensive application & workflows
    c. AI/ML/DL
3. Summary

High performance computing (HPC)

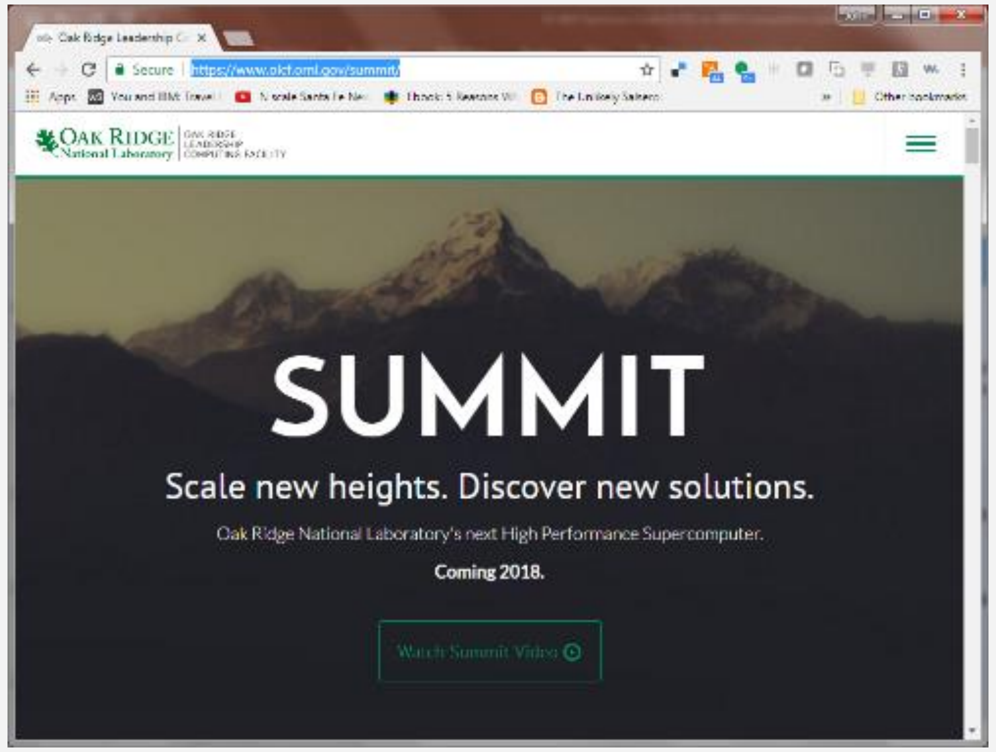
- HPC is the Big Data of the 1980s/1990s. HPC always had the problem that it requires fast access to a lot of data.
- Over the time IBM made enhancements to Spectrum Scale to keep up to date with new technology (e.g. IB EDR, RoCE, NVMe, SSD) and new workloads (e.g. small files) to keep up to date for customers computing needs.
- Nowadays Analytics/AI/ML/DL is everywhere. It is a Big Data Problem, too.
- Scaling and performance enhancements for HPC help Analytics and other use cases.
- Enhancements for other use cases help HPC, e.g., the Spectrum Scale HDFS connector enables HPC customer to spin-up and terminate Hadoop or Spark clusters on their existing super computers like any other HPC job.

- **Computer cluster (10s-1000s of nodes)**
- **NFS and other protocols to ingest data and to access results**



Performance engineering matters

IBM Spectrum Scale



https://www.olcf.ornl.gov/summit/

Imagine you need to meet these goals:

- 2.5 TB/sec single stream IOR as requested from ORNL
- 1 TB/sec 1MB sequential read/write as stated in CORAL RFP
- Single Node 16 GB/sec sequential read/write as requested from ORNL
- 50K creates/sec per shared directory as stated in CORAL RFP
- 2.6 Million 32K file creates/sec as requested from ORNL

IBM Spectrum Scale innovations have delivered these requirements

Storage for the world’s most powerful supercomputers

IBM Spectrum Scale

OAK RIDGE National Laboratory

World’s most powerful supercomputer

Summit System

- **4608 nodes**, each with:
  - 2 IBM Power9 processors
  - 6 Nvidia Tesla V100 GPUs
  - 608 GB of fast memory
  - 1.6 TB of NVMe memory
- **200 petaflops** peak performance for modeling and simulation
- **3.3 ExaOps** peak performance for data analytics and AI

IBM Spectrum Scale
IBM Elastic Storage Server
2.5 TB/sec throughput to storage architecture
250 PB HDD storage capacity

Lawrence Livermore National Laboratory

World #2 supercomputer

Sierra System

- **4320 nodes**, each with
  - 2 IBM Power9 processors
  - 4 Nvidia V100 GPUs
  - 320 GB of node memory
  - 1.6 TB of NVMe memory
- **IBM Spectrum Scale**
- **IBM Elastic Storage Server**

125 petaflops peak performance
154 PB HDD storage capacity

Outline

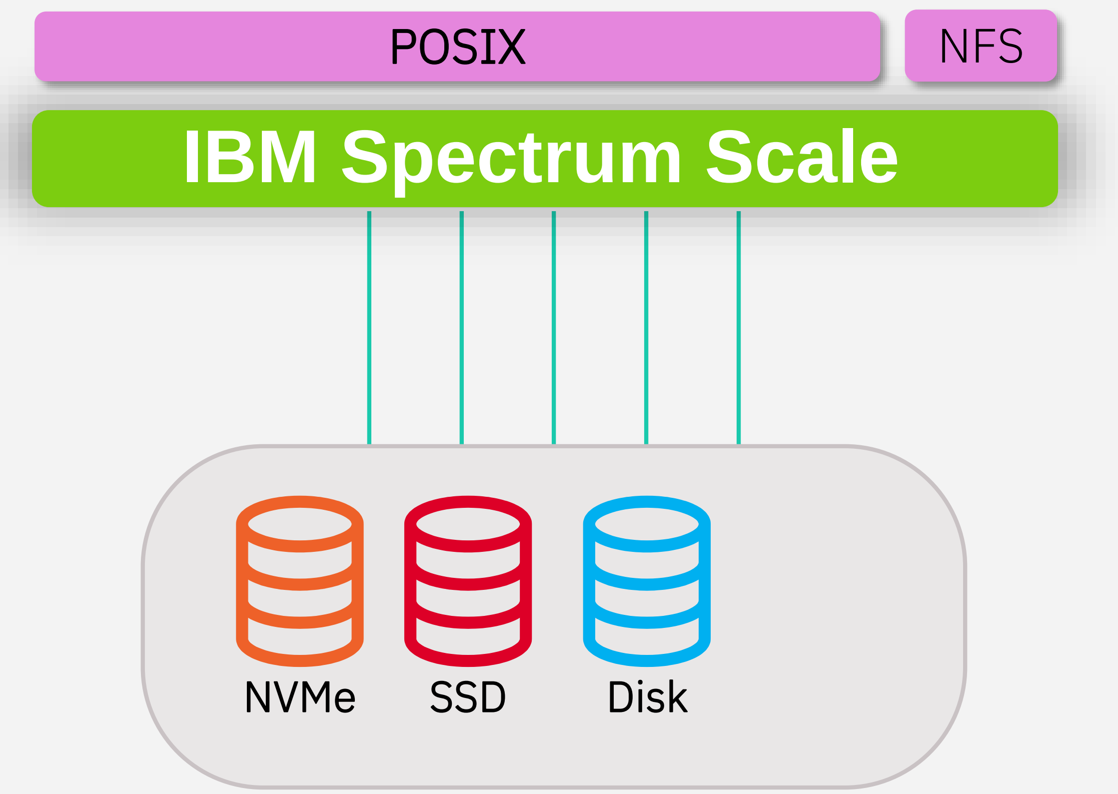
1. What is IBM Spectrum Scale?
    a. Evolution
    b. Key concepts
2. **Primary Use Cases**
    a. High performance computing (HPC)
    b. **Data intensive application & workflows**
    c. AI/ML/DL
3. Summary

