The Unmatchable ROI of Managed Apache Kafka® Services

Overview

We live in a world where data and data analytics is growing at an unimaginable pace. To keep up with this growth, real-time data streaming is now becoming the standard technology method through which organizations receive and process that data.

Apache Kafka® is quickly becoming the technology of choice for building real-time streaming data pipelines. Kafka is now considered to be the leading streaming and queuing technology for large scale, always-on applications, and event-driven organizations. Apache Kafka comes with built-in features of horizontal scale, high-throughput, and low-latency.

:instaclustr
Now part of **Spot by NetApp**

Why Managed Services?

Managing a traditional dataset and managing the movement and ingestion of event data are two very different and separate activities. While messaging queues have been in use for a long time and are being used for point-to-point communication between applications that need to share data continuously, traditional messaging streaming technology is difficult to scale. This problem gave birth to Kafka. Kafka is different from other messaging systems—it is a perfect blend between file system, messaging system, and database. The technology architecture and implementation of Kafka make it highly reliable and highly available, enabling stream processing applications to utilize geographically distributed data streams.

Kafka streaming needs to have an agile and flexible deployment model that allows for changes to pipeline configuration, implementation, and management in order to maintain the timely delivery of data. When deploying a crucial technology such as Kafka there are expectations of continuous availability, high performance, and rapid scale. The reality is that meeting these expectations can be challenging through the complexity of the technology and the management ecosystem.

When deciding on deploying Apache Kafka as a crucial part of your technology stack you have to to consider capability management requirements. Is it the right decision to build the necessary in-house expertise, or is the right strategy to outsource the responsibility to a Managed Service Provider (MSP)?

This is a major decision that can have serious consequences. An MSP has already made the necessary investments to build world-class expertise and infrastructure so that you can focus on building world-class applications. With the emergence of new technologies, managed services today have moved away from the early days of remote monitoring and management of servers, networks, and infrastructure, to be more of a strategic business, imperative for improving operations, addressing risks, and reducing operational costs. With the rising popularity of MSP, it is a natural consideration for many organizations either looking for mission-critical data deployment or choosing a fast, scalable, and distributed streaming platform to build the next generation application.

To help you make the decision to either build in-house capability for Kafka operations or to outsource, we believe there are three broad considerations: **cost, time, and business risk.**

Financial Considerations

Thousands of companies are using Apache Kafka to build streaming applications and real-time data pipeline. Despite its popularity, Kafka is difficult to deploy without engaging an expert. To manage Kafka within your team, you will need to add the critical skills for managing this technology.

The knowledge of fundamental concepts of Kafka development like topics, partitions, replication, brokers, producers, consumer groups, and more is just a beginning. Your team will need to integrate Kafka into your existing ops infrastructure, develop and test operational procedures, and then learn in production as unforeseen issues arise.

To achieve all this internally, your company should be willing to invest, at a minimum, $400,000 to $500,000 a year which accounts for building the necessary expertise, infrastructure, and processes. Alternatively, you could outsource this critical operation to those that have already made such an investment. The chart below shows that, for the first year, a company can expect to spend at least $400,000 to $450,000 to operationalize Kafka. You would also be spending nearly as much for each year after.

Demand for Kafka expertise currently far outweighs the supply and you will need to account for higher recruiting fees. Plus, each year, you can expect the salaries for your core Kafka team to grow by 10% or more just to keep them in your company and not succumb to poaching.

Compare this with the cost of outsourcing to a MSP that charges an average of $60,000 to $70,000 per year to expertly build, deploy, manage, scale, and monitor your Kafka pipeline—*which is less than one-fifth of the cost*.

Year 1	Base ($)	Burdened ($)
Kafka Expert	200,000	236,000
Support 1	90,000	106,200
Recruiting Cost	58,000	200,000
Equipment Cost	30,000	200,000
Other	5000	200,000
Total First Year		435,200

Timeline Considerations

As has been indicated in the above section, there are few Kafka experts who have completed enough Kafka implementations to have gained sufficient expertise. A typical timeline for in-house deployments run between 4 and 6 months, the time it takes to fully

design and implement a product-ready Kafka environment.

Compare that with an MSP who can get you up and running with experts in a matter of 2-3 weeks—again a high magnitude difference!

Unlike the cost consideration—which is mostly operational in nature—this has strategic implications. In today's ever-changing customer needs, the ability to develop rapidly or to quickly release new features that can keep pace with growing customer needs must be matched with an ability to monitor 24x7 and rapidly make necessary changes to your Kafka deployment. Without continuous monitoring of the ecosystem, your very own agile development competencies could bring you to a crawl.

Overall Business Considerations

A third, and perhaps the most important, consideration is the risk of getting it all wrong. Apache Kafka can be an inscrutable beast for the uninitiated. If you dive in and just try to "wing it," you are likely to make mistakes. While initial development with Kafka can be straightforward, operationalizing and dealing with issues in production is a much more complex undertaking. Working with a trusted, experienced partner reduces risks and ensures your project stays on the right path.

Summary Comparison

After going through the 3 important factors of cost, timeline, and overall business risk between deploying in-house, and outsourcing to an expert Kafka Managed Services Provider, you can decide for yourself what works for you.

The Right Provider

How to Choose an MSP for Apache Kafka

1) Integration Ready and Effortless Deployment

An MSP should help you with simple provisioning of your Kafka cluster so that you don't have to work too hard. Kafka has an active ecosystem that helps for easier integration of different applications and service with Kafka. The right partner should be willing to make a continuous investment in products and tools to ease deployment and management so that you get fault-tolerant and durable message collection and processing available on the cloud of your choice.

2) Experience With Open Source Technology and Industry Recognized Security Protocols

Kafka is rarely if ever the sole open source product in an application stack so you need to ensure that your managed service player is an expert for open source technologies. The provider should be able to bring expertise in the various aspect of data management —store, search, stream, analyze, and explore so that you can use them for one or all your requirements as you build your application.

You would also need to ensure that the managed service platform you opt-in for has an external audit assurance that ensures that the security and availability of your cluster is to industry best practice. You would need to look for the security program that includes security and availability considerations in the design. Additionally, the cluster should be continually reviewed, tested, and monitored and the provider should look at providing a suitable response in case of any troubles.

3) Experienced Team With 24x7 Expert Support

Before zeroing down you may want to look if they have enough people on the ground to manage your Kafka cluster. One of the advantages of having a managed player on your side is that you have 24x7 support irrespective of the location of your business.

4) Transparent Pricing Model

Open source means no licensing cost. Your managed service platform should be able to provide you with a pricing model for both on-premises and off-premise integration. The pricing should be transparent right from the beginning till the end so that you are aware of the cost you would pay as your business grows over time.

5) Required Service Level Agreements

Ensuring the Service Level Agreements (SLAs) based on the size of a cluster would help ensure a consistent level of performance and availability and peace of mind.

Kafka can solve a lot of problems and the right managed service provider can help you get the performance and scaling benefits of Kafka without having to do it alone.

instaclustr
Now part of Spot by NetApp
www.instaclustr.com   info@instaclustr.com   @instaclustr   5

About Instaclustr

Instaclustr helps organizations deliver applications at scale through its managed platform for open source technologies such as Apache Cassandra®, Apache Kafka®, Apache Spark™, Redis™, OpenSearch®, PostgreSQL®, and Cadence®.

Instaclustr combines a complete data infrastructure environment with hands-on technology expertise to ensure ongoing performance and optimization. By removing the infrastructure complexity, we enable companies to focus internal development and operational resources on building cutting edge customer-facing applications at lower cost. Instaclustr customers include some of the largest and most innovative Fortune 500 companies.