

ArcaStream

High Performance, Data Aware Software-Defined Storage

Highlights

High Performance Parallel File System with Limitless Scale

IBM Spectrum Scale enhanced with added ease of use, collaboration, tiering, cloud integration and relentless support – on open standard hardware – greatly reducing costs

Transparent Data Tiering to Object, Tape, and the Cloud

Reserves high speed storage for work in progress, moves everything else to low cost archive, while maintaining a single view of all data

Effortless Cloud Bursting

Painlessly cloudburst onto unlimited nodes, minimizing data transfer and maximizing data locality

Simplified Management and Workflow Automation

Easy Python API programming interface removes complexity and provides 'DevOps' style management capabilities for systems administrators and pipeline developers – automating management and workflows

Complete Data Awareness and Lightning Fast Search

Intuitive user defined metadata harvesting, tagging and easy search for data – no more tree walks or lost data

Advanced Analytics

Real time performance, history and trending with statistics reporting about data and its contents through a simple and intuitive graphical interface

Multi-protocol

SMBv2/3, NFS, FTP, HTTP, S3, HDFS and POSIX parallel file system client for extreme performance

The Scientific Data Deluge

Increased sensor resolution from ever growing numbers of sequencers, cameras, microscopes, telescopes, scanners and instruments of all types – along with higher scale and fidelity simulations, are driving a deluge in scientific data that organizations are tasked to process, analyze, store, share and preserve.

ArcaStream Accelerates Scientific Workflows

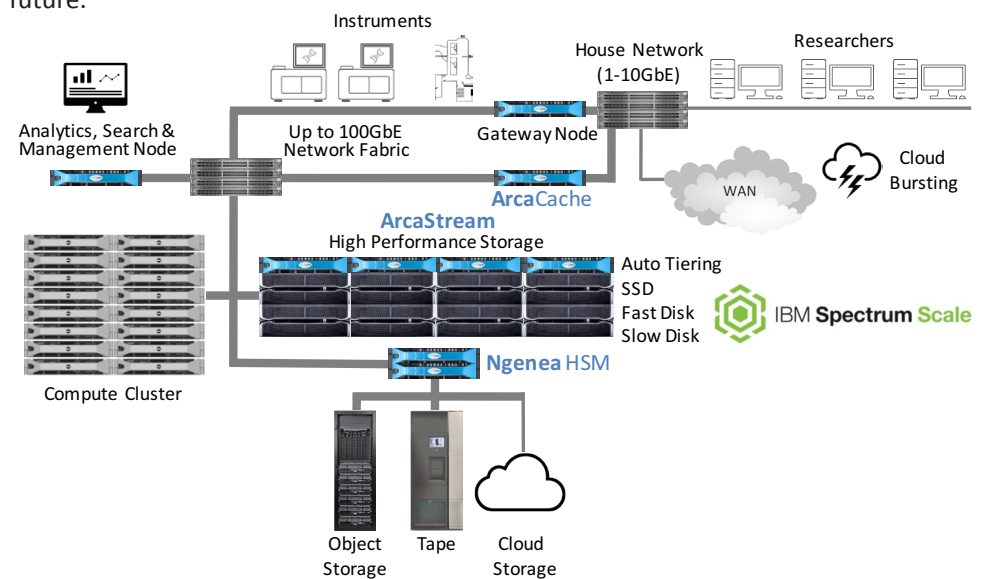
ArcaStream provides high performance, data aware, software-defined storage and networking solutions specifically designed to accelerate the world's most challenging scientific data workflows. ArcaStream high performance storage combines flash, disk, tape, and cloud storage into a unified system that's higher performing, limitless in scale and lower cost than traditional solutions. Data moves seamlessly through various tiers of storage – from fast flash to cost-effective, high capacity object storage, all the way out to the cloud – depending on how frequently it needs to be accessed. This allows research organizations to store valuable data more intelligently and economically.

Tightly Integrated with the Cloud

In addition to transparent data tiering to the cloud, ArcaStream storage facilitates effortless, highly efficient Cloudbursting onto unlimited compute nodes when needed, minimizing data transfer and maximizing data locality between data stored on premises and cloud computing resources. Results are automatically transferred directly back to on-premises storage with no manual intervention.