Data intensive applications

IBM Spectrum Scale

- Based on GPFS, a robust, fast and mature parallel file system
- Type 1: Multiple tightly coupled instances of the same application running on multiple servers
  - Need: Fast shared filesystem for concurrent access to the same set of data
  - Examples:
    - IBM DB2
    - SAS
- Type 2: Multiple isolated or loosely coupled instances of the same application running on multiple servers
  - Need: File system virtualization layer that flexibly provisions fast file storage to each application instance
    - IBM Spectrum Protect
    - SAP HANA

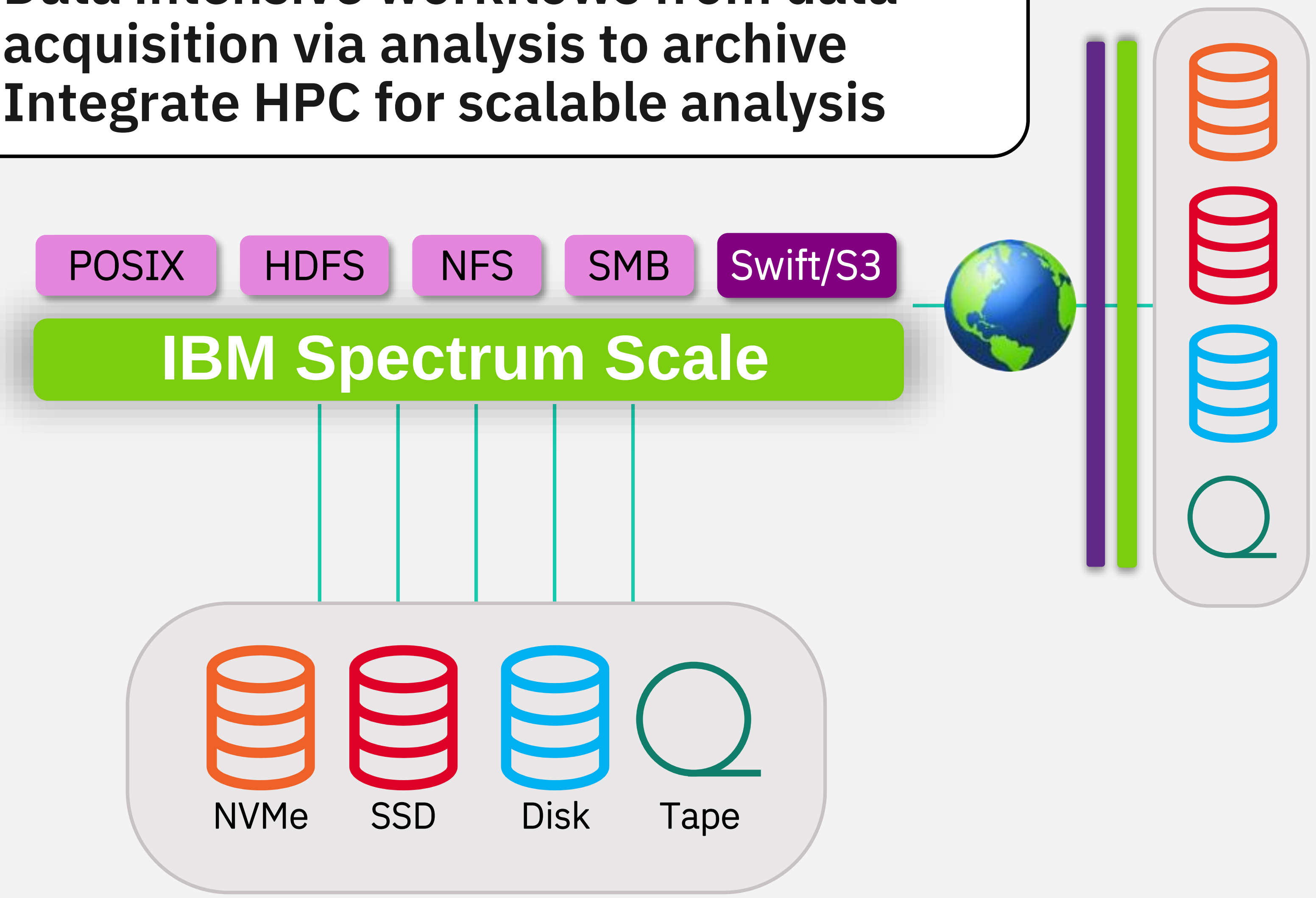
- **Application farm that benefits from filesystem with scalable performance**
- **Data access is typically via applications**



Data intensive workflows

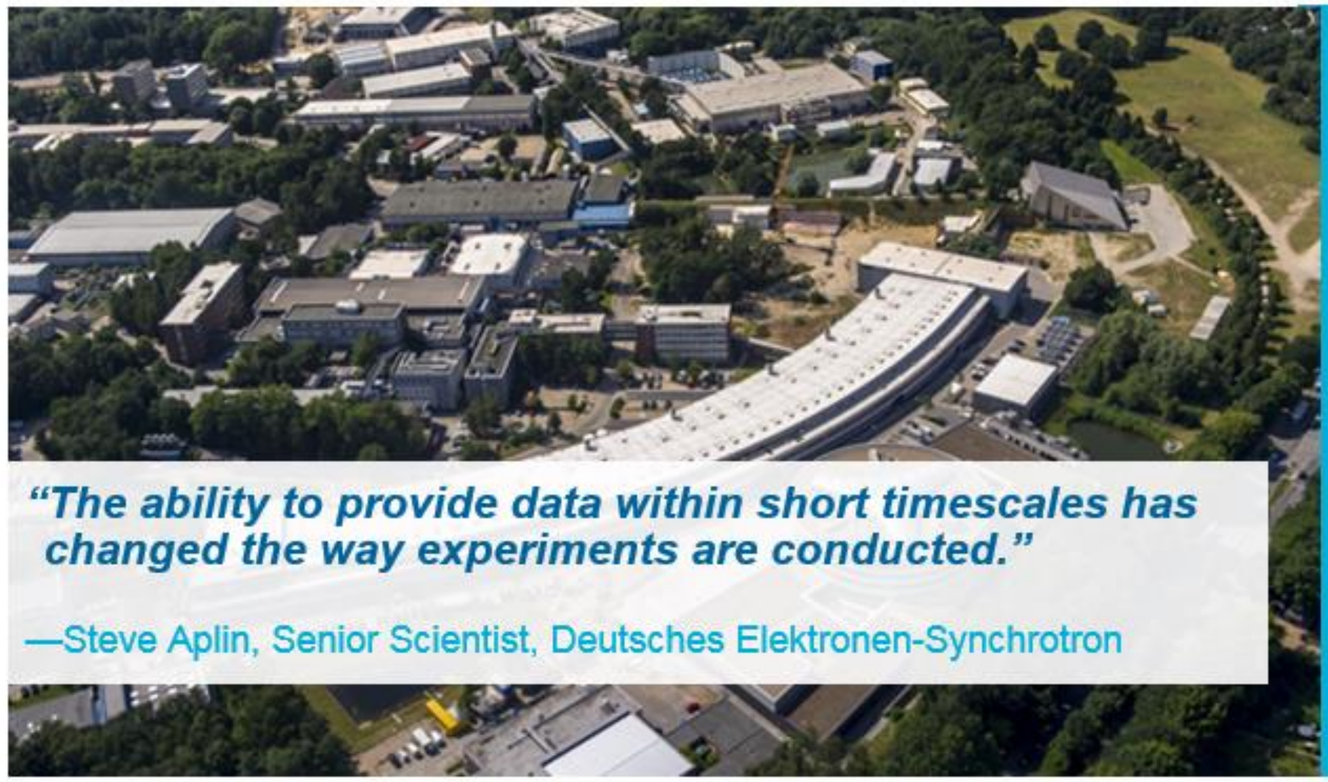
- Based on GPFS, a robust, fast and mature parallel file system
- Instruments and sensors like high-speed cameras, genome sequencers and super microscopes generate huge amounts of data that require HPC-like infrastructure to store and analyze the acquired measured data
- Spectrum Scale enables scientists to seamlessly integrate HPC-like infrastructure into their experiments and into their workflows to get timely insight in new data sets
- The built-in support for multi-protocol eliminated the need to copy data for workflows that for instance ingest data via object, clean data via HDFS, analyze via POSIX and provide results via NFS or SMB

- **Data intensive workflows from data acquisition via analysis to archive**
- **Integrate HPC for scalable analysis**



IBM

DESY



"The ability to provide data within short timescales has changed the way experiments are conducted."

