IDC

White Paper

Empowering Developers Through Cloud Services

Sponsored by: Red Hat

| Stewart Bond | Al Gillen | Lara Greden | Amy Machado |

September 2021

IDC OPINION

The enterprise is under tremendous pressure when it comes to modern application deployment. The industry faces a shortage of developers while seeing a surge in requests for complex application development. Resources are already squeezed, and companies want to free up their development teams to focus on the higher-order design decisions and execution to deliver business outcomes.

Today's applications require and consume services such as data streaming, data science, machine learning (ML) capabilities, and API management; these are some of the most complex to configure at scale, and they need to work together seamlessly. Ease of access, simplicity of use, and consistent deployment across hybrid cloud platforms are key for empowering developers to productively create complex applications that scale and run where they are needed.

While open source solutions like Apache Kafka for data streaming may seem easy to implement and use in development and test scenarios, complexity enters the scene when applications move to production, as the Kafka implementation needs to scale and be resilient and fault tolerant. This is where many developers who have created solutions on top of Kafka can struggle.

Infrastructure was the first layer of opportunity to abstract for developers in platform-as-a-service (PaaS) offerings. Application and data services are quick to follow, with the same set of benefits to deliver – abstraction that delivers consistency, helping enable high-throughput applications to run securely and reliably. With a PaaS solution that includes application and data services, organizations no longer need to commit valuable resources to deploying, configuring, and running those underlying services. This is especially important with open source technologies, which developers love to use because of the innovation benefits. The cloud services model enables organizations to shift those valuable resources to making software that competitively differentiates, brings in revenue, and improves business operations – empowering developers to do more of what they want to be doing.

SITUATION OVERVIEW

Modern applications need to work across the hybrid cloud (i.e., a mix of infrastructures including on premises, private cloud, public cloud, and edge). To handle that increasingly complex mix efficiently and at the scale required, development and IT operations teams need consistent deployment environments that support applications across those environments.

Furthermore, open source software innovations have significantly contributed to the growth in use of the cloud, enabling a faster rate of innovation and improved developer productivity, according to IDC research (see Figure 1). For example, Kubernetes orchestration and container packaging are important aspects of deploying and managing modern workloads in a hybrid cloud environment. While innovative developers are advocates for use of open source technologies, often themselves contributing firsthand to their development, roadblocks often come up in production. DevOps and IT operations teams, charged with the complex task of deploying, configuring, and running open source technologies as part of production-grade applications, can block their usage and/or struggle to maintain the reliability and uptime expected by customers and the business. Making sure that open source technologies meet enterprise-grade standards for uptime and reliability is key to empowering developers to make innovative software that meets business needs.

FIGURE 1

Top Benefits Expected from Open Source Software in the United States

Q.   What are the benefits your organization gets or expects it will get from using open source software?
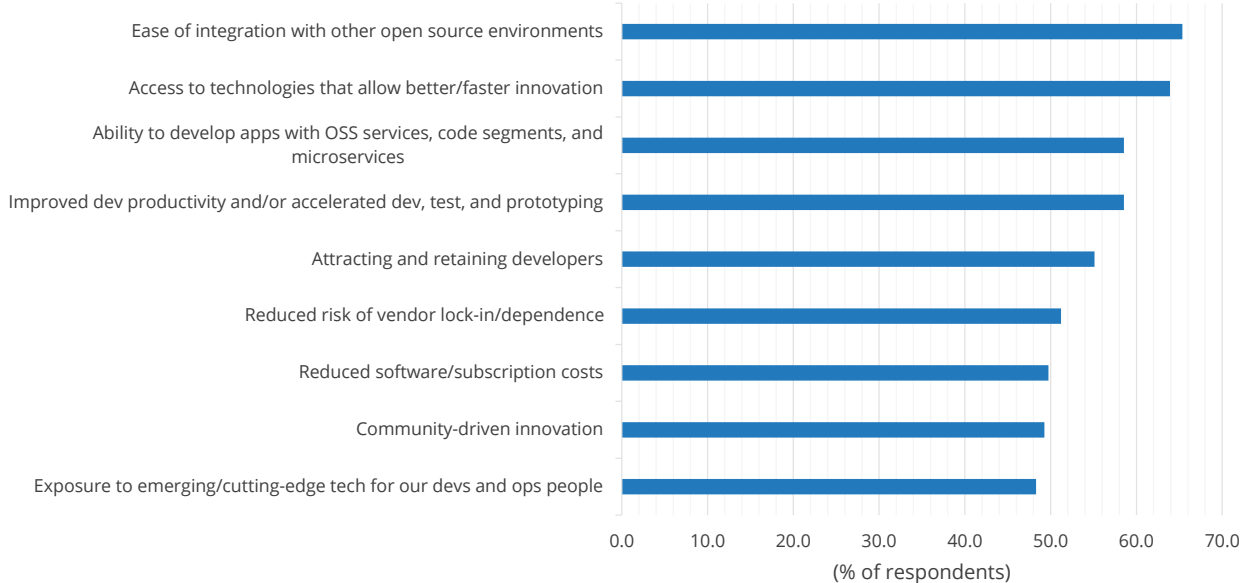
n = 401

