# **Cloud Optimization Tools: Overview**

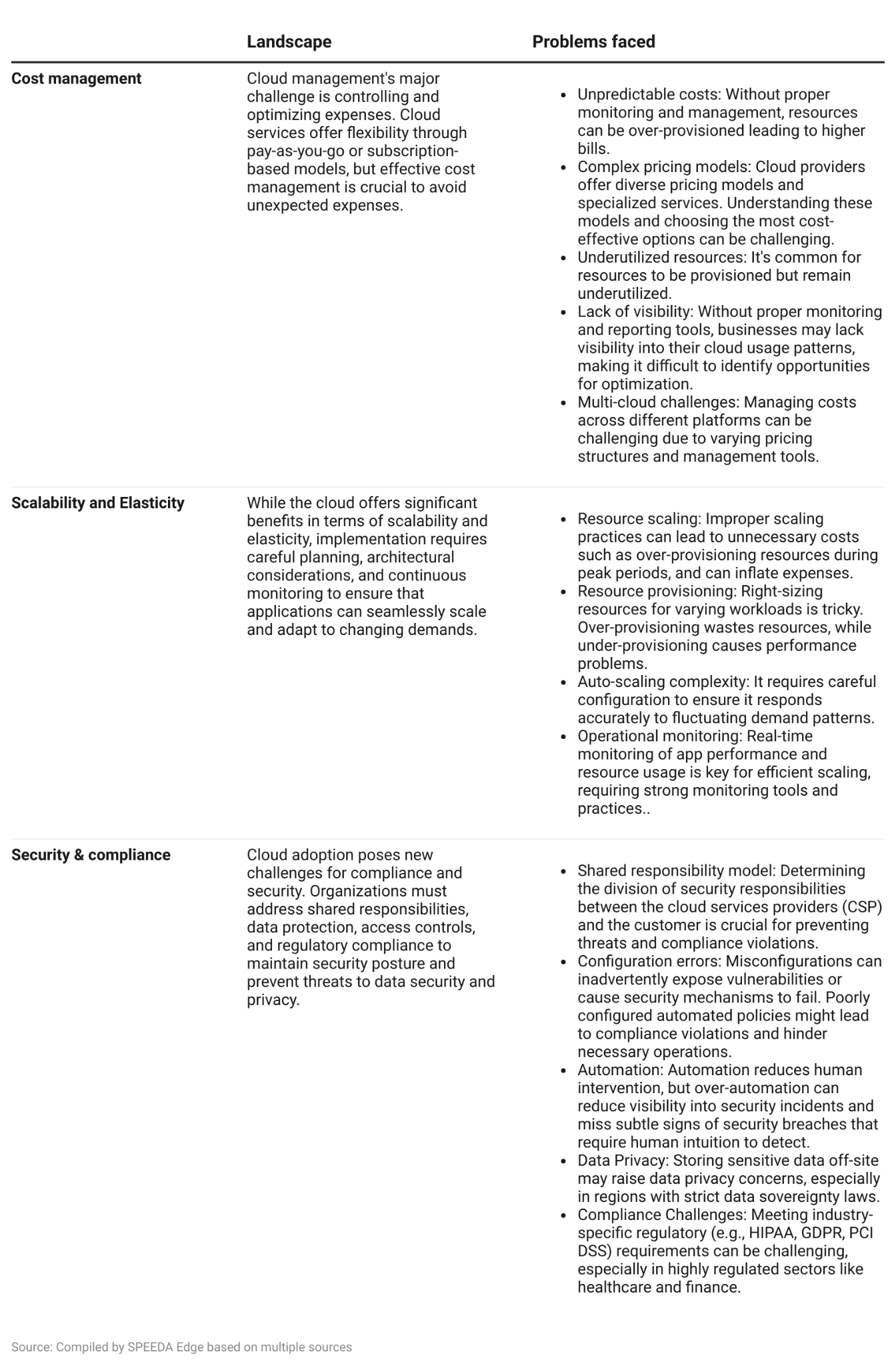
## 

## **Scaling smartly and limiting wastage: Cloud optimization for efficient adoption**

Cloud optimization tools leverage AI/ML and data analytics to enable organizations to harness the power of their cloud environments while minimizing wastage. Organizations are rapidly moving their operations to the cloud to take advantage of its benefits, which include cost-effectiveness, scalability, flexibility, speed, and agility. However, cloud adoption has introduced operational challenges such as cost overruns, resource under-utilization, multi-cloud architecture, varying pricing and resource types, configuration rules, industry regulations, and security threats. Cloud optimization tools help organizations overcome these challenges by fine-tuning resource allocation, optimizing workload distribution, and navigating multi-cloud deployments.

