

2021 Kubernetes & Big Data Report

The Complexity and Scale of Kubernetes Solutions Highlight the Need for Observability and Management

Executive Summary

This survey was conducted to understand how companies are attempting to meet their big data needs through the use of Kubernetes.

Kubernetes has become the cloud-native standard. However, it is a complex technology, and companies are still struggling to use it properly and effectively. Adding engineers is one way to solve the problem but doesn't always address performance problems.

When IT doesn't have visibility into Kubernetes performance, optimal performance and cost containment are hard to achieve. Automation is key to optimizing performance.

This survey was conducted online, in March of 2021, among 800 participants from a range of industries. A majority (72%) of respondents worked at companies with between 500 and 5,000 employees.

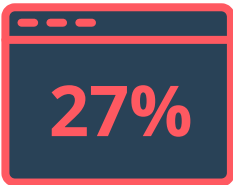
Key Findings:



77% of respondents are embracing Kubernetes because they want to **improve resource utilization to reduce cloud costs.**



The biggest challenges with moving to Kubernetes are **"initial deployment and migration."**

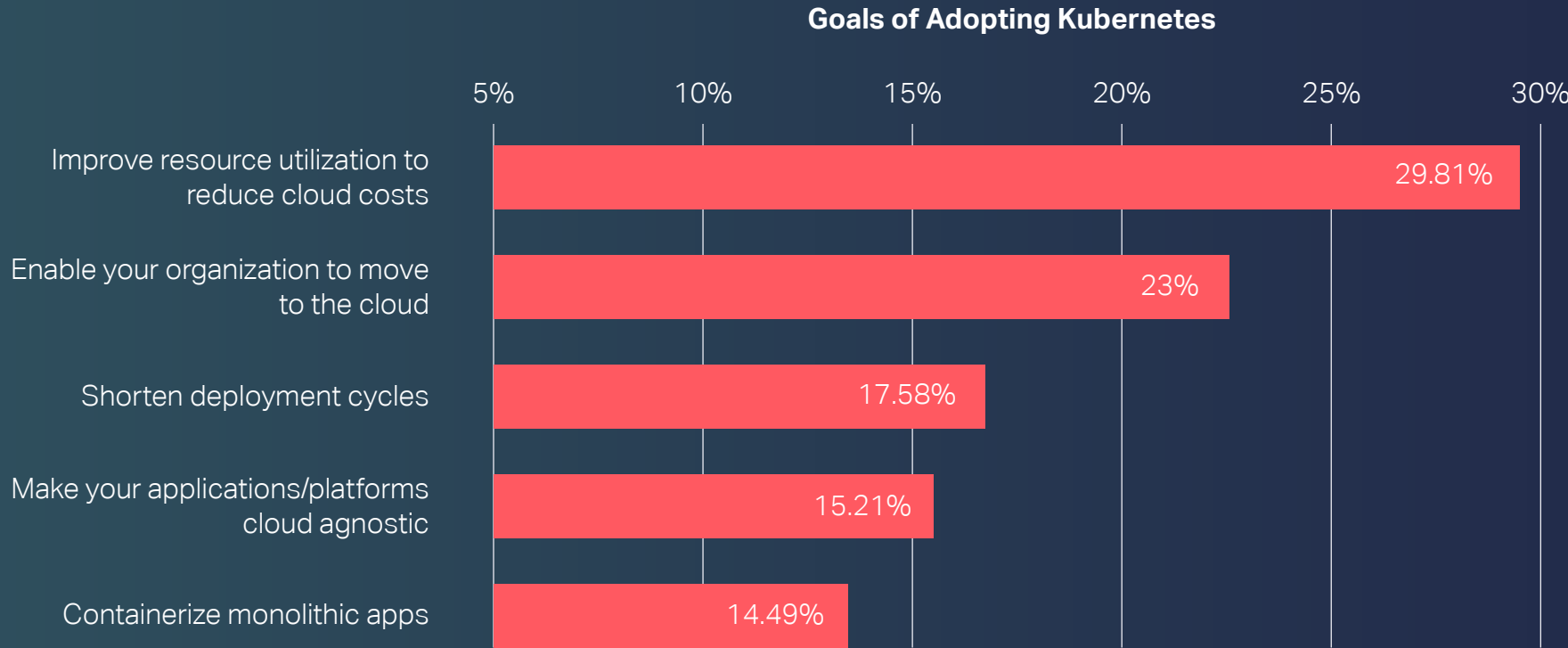


The largest percentage of companies (27%) monitor and measure application/workload performance using a **manual method or homegrown solution.**

What are your goals in adopting Kubernetes?