—Steve Aplin, Senior Scientist, Deutsches Elektronen-Synchrotron

Business challenge

Research center Deutsches Elektronen-Synchrotron (DESY) found that increasingly resource-intensive experiments was affecting storage system performance, limiting research. How could the organization handle over five gigabytes of data streaming into its computing center every second?

Transformation

With a flexible, high-performance storage solution from IBM, DESY can meet growing demand cost-effectively. Scientists can now start analyzing the data in just a few minutes, instead of days, accelerating ground-breaking research.

In production since 2015!

https://www.youtube.com/watch?v=JLCj4jQI3q8

Business benefits:

Ensures

DESY can easily maintain a multi-PB library of research data to meet growing demand and remain an attractive research destination

Rapid

access to millions of data points accelerates research and helps lead to breakthroughs

Increases

administration efficiency with automated data management, improving DESY's service offering

DESY

Making the next breakthrough in scientific research possible with the latest in storage innovation

DESY, Deutsches Elektronen-Synchrotron, is a national research center in Germany that operates particle accelerators and photon science facilities used to investigate the structure of matter. DESY is housed in Hamburg and Zeuthen, Germany, and attracts over 3,000 scientists from over 40 countries annually.

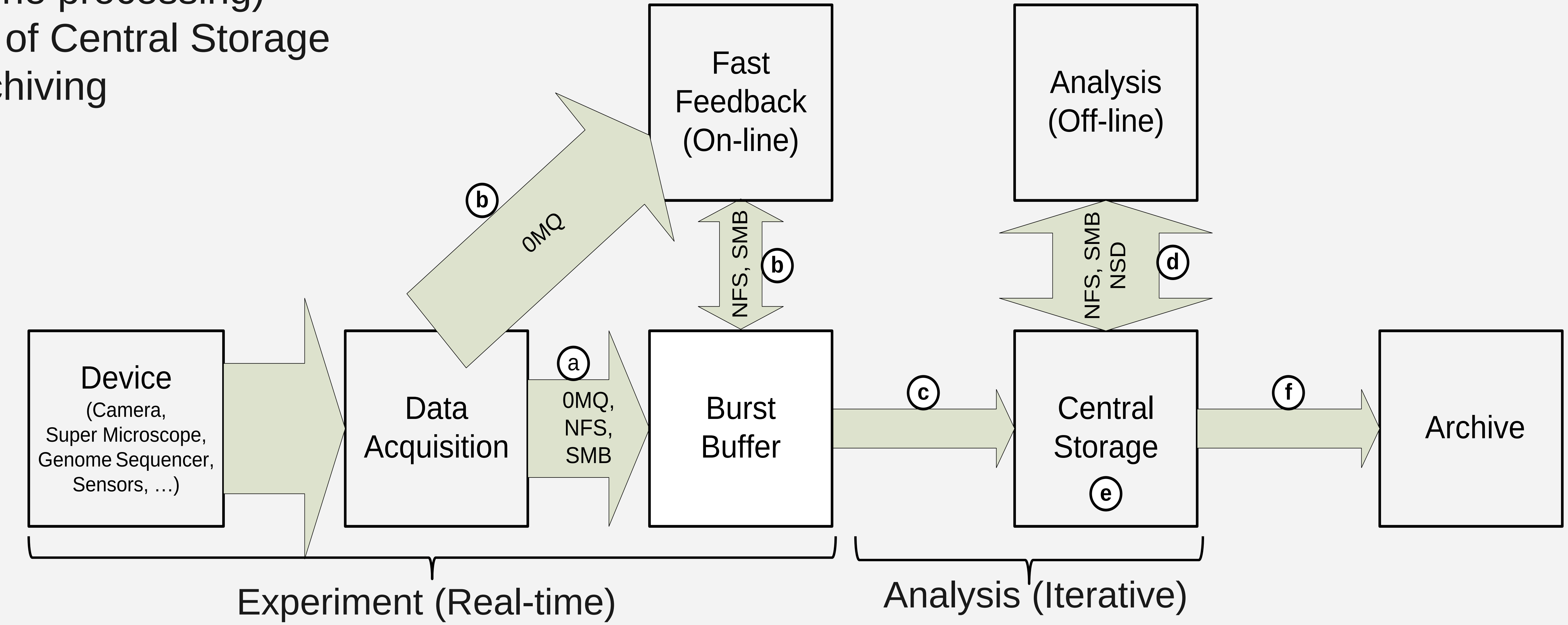
Solution components

- IBM® Spectrum Scale™
- IBM Spectrum Scale RAID
- IBM Elastic Storage™ Server GS1
- IBM Elastic Storage Server GL4 and GL6
- IBM Power® S822L
- IBM Systems Lab Services

Share this

Typical Workflow for Data Intensive Science

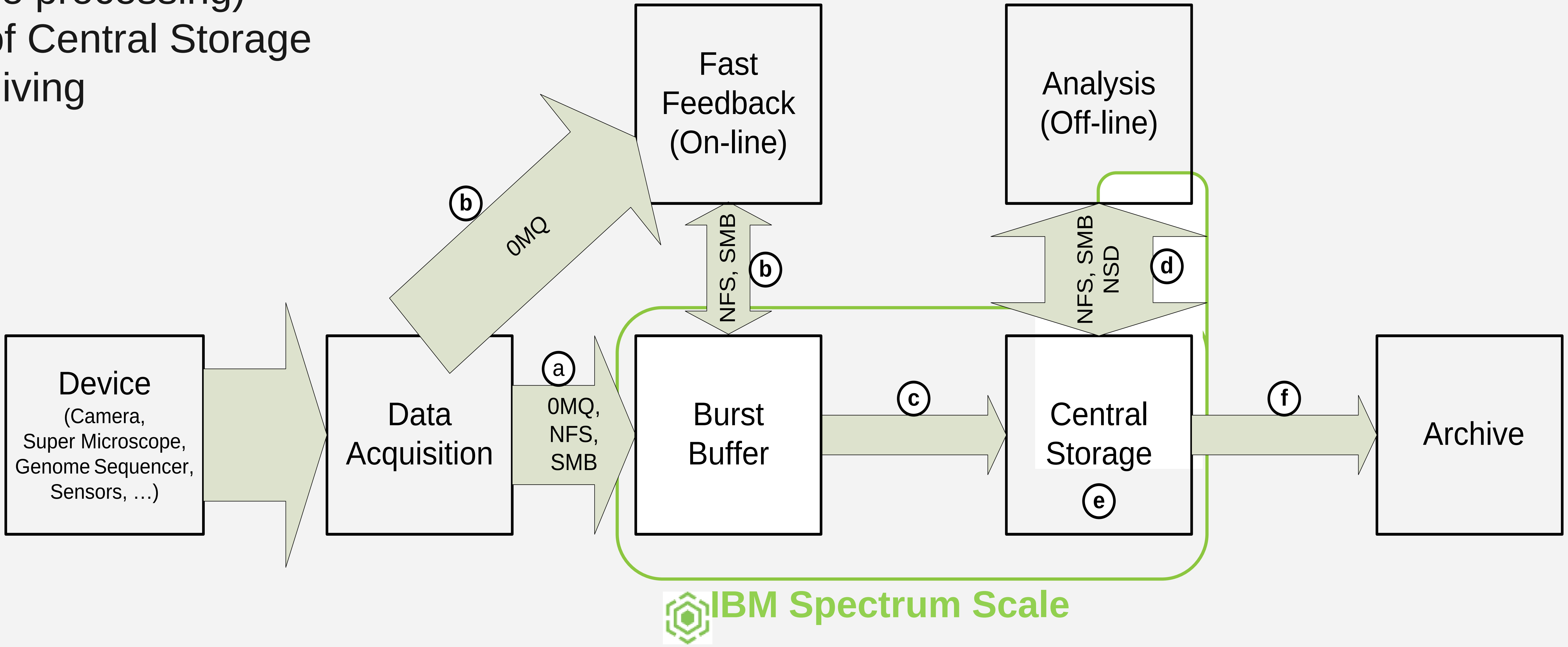
a) Real-time data ingest (data acquisition)
b) Visualization and near real-time analysis (online processing)
c) Data movement from Burst Buffer to Central Storage
d) Deep analysis (offline processing)
e) Data management of Central Storage
f) Long-term data archiving