### 

### **Key complications of cloud adoption**

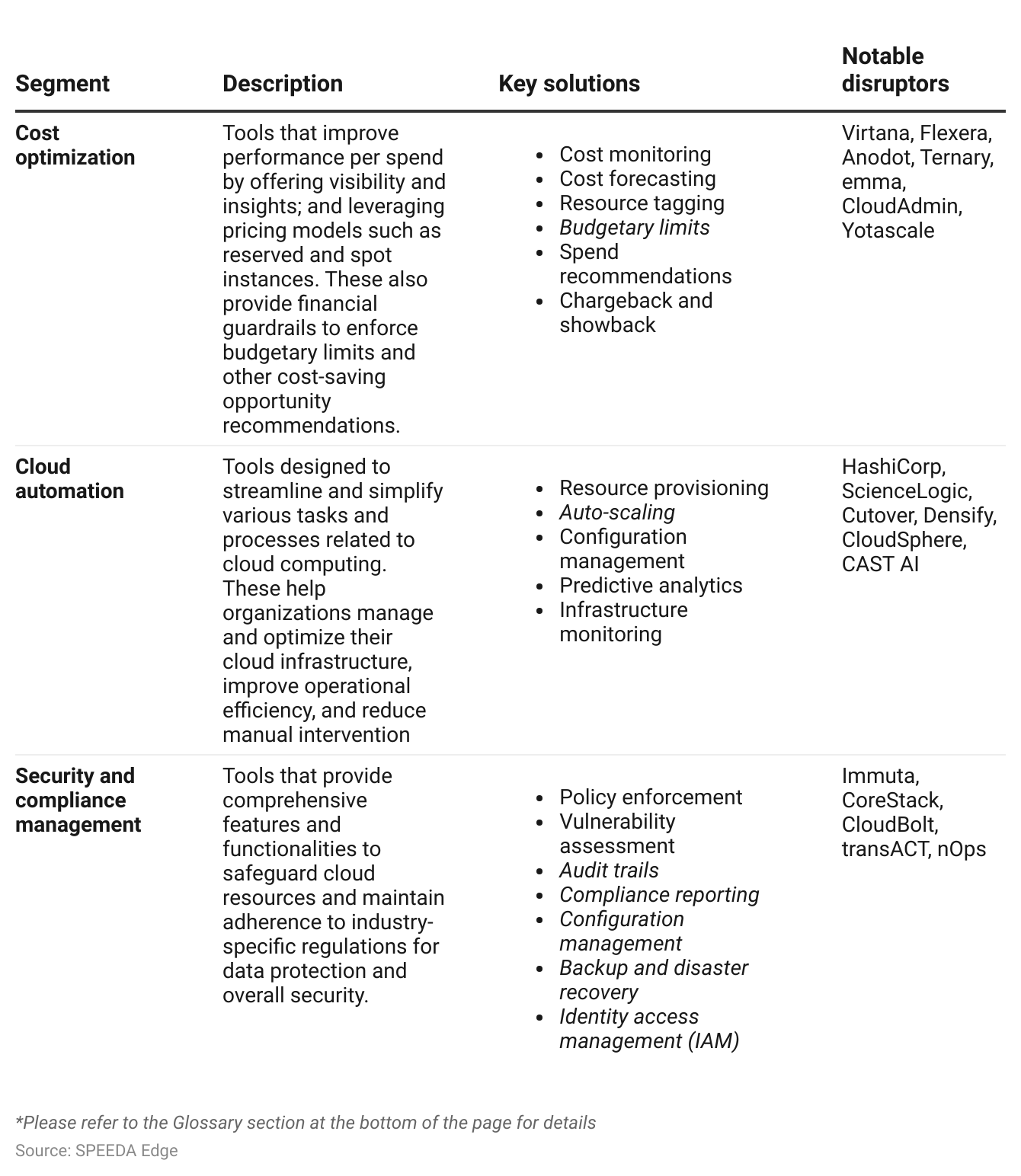


Cloud optimization encompasses a broad range of solutions to enhance the efficiency, performance, security, and cost-effectiveness of cloud-based resources and services. **This industry hub focuses on tools that provide cost optimization, cloud automation, and security and compliance management.**

There are also three main sources of cloud optimization tools:1) cloud provider software (such as AWS Cost Explorer, Amazon Cloud Formation, Azure Automation, and Azure Cost Management), 2) custom in-house solutions, and 3) third-party tools. Cloud provider tools are easy to access and cost less but are limited to platform-specific optimizations, while in-house solutions are developed to address company-specific issues but require additional time and capital. **The focus of the hub would be third-party tools that cover multi-cloud workload management** (i.e., tools that manage and optimize services from multiple cloud service providers).

## 

## **Segmental overview**



## **Rapid AI/ML developments enable advanced cloud optimization solutions**

AI/ML algorithms provide deep visibility and insights into complex infrastructure components such as configurations, application stacks, and storage during cloud migration and setup.

**AI analyzes historical usage data and identifies trends to predict resource demands**, **minimizing overprovisioning and reducing unnecessary costs.** AI tools monitor resource usage in real time, identifying idle or underutilized resources for prompt scaling down or termination, resulting in significant cost savings. AI's anomaly detection capabilities identify irregular patterns based on historic usage, aiding in the early detection of potential cost spikes, compliance issues, and security breaches.

**AI also simplifies resource optimization and operations automation by analyzing vast amounts of data generated by cloud services and identifying performance bottlenecks, resource shortages, and potential failures in real time.** [AI-driven chatbots and virtual assistants](https://www.alibabacloud.com/solutions/ai-chatbots#J_7877782360) provide instant responses to user queries, troubleshoot common issues, and offer self-service solutions, reducing the need for human intervention in routine tasks. By harnessing AI, organizations achieve better operational efficiency, minimize downtime, and enhance the overall reliability and responsiveness of their cloud optimization tools.

**Moreover, AI/ML tools strengthen security posture and maintain compliant infrastructure by scanning entire ecosystems for vulnerabilities and potential threats and recommending rectification strategies.** These tools ensure compliance with industry regulations and standards using automated compliance checks to identify potential violations and deploy proactive measures to fix non-compliant actions.