Companies are eager to reduce cloud costs, or at least control them. Kubernetes allows you to meld workloads and share resources more efficiently.

Companies also see Kubernetes as a way to smooth out their adoption of the cloud. Kubernetes can be a common plane where teams don't have to learn three different cloud vendors' sets of utilities or resource management frameworks. The biggest challenge identified was improving resource utilization with the goal to reduce cloud costs.

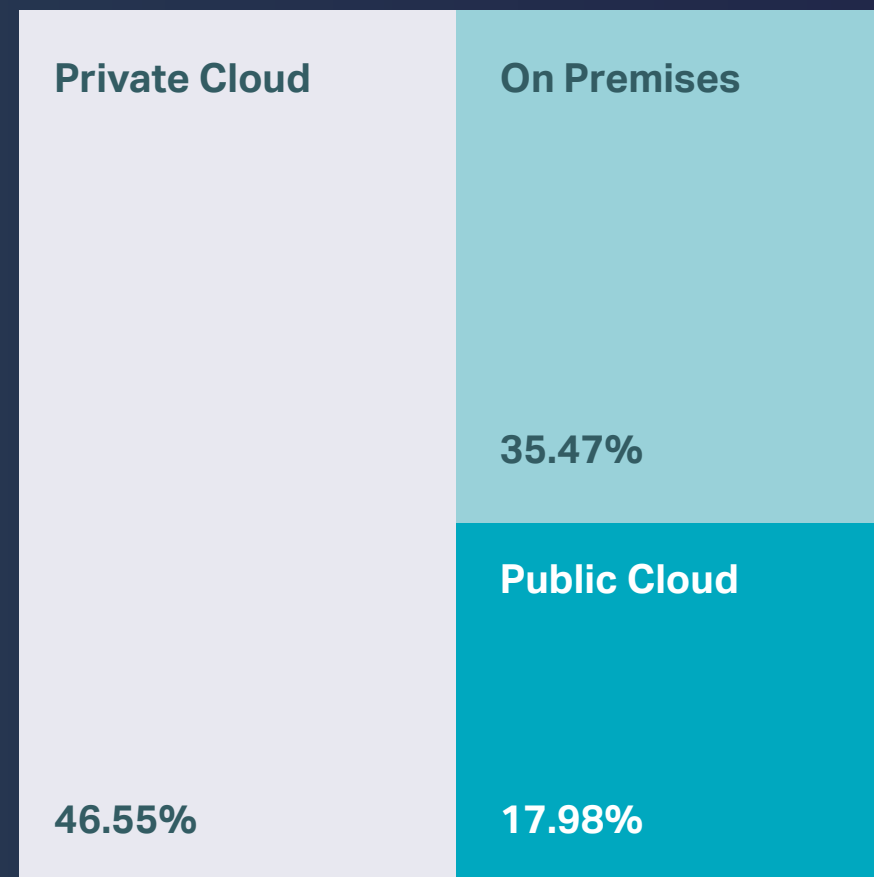


Where are you using Kubernetes containers today?

Some people might be surprised and expect public cloud to be on top, but that's not the case.

But when people are moving to Kubernetes, they want to get their heads around it in an environment they can experiment in. Spinning something up in your own private cloud gives you the ability to learn and securely adopt the technology before you move into those public cloud and off-prem environments.