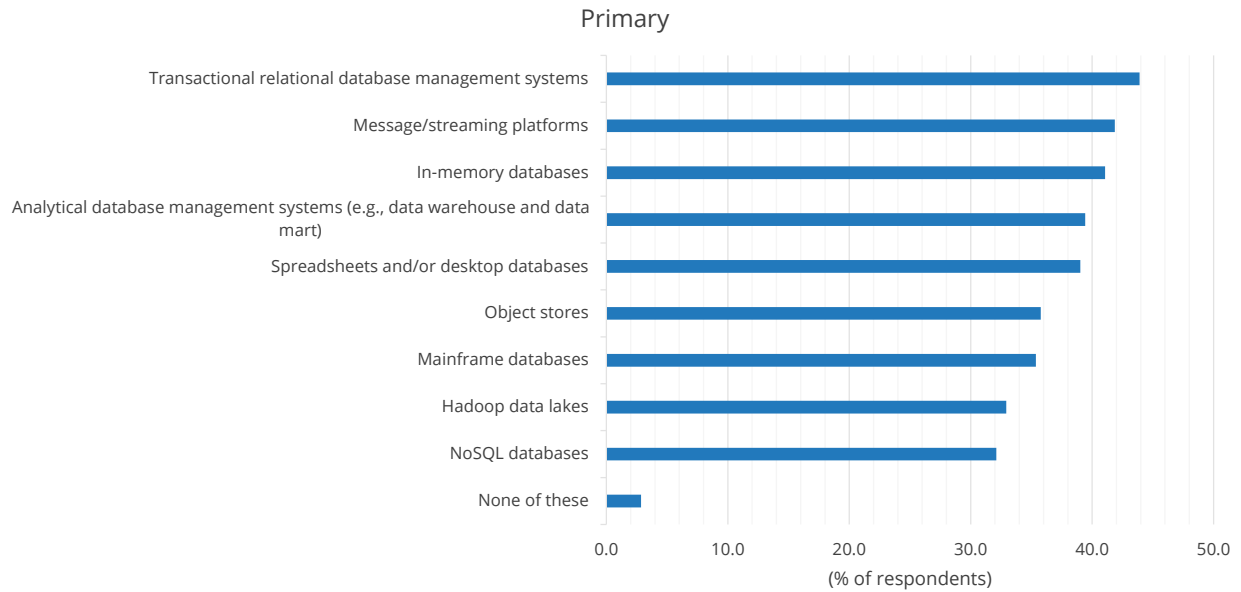Base = U.S. companies that use or plan to use OSS in their environments

Source: IDC's *Open Source Software Use Survey,* March 2021

One open source area particularly relevant today is data streaming. There is a major shift in the industry toward and expectation for immediate, relevant experiences and communication. Real time feels like the norm, not the exception. Streaming platforms are of high value and usage in the enterprise data environment, and developers are under pressure to incorporate real-time streaming data in a broad set of applications. Organizations overall are then challenged with incorporating these new technologies reliably and consistently (see Figure 2).

