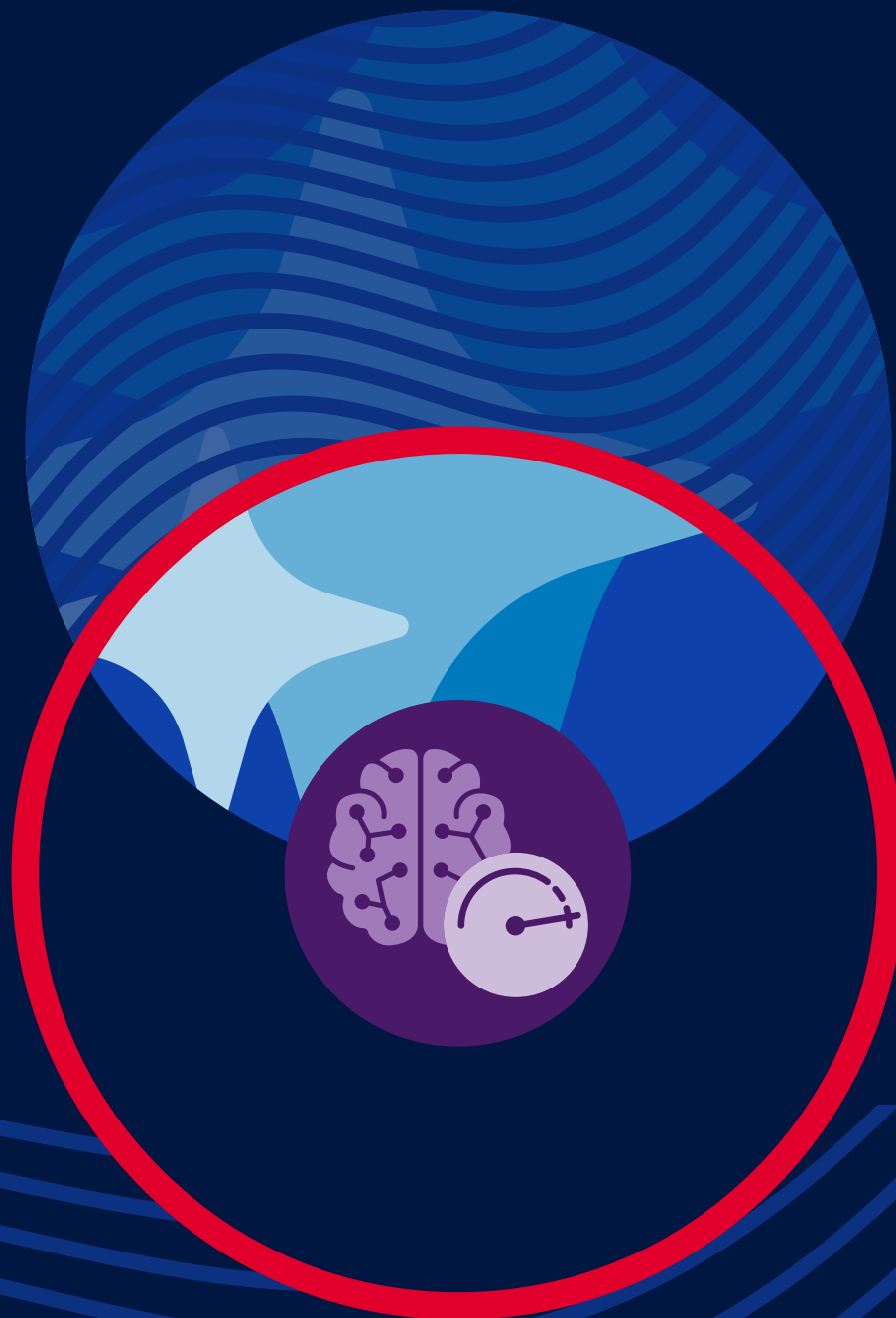




Accelerate AI model training

Build efficient and secure data pipelines for AI workloads with advanced traffic management for data ingestion from F5 and AWS.



