INTELLIGENCE UPDATE

Cloud cost savings depend on application design

 Douglas Donnellan  Owen Rogers

12 Mar 2025

Enterprises continue to choose the cloud for many of their workloads to improve scalability and lower costs. Once in the cloud, however, many enterprise operators are surprised by additional costs or find that savings are lower than expected.

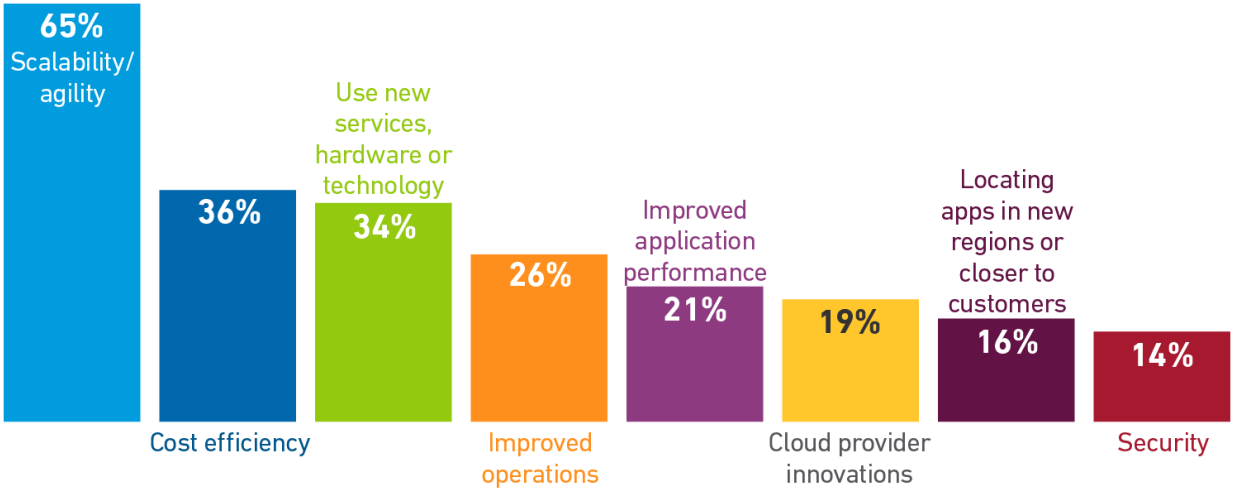
These cost overruns are primarily driven by how cloud-based applications scale, increasing resource consumption to meet higher demand. At the same time, applications that are not designed for dynamic scaling may be unable to relinquish resources during periods of low demand, leading to unnecessary and significant cloud costs.

Despite these challenges, enterprise use of the cloud remains strong — and there are signs that cloud adoption is increasing. Results from the Uptime Institute Capacity Trends and Cloud Survey 2024 (conducted from October through December 2024) reveal that most enterprises (69%) report at least some usage of public cloud (up from 62% in 2023) and that nearly two-thirds report moving some production applications from on-premises or colocation to the public cloud in the past 12 months.

The survey also found that operators view the primary benefit of moving applications to the public cloud as being application scalability/agility (see **Figure 1**). While this benefit influences enterprises to increase cloud adoption, scalability issues can lead to unexpected cost overruns and cause operators to repatriate some workloads.

Figure 1 Scalability and agility are top benefit of cloud migration

Which of these are the primary benefits behind moving applications from your own data center or a colocation facility to the public cloud? Choose no more than two. (n=110)