➔ Scientists need access to data during each stage of the workflow

Typical Workflow for Data Intensive Science (continued)

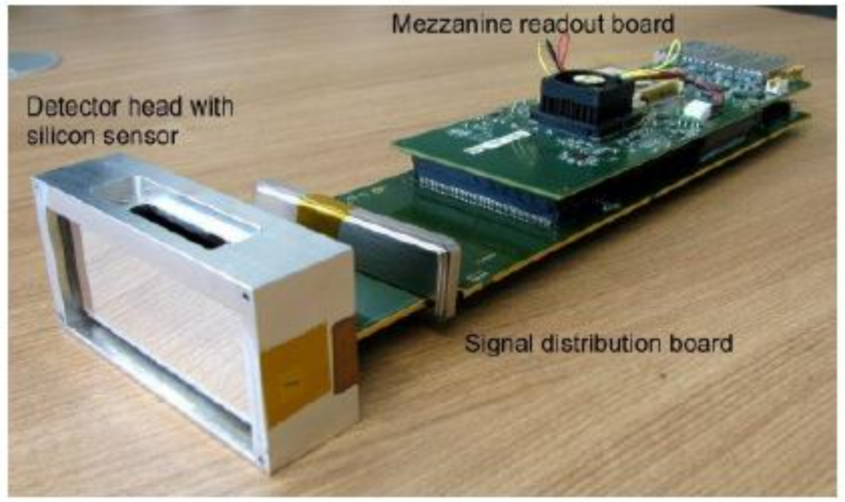
a) Real-time data ingest (data acquisition)
b) Visualization and near real-time analysis (online processing)
c) Data movement from Burst Buffer to Central Storage
d) Deep analysis (offline processing)
e) Data management of Central Storage
f) Long-term data archiving



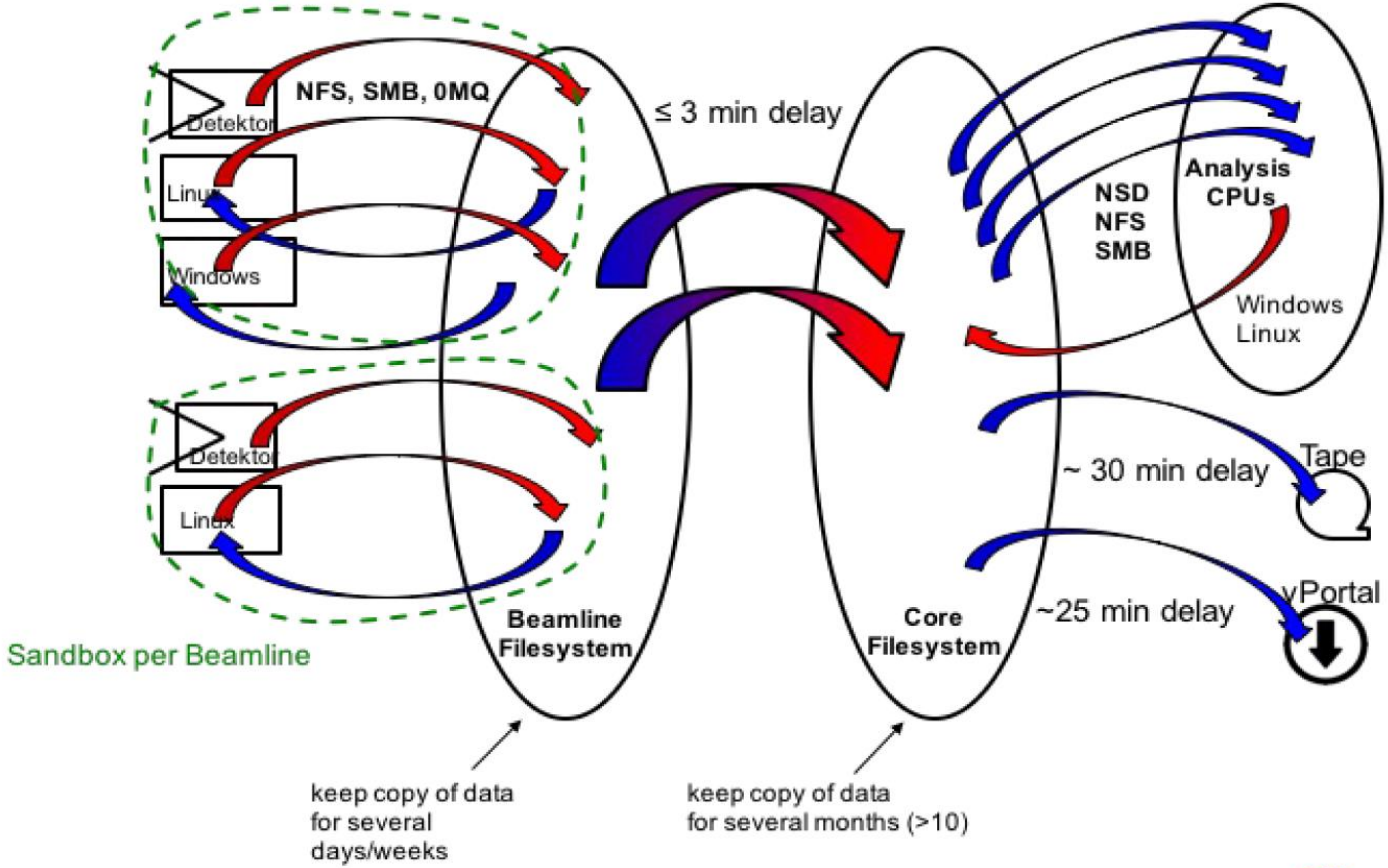
➔ Scientists need access to data during each stage of the workflow
➔ IBM Spectrum Scale has proven to support this workflow

Current and Future Detector Rates

- Detectors exceeded capabilities of prev. system:
  - Pilatus 300k: 1,2 MB Files @ 200 Hz
  - Pilatus 6M: 25 MB files @ 25 Hz
7 MB files @ 100 Hz
  - PCO Edge: 8 MB files @ 100Hz
  - PerkinElmer: 16 MB + 700 Byte files @ 15 Hz
  - Lambda: 60 Gb/s @ 2000 Hz (Future)
  - Eiger: 30 Gb/s @ 2000 Hz (Future)
- GPFS is now used to handle those rates
  - SMB/NFS sufficient for current detectors
  - Future detectors need new methods