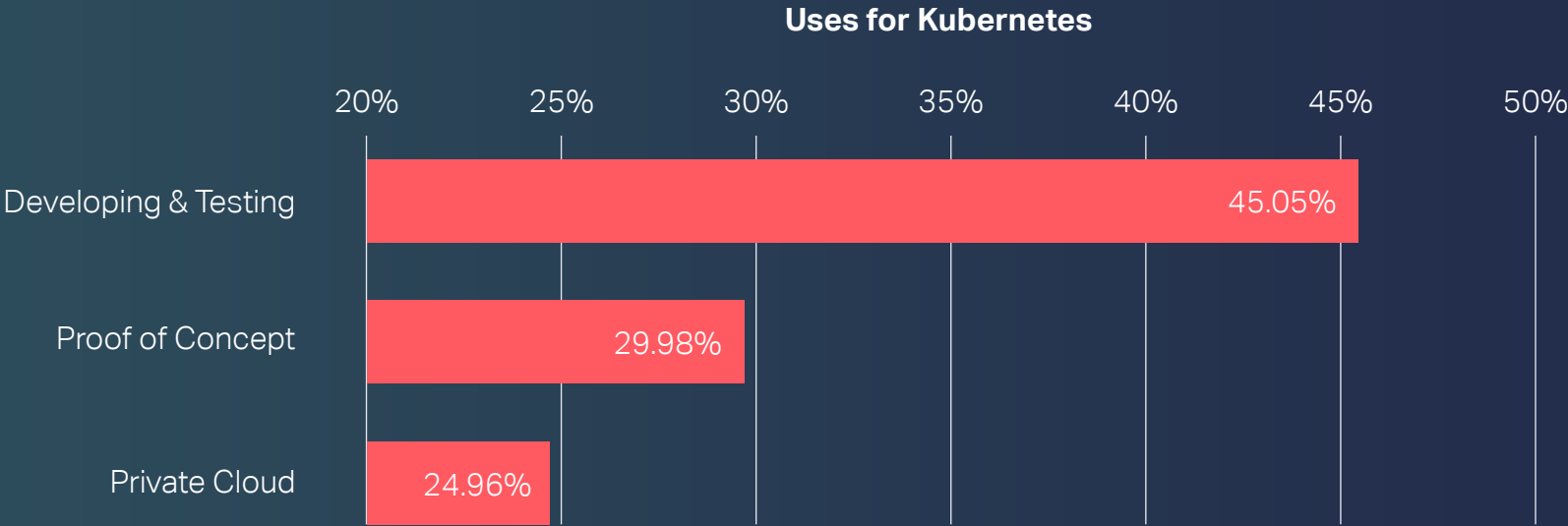
Moving between clouds means moving between different toolsets. To avoid friction, companies should seek out a common overarching view of their workloads.



How are you using Kubernetes today?

Developing and testing are common uses for Kubernetes because users are moving production workloads into a new resource management framework. With that, comes the necessary exploration and testing of new technologies and vendor options. In addition, many companies are dealing with vendors who are improving things dynamically and on the go. This makes testing even more important.

