**AMAZON MSK**

**Amazon Managed Streaming for Apache Kafka (Amazon MSK)**

**What is Amazon MSK?**

Amazon MSK is a fully managed service that makes it easy for you to build and run applications that use [Apache Kafka](https://aws.amazon.com/streaming-data/what-is-kafka/) to process streaming data. Apache Kafka is an open-source platform for building real-time streaming data pipelines and applications. With Amazon MSK, you can use native Apache Kafka APIs to populate data lakes, stream changes to and from databases, and power machine learning and analytics applications.

**Benefits**

1. **Fully Managed**

The responsibility of managing Apache Kafka is entirely taken by Amazon MSK making us developers focus on development and migration of code. Amazon MSK manages the provisioning, configuration, and maintenance of Apache Kafka clusters and Apache ZooKeeper nodes for you. Everything from metrics maintenance to managing kafka cluster this responsibility I taken over by Amazon MSK.

1. **Elastic stream processing**

Amazon MSK can fully manage the scaling of data streams without develop having to worry about the increase load which might impact performance. As an advantage it is an opensource streaming processing framework.

1. **Highly available**

Amazon MSk ensures multi AZ replications within an AwS region thereby making it highly available system. If a cluster comes down it automatically makes another replica cluster available.

1. **Highly secure**

Amazon MSK provides multiple levels of security for your Apache Kafka clusters including VPC network isolation, AWS IAM for control-plane API authorization, encryption at rest, TLS encryption in-transit, TLS based certificate authentication, and supports Apache Kafka Access Control Lists (ACLs) for data-plane authorization. So one does not have to worry about the data security that is being transferred over network.

1. **Fully compatible**

It is fully compatible with build in tools such as [MirrorMaker](https://cwiki.apache.org/confluence/pages/viewpage.action?pageId=27846330), [Apache Flink](https://flink.apache.org/), and [Prometheus](https://prometheus.io/).

**CASE STUDY Poshmark**

Poshmark is a social commerce marketplace where people in the United States can buy and sell new or used clothing, shoes, and accessories. Through technology, their mission is to build the world’s most connected shopping experience, while empowering people to build thriving retail businesses.

They have opted to Kafka because they have a growing business, they need to fulfil the dynamically increasing customer demand. Amazon MSK has made it easy to setup, maintain and scale Kafka clusters, enabling an end-to-end ingestion pipeline supported by a fully managed service. Now the real time increasing customers can be addressed without having to worry about the kafka cluster servers.

**CASE STUDY Compass**

Compass is a real estate platform, pairing the industry’s top talent with technology to make the search and sell experience intelligent and seamless. It supports the entire buying and selling workflow. Apache Kafka provide its agents with quick access to fresh and accurate data from hundreds of real estate data sources nationwide. And as we all are aware buying and selling of property will not go down anytime soon, in fact it will increase with increasing population and hence the growing business. They need to support the expanding customer needs. As stated by their CTO Joseph Sirosh “By using Amazon MSK we no longer need to spend effort on ensuring data durability, cluster availability, and scalability, and can now focus on building applications. Amazon MSK relieves our team from the burden of operating and maintaining Kafka and at the same time, we experience less downtime as our clusters are fully-managed.” They can now fully focus on their development perspective while the entire management of Kafka clusters is taken care by Amazon MSK.

**CASE STUDY ZipRecruiter**

ZipRecruiter is a job search platform that supports many customers and direct them to available job opportunities in their localities. They use AI driven technologies to search for available jobs.

They are using it for logging their events which is approx. 6 billion per day transactions. Hence maintaining and managing the kafka clusters is a prominent task and to focus more on product development for providing their customers with a rich experience they turned to Amazon MSK for managing kafka clusters.

**CASE STUDY Delhivery**

This is India’s one of the biggest Courier delivery service. And considering the population of India it is a huge business. To ensure and manage the customer request earlier they were using kafka clusters, but it adds on to additional management task. Now they have moved on to Amazon MSK for easier management of kafka clusterd without giving an overhead. As stated by their CTO Kapil Bharti “Delhivery uses Apache Kafka as a critical component in their pipeline to collect, store, and logically swivel metadata of First, Mid, and Last mile operations. As we scaled our logistical operations, significant effort was spent managing the infrastructure complexities within our Apache Kafka clusters. Amazon MSK not only helped us offload infrastructure overhead, but we also maintained high throughput and performance for our business-critical metadata pipelines in a more secure and reliable manner.” Amazon MSK has significantly help them in releasing from management task.