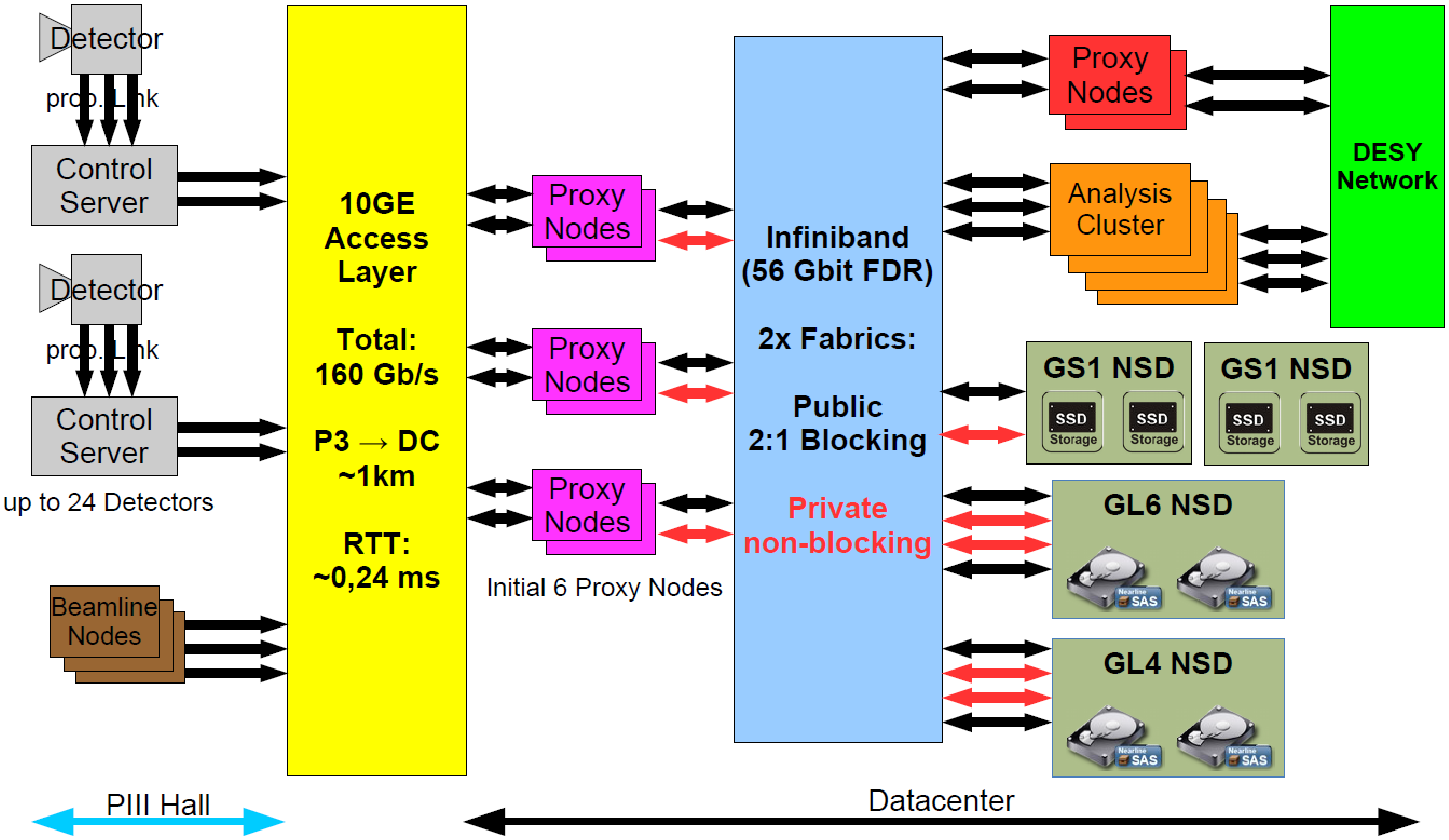


from the cradle to the grave



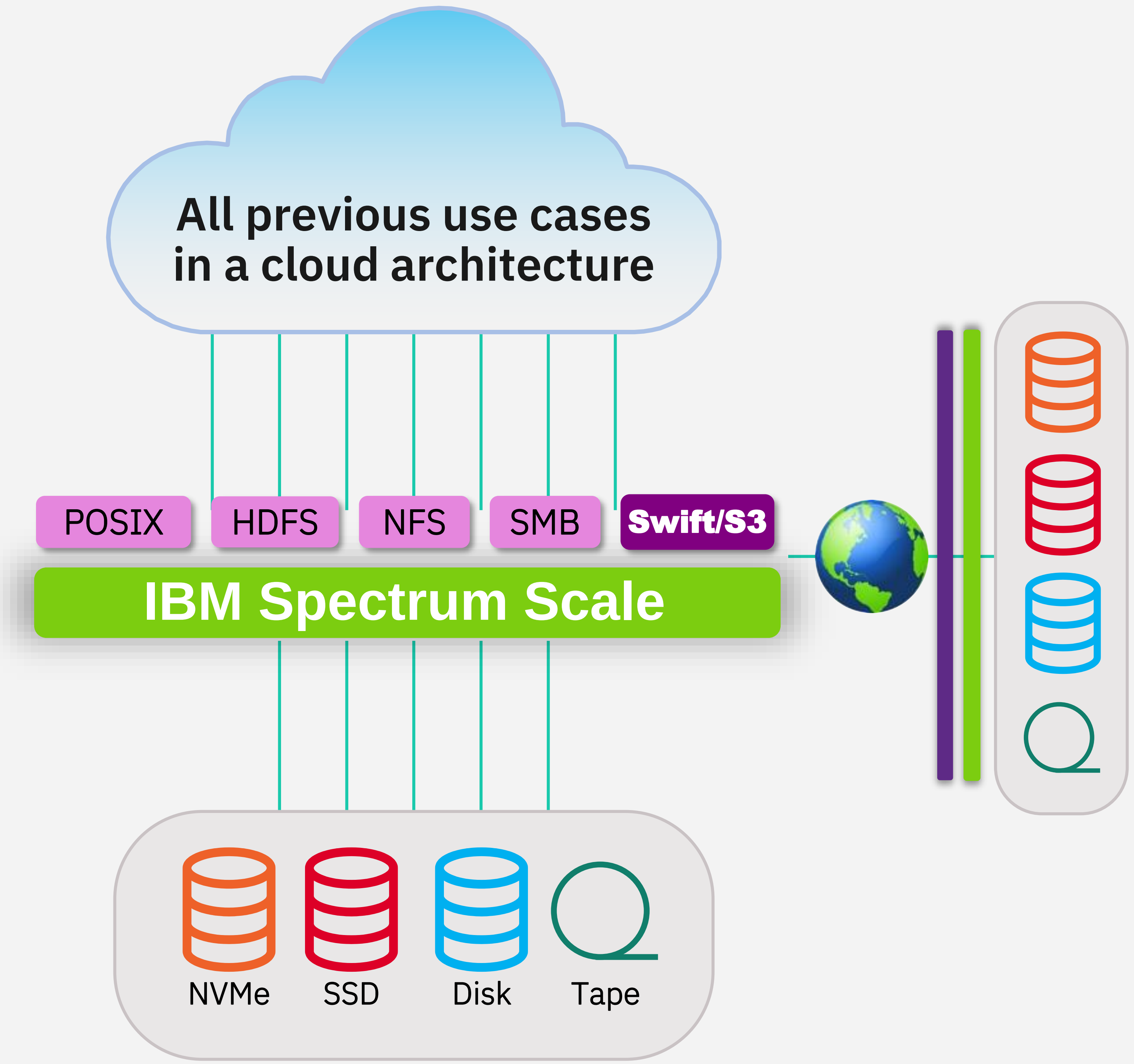
Martin Gasthuber | GPFS for Data Taking & Analysis at Petra III and European XFEL | 4/13/16 | **Page 6/28**