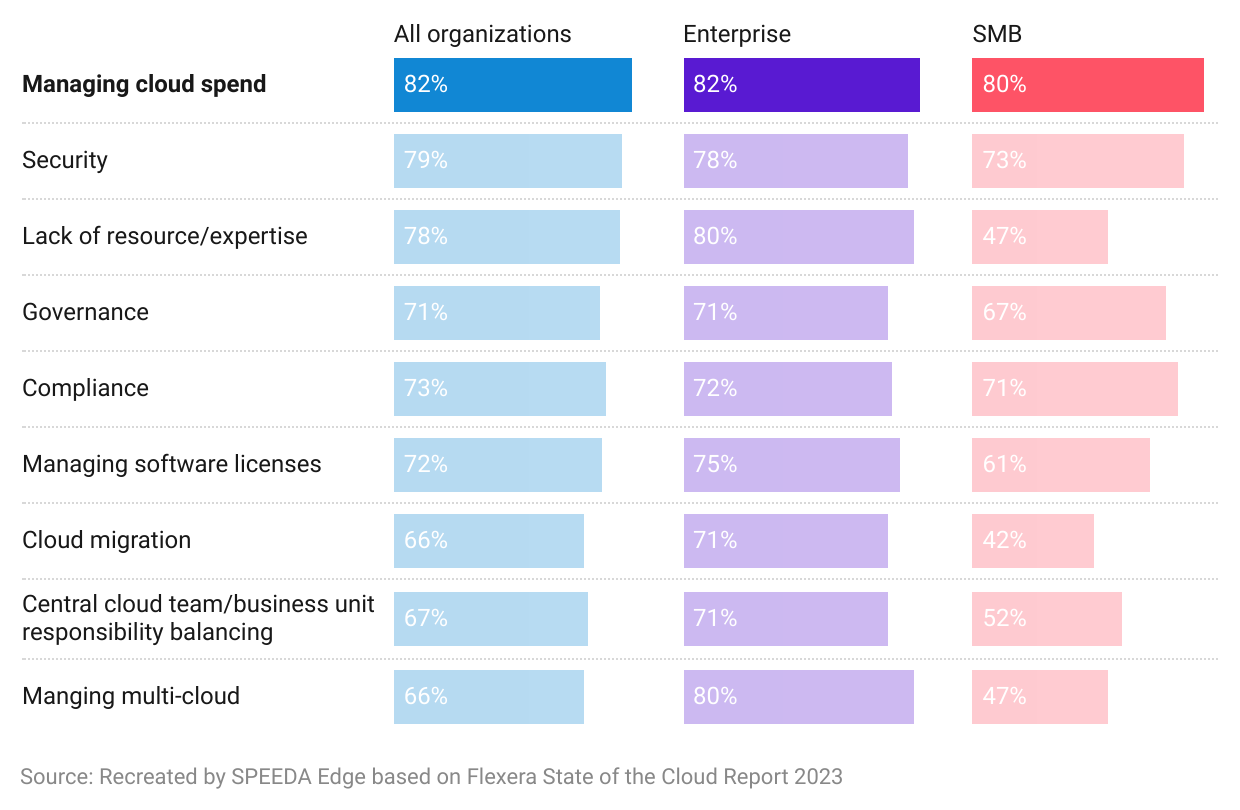
# 

# **Demand drivers**

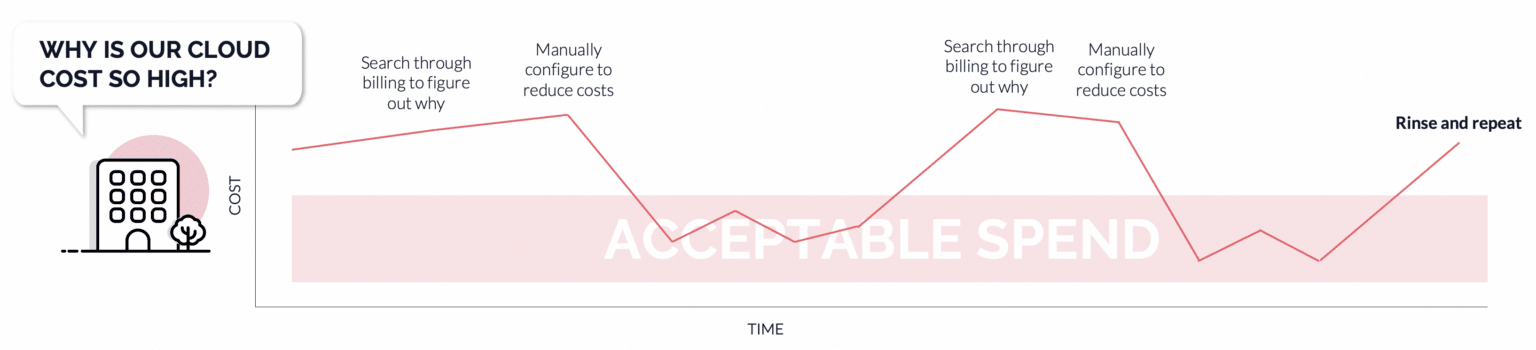
## **1. Rising cloud expenses and the emergence of FinOps**

According to the findings of the [Flexera 2023 State of the Cloud](https://www.flexera.com/blog/cloud/cloud-computing-trends-flexera-2023-state-of-the-cloud-report/) report, on average, public cloud expenditures exceeded budget by ~18%, up from the previous year's ~13%. Additionally, ~30% of respondents anticipated an increase in spending over the following 12 months. According to the same survey, managing cloud spend was the biggest challenge for organizations of all sizes in 2023.