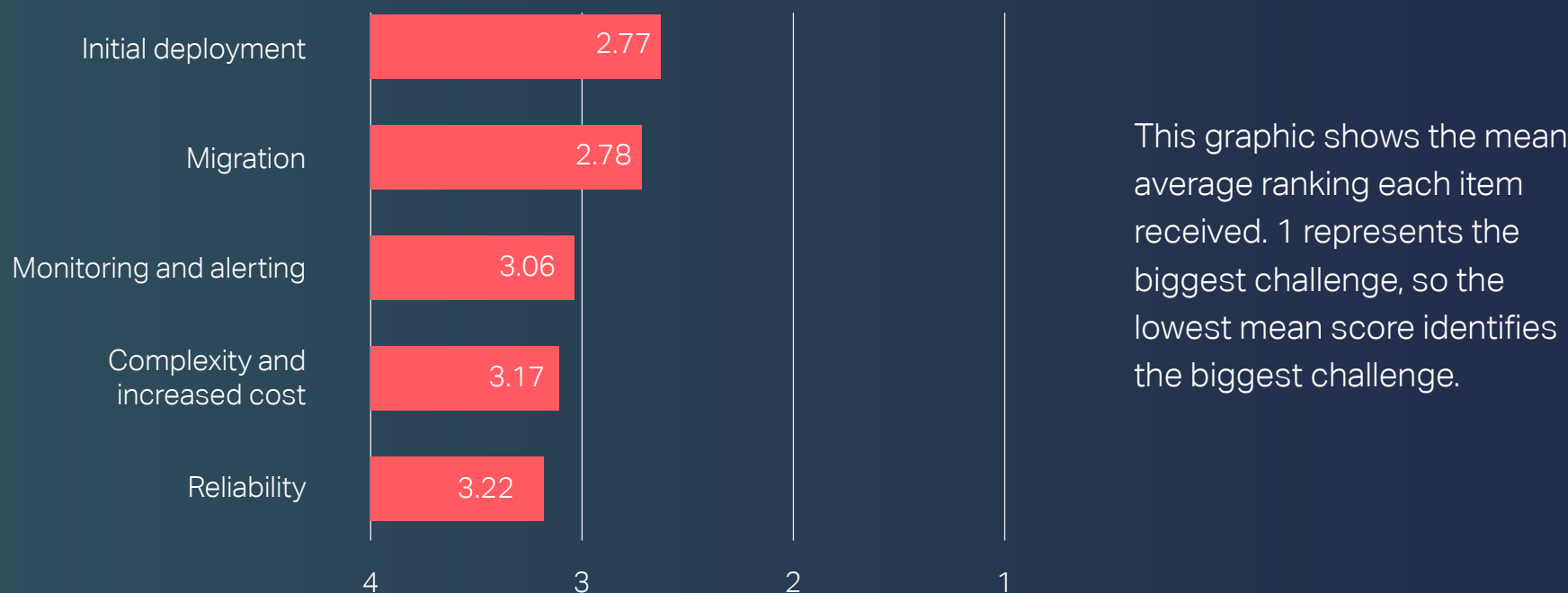
With this testing, observability is key. Teams need to know: Are these development and testing efforts even successful? Where are you now compared to last month?



What are your biggest challenges in moving to Kubernetes?

Kubernetes is a new technology, and everyone is to some extent still figuring it out. Migration is the biggest challenge, and porting hundreds or thousands of apps over to a new framework can be daunting.

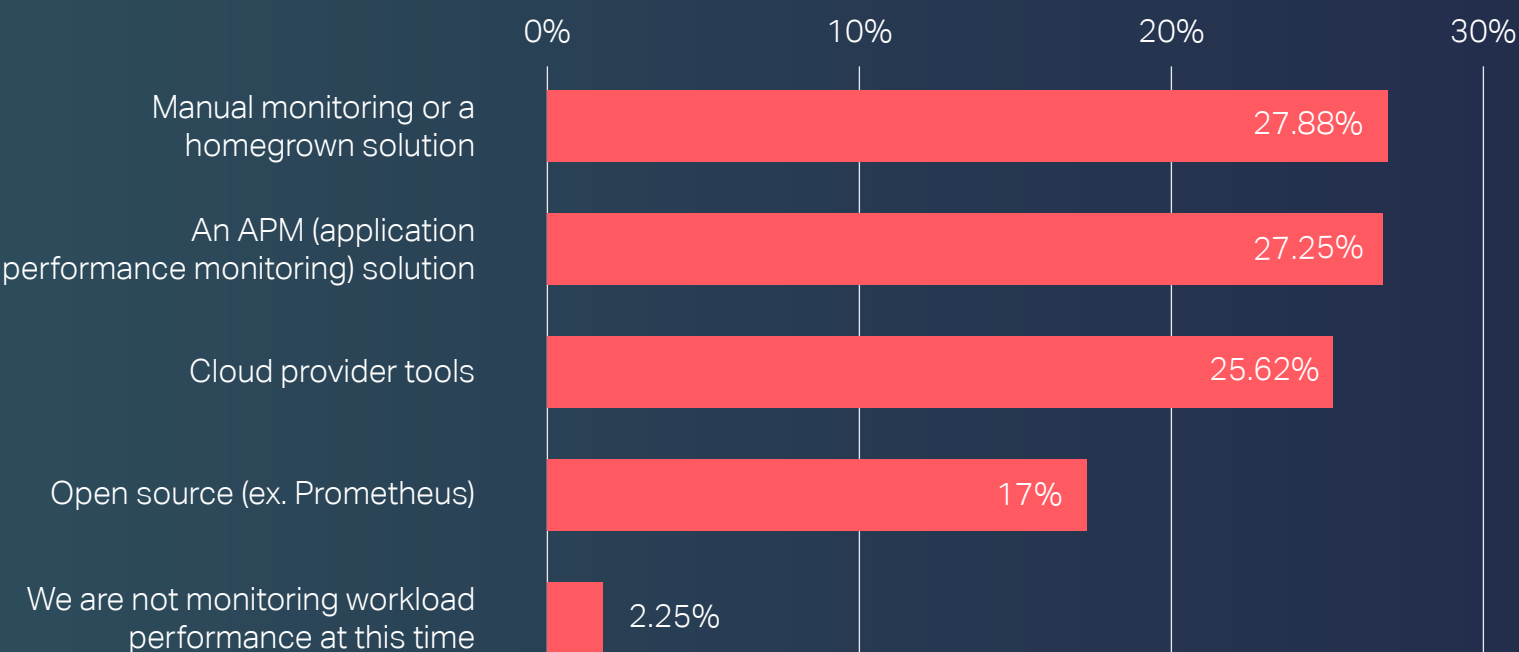
In this dynamic landscape, companies need stable monitoring and alerting technologies that can give them the visibility they need to ensure they are successfully porting apps to Kubernetes.



