

Service Name

Amazon MSK

Service Overview

Amazon Managed Streaming for Apache Kafka (Amazon MSK) is a fully managed service that makes it easy to build and run applications using Apache Kafka to process streaming data. MSK eliminates the operational overhead of managing Apache Kafka infrastructure by automatically creating and managing broker and ZooKeeper nodes, enabling developers to focus on building applications. The service is compatible with open-source Apache Kafka versions and supports existing applications, tooling, and plugins. MSK provides control-plane operations for cluster management while allowing users to leverage Apache Kafka data-plane operations for producing and consuming data. It offers both provisioned and serverless deployment options to meet varying workload requirements.

Key Use Cases

- Real-time data streaming and processing: Capturing and processing application log files, website clickstreams, and IoT sensor data in real-time
- Event-driven microservices architecture: Building decoupled, scalable applications where services communicate through event streams
- Data ingestion for analytics: Streaming data to data lakes and warehouses for real-time analytics and machine learning
- Financial transaction processing: Tracking and processing financial transactions with low latency requirements
- Content delivery and media streaming: Distributing real-time content updates and media streams to applications

Features

- **MSK Provisioned:** User-controlled broker scaling with Standard and Express broker options for optimized performance and cost
- **MSK Serverless:** Fully managed, auto-scaling Kafka clusters without capacity planning or infrastructure management
- **Multi-AZ Deployment:** High availability with automatic failover and data replication across multiple Availability Zones
- **MSK Connect:** Fully managed Kafka Connect service for streaming data between Kafka and external systems
- **MSK Replicator:** Cross-region and cross-cluster data replication for disaster recovery and multi-region architectures
- **Enterprise Security:** Encryption at rest and in transit, IAM authentication, VPC isolation, and granular access control

- **Monitoring and Observability:** CloudWatch metrics, JMX metrics export to Prometheus, and broker log streaming
- **Tiered Storage:** Cost-optimized storage with automatic data tiering to reduce storage costs while maintaining performance
- **Auto Scaling:** Automatic broker and storage scaling based on demand for both provisioned and serverless clusters

Value Proposition

Amazon MSK eliminates the undifferentiated heavy lifting of managing Apache Kafka infrastructure, allowing organizations to focus on application development rather than cluster operations. MSK provides superior operational efficiency by handling patching, scaling, security, and failure recovery automatically, reducing management overhead by up to 80% compared to self-managed solutions:

- The service offers flexible deployment options with both provisioned and serverless models to optimize costs based on usage patterns
- MSK ensures enterprise-grade security with built-in encryption, IAM integration, and VPC isolation without additional configuration complexity
- Deep AWS service integrations enable seamless data pipelines with Lambda, EMR, Flink, S3, and other services, accelerating time-to-market for streaming applications

Service Integration

Amazon MSK integrates deeply with the AWS ecosystem to enable comprehensive data streaming architectures. Core integrations include AWS Lambda for serverless event processing with automatic scaling and built-in event source mappings, Amazon EMR and Amazon Managed Service for Apache Flink for large-scale stream processing and analytics, Amazon S3 for data lake storage through Firehose integration, and AWS Glue for ETL operations on streaming data:

- Additional integrations encompass Amazon CloudWatch for monitoring and alerting, Amazon VPC for secure network isolation, AWS IAM for authentication and authorization, Amazon EventBridge Pipes for event routing, and Amazon Athena for querying processed data
- These integrations eliminate the need for complex custom connectors and provide a unified, managed experience across the entire data streaming pipeline

Onboarding and Offboarding

Getting started with Amazon MSK involves a straightforward process through the AWS Management Console, CLI, or APIs. Customers begin by creating a VPC and subnets, then provision an MSK cluster by selecting either provisioned or serverless deployment options:

- The service provides step-by-step tutorials including cluster creation, IAM role configuration for topic access, client machine setup, and topic creation
- MSK supports standard Apache Kafka clients and tools, enabling existing applications to connect with minimal changes
- The getting started guide includes producing and consuming data samples, monitoring setup through CloudWatch, and cleanup procedures
- AWS provides comprehensive documentation, hands-on labs, and migration guides to accelerate adoption
- For organizations with existing Kafka deployments, MSK offers migration tools and best practices to transition workloads seamlessly

Getting Started Guide: <https://docs.aws.amazon.com/msk/latest/developerguide/getting-started.html>

Data Removal

Data removal from Amazon MSK requires deleting topics, consumer groups, and cluster resources through the Kafka APIs or AWS Management Console. When offboarding, customers should first stop all producers and consumers, then delete topics containing sensitive data using Kafka administrative commands. The MSK cluster itself can be deleted through the AWS Console, CLI, or API, which removes all associated broker nodes, ZooKeeper instances, and cluster metadata. Storage volumes are automatically deleted when the cluster is terminated. For compliance requirements, customers should ensure proper data retention policies are configured and verify complete data deletion through AWS CloudTrail logs.

