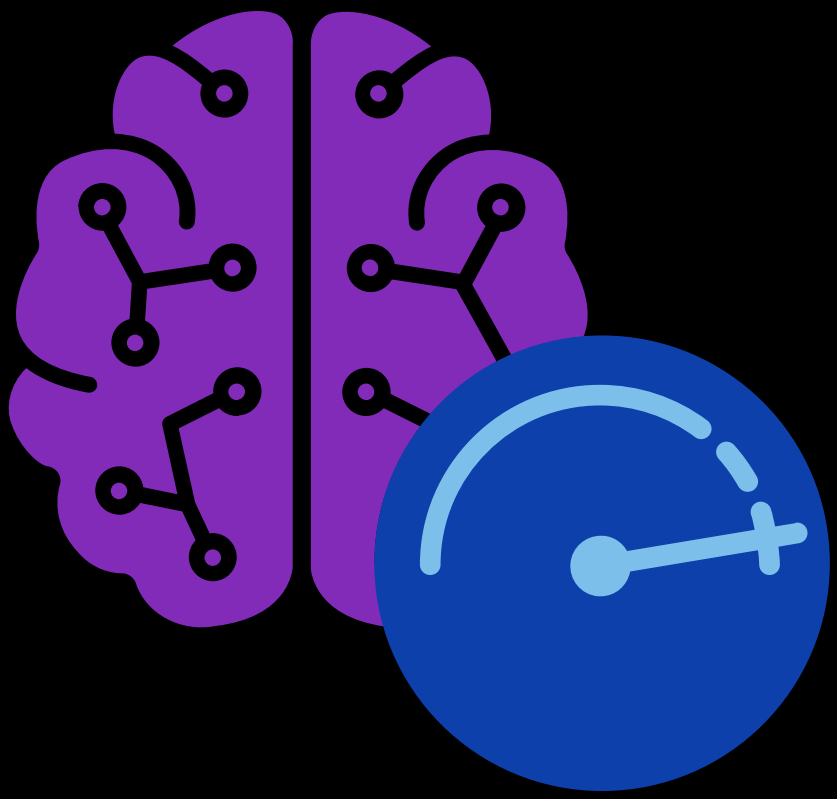
SOLUTION OVERVIEW

Accelerate AI data delivery for enhanced S3 storage with F5 and Dell

F5® BIG-IP® drives more value from Dell ObjectScale S3-compatible storage with intelligent, global traffic management and advanced data protection.



Key benefits

Achieve seamless horizontal scaling
Add more S3-compatible storage clusters while maintaining consistent, high-speed performance.

Reduce costs and accelerate data processing
Optimize workflows while improving data availability for high-value, real-time AI applications.

Simplify AI security and governance
Centralize control for distributed AI and protect sensitive models, data, and the organization's reputation.

Enable resilient performance
Keep mission-critical AI applications online with a combination of global and local DNS services.

Managing AI data flows is the key to operational efficiency

For AI initiatives, data is the lifeblood that powers model training, fine-tuning, and inferencing. However, as enterprise AI scales up, organizations need to securely connect more data sources across distributed, multicloud environments. The process of ingesting vast amounts of data stored in S3 deployments across these distributed environments creates a relentless influx that's difficult to manage, resulting in bottlenecks that impact AI training and inference efficiency.

To succeed, enterprises require data delivery solutions that can scale seamlessly while enabling high performance and resilience. When paired with generic load balancing or security solutions, even the most advanced multi-zone or multi-cluster deployments can be slowed down, creating unnecessary friction in data delivery. These slowdowns not only diminish operational efficiency but also increase the risk of non-compliance with SLAs and regulatory requirements.

Modernized infrastructure and AI performance needs
Cost-effective, scalable, S3-compatible storage is a requirement for ensuring data is readily available to support AI workflows. Many organizations already take advantage of Dell ObjectScale as a container-based, software-defined storage platform that's purpose-built for large-scale data collection, backup, and analytics. But elastic storage isn't the only requirement.