### **Top cloud challenges in 2023 by organization size**



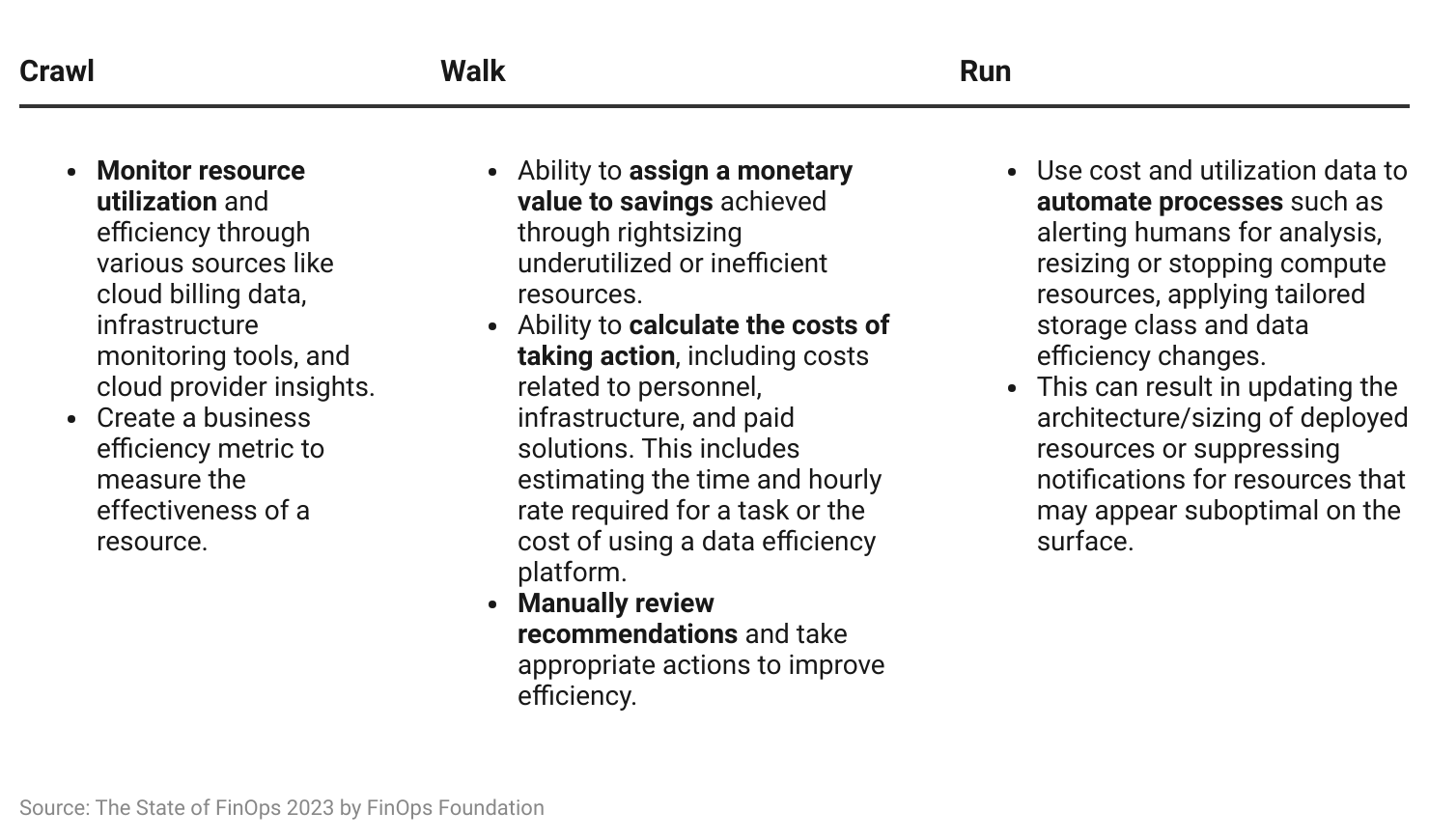
Companies often grapple with unforeseen cloud expenses because they primarily adopt a reactive approach. Consequently, when a usage surge inevitably occurs, companies face the challenge of labor-intensive manual processes to restore costs to acceptable levels. By employing cloud (cost) optimization tools, companies identify inefficiencies and potential savings, set quotas and spending limits, and automate power schedules and expiration dates to ensure that cloud spending is optimized without the need for constant manual intervention.



Source: CloudBolt Software

Additionally, FinOps is a framework that optimizes the financial management of cloud resources, aligning cloud usage with business goals, reducing costs, and enhancing resource efficiency. Before FinOps, manual efforts and basic tools were used to gain visibility into cloud spending but fell short of scaling effectively. FinOps [adoption](https://www.finops.org/framework/capabilities/utilization-efficiency/) has been growing in organizations around the world, including among a significant number of Fortune 500 companies. This has led to the use of specialized cloud optimization tools that provide real-time insights, cost visibility, and optimization recommendations based on FinOps principles.