How do you monitor and measure application/workload performance for Kubernetes?

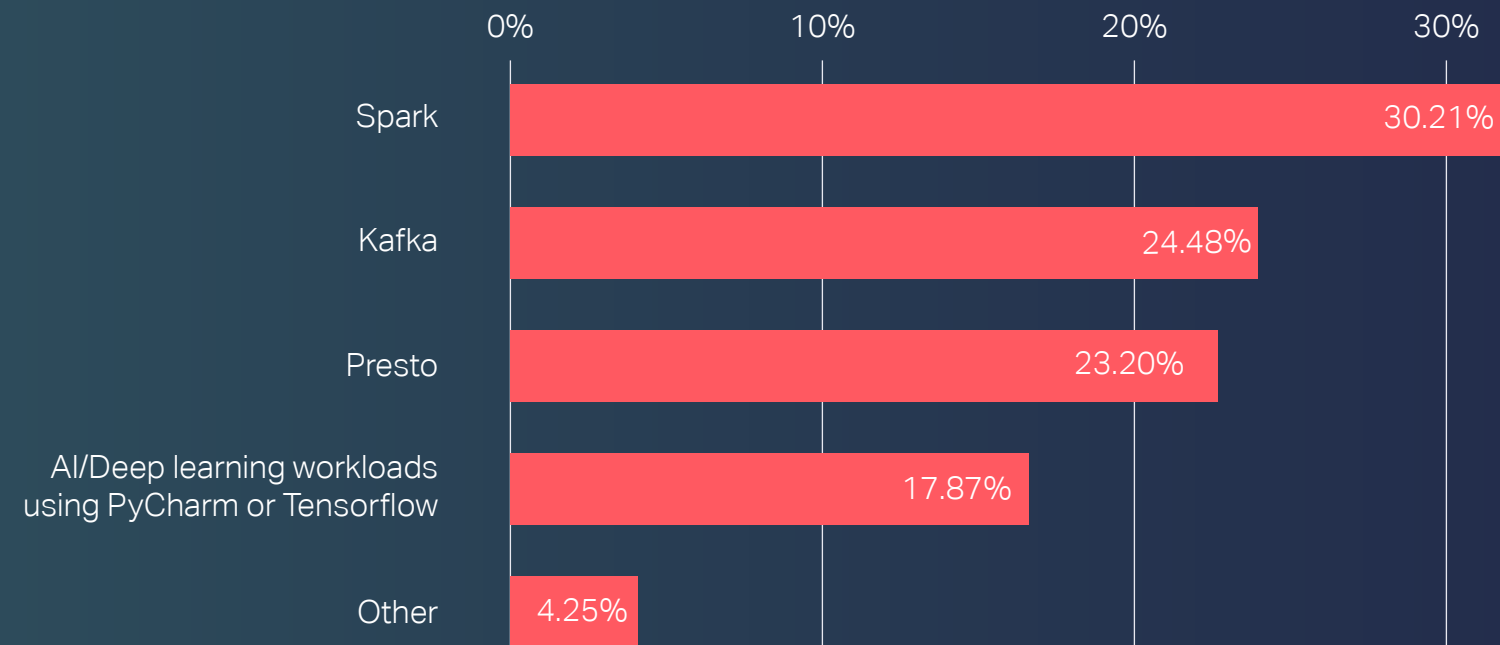
Kubernetes is extremely complicated. Manual monitoring cannot keep up, and proprietary solutions are unlikely to be up to the task. It's always tempting, faced with a new tool like Kubernetes, to use a homegrown solution that already exists.