Contents

3	The data challenge in AI model training
4	Native AWS data ingestion services
5	Enhancing AI data transfers
6	Secure multicloud connectivity
7	Performance and cost optimization
8	Consistent security controls
9	Accelerate your AI initiatives
10	Streamline AI model training with F5 and AWS

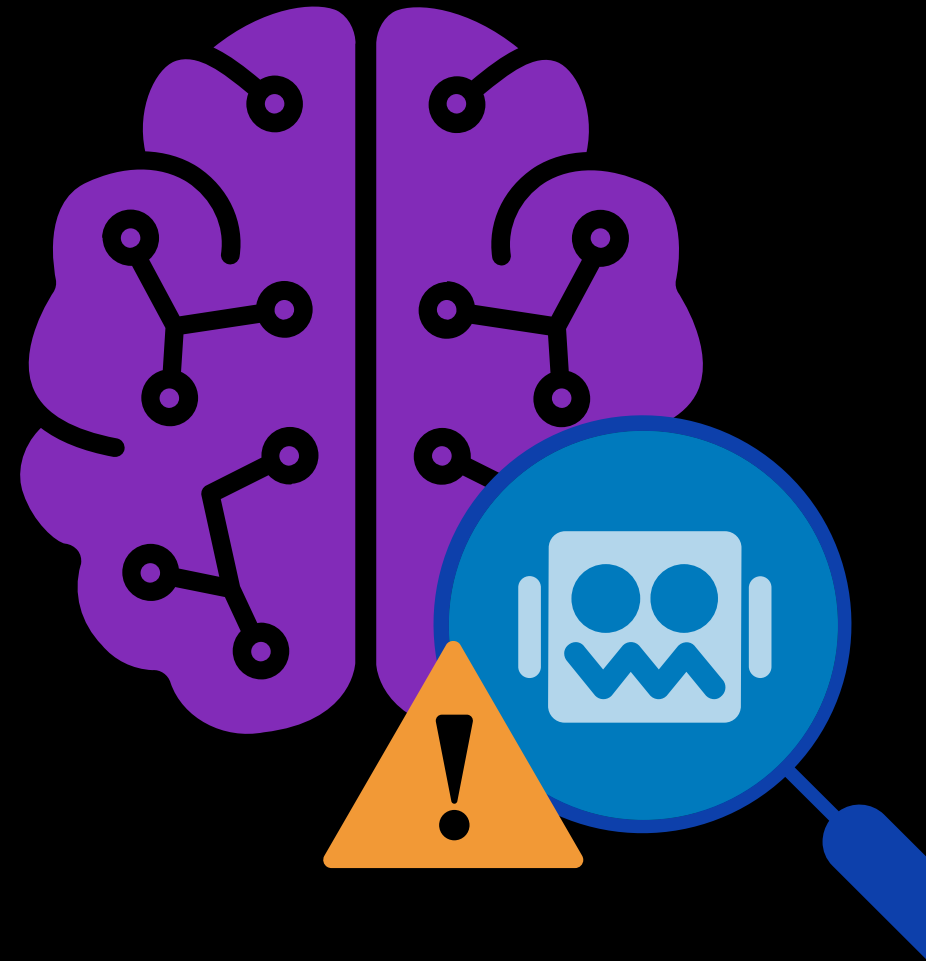
The data challenge in AI model training

Training or refining an AI model requires moving massive amounts of data to your AI infrastructure. With models requiring terabytes or even petabytes of training data, a secure and reliable network infrastructure is needed to support efficient training without incurring high costs or security risks.

Data used to train AI models frequently includes not just text but also images, audio, and videos, increasing the volume and complexity of data traffic. As most organizations now operate hybrid cloud architectures, their data sources are often spread across multiple environments, including on-premises data centers, public and private cloud storage, and edge locations that may be geographically diverse.

Efficient GPU utilization (a key metric for AI training cost efficiency) requires a constant data stream. When data ingestion bottlenecks occur, expensive GPU resources sit idle, resulting in both higher costs and slower model training. Organizations need data pipelines that can consistently feed into AI infrastructure with speed and efficiency to maximize their investments.

As training data volumes increase, organizations can face unexpected higher transfer costs to move data across distributed environments. Sophisticated traffic management is needed to control both GPU and transfer costs, accelerate model training, and protect sensitive or proprietary data.



Native AWS data ingestion services

AI developer services from AWS simplify creating and training AI models, offering model tuning and foundation models on serverless infrastructure. Moving training data to AWS storage provides a scalable way to maximize GPU efficiency and help control costs.