ASAP³ Architecture



Cloud infrastructures

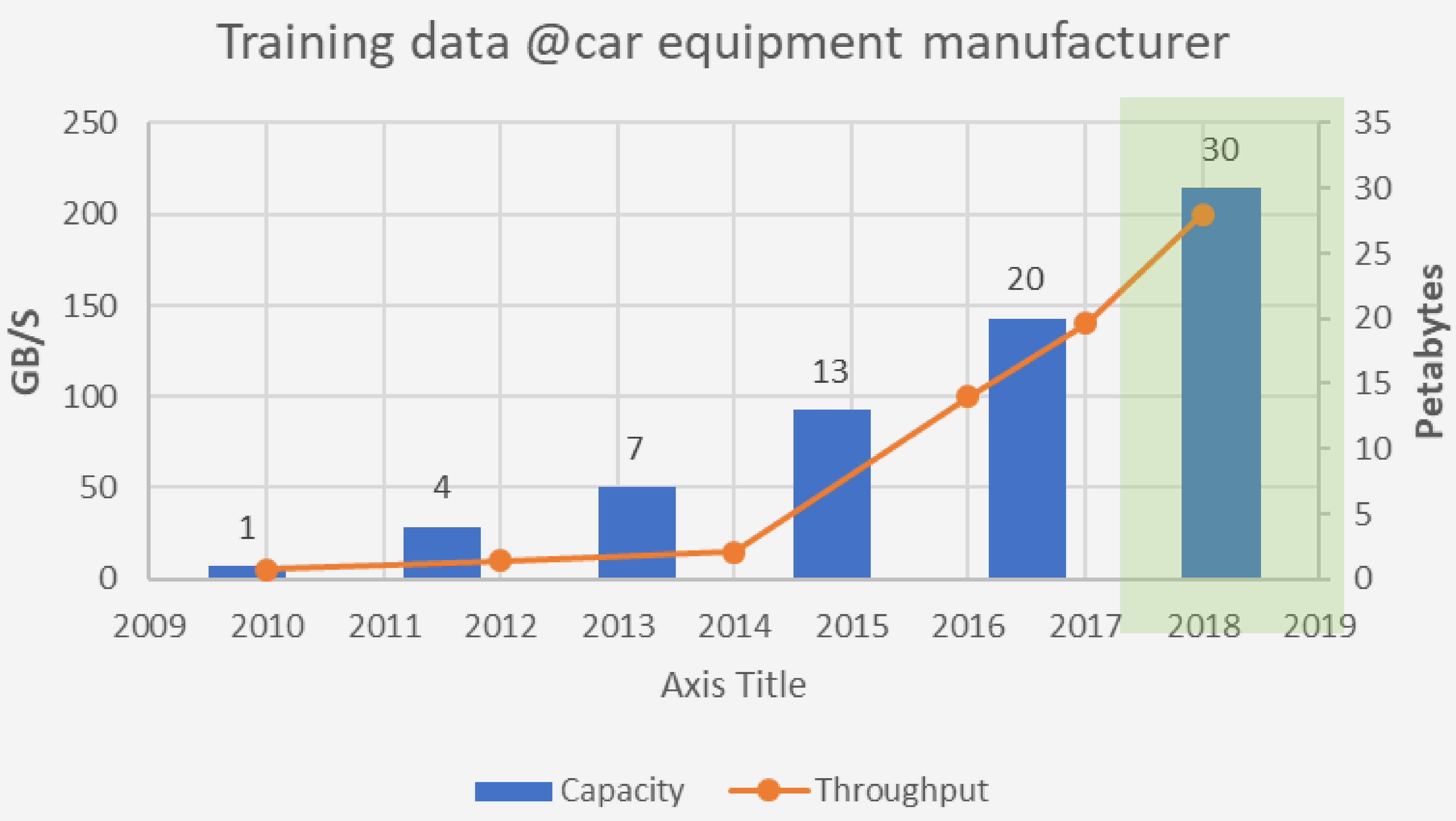
- Pervasive Computing and Cloud is driving the development of new technologies such as object storage, virtual machines and containers
- Those technologies get increasingly adopted in traditional enterprise data centers, in HPC environments and for Analytics/AI/ML/DL
- IBM makes enhancements in Spectrum Scale to integrate in cloud architectures such as
  - Data access via object protocols
  - Object storage as tier for cold data
  - Plug-ins to map directories into containers
  - Ready-to-use templates to run Spectrum Scale on AWS



Outline

1. What is IBM Spectrum Scale?
    a. Evolution
    b. Key concepts
2. Primary Use Cases
    a. High performance computing (HPC)
    b. Data intensive application & workflows
    c. AI/ML/DL
3. Summary

Case Study: Training data for autonomous driving development



➔ Increase in data volume triggers increase in required bandwidth.
➔ Data workflows need to be automated.
➔ NAS with SMB/NFS does not provide scalable performance.

Storage for AI/ML/DL & data intensive science

	Capacity	Performance	Scalable Performance
Flash storage	☒	☑	☒
Object storage	☑	(☑)	(☑)
NAS	☒	☒	☒
Scale-out NAS	☑	☒	☒
Parallel Filesystem	☑	☑	☑

Storage for AI/ML/DL & data intensive science, but …

	Capacity	Performance	Scalable Performance	Interactive Access
Flash storage	☒	☑	☒	☑
Object storage	☑	(☑)	(☑)	(☑)
NAS	☒	☒	☒	☑
Scale-out NAS	☑	☒	☒	☑
Parallel Filesystem	☑	☑	☑	(☑)

Storage for AI/ML/DL & data intensive science, but …

AI/ML/DL, Data Intensive Science

	Capacity	Performance	Scalable Performance	Interactive Access
Flash storage	☒	☑	☒	☑
Object storage	☑	(☑)	(☑)	(☑)
NAS	☒	☒	☒	☑
Scale-out NAS	☑	☒	☒	☑
Parallel Filesystem	☑	☑	☑	(☑)

HPC

Storage for AI/ML/DL & data intensive science, but …

AI/ML/DL, Data Intensive Science

	Capacity	Performance	Scalable Performance	Interactive Access
Flash storage	☒	☑	☒	☑
Object storage	☑	(☑)	(☑)	(☑)
NAS	☒	☒	☒	☑
Scale-out NAS	☑	☒	☒	☑
Parallel Filesystem	☑	☑	☑	(☑)

HPC

➔ A parallel filesystem is a good foundation for AI/ML/DL & Data Intensive Science.
➔ Interactive access and data ingest need to be architected carefully.

