Hitachi Content Software for File: High-Performance Storage for AI and ML

Faster Access to More Data at Lower Costs
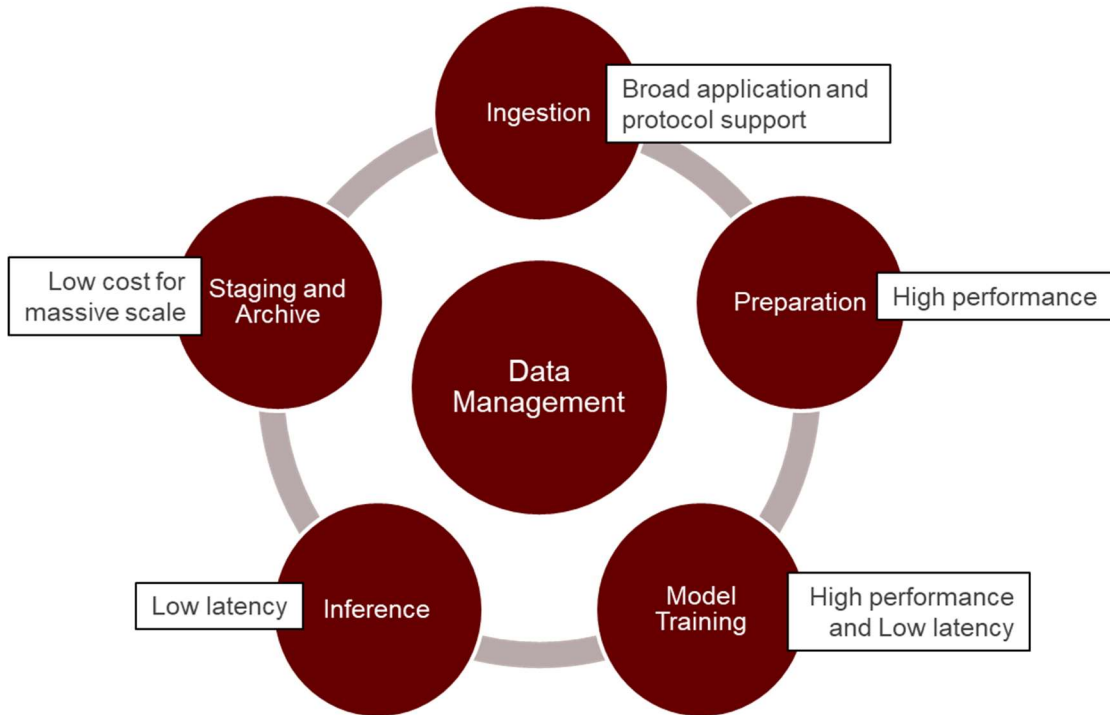
SOUTION PROFILE

In the race to make the most of artificial intelligence, machine learning and high performance computing based analytics, the winners are those who can run more models with more complex algorithms and do it faster than their competitors. To win, you need faster access to more data at a lower cost than your competition. To do so, your applications need more data, but storage performance has been left behind. While graphics processing units have shrunk compute infrastructure by 40%, the data they process has grown by 50%. Investments in compute resources and networks often sit idle, waiting for data. Response times skyrocket, but adding more compute is ineffective as legacy storage can't scale while maintaining high performance.

Unique Requirements at Every Workflow Stage

Each workflow stage has unique compute, storage, and networking needs. This leads to silos and creates data management challenges which drives up costs and time to results. Ingestion requires the ability to collect data from diverse sources over multiple protocols. Preparation demands high performance while inference needs low latency, and model training requires both. Across all these stages is the need for massive scale and automated data management at a cost that won't break the bank.

Each workflow stage has unique compute, storage and networking needs. If each stage has its own silo of infrastructure, this creates cost, complexity and management challenges that you cannot afford.

Hitachi Content Software for File: A Single, Integrated Storage Solution for All Stages

Hitachi Content Platform for File is a high-performance storage solution for AI, ML, analytics and other GPU accelerated workloads. It gives you the speed of a distributed file system (DFS) with the capacity and cloud capabilities of an object store. Its support for file and object protocols makes data ingestion easy. The DFS provides both high performance and low latency for data preparation, model training and inference. The object store provides massive storage capacity at a lower cost and offers powerful, data management automation driven by metadata.
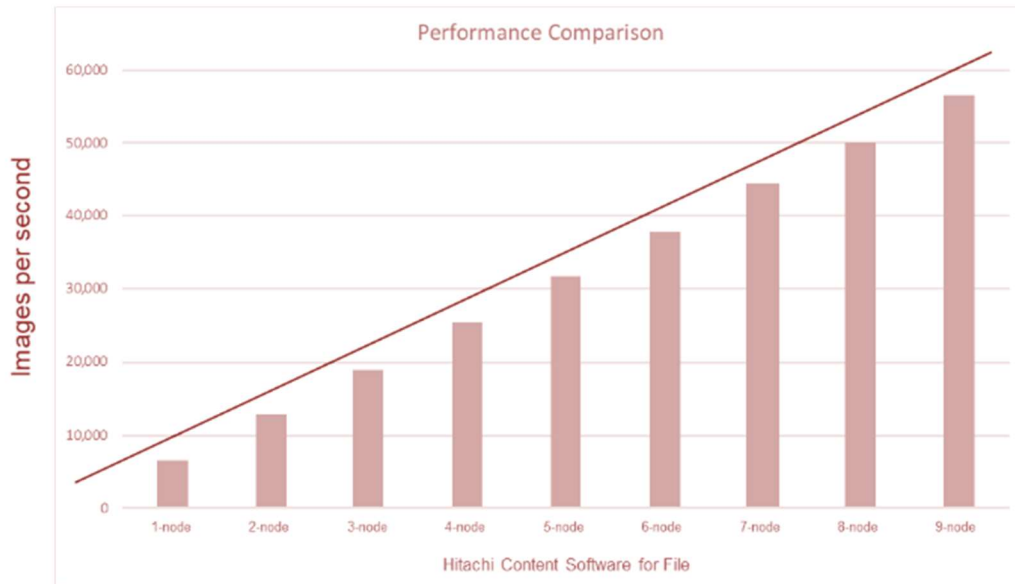
This unique integration of file and object storage offers you a single solution for an appliance-like experience. It creates a single pool of capacity with independent scale of compute and storage capacity with intelligence metadata-based data management automation to move data between on-premises and public cloud storage for cost, compliance and business continuity purposes.