FIGURE 2

The Importance of Data Streaming Services

Q. *Thinking of your data environment, select all the data management technologies in scope for the DataOps solutions that are being used to store data in public cloud.*

Primary

| Category	% of respondents
Transactional relational database management systems	~44
Message/streaming platforms	~42
In-memory databases	~41
Analytical database management systems (e.g., data warehouse and data mart)	~39
Spreadsheets and/or desktop databases	~39
Object stores	~36
Mainframe databases	~35
Hadoop data lakes	~33
NoSQL databases	~32
None of these	~3

(% of respondents)

n = 401

Source: IDC's *DataOps Survey*, December 2020-January 2021

IDC research indicates that Kafka holds a predominant position in the data management technology landscape for sharing event data across multiple applications and environments. As an open source technology, it has become popular with application and data engineers that have event distribution requirements in their applications and data products. Having been originally developed at LinkedIn, Kafka is built for scale, performance, and resiliency, but achieving these qualities of service requires complex configuration and management of Kafka clusters. As such, delivering Kafka at scale can slow down the velocity of the development and operations teams in CI/CD processes. Developers love the abilities Kafka brings to creating event-driven applications, but the complex administration required in self-managed instances brings friction that impedes the ability for developers to use Kafka. PaaS solutions help remove that friction and enable innovation (see Figure 3).

FIGURE 3

Primary and Secondary Event Movement Technologies

*Q.   What are the primary technologies used for real-time or event data movement
         and management?*

Q.   Select any technologies being used for real-time or event data movement and management.
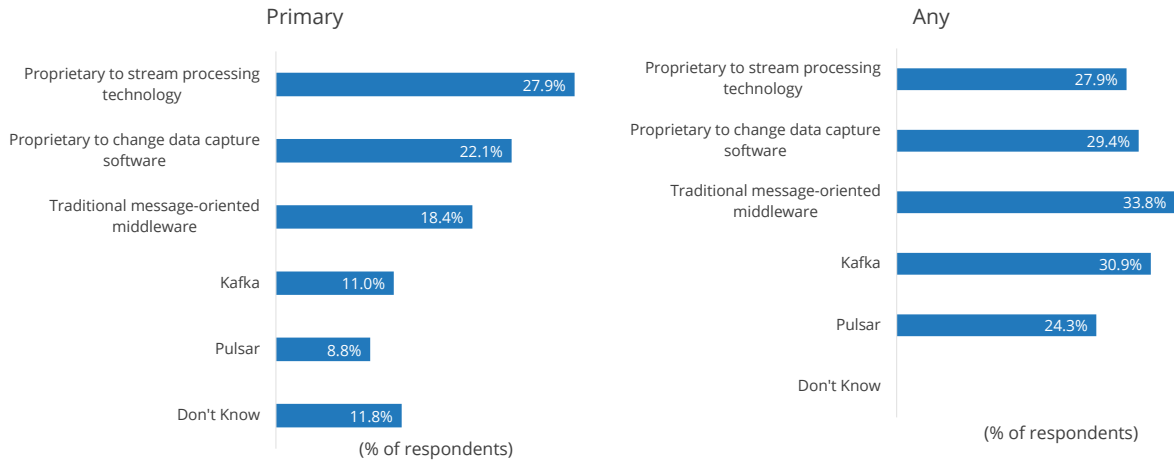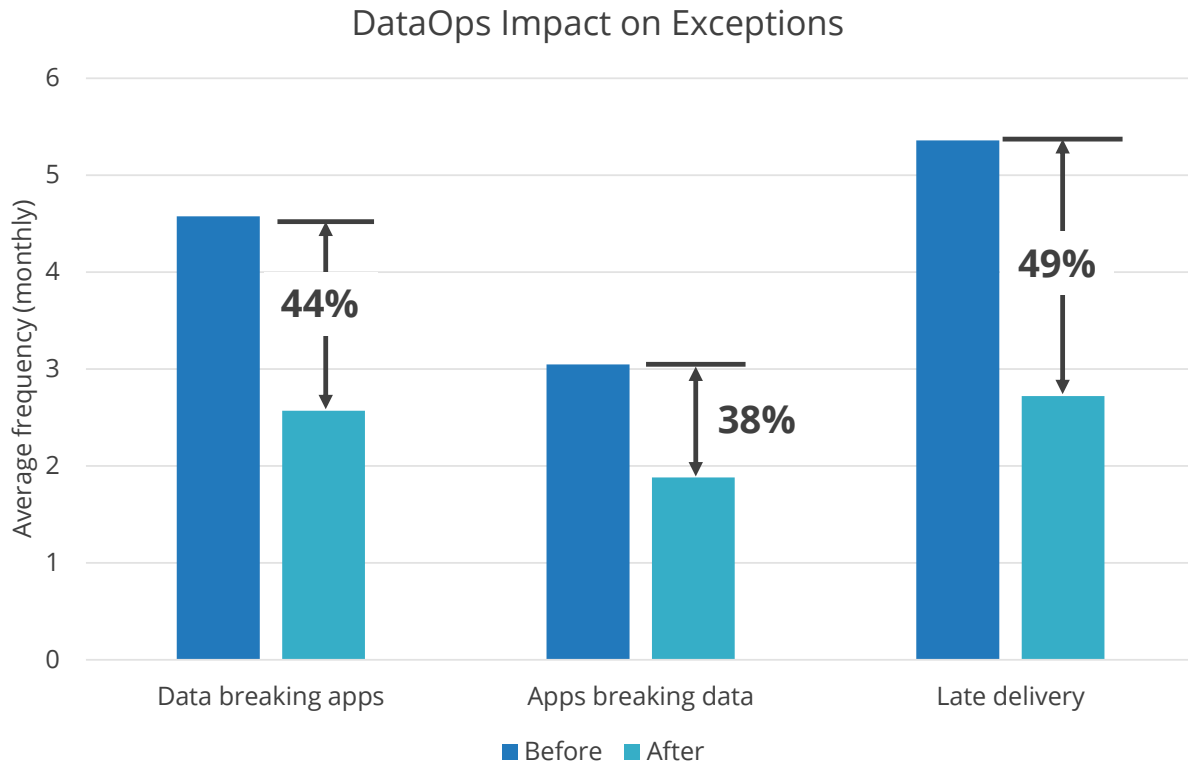