Amazon offers a number of data ingestion and migration services to bring external data to AWS.



Amazon SageMaker Pipelines is a serverless workflow orchestration service that integrates with other AWS data services to automate the model training workflow and can scale to accommodate thousands of concurrent workflows.



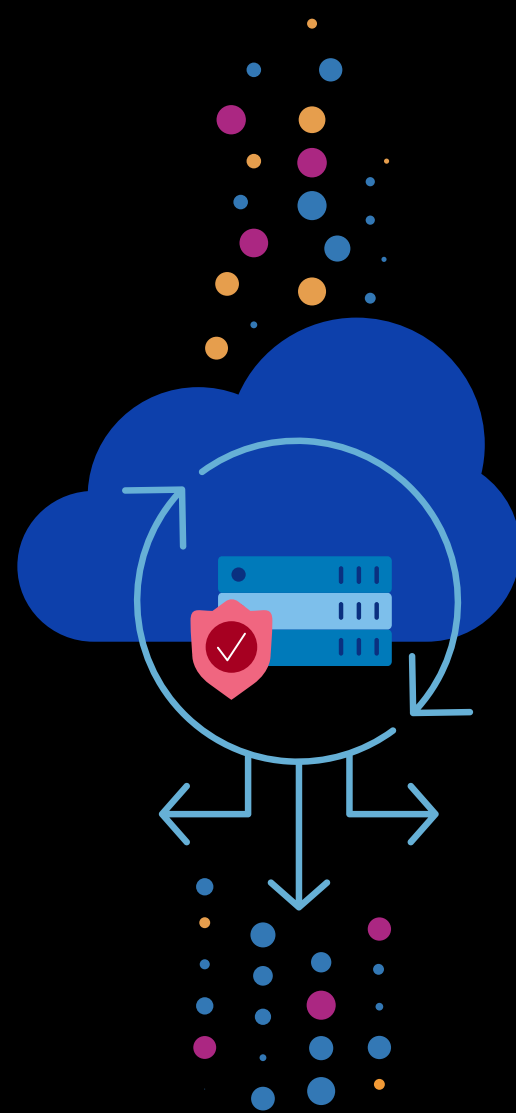
Amazon Data Firehose is designed to deliver real-time streams into data lakes, warehouses, and Amazon Simple Storage Service (S3) from over 20 AWS services.



AWS Database Migration Service (DMS) can maintain a continuous flow of data from source databases to AWS, supporting both one-time migrations and ongoing replication.



AWS DataSync moves large datasets between on-premises storage systems and AWS, and the **AWS Transfer Family** complements this by supporting file transfers using standard protocols.



The large volume of traffic used in model training requires load balancing to maintain a steady stream of data for ingestion and processing. AWS offers multiple load balancers designed for applications, networks, and appliances. While these services provide essential capabilities, organizations often need additional features for complex AI data ingestion scenarios, such as hybrid environments or stringent security requirements. F5 complements these native AWS capabilities with multicloud networking, intelligent traffic management, and consistent security.

Enhancing AI data transfers

If you need to connect multiple external sources to your AWS infrastructure, F5 can extend networking, load balancing, and security across hybrid multicloud environments. A leader in application delivery and security, F5 extends AWS services through:



Secure multicloud connectivity

F5® Distributed Cloud Network Connect and the F5® Global Network provide secure, high-performance connectivity between data sources and AI services.



Performance and cost optimization

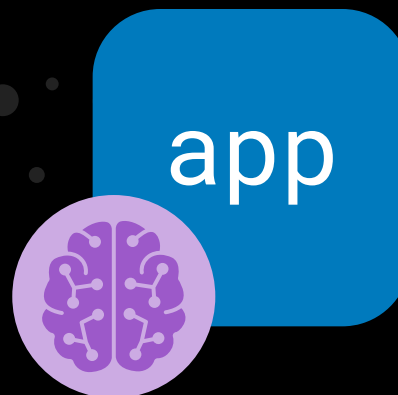
Intelligent routing and traffic optimization help control costs while maintaining the performance needed for effective model training.



Consistent security controls

Protect valuable data with encrypted traffic inspection, DDoS protection, web application firewalls, and API security across environments.