Hitachi Content Software for File Features and Benefits

Distributed File System

Hitachi Content Software for File is 3 times faster than local flash drives and 10 times faster than a traditional all-flash array, allowing it to fully saturates compute resources for great efficiency. Performance scales linearly as you add nodes, so you get the most out of your compute resources.

Blazing speed for faster results.

Object Storage

Hitachi Content Software for File effortlessly presents PB data sets and seamlessly handles large & small files. This single data lake allows federated search and audit, reduces costs with faster e-discovery, shrinks audit times, lowers the risk of missing relevant data and easily adapts to fast-changing global regulatory requirements.

Massive capacity for more accurate analysis.

Integrated Solution

Hitachi Content Software for File simplifies deployment and management. It eliminates silos and multiple copies by providing single storage platform with built-in backup and disaster recovery for the entire data pipeline with seamless movement of data and full elasticity with public cloud.

Lower cost of ownership.

Financial Services Features and Benefits

Performance

Roughly 90% of all public market trading is done by quantitative means, but humans can only handle about 10 technical indicators across small datasets. With AI/ML supported by Hitachi Content Software for File, you can process hundreds of technical indicators across datasets that are thousands of times larger.

Faster, more accurate data analysis.

Data Classification and Governance

The banking and finance industries are among the most regulated environments in the world. The more silos of information you maintain, the harder it is to ensure proper governance as regulations change and produce the right data when the auditors show up. The object store in Hitachi Content Software for File is one of the best archiving and compliance solutions on the market.

Better understand how data is classified for the purposes of safeguarding it.

One Solution for Diverse Data

Structured data is great and provides lots of insights. However, 80%-90% of your data is semi-structured or unstructured data. That's a lot of value that isn't being realized. With Hitachi Content Software for File, these datasets can be tapped to improve the results of your analyses.

Leverage data such as text, voice and system logs, among others.

Life Sciences Features and Benefits

Scale

Data growth is overwhelming. With data doubling every 7 months, we're looking at roughly 40 exabytes by 2025, which will put massive strain on research budgets. Hitachi Content Software for File brings highly efficient, low cost storage with hybrid cloud capabilities to keep the cost of storage to a minimum while keeping it readily accessible.

Keep more data at a lower cost.

Protection and Governance

Managing and protecting data at this scale is a nightmare. It is too big for traditional backup but still needs to be secure, properly retained and protected. Hitachi Content Software for File brings built-in data protection, disaster recovery, and automated data governance features.

Ensure your data is safe and secure.

One Solution for Diverse Data

Multiple, highly specialized tools for genomics, proteomics, metabolomics, bio imaging, neuro research and more mean supporting a wide range of data profiles. Some create and delete thousands of small files. Others spend 80% of their file activity on just a few files. Hitachi Content Software for File can support diverse data types, access patterns and applications.

Avoid silos of infrastructure so you can better leverage research budgets.

The Best of Object and File

Hitachi Content Software for File unites one of the fastest distributed file systems with one of the best object stores and delivers it through a single vendor with a single point of contact for support.

Key Data Points

- High performance, POSIX compliant distributed file system.

- Scalable, flexible, intelligently automated object storage.

- Integrated appliance-like experience for ease of deployment and management.

- One point of contact for straightforward support.

All from a trusted vendor so you can jumpstart or accelerate your artificial intelligence, machine learning, analytics, and other GPU accelerated workflow projects.

Take Your Next Steps Today

In this new race to exploit artificial intelligence and machine learning, the winners are the companies that can run more models, with more complex algorithms, faster than the competition. You can't afford to wait for internal science projects to bear fruit.

- **Ask your representative for a meeting.** Our sales and technical teams can deliver a presentation or have a working whiteboard session to help you understand how Hitachi Content Software for File can benefit you.

- **Schedule a demonstration.** See how Hitachi Content Software for File works and get answers to your deepest questions.

- **Visit our website to learn more.** Get all the latest and greatest content for Hitachi Content Software for File.

85%

of infrastructure and operations (I&O) leaders are looking to use artificial intelligence (AI) in their infrastructures during next two years. AI workloads are diverse and some are fundamentally different from any other workload the organization may have run in the past. Although interest in leveraging AI applications is on the rise, I&O leaders are often unprepared to address storage requirements and data management challenges for growing datasets of large-scale machine learning (ML) deployment.

Source: Polling Results, AI Research Circle, January 2020

[Learn more](#) about Hitachi Content Software for File and other object storage solutions.

Hitachi Vantara Federal

February 2021