Adoption of Data Intensive Science

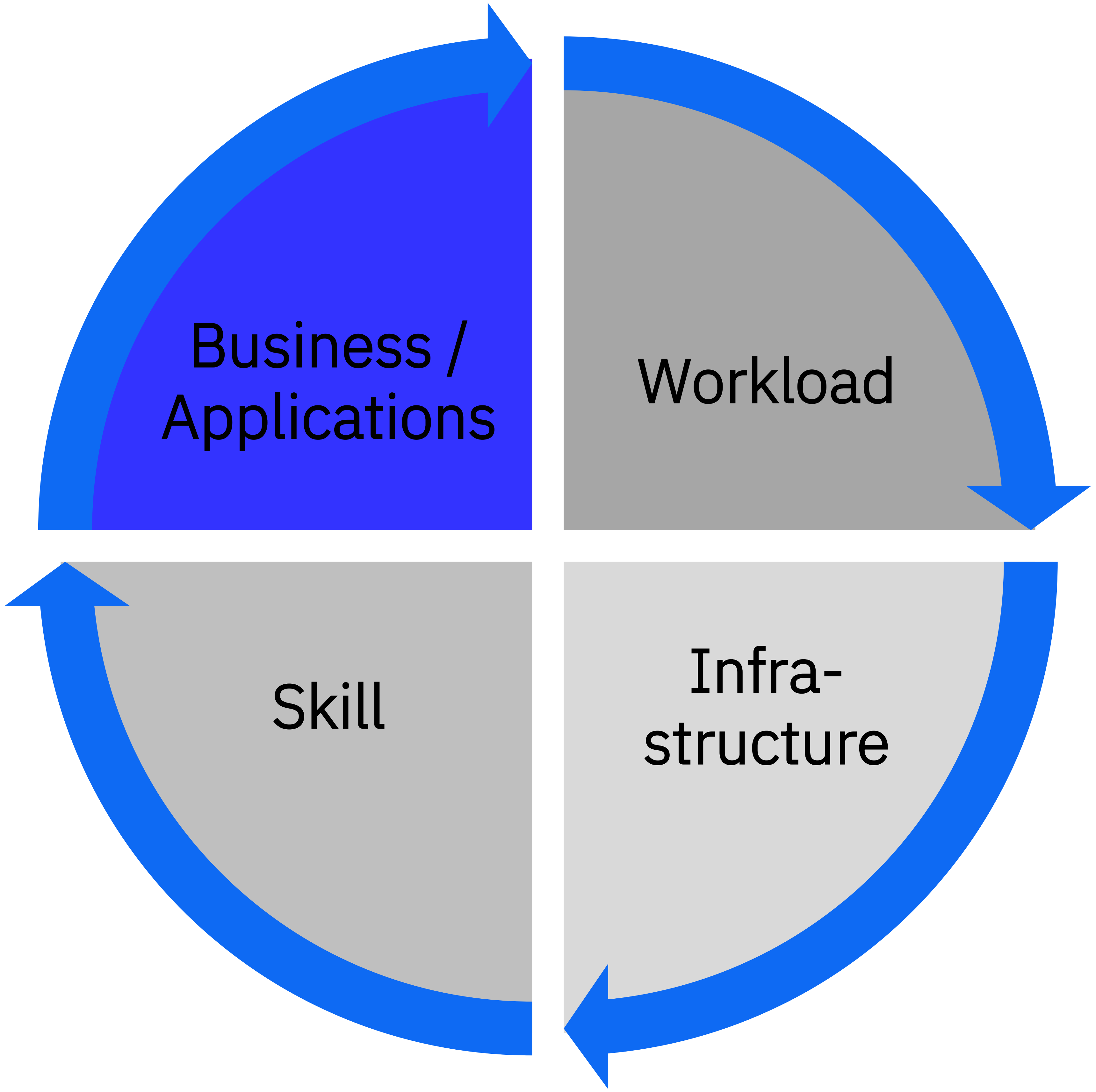
- Well established in a very few fields
  - High Energy Physics
  - Astronomy
  - Oil & Gas
- Some fields are forced to adopt quickly
  - Life science
  - Autonomous Driving
  - AI / ML / DL

➔ "Physicians are not physicists!"

Contrasting file-based workloads

		Parallel File System (POSIX)	Network Attached Storage (NAS)
Workload	**Applications**	Broad range of scientific applications, big data and analytics, ML/DL, parallel applications	Broad range of office applications, roaming profiles, etc.
	Scalable Performance	High (large data sets, fast metadata operations, high throughput, low latency)	Medium/Low (average performance and scaling needs)
	Consistency	Strict (Node see updates from remote nodes immediately)	Eventual (Node may see updates from remote nodes after a delay)
Infrastructure / Features	**Access to clients**	Controlled (Limited number of privileged users)	Wild west (End user have root access to laptops, etc.)
	Client OS Interoperability	Limited (number of operating systems, number of versions, number of architectures)	Flexible (Broad range of different OS versions including very old OS versions and architectures)
	Predominant Client OS	Linux	Linux, Windows, macOS
	Protocol	Proprietary (e.g., Spectrum Scale NSD)	Standard (NFS, SMB)
	Number of clients	Thousands (<16k for Spectrum Scale)	Tens of thousands
	Network	Private Cluster Network	Shared Data Center Network
Skills	**Deployment Model**	Software Defined Infrastructure	Hardware Appliance
	Client Software	Additional software package for access to parallel filesystem	NFS and or SMB are included in the operating system
	Admin Skills	System administrators (Deep skills in Linux, networking, system software, etc.)	Storage administrators (Mostly management of storage appliances)

Choosing the right solution



- The **business requirements** determine the required **applications**.
- The **applications determine** the generated **workload**.
- The **workload determines** the required **infrastructure**.
- The **infrastructure determines** the required **skills**.
- The available **infrastructure and skills determine** the capability to support the **business**.

Approach option – Cloud

MARKETS BUSINESS INVESTING TECH POLITICS CNBC TV

Zuckerberg says Amazon cloud bill for his philanthropy is sky high: 'Let's call Jeff up and talk about this'