### **Stages of FinOps adoption**

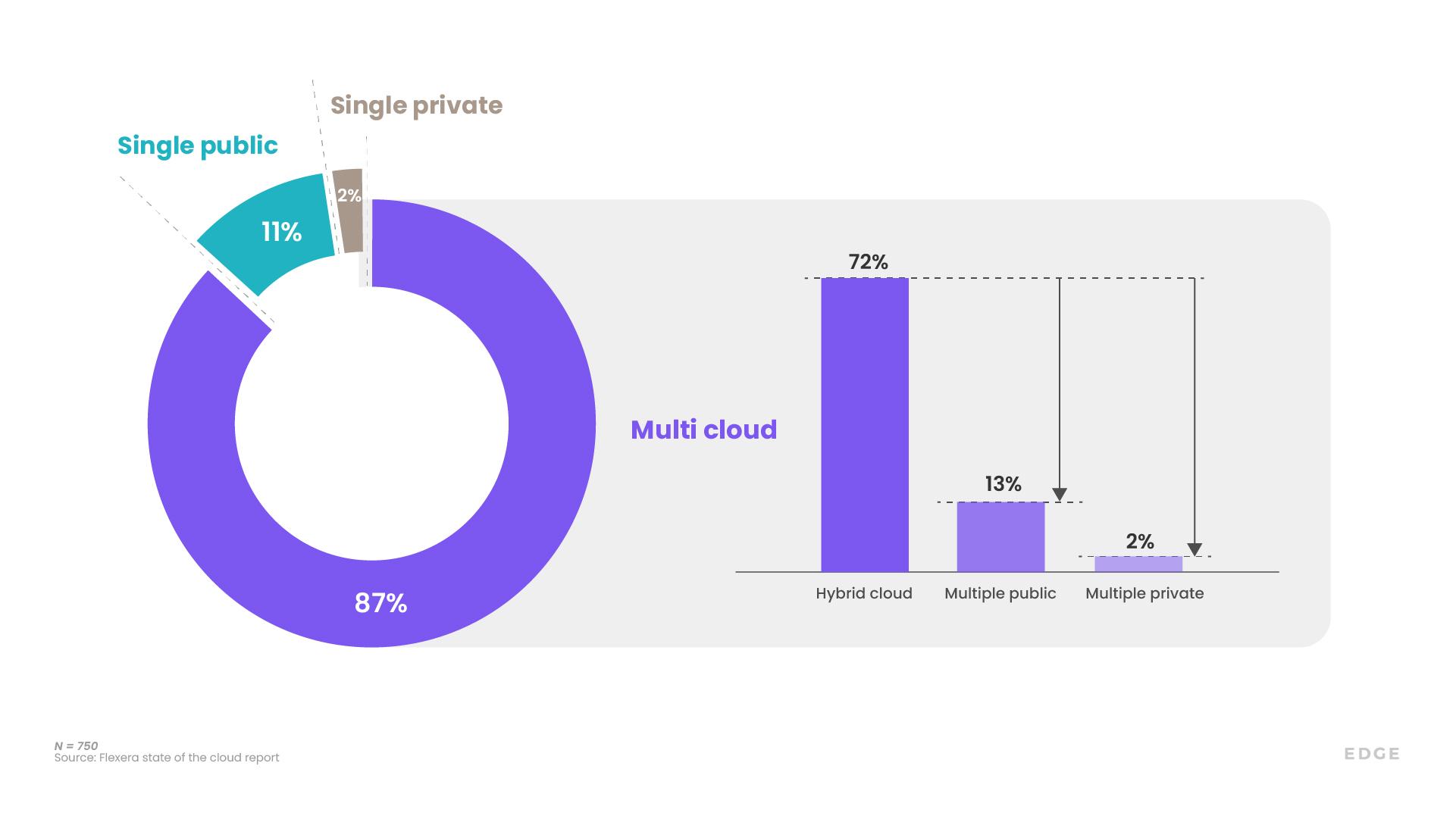


## **2. Rapid cloud migration and complexity of multi-cloud and container environments**

Companies are migrating workloads from on-premise infrastructure to the cloud. [Gartner](https://www.gartner.com/en/newsroom/press-releases/2022-02-09-gartner-says-more-than-half-of-enterprise-it-spending) predicts that enterprise spend on cloud migration will increase at a CAGR of 11.5% to USD 1.8 trillion by 2025 from USD 1.3 trillion in 2022.

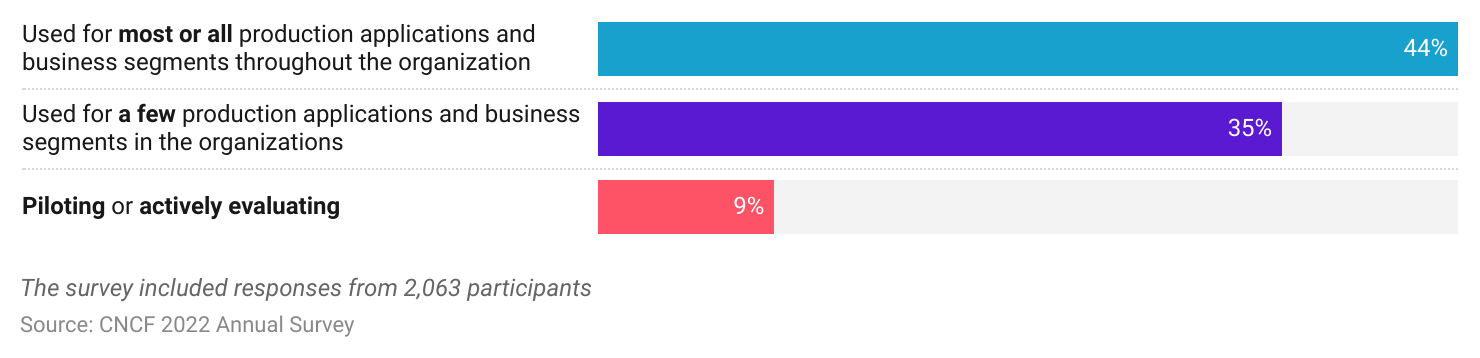
Moreover, ~87% of [Flexera 2023 State of the Cloud](https://www.flexera.com/blog/cloud/cloud-computing-trends-flexera-2023-state-of-the-cloud-report/) survey respondents stated that they use multi-cloud architecture. This introduces challenges like increased complexity in management and governance, the potential for higher costs due to inter-cloud data transfer and redundancy, interoperability issues, and the need for specialized expertise in multiple cloud platforms.

### **Multi-cloud: The most preferred deployment method**

****

Moreover, as cloud landscapes grow, [container usage](https://www.cio.com/article/220563/more-enterprises-are-using-containers-here-s-why.html) has also risen significantly in recent years, offering advantages in portability, scalability, and efficiency compared with traditional virtual machines (VM). The dynamic nature of containers warrants meticulous management and monitoring to ensure cost-effectiveness. Varying price models for containers by CSPs add to these challenges. Cloud optimization tools offer simplified solutions to manage these complexities while optimizing resource utilization.