But with Kubernetes, more custom solutions are required. General-purpose APM won't cut it; companies need tools purpose built for big data workloads on Kubernetes.



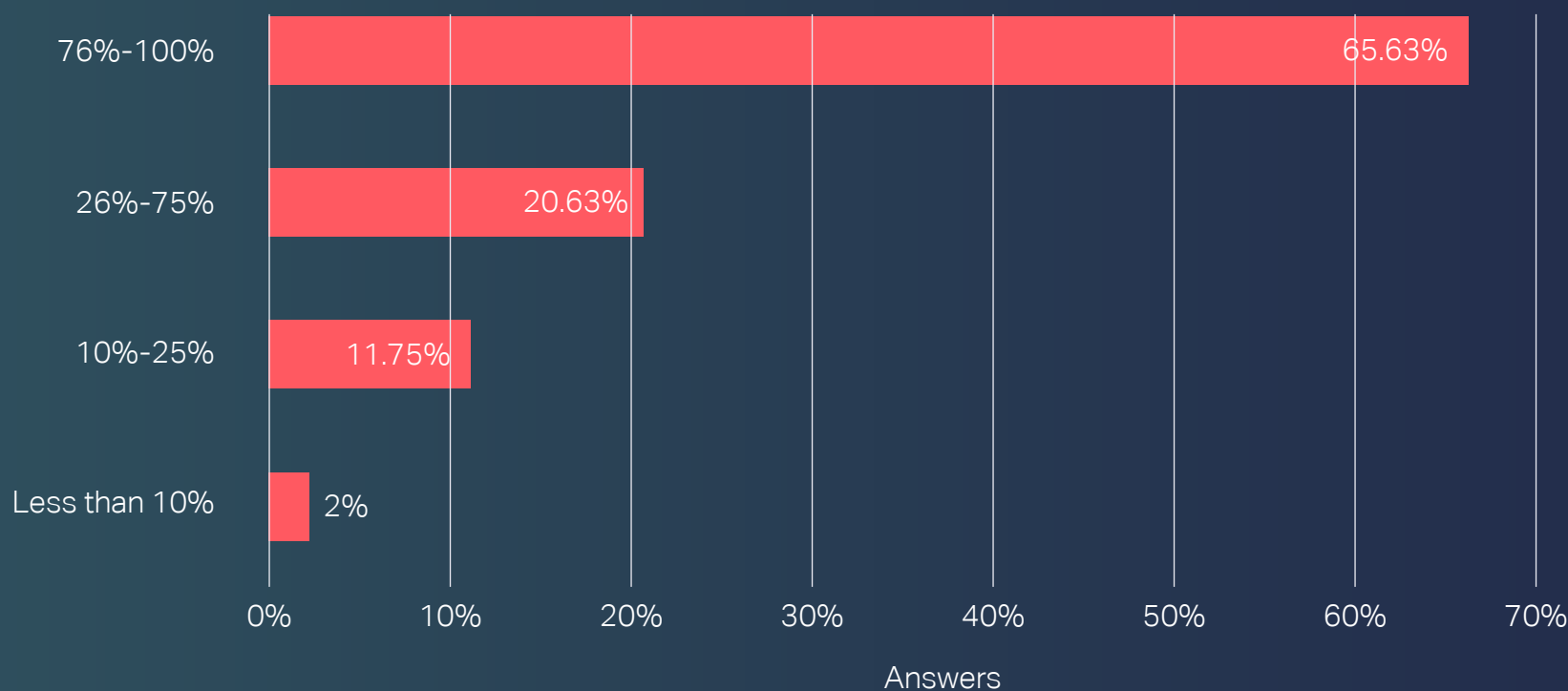
What kinds of applications/workloads are you running on Kubernetes?