Backup/Restore and Disaster Recovery

Amazon MSK provides disaster recovery capabilities primarily through MSK Replicator for cross-region replication and built-in multi-AZ deployment for regional resilience. The service automatically replicates data across Availability Zones within a region for high availability. For cross-region disaster recovery, MSK Replicator continuously replicates topics to secondary regions with configurable replication settings. MSK integrates with AWS Backup for operational backup scenarios, though Kafka's distributed nature makes traditional backup less critical. Recovery objectives are supported through automated failover mechanisms and the ability to quickly provision replacement clusters from replicated data in alternate regions.

Backup Capabilities

Amazon MSK does not integrate directly with AWS Backup service due to Kafka's distributed streaming architecture where data durability comes from replication rather than traditional backups. Instead, MSK provides data protection through automatic multi-AZ replication within clusters, configurable topic-level replication factors, and MSK Replicator for cross-region data copying. The service focuses on continuous data

availability and streaming resilience rather than point-in-time backup recovery, which aligns with Apache Kafka's design principles for handling streaming data workloads.

Disaster Recovery

Amazon MSK does not directly integrate with AWS Elastic Disaster Recovery (DRS) as it is designed for streaming data rather than traditional disaster recovery scenarios. Instead, MSK provides disaster recovery through MSK Replicator for cross-region replication, multi-AZ deployment for zone-level resilience, and automatic cluster recovery capabilities:

- Organizations implement disaster recovery by deploying MSK clusters in multiple regions, using MSK Replicator to maintain data synchronization, and implementing application-level failover logic to redirect traffic during regional impairments
- This approach is specifically designed for streaming workloads that require continuous data flow rather than point-in-time recovery

Recovery Objectives

Amazon MSK helps customers meet Recovery Point Objective (RPO) and Recovery Time Objective (RTO) requirements through multiple mechanisms. For RPO, MSK provides configurable replication factors and MSK Replicator for cross-region replication with minimal data loss potential. Multi-AZ deployment ensures sub-second RPO within regions through synchronous replication. For RTO, MSK offers automatic broker replacement and fast cluster recovery typically within minutes. MSK Replicator enables quick failover to secondary regions, and the serverless option provides instant scaling without capacity planning delays, supporting aggressive RTO targets for business-critical streaming applications.

Service Constraints

Service Quotas: <https://docs.aws.amazon.com/msk/latest/developerguide/limits.html>

FAQ: <https://aws.amazon.com/msk/faqs/>

Technical Requirements

Service Documentation: <https://docs.aws.amazon.com/msk/latest/developerguide/what-is-msk.html>

User Guide: <https://docs.aws.amazon.com/msk/latest/developerguide/getting-started.html>

API Reference: <https://docs.aws.amazon.com/msk/1.0/apireference/>

Pricing Overview

Please see the AWS UK G Cloud 15 Pricing Document accompanying this service in the Digital Marketplace.

Public Pricing: <https://aws.amazon.com/msk/pricing/>

Cost Optimization

Amazon MSK offers unique cost optimization through Express brokers that provide higher throughput and faster scaling at optimized costs compared to Standard brokers. MSK Serverless eliminates idle resource costs by charging only for actual usage without minimum capacity requirements:

- Tiered storage automatically moves older data to cost-effective storage tiers while maintaining access performance
- The service provides granular pricing models including separate charges for compute, storage, and data transfer, enabling fine-tuned cost control
- MSK Connect and MSK Replicator offer usage-based pricing for data integration and replication needs
- Organizations can optimize costs by selecting appropriate broker types, configuring retention policies, and leveraging auto-scaling capabilities to match capacity with demand patterns

Savings Considerations

Primary cost savings strategies for Amazon MSK include selecting MSK Serverless for variable workloads to eliminate over-provisioning costs and using Express brokers for predictable workloads requiring higher performance. Implementing proper data retention policies and leveraging tiered storage can significantly reduce storage costs over time:

- Organizations should optimize partition counts and replication factors based on actual durability requirements rather than using maximum settings
- Cross-region data transfer costs can be managed by strategic placement of clusters and careful planning of MSK Replicator usage
- Regular monitoring of CloudWatch metrics helps identify underutilized resources and opportunities for downsizing
- Consolidating multiple small clusters into larger, more efficient deployments can reduce per-broker overhead costs while maintaining isolation through topic-level access controls