In distributed environments, AI also depends on intelligent load balancing that can manage local and global traffic, deliver SSL encrypt/decrypt security for data in flight, and deliver data where it's needed most, such as storage clusters, AI models, and retrieval-augmented generation (RAG) databases. Integrated security and manageability controls are also a linchpin to ensuring data privacy and safeguarding against advanced attacks and AI-powered threats.

By overcoming these obstacles, organizations can make better use of their modernized infrastructure and keep costs and sprawl under control even as they add more data, models, and services into the mix.

Key features

Balance workloads strategically
Distribute S3 API traffic evenly across ObjectScale nodes using real-time metrics to gauge server health.

Reduce latency and improve throughput
Streamline high-volume S3 operations with geolocation-based routing, SSL offloading, TCP connection reuse, and DNS caching.

Safeguard sensitive data anywhere
Protect applications with behavioral analytics, DDoS mitigation, application-layer encryption, threat intelligence services, and API security.

Choose from flexible options
Deploy in the cloud or leverage on-premises or colocated BIG-IP appliances to train AI with tighter security and control.

Secure, intelligent, adaptive AI traffic management for ObjectScale clusters

The F5 Application Delivery and Security Platform (ADSP), which features F5 BIG-IP, provides a scalable, secure, high-performance platform for managing AI data traffic that synergizes with and amplifies the value of ObjectScale deployments. F5 ADSP simplifies connecting distributed data sources, AI models, and ObjectScale storage clusters and applies intelligent traffic management and policy-driven security to ensure AI pipelines are reliable, predictable, compliant, and governable.

Adaptive traffic management accelerates AI across regions
F5 BIG-IP Local Traffic Manager™ (LTM), along with iRules, intelligently distributes incoming S3 API traffic—including PUT, GET, DELETE, and LIST—across multiple ObjectScale nodes or clusters. The platform balances workloads based on real-time metrics including connection count and server health to ensure data availability, prevent network congestion, and keep everything running smoothly even as organizations add more storage clusters. Additionally, iRules can be used to optimize read/write traffic routing, enabling fine-grained control that further enhances performance and resource efficiency.

F5 BIG-IP DNS delivers global server load balancing to direct client requests based on geolocation, latency, and availability, enabling intelligent traffic routing in hybrid/multi-site deployments across regions. The solution optimizes performance through features such as SSL offloading, TCP connection reuse, and DNS caching, all of which help reduce latency and improve throughput for high-volume S3 operations.

Robust security protects sensitive data and AI models at scale
BIG-IP acts as a single control point for all AI data ingress and egress, inspecting and enforcing policies to block data poisoning, prevent unauthorized access, discover and lock down API endpoints, and stop model exfiltration before it happens. With integrated behavioral analytics and layer 7 distributed denial-of-service (DDoS) mitigation, F5 BIG-IP Advanced Web Application Firewall (WAF)® filters out legitimate user traffic from malicious bots to further protect AI services and keep them responsive.

Multiple layers of resiliency ensure AI service availability
Together, BIG-IP LTM and BIG-IP DNS provide continuous health monitoring and automatic failover, ensuring that object storage services remain accessible even during node or site failures. For more redundancy, organizations can add SaaS-based F5 Distributed Cloud DNS as a hybrid, authoritative DNS solution that supports global reach. If the SaaS-based DNS goes down, BIG-IP DNS functions as the backup to ensure continuity for online services.

Mitigating cost with on-premises BIG-IP appliances

Organizations can unlock additional cost efficiencies by deploying BIG-IP with F5 rSeries and F5 VELOS® appliances. These solutions provide robust colocation or on-premises alternatives that mitigate the high cost and limitations associated with cloud-based AI training, while keeping sensitive training data more secure in a tightly controlled, private environment.

Consistent performance and scalability for AI data delivery

By providing intelligent traffic management to seamlessly move data across clusters and clouds, F5 solutions augment ObjectScale deployments to improve both data availability and resiliency for AI services. Unlock low-latency connectivity and in-depth security to support real-time workflows for competitive, in-demand AI use cases with F5 and Dell.

Contact F5 to get started or learn more at f5.com/dell.