More and more big data applications are moving from legacy apps to Spark, then to Spark on Kubernetes. But, Spark is prone to memory inefficiencies. This means there is a danger of carrying these challenges over to the Kubernetes platform.



What percentage of your big data workloads will be on Kubernetes by the end of 2021?

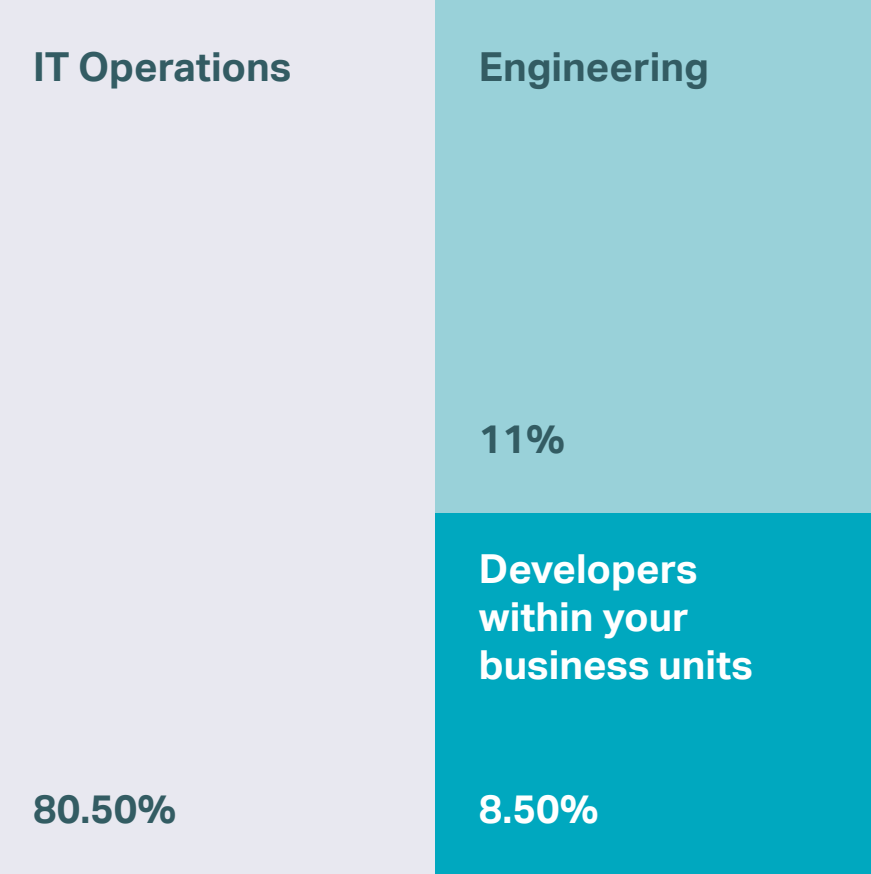
While we are in the early adoption phase for Kubernetes, developers are migrating key applications and workloads sooner than anticipated. Kubernetes promises that it will make it easier to move to the cloud, and companies are putting it to the test. This chart shows that we can expect there to be significant growth in 2021 with workloads moving to production.



In your organization, who is responsible for deploying Spark and other big data applications built on Kubernetes?

Despite some in the industry believing that developers will play a bigger role in big data and Kubernetes, IT operations teams will also play an important role in managing these applications.

Assuming Kubernetes delivers on its promise, developers will be able to focus more on their applications and worry less about the infrastructure.



IT operations teams and developers can't afford runaway costs and missed SLAs. Pepperdata makes sure cloud agility doesn't lead to trouble by automatically optimizing big data resources to reduce costs and track resource usage for clear accountability and increased efficiency. Users get powerful insights while controlling costs and improving application performance to achieve their business requirements. In addition, users gain insight for planning, debugging, and troubleshooting, plus the confidence that their applications and workloads will meet SLAs.

Learn how you can start optimizing today.

[START A FREE TRIAL](#)



pepperdata®

www.pepperdata.com