### **Container usage within organizations**

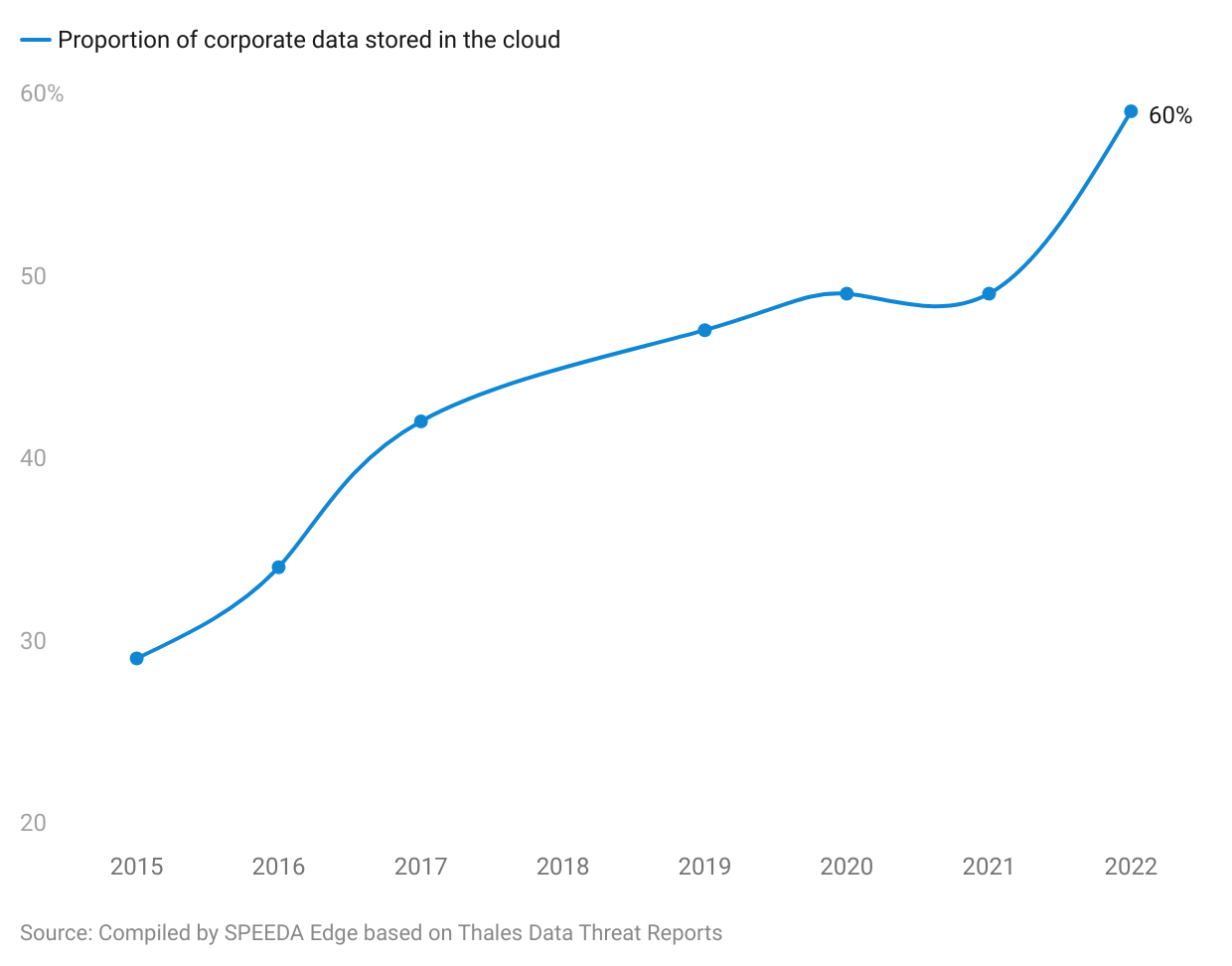


## **3. Growth in cloud storage and underlying risks**

Storing data in the cloud has become increasingly appealing for enterprises, driven by a multitude of advantages that cloud storage offers over traditional on-premises solutions. This shift is underscored by a substantial increase in the number of enterprises opting for cloud storage, as evidenced by a two-fold growth from 2015 to 2022.

### 

### **Cloud data storage witnesses two-fold growth**



The growth of cloud data storage has given rise to data privacy concerns and regulations. Adhering to regulatory standards such as the General Data Protection Regulation (GDPR), Federal Risk and Authorization Management Program (FedRAMP), Payment Card Industry Data Security Standard (PCI DSS), and Health Insurance Portability and Accountability Act (HIPAA) has become vital in the modern landscape. Cloud compliance challenges include visibility into hybrid and multi-cloud workflows, certifications and attestations, and data residency issues.

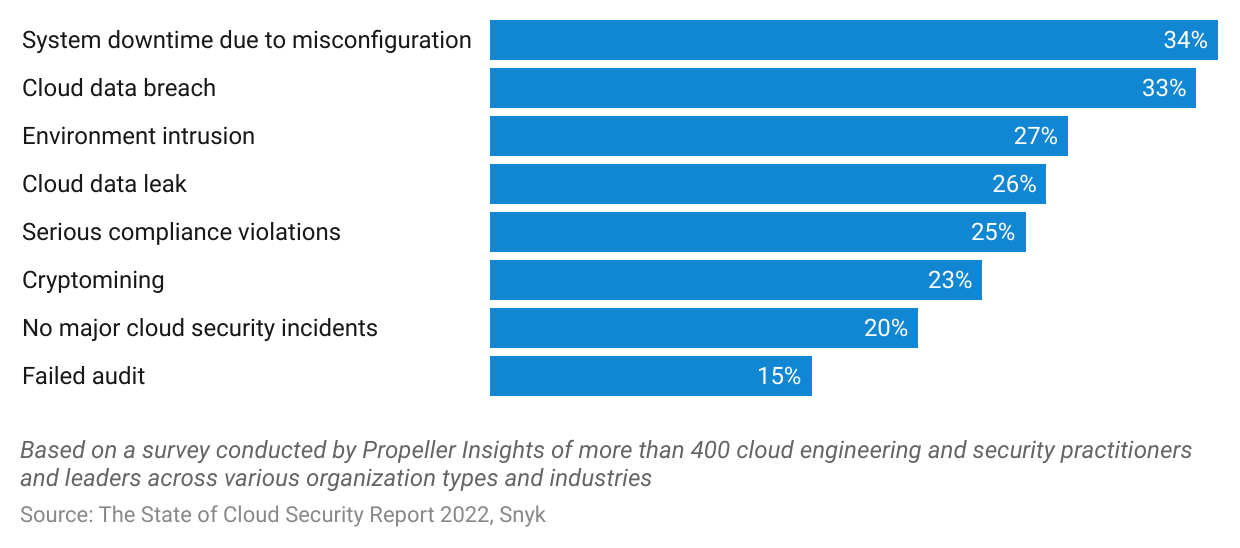
CSPs employ the shared responsibility model for compliance but taking proactive measures to ensure data protection is crucial. Security and compliance optimization platforms provide pre-integrated policy packs, continuous scanning, policy-as-a-code frameworks, and automated alerts. Cloud optimization tools provide centralized management interfaces, automated policy enforcement, audit trails and logs, continuous monitoring and reporting, pre-defined compliance templates, and automated remediation.