The Software Defined Economic Advantage

ArcaStream deploys its solution on software-defined, open standard hardware, with no vendor lock-in and standard data protection services – providing organizations tremendous economic advantage and purchasing power, now and into the future.



ArcaStream End-to-end Collaborative Scientific Workflow Architecture

Advanced ArcaStream Software

- Centralized control of all file system components from a single interface
- Pythonic and REST-compliant toolset can be used to automate Spectrum Scale operation and create customized web interfaces
- Vastly reduced complexity for otherwise difficult to manage file system components
- Monitoring, alerting, search and analytics
- High performance tiering to object storage with full path retention so object is directly accessible by cloud applications, enhancing workflow
- Seamless cloudbursting
- Cost effective, simplified Disaster Recovery via remote ArcaStream asynchronous replication

Unified Services and Support

- “Disk to desktop” holistic solution understanding
- Direct access to knowledgeable expertise – no “briar patch” support
- Hybrid managed service: regular maintenance tasks such as firmware and upgrades are handled by ArcaStream support, rather than the customer’s technical staff
- Alerts generated by the system are proactively investigated and escalated by ArcaStream support
- Single point of contact multi-vendor incident management ensures SLA’s and top-quality support experience is delivered
- Automated weekly health checks

Comprehensive Solution Design

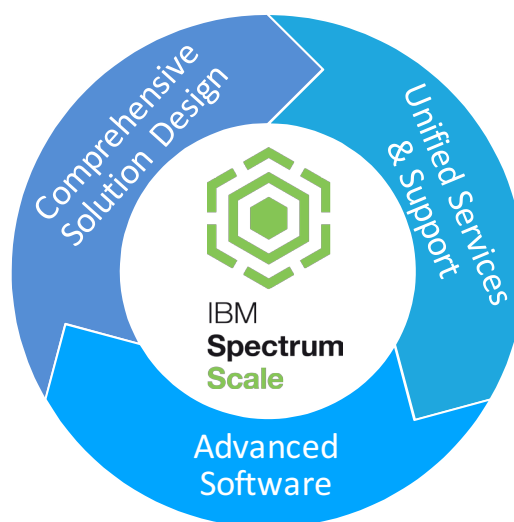
- Extensive, in-depth analysis of existing environment focusing on workflow requirements, storage, network, and applications
- Identify issues and bottlenecks
- Design solution architecture including hardware, software configuration and procurement guidance
- Integrated system build and implementation

Under the Hood: What makes ArcaStream Storage So Capable

ArcaStream storage is powered by IBM Spectrum Scale, a proven high-performance parallel file system trusted by thousands of organizations worldwide. IBM Spectrum Scale can easily manage petabytes of data and billions of files, all under a single global namespace. IBM Spectrum Scale stripes data in parallel to multiple disks across multiple servers delivering extreme, scalable performance – and supports the newest low-latency, high-bandwidth NVMe flash technology. IBM Spectrum Scale scatters data across spinning disk, removing the impact of fragmentation on performance, so there is no degradation in performance as the system fills – as is common in many file systems.

ArcaStream Enhancements to IBM Spectrum Scale

ArcaStream builds on top of IBM Spectrum Scale, delivering innovation across three areas: advanced software that expands IBM Spectrum Scale’s abilities, making it more useful and easier to deploy and manage; comprehensive solution design, to make sure customer workflow requirements and integration goals are met and exceeded; and all-inclusive, expertly delivered “disk to desktop” unified services and support.



ArcaStream Enhancements to IBM Spectrum Scale

An Industrial Strength, Robust, Easy to Manage Storage Platform

The result is an easy to manage, scalable, high performance storage system with centralized control of all file system components from a single interface employing graphical interfaces for common tasks, and Python and REST APIs for file system automation and workflow integration. ArcaStream storage monitoring, alerting and performance analytics are tailored to the exact needs of the administrative workflow and empower the organization with unprecedented insight into data usage and trends.

A Versatile Architecture that Grows with Future Needs

The ArcaStream architecture can scale in capacity, performance and capability – on premises and in the cloud – linearly and independently, by simply adding components to the existing infrastructure. And since ArcaStream storage is software-defined on open standard hardware – with no vendor lock-in, new technology like NVMe and the latest object storage or cloud integration can be easily added from multiple vendors down the road – ensuring economic architectural longevity that’s adaptable to meet the most demanding scientific data and workflow challenges the future brings.