n = 401

Source: IDC's *DataOps Survey*, December 2020-January 2021

To manage the complexity of modern data environments, in support of both operational and analytical workloads such as data science and AI/ML, organizations have been applying CI/CD methods of DevOps to data and adding statistical process control to monitor the health of data in pipelines. This approach has become known as DataOps, a combination of technologies and methods with a focus on quality for consistent and continuous delivery of data value. DataOps is connecting producers and consumers of data to extract value for any data-driven outcome, including analytical models, dashboards, algorithmic results, and operational or strategic business decisions. DataOps is still a relatively new discipline, but organizations that have adopted it are reaping the benefits, as illustrated in Figure 4. PaaS solutions that help manage the underlying complexity of running modern data environments empower developers to keep focusing more time on building software, as opposed to fixing problems and delivering late.

FIGURE 4

DataOps Benefits

Q.   On average, how many times were the following exceptions or errors occurring monthly in data and analytics deliverables before DataOps solutions were implemented?

DataOps Impact on Exceptions

n = 401

Source: IDC's *DataOps Survey,* December 2020–January 2021

FUTURE OUTLOOK

A growing set of value-added services that extend a PaaS environment are the next layer of opportunity for streamlining your development toolchain, including API management, event streaming data, and data science capabilities. From a technical perspective, these services and the applications they support are not discrete buckets, but rather intertwined ecosystems. The ability to abstract the integrations and make them an "already taken care of" part of the infrastructure is a leading set of capabilities that PaaS solutions can offer. Empowering developers with those capabilities frees them up for more rapid development and experimentation — doing more of what they love.

Enterprises will continue to increase their use of open source software, both commercially supported and community supported. Developers themselves are driving the growth for commercially supported open source software, including development tools and data management. Because commercially supported open source software delivered via PaaS means that the vendor delivers the ability to scale

and be resilient and fault tolerant, it brings more freedom to developers to focus on the higher-order value of creating applications that deliver value.

Furthermore, the economic benefit of cloud and the associated managed cloud service, or as a service, pricing model allows organizations to invest less up front in infrastructure and pay less overall in operation. At the same time, they receive the SLA benefits of uptime and the ability to devote developers to more strategic roles.

If there is one constant, it is change, and the rate of change has been getting faster. Streaming data solutions enable the ability for applications and analytics, including AI and ML, to operate in real time – on what is happening now, not what happened a few hours ago, yesterday, or last month. IDC expects developer demand to incorporate new technologies, such as streaming data solutions, into their PaaS environments to continue to increase.