These capabilities work together to accelerate data ingestion and AI model training. The combination of F5 and AWS lets you build data pipelines that are high-performing and secure, efficient, and optimized. The following sections explore each of these areas in detail.



Secure multicloud connectivity

To connect training data sources spread across hybrid multicloud environments, F5® Distributed Cloud Services provide secure connectivity at both the network and application layers, simplifying data movement across complex environments. Integration between F5 and AWS services ensures seamless multicloud connectivity and simple deployment through the AWS Marketplace.

F5 Distributed Cloud Network Connect enables secure Layer 3 connectivity between training data sources and AWS. This allows you to:

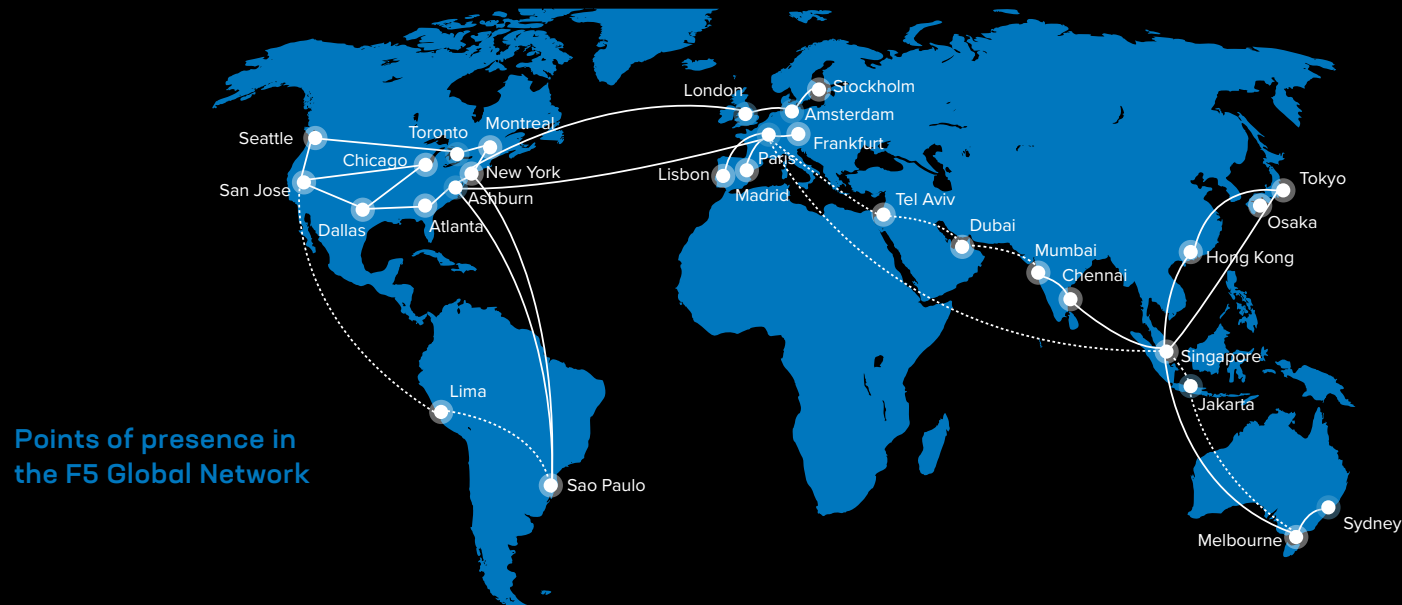
- Connect cloud-based and on-premises storage directly to AWS with one-click provisioning
- Protect data with encrypted site-to-site connectivity and security service insertion
- Enable a consistent data flow for training workloads with end-to-end visibility
- Add Layer 7 service networking with F5® Distributed Cloud App Connect
- Increase AI inference accuracy by connecting data sources for retrieval-augmented generation (RAG)

Multicloud networking from F5 can connect sites over the internet, via a private backbone, or through the private F5 Global Network to:

- Secure transit between environments with global DDoS protection
- Enable high-performance data movement
- Keep sensitive data off the public internet
- Provide global coverage with 15+ Tbps of peered capacity

Using NetApp storage?

F5 securely connects NetApp data stores in the cloud and on premises through multiple protocols, including NFS, SMB, and Amazon S3 APIs, providing flexible and protected options for data access.