UPTIME INSTITUTE CAPACITY TRENDS AND CLOUD SURVEY 2024

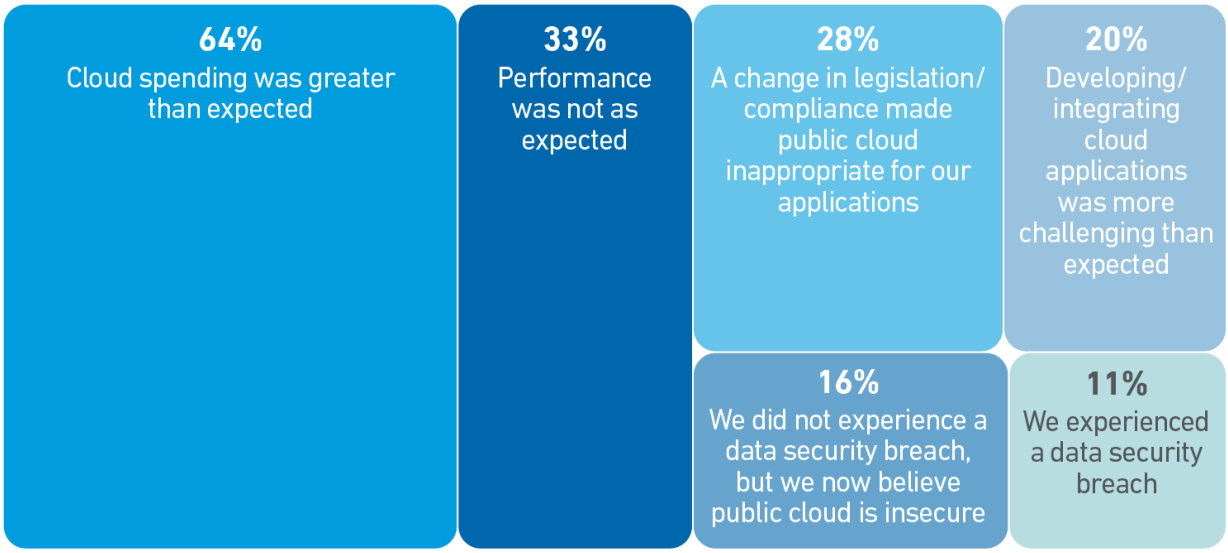
For example, 42% of enterprises using the cloud report moving production applications back to on-premises or colocation data centers in the past 12 months. Of this group, unexpectedly high cloud costs is cited as the top cause (see **Figure 2**). This response rate suggests that enterprises are still facing challenges when evaluating which venue is the best and most cost-efficient for specific applications.

Whether scalability leads to cost savings or overruns largely depends on the nature of the application:

- Applications designed to be scalable will consume more resources when required due to a rise in demand. These resources, charged per unit, add up to higher costs but can deliver greater value from the application, such as an improved experience for the user.
- Applications not designed to scale cannot shrink during periods of low demand, leading enterprises to pay for unused cloud resources. Enterprises may lift and shift these applications from on-premises infrastructure to the cloud to avoid costs from rearchitecting for scalability, but this often leads to higher costs over time.

Figure 2 Cloud repatriation is largely driven by high costs

What best describes your organization's reason for moving applications away from public cloud? Choose no more than two. (n=61)



UPTIME INSTITUTE CAPACITY TRENDS AND CLOUD SURVEY 2024

In addition to scalability challenges, many enterprises underestimate how rapidly cloud costs can escalate due to increased data usage, egress fees and inefficient architectures. Applications may also accumulate data over time, often automatically or inadvertently, resulting in high storage costs. These factors contribute to enterprises reassessing their cloud strategies and, in some cases, moving workloads back on-premises.

However, this repatriation rate does not suggest enterprises are leaving the public cloud entirely. Only one in four enterprises that repatriated workloads report a significant reduction in overall cloud usage — a rate consistent with data from 2023 (see *Capacity expands rapidly, but complexity is challenging*).

The migration of workloads between venues indicates that enterprises are evolving their hybrid IT strategies based on cost, performance, and business needs rather than treating the cloud as an all-or-nothing decision. However, it also suggests that enterprises are not fully assessing the benefits and costs of the public cloud. Operators who carefully consider which approach to take for migrating to the public cloud (lift and shift versus modernization) stand to maximize scalability benefits and save costs.

ABOUT THE AUTHORS

Douglas Donnellan

Douglas is a Research Associate at Uptime Institute covering sustainability in data centers. His background includes environmental research and communications, with a strong focus on education.

ddonnellan@uptimeinstitute.com

Owen Rogers

Dr. Owen Rogers is Uptime Institute's Sr. Research Director of Cloud Computing. Dr. Rogers has been analyzing the economics of cloud for over a decade as a product manager, a PhD candidate and an industry analyst. Rogers covers all areas of cloud, including economics, sustainability, hybrid infrastructure, quantum computing and edge.

orogers@uptimeinstitute.com

About Uptime Institute

Uptime Institute is the Global Digital Infrastructure Authority. Its Tier Standard is the IT industry's most trusted and adopted global standard for the proper design, construction, and operation of data centers – the backbone of the digital economy. For over 25 years, the company has served as the standard for data center reliability, sustainability, and efficiency, providing customers assurance that their digital infrastructure can perform at a level that is consistent with their business needs across a wide array of operating conditions.

With its data center Tier Standard & Certifications, Management & Operations reviews, broad range of related risk and performance assessments, and accredited educational curriculum completed by over 10,000 data center professionals, Uptime Institute has helped thousands of companies, in over 100 countries to optimize critical IT assets while managing costs, resources, and efficiency.