## **4. Rising risks of cloud security incidents**

Operating in the cloud exposes enterprises to numerous security vulnerabilities, and according to Snyk, [80% of organizations](https://snyk.io/news/snyks-state-of-cloud-security-report-reveals-organizations-have-experienced-severe-cloud-security-incidents/) have experienced a significant cloud security threat in 2022. Moreover, Checkpoint’s 2023 Cloud Security Report revealed that [76% of respondents](https://www.checkpoint.com/press-releases/cloud-security-threats-remain-rampant-check-point-survey-reveals-heightened-concerns-for-76-of-organizations-amid-48-increase-in-cloud-based-network-attacks/) have heightened concerns about cloud security, while cloud-based network attacks surged by 48% YoY in 2022.

Misconfigurations in cloud environments can cause vulnerabilities, security mechanism failures, and system crashes, which affect operations negatively. 34% of cloud security incidents are related to misconfigurations according to [The State of Cloud Security Report 2022](https://resources.snyk.io/state-of-cloud-security) by Snyk. Poorly configured automated policies can trigger compliance violations, which could lead to legal action, fines, and loss of trust.

### **Configuration issues top serious cloud security incidents faced by organizations**



Moreover, the use of multiple CSPs can potentially increase an enterprise’s attack surface (the total points or pathways through which an attacker can potentially exploit a system, network, or application), making them more vulnerable to security threats. While all major CSPs provide built-in tools for access management and general maintenance of security posture, this presents the issue of fragmented access policies and low visibility. A study conducted by Strata Identity has revealed that fragmented access policies are a [key security concern for enterprises](https://www.businesswire.com/news/home/20230821223452/en/Strata-Identity-Study-Reveals-Fragmented-Access-Policies-are-Top-Security-Concern-for-Multi-Cloud-Enterprises), with over 75% of enterprises lacking knowledge of their access policies.

Cloud optimization tools deliver unified solutions to these multifaceted security threats through robust encryption, access controls, continuous monitoring, and security best practices, ensuring the safety and privacy of cloud infrastructure.

## **5. Higher cloud resource wastage**

Cloud infrastructure requires various resources to work together efficiently, as inefficient resource utilization can lead to increased costs, decreased performance, and negative impacts on overall efficiency. This can occur due to over-provisioning, underutilization, and improper management practices, resulting in cloud resource wastage.

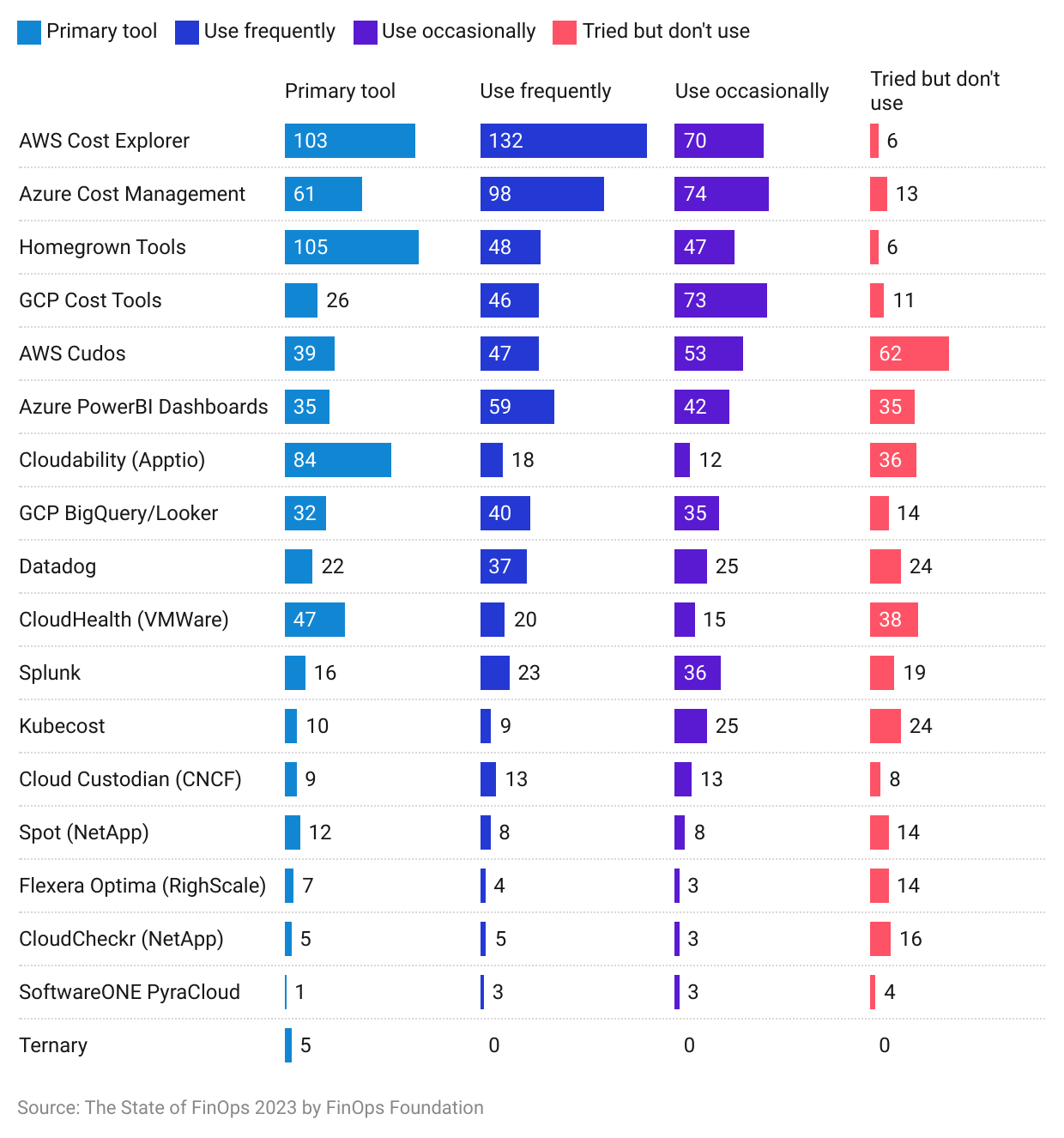
A survey by Forrester Consulting stated that [94% of companies](https://www.hashicorp.com/state-of-the-cloud) reported avoidable cloud waste (cloud spending), with overprovisioned resources and idle or underutilized resources contributing the most toward this. Meanwhile, in the Flexera State of the Cloud 2023 report, respondents self-estimated [28%](https://www.flexera.com/blog/cloud/cloud-computing-trends-flexera-2023-state-of-the-cloud-report/) of total cloud spend as waste. Cloud optimization tools offer a range of capabilities that help organizations tackle challenges associated with resource waste.