Performance and cost optimization

Moving large volumes of training data to your AI model requires sophisticated traffic management to maintain performance and reliability in the wake of multiple source locations and fluctuating volumes. F5® BIG-IP® Virtual Edition (VE) provides advanced load balancing and Layer 4-7 traffic optimization capabilities on AWS that go beyond basic distribution to maximize infrastructure investments.

Reduce costs through intelligent traffic management

With F5 traffic management, you can efficiently ingest data into AI models on AWS and continuously optimize data pipelines with monitoring and analytics. This limits optimize GPU usage and lowers data transfer fees while accelerating model training and reducing cost.

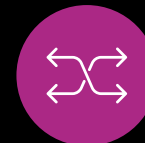


KEY TRAFFIC MANAGEMENT FEATURES INCLUDE:



Advanced load balancing

- TCP optimization tuned for large data transfers
- Server health monitoring ensures reliable data flow
- Intelligent routing based on server capacity maintains high throughput



Traffic optimization

- Efficient handling of large-scale data transfers maximizes GPU utilization
- SSL offloading removes encryption processing overhead
- FastL4 and Fast HTTP profiles increase speed

Consistent security controls

Keeping your data secure becomes more complicated in hybrid multicloud environments. You need security that works consistently everywhere to protect your applications and data. F5 security solutions in BIG-IP and Distributed Cloud Services provide broad protection that complements AWS security capabilities.

KEY SECURITY FEATURES INCLUDE:



End-to-end protection

- Integrated security controls and visibility across any environment
- DDoS mitigation for AI workloads
- Web application and API protection
- Full SSL/TLS inspection



Zero trust access control

- Centralized authentication and authorization
- Consistent policy enforcement across environments
- Protection against unauthorized data access



Compliance support

- Encryption to protect data in transit
- Audit and logging capabilities to demonstrate compliance
- Support for regulatory requirements

By integrating security with multicloud networking and traffic management, you can protect AI investments and data sources more easily. Consistent security policies and controls across all environments provide the confidence to transfer sensitive training data without risking a data breach or compliance violation.

Accelerate your AI initiatives

By combining F5's expertise in application delivery and security with AWS secure cloud infrastructure and AI development services, you can build efficient, secure data pipelines for your AI initiatives. This integrated approach helps you:



Maximize infrastructure value

- Optimize GPU utilization through consistent data delivery
- Reduce costs with efficient traffic management
- Free up resources by offloading SSL processing



Simplify complex environments

- Connect data sources quickly and securely
- Reduce networking complexity across hybrid multicloud environments
- Enable consistent operations and management



Enhance security posture

- Implement consistent security controls to protect valuable training data
- Meet and demonstrate security compliance requirements
- Get end-to-end visibility



Enable business innovation

- Accelerate AI model development
- Scale data pipelines confidently
- Support diverse training data sources and RAG

Leverage these benefits to accelerate your AI initiatives while maintaining security and controlling costs. Together, F5 and AWS provide the foundation needed to build enterprise-grade AI infrastructure, train models, and build secure AI apps that power your business.

Streamline AI model training with F5 and AWS

F5 and AWS deliver solutions to efficiently ingest data to develop and train AI models. Backed by a robust portfolio and over a decade of joint engineering and innovations, F5 and AWS provide a faster path to innovation while protecting your sensitive data and minimizing cost.

Ensure AI data ingestion is always:



Secure



Fast



Optimized

THE F5 AND AWS PARTNERSHIP

10+

years of collaboration

Over 1K

joint customers

Competencies for containers, networking, and security

Service validations for AWS WAF, Amazon CloudFront, AWS Outposts, and Linux

ABOUT F5

BRINGING A BETTER DIGITAL WORLD TO LIFE

F5, Inc. (NASDAQ: FFIV) is the global leader that delivers and secures every app. Backed by three decades of expertise, F5 has built the industry's premier platform—F5 Application Delivery and Security Platform (ADSP)—to deliver and secure every app, every API, anywhere: on-premises, in the cloud, at the edge, and across hybrid, multicloud environments. F5 is committed to innovating and partnering with the world's largest and most advanced organizations to deliver fast, available, and secure digital experiences.

Together, we help each other thrive and bring a better digital world to life.

For more information, go to f5.com.

Learn more about F5 solutions for AWS at f5.com/aws.

