

Enterprises are looking for a one-stop managed platform solution for open source data technologies. With the data growth of AI and cloud-native applications, IT teams must simplify their underlying data infrastructures to support business objectives.

Solving Open Source Complexity with a Managed Data Infrastructure Platform

June 2024

Written by: Jevin Jensen, Research Vice President, Infrastructure and Operations

Introduction

The adoption rate of open source by enterprises and DevOps teams is rising. A pivotal point of reference is the Open Source Initiative's (OSI's) recent survey, which revealed that over 67% of enterprises increased their use of open source software (OSS) in 2023, with many reporting significant growth. IDC's research found that 96% of companies now utilize open source. This surge in open source adoption and in multiple public cloud providers increases complexity for IT. In fact, companies tell IDC they utilize a median of three public hyperscalers.

The potential risks of overlooking the growth of OSS-based applications are becoming more apparent as adoption increases. IDC's research, as Figure 1 shows, indicates that 67% of businesses consider OSS applications to be important or critical to their business. A prime example is the emergence of the AI-enabled applications that enterprises are developing. The growth of new applications is accelerating as DevOps teams continue to support and expand upon the digital transformation many companies initiated several years ago. This rapid acceleration is the second important factor that leaders must be vigilant about. The number of new applications is increasing at an unprecedented rate, and it is easy to overlook the resilience and scalability of mission-critical OSS-enabled applications.

The exponential growth of data and applications is the final factor significantly impacting today's businesses. Data growth is compounding, with IDC estimating growth of 23% in 2024. In the future, AI-enabled applications will likely require even more data to enhance the accuracy of their models, and they will retain data for more extended periods to drive improved responses. The global expansion of data continues to drive the need for better management and governance. New data technologies, such as Cassandra, Kafka, Spark, and OpenSearch, are

AT A GLANCE

KEY STATS

According to IDC research:

- » 96% of businesses surveyed utilize open source. Growth is accelerating.
- » Data in 2024 will grow by an expected 23%, reaching 163ZB.
- » 67% of businesses ranked OSS as critical or important to their line of business.
- » In a survey on their use of open source, DevOps teams reported high adoption rates because of cost and flexibility advantages.
- » New container-based applications and AI use cases are driving enterprise data rates higher.

WHAT'S IMPORTANT

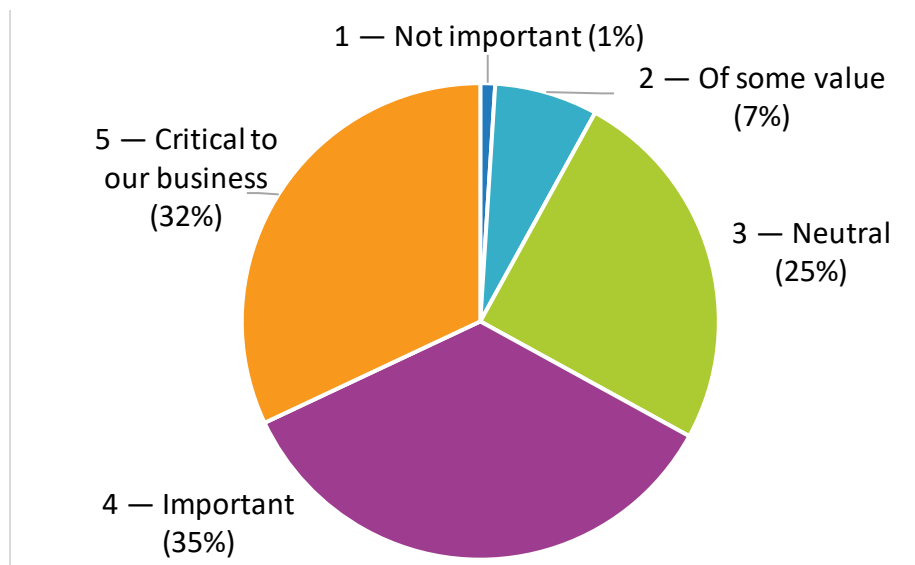
High adoption of open source solutions means IT teams need a standard way to handle data infrastructure. Higher data consumption and retention means enterprises need a data strategy. Modernizing applications with AI or deploying new cloud-native apps drives the need for a managed solution to handle data infrastructure.

emerging and require a platform approach to support these OSS solutions. Many companies are grappling with the ongoing workforce skill set challenges, causing SRE and CloudOps teams to struggle to maintain this data and new applications.

Today, these three trends are converging: the growth of OSS adoption, the dramatically increasing number of new applications, and the continued expansion of data/data technologies. This confluence means IT teams must act today to deliver business value and keep mission-critical OSS applications resilient. However, limited staffing and adaptation skills pose a challenge.

FIGURE 1: **Importance of Open Source Software to Business**

Q How critical is open source software to your company's operations?



Note: Scores are based on a scale of 1 to 5, where 1 = not important and 5 = critical to our business.

Source: IDC, 2024

Benefits

One way to address these three converging factors along with workforce challenges is to leverage a managing platform for OSS database technologies. Such a platform will typically be delivered via software as a service for quicker implementation. The OSI reports that databases and data technology are enterprises' most important category for new OSS investments. A managed data platform offers scalability, resiliency, and support for mission-critical applications. Following are some of the other advantages and benefits of a managed OSS data platform:

- » **Managed services for open source technologies:** OSS provides a single platform to access powerful open source data tools without adding complexity.