CHALLENGES AND OPPORTUNITIES

For developers to create applications for multicloud environments, the major challenge lies in inconsistency across complex technology ecosystems. The opportunity therein lies in abstraction of those complex technology ecosystems to reduce friction for developers and enable high availability of applications in production. Further:

- **Challenge:** Supporting infrastructure, especially the cutting-edge technology found in open source software, takes resources away from other development and DevOps activities.
- **Opportunity:** Managed cloud services allow access to cutting-edge technology without having to maintain local management and administration. Using managed services allows your talent to focus on developing for business outcomes and not have to be experts in managing the requisite infrastructure. Given the complexities of managing and administering Kafka and Kubernetes across the entire application estate, managed services for these technologies help organizations reduce risk and allow developers to focus on what they do best: use cutting-edge technology to get great solutions to market faster.
- **Challenge:** The complete architecture to deliver cloud-native, event-driven, and data science-enabled solutions includes other technologies such as Apache Camel, Debezium, Apicurio, and Quarkus and ML notebooks and ML libraries such as JupyterLab.
- **Opportunity:** Streamline the integration of other services with the Kafka service, API management, and ML models across multiple clouds.
- **Challenge:** Development teams want consistency with other services such that when a new service is added, you can still have uniform monitoring, metrics, and alerts.
- **Opportunity**: Need the workloads to connect to topics across hybrid cloud environments. Service bindings enable that ability to create a consistent experience across public and private clouds.
- **Challenge:** Business has always operated in real time, but it was technology that forced the batch paradigm. Technology such as Kafka is allowing organizations to break out of the batch paradigm, but not without service management headaches.
- **Opportunity:** Enable real-time data streaming by leveraging a managed cloud service for Apache Kafka, allowing for scalability, resiliency, and fault tolerance.

RED HAT CLOUD SERVICES

Red Hat provides the Red Hat OpenShift platform as a managed cloud service to meet enterprise requirements for using open source software, particularly security and compliance concerns, while helping improve developer productivity and experiences.

Now, the company is adding another layer of abstraction by adding application and data services:

- Streams for Apache Kafka
- API management
- Data science capabilities

Red Hat OpenShift Streams for Apache Kafka is a dedicated Apache Kafka cluster delivered in the customer's location of choice. It streamlines the developer experience, is delivered as a service, and managed by a Red Hat SRE.

Red Hat OpenShift Streams for Apache Kafka can be the data broker that sits at the center of a Kafka ecosystem, receiving and delivering events across third-party systems, APIs, and data science tools, enabled by a schema registry and connectors and native integration with OpenShift.

Streams for Apache Kafka and OpenShift can provide seamless Kafka operations across hybrid cloud environments. It is easily accessible for developers, provides quick-start guides, and provides a developer-first interface.

API management provides a consistent and streamlined developer user experience to build, monitor, share, and secure APIs. It also allows for easy discovery and access of existing APIs for faster and smarter workflow. And once the data is gathered and prepared, data scientists can use the Red Hat OpenShift Data Science managed cloud service to train models, put models into production with CI/CD capabilities, and monitor the performance of models in production for drift.

CONCLUSION

There are three tactical benefits to empowering developers via cloud services:

- Managed cloud services help enable development teams to get out of the business of infrastructure administration (e.g., continually respecifying resource definitions) and instead focus on developing features that bring value to end users and the business.
- Dealing with multicloud and hybrid cloud environments, including edge deployments, is complex and therefore necessitates a platform that abstracts that complexity and brings consistency to application and data services, allowing developers to be more productive and allowing the applications to be resilient and fault tolerant.
- Developers are happiest when the various services they need to create applications actually work together and provide a consistent user experience.

The next frontier of reducing friction for developers lies in abstracting the application and data services they use to create applications. Data is only getting faster, as the demand for organizations to respond to changes and make decisions in real time becomes the norm, not the exception. By using a managed cloud service for real-time data event distribution and processing, organizations can shift resources from dealing with the inherent risks involved with the management of the underlying infrastructure and instead move to the higher-level value proposition of empowering developers to accelerate transformation.

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1,100 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For 50 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.

Global Headquarters

140 Kendrick Street
Building B
Needham, MA 02494
USA
508.872.8200
Twitter: @IDC
blogs.idc.com
www.idc.com