PUBLISHED THU, OCT 10 2019•2:49 PM EDT | UPDATED FRI, OCT 11 2019•11:51 AM EDT

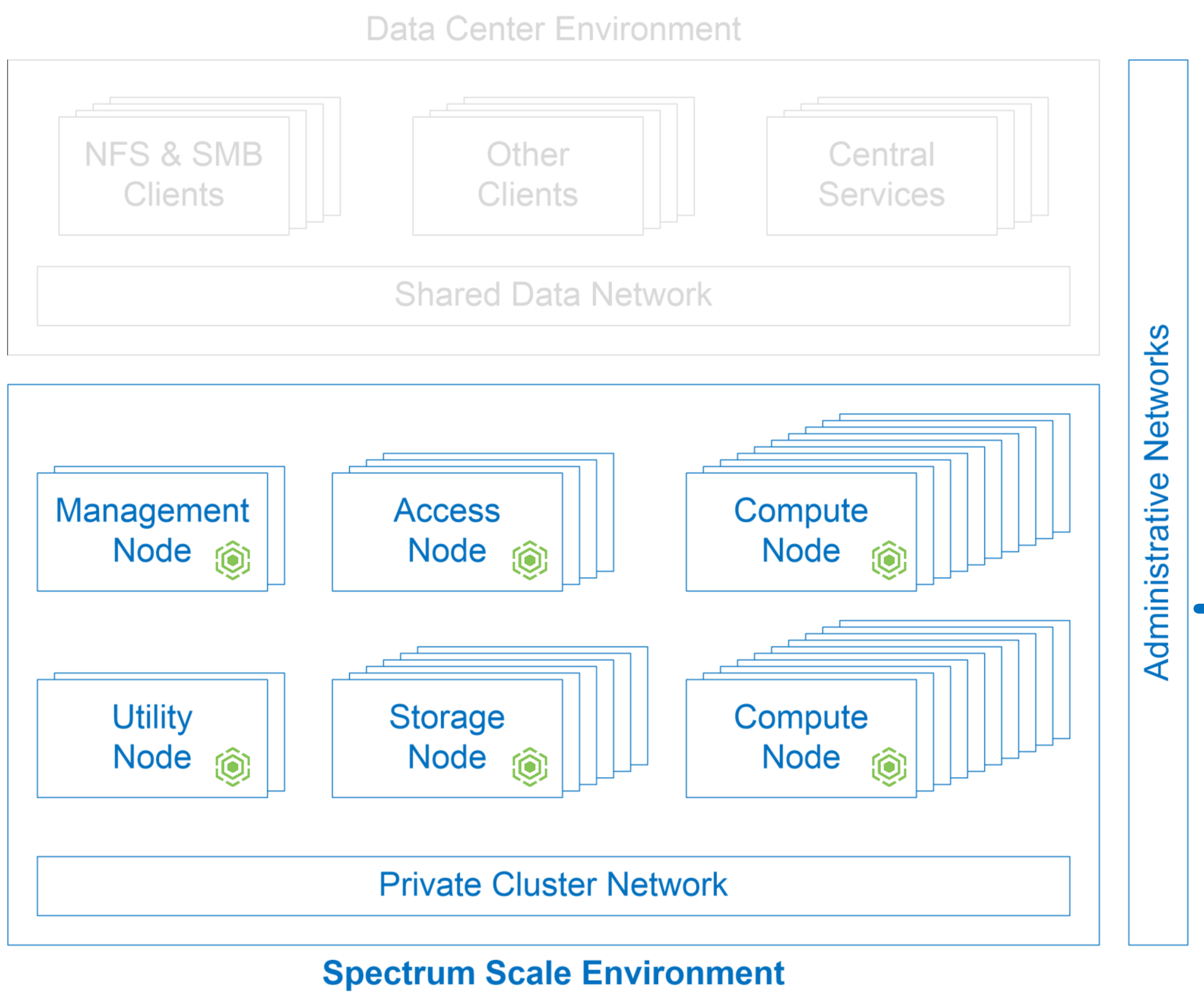
Christina Farr
@CHRISSYFARR

SHARE

KEY POINTS

- Mark Zuckerberg said in a discussion at the Chan Zuckerberg Initiative that compute costs are a big hurdle for scientists.
- Genomic data and analysis require massive IT and storage capabilities.
- Zuckerberg joked that he should bring it up with Amazon CEO Jeff Bezos, "Let's call up Jeff and talk about this."

Approach option – HPC



- ➔ The Shared Data Network provides remote access to the Spectrum Scale environment.
- ➔ The Private Cluster Network connects all components of the Spectrum Scale environment.

Compute Nodes (NSD Clients)
- Run applications to access and analyze data stored in one or more Spectrum Scale filesystems
- Most nodes of a Spectrum Scale environment are Compute Nodes.

Storage Nodes (NSD Server)
- Provide the storage capacity for the Spectrum Scale filesystems

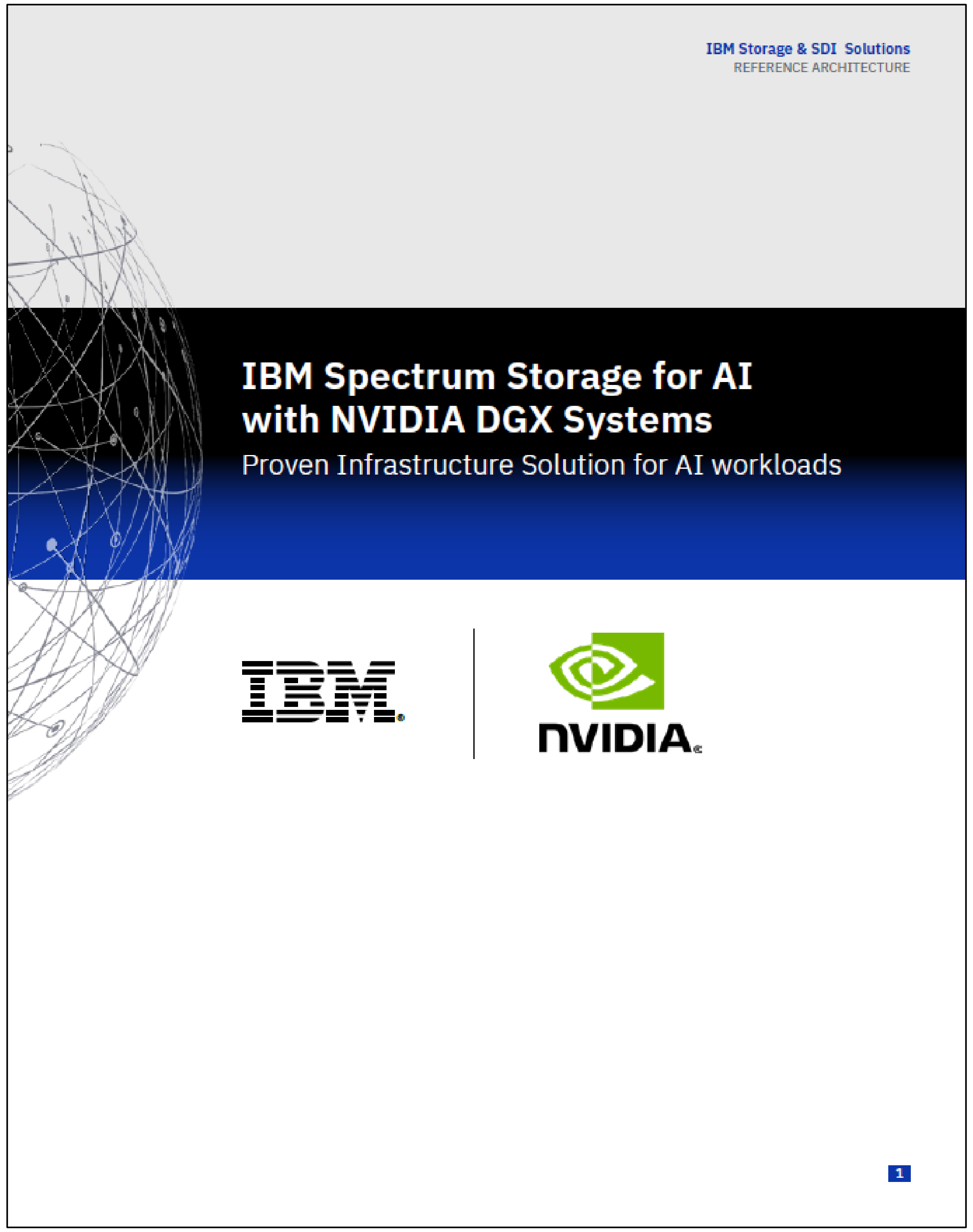
Data Access Nodes (Remote & Local Access)
- Access to Spectrum Scale filesystems using protocols like NFS, SMB, HDFS and Object

Utility Nodes (Data Management Nodes)
- Dedicated nodes for heavy-weight data management tasks such as backup, tiering, hybrid cloud workflows.

Management Nodes
- Provides administration services (e.g., Spectrum Scale GUI, Zimon Collector, Compute Cluster Login Node, Compute Cluster Management Node).

Spectrum Scale for Data Intensive Science

IBM Spectrum Scale



Discussion points:

- Workload scheduling
- Cloud support
- Container support
- OpenShift
- Global workflows
- Tiering to object and tape
- Eliminate root access
- REST API
- Ansible

Success Factors

Successful deployments of Spectrum Scale and Spectrum Scale Protocols depend on

- System administrators (Deep skills in Linux, networking, system software, etc.)
- End-to-end skills to architect, implement, operate and troubleshoot the whole Spectrum Scale Environment including software, servers, storage and networks as well as additional functions such as backup, workload scheduling and monitoring
- Follow HPC best practices
- Close collaboration with users
- Availability of low latency and high throughput Private Cluster Network
- The right workload

➔Start with a small environment and use elementary features only.
➔Acquires skill in a stable production environment.
➔Incrementally grow environment and adoption of advanced features.

Outline

1. What is IBM Spectrum Scale?
   a. Evolution
   b. Key concepts
2. Primary Use Cases
   a. High performance computing (HPC)
   b. Data intensive application & workflows
   c. AI/ML/DL
3. **Summary**

Summary

- Spectrum Scale is based on GPFS, a robust, fast and mature parallel file system.
- The filesystem of the largest super computers are build on Spectrum Scale.
- Spectrum Scale’s built-in parallelism enables a data layer that meets the performance and scaling requirements of data intensive applications and workflows such as Big Data, Analytics and AI/ML/DL.
- Spectrum Scale’s built-in support for POSIX, NFS, SMB, HDFS and object accelerates workflows that require multiple access methods.