- » **Expansive support and environment monitoring:** To be proactive, enterprises need 24 x 7 support and monitoring services for mission-critical data systems.
- » **Modern data architectures:** Being proactive means leveraging event-driven systems that scale easily and quickly.
- » **Cybersecurity and regulatory compliance:** Encryption and comprehensive security are critical to secure data and meet regulatory requirements such as privacy. Areas include GDPR, SOC 2, ISO 27001, ISO 27018, HIPAA, and PCI.
- » **Platform scalability and performance:** Data and application growth require the large-scale handling of workloads with automation.
- » **Analytics:** SRE, CloudOps, and DevOps teams need dashboards and reporting to understand the performance and health of their data environment for planning and incident resolution.

Digital enterprises require high availability and resiliency. If data services are not maintained at the highest levels, they impact the customer experience and revenue of the digital business. Digital enterprises that leverage a managed data platform can see the following additional benefits:

- » **Integration:** Improved support for multiple clouds and various OSS products and additional integration with other solutions are essential to maintaining flexibility.
- » **Cost efficiency:** A hosted, managed solution means existing staff can support a growing business and keep capital costs low. Open source is already one of the most cost-effective options for businesses.
- » **Simplified operations:** Automated patching, maintenance, and backup of large data sets allow IT operations teams to focus on value-added activities.

Considerations

Many companies are interested in managed platform solutions for open source data technologies. These primarily focus on streamlining data management in open source solutions such as Cassandra, Kafka, PostgreSQL, Spark, OpenSearch, and Valkey. A platform approach offers managed services for deploying, managing, and scaling these open source technologies in the cloud. Platforms often include provisioning, monitoring, maintenance, and support. By allowing organizations to focus on building and supporting applications rather than the underlying data infrastructure, teams can deliver higher business value. Enterprises aim to bolster their cloud-native data management capabilities by leveraging a provider's expertise in open source data technologies and managed services.

Organizations that want the latest and greatest features from open source may need to wait for the managed provider to finish testing and upgrades. This third-party testing adds stability and peace of mind for many companies, but some cloud-native companies should be especially aware of this consideration. In addition, IT leaders must identify, understand, and appropriately budget for the platform's cost drivers. Engage FinOps teams as necessary to plan these costs and chargebacks.

Some managed platforms offer a range of supported options, but not every option is available on every open source solution. Enterprises must review options carefully while verifying data sovereignty requirements for their operating

countries before selecting a managed data platform partner to confirm they meet the organization's needs. Finally, data workload migrations like cloud migration can be complex and time-consuming. Speaking to other customers and asking about migration tools and time frames are critical.

Trends

Essential trends for enterprises include an increase in data growth, which will continue and possibly rise. Executives are prioritizing investments in open source data technologies to increase the business value of this data expansion. New applications, especially ones that embed AI capabilities, will increase this growth rate. IDC predicts the total global data for 2024 at 163ZB.

Skilled workforce issues continue to challenge IT teams. Enterprises report that filling a skilled OSS or data technology role may take months. Companies looking to fill critical roles in development and AI will struggle to fill data infrastructure and OSS positions simultaneously. Line-of-business owners are pushing for new applications to continue their journey toward digital business, so considering a managed OSS data platform is an increasingly common strategy.

Conclusion

A confluence of factors will soon challenge IT's ability to support business objectives. These factors include the increasing usage of OSS technology, the brisk growth of new applications (including AI), and the enterprise's expanding growth rate of data/data technologies. This confluence, coupled with the mission-critical nature of OSS and skilled workforce issues, means IT leaders must act today. A comprehensive strategy to address these factors should include a managed OSS data platform in which third-party experts provide support, testing, resiliency, and scalability. This approach frees IT teams to increase business value without sacrificing quality or agility.

IT leaders must prepare now for a confluence of factors. A managed OSS data platform can address many of these factors today.

About the Analyst



Jevin Jensen, Research Vice President, Infrastructure and Operations

Jevin Jensen is the research vice president covering IDC's Intelligent CloudOps Markets service. He covers infrastructure as code, GitOps, IT infrastructure automation, cloud cost transparency, FinOps, hybrid cloud/public cloud/multicloud management platforms, and edge management.

MESSAGE FROM THE SPONSOR

Instaclustr helps organizations deliver applications at scale through its managed platform for open source technologies such as Apache Cassandra, Apache Kafka, Apache Spark, Valkey/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.

Learn more: www.instaclustr.com.

IDC Custom Solutions

The content in this paper was adapted from existing IDC research published on www.idc.com.

This publication was produced by IDC Custom Solutions. The opinion, analysis, and research results presented herein are drawn from more detailed research and analysis independently conducted and published by IDC, unless specific vendor sponsorship is noted. IDC Custom Solutions makes IDC content available in a wide range of formats for distribution by various companies. A license to distribute IDC content does not imply endorsement of or opinion about the licensee.

External Publication of IDC Information and Data — Any IDC information that is to be used in advertising, press releases, or promotional materials requires prior written approval from the appropriate IDC Vice President or Country Manager. A draft of the proposed document should accompany any such request. IDC reserves the right to deny approval of external usage for any reason.

Copyright 2024 IDC. Reproduction without written permission is completely forbidden.

IDC Research, Inc.
140 Kendrick Street
Building B
Needham, MA 02494, USA
T 508.872.8200
F 508.935.4015
Twitter @IDC
blogs.idc.com
www.idc.com