# **Risks to growth**

## **1. Shift toward homegrown tools**

Enterprises have begun opting for homegrown tools to meet their cloud optimization needs, as they seek greater control, customization, and cost efficiency in their cloud operations. According to the FinOps Foundation’s The State of FinOps 2023 report, homegrown tools have risen to the [top spot](https://data.finops.org/#4914) as the primary FinOps tool used, up from the fifth place in 2022, and now rank as the [fourth](https://data.finops.org/#4944) most utilized solution, behind CSP native tools. This shift is driven by several factors. First, many organizations have unique requirements and specific business processes that third-party solutions may not fully address. Building their own tools allows them to tailor solutions to their exact needs. Second, enterprises want to reduce their dependency on third-party vendors and mitigate the risk of vendor lock-in. By developing in-house solutions, they can have more autonomy over their cloud infrastructure. Lastly, cost optimization is of paramount concern and custom tools can be designed to maximize efficiency and minimize unnecessary spending.

### **Homegrown tools are catching up to CSP native tools**



## **2. Concerns over privacy and regulatory compliance**

Cloud optimization involves monitoring and analyzing usage data, which could raise privacy concerns or conflict with regulatory compliance requirements, especially in industries with strict data protection rules. [Data residency and sovereignty regulations](https://www.seagate.com/in/en/blog/data-sovereignty-and-cloud-computing/) require that certain types of data remain within specific geographic boundaries or jurisdictions. When using cloud optimization tools, data might be stored, processed, or transmitted across different regions or countries, which might violate such rules.

Moreover, many regions have strict data protection regulations, such as the European Union's [General Data Protection Regulation (GDPR)](https://www2.deloitte.com/nl/nl/pages/risk/articles/cyber-security-privacy-gdpr-update-the-impact-on-cloud-computing.html). These regulations require businesses to obtain explicit consent from individuals before collecting and processing their data and grant individuals various rights over their data. Implementing the necessary mechanisms to ensure compliance, such as setting data retention periods, access controls, and data portability, can be challenging within the context of cloud optimization tools.

## **3. Complexity and learning curve**

Implementing cloud optimization solutions can be complex, requiring a learning curve for both IT teams and business stakeholders. Complexity might lead to slower adoption or implementation challenges. Companies need to consider the learning curve for the cloud optimization tool they intend to use and if there are skilled personnel to operate them at an optimal level.

Effective cloud optimization tool usage requires continuous monitoring, analysis, and decision-making. In the absence of skilled manpower, organizations may fail to capitalize on the platform's automation capabilities, missing out on opportunities to dynamically scale resources, reduce costs, and ensure compliance. Furthermore, the lack of skilled personnel can lead to inefficient troubleshooting and problem resolution. Technical issues and challenges that arise within the optimization tool or the broader cloud environment might not be effectively addressed, leading to prolonged downtime, decreased productivity, and potential security vulnerabilities.

With the [looming shortage of cloud labor and skills](https://www.spglobal.com/marketintelligence/en/news-insights/research/closing-the-cloud-skills-gap-a-perennial-problem-for-businesses), companies might find it difficult to hire skilled and qualified personnel, deterring them from using cloud optimization tools and opting for managed services provided by CSPs.

# 

# **Glossary**

| **Term** | **Description** |
| --- | --- |
| Access management | Processes and techniques used to control and manage the permissions and privileges granted to users, applications, and services accessing cloud resources |
| Auto-scaling | The process of automatically adjusting the number of compute resources based on changing workload demands. |
| Automation | Using scripts, tools, and workflows to automate various cloud management tasks, such as provisioning, scaling, and configuration management. |
| Budgetary limits | Predefined spending thresholds set by organizations to control and manage their cloud expenses effectively. |
| Cloud cost allocation | The process of attributing cloud costs to specific projects, departments, or teams, enabling better budgeting and resource optimization. |
| Cloud governance | Predetermined set of policies and rules used to ensure adherence to industry and company standards. |
| Cluster | Group of interconnected servers or nodes that work together to manage and deploy containerized applications. |
| Containers | Lightweight, portable, and self-contained environments that package an application and its dependencies, allowing it to run consistently across different computing environments. |
| FinOps (Financial Operations) | A concept established by the FinOps Foundation (an offshoot of the Linux Foundation) that combines financial and operational processes to optimize cloud spending. |
| Guardrails | Proactive control measures that are deployed to maintain compliance, budgetary control, and security posture |
| Identity access management (IAM) | Security framework that governs and manages user access to resources to ensure security and compliance. |
| Idle resources | Cloud resources that are provisioned but not actively utilized. |
| Instances | A specific configuration of virtual machine resources, including CPU, memory, and storage, optimized for different types of workloads. |
| IT chargeback | A cost allocation method where IT costs are billed to individual departments or business units based on their actual usage of IT resources and services. |
| IT showback | A financial model that provides departments or business units with visibility into their IT resource consumption and associated costs but doesn't involve actual billing. |
| Load balancing | Distributing incoming network traffic across multiple compute resources, ensuring even distribution of workloads and improved availability. |
| Multi-cloud | Using multiple cloud service providers simultaneously to access a variety of cloud services and resources |
| Reserved instances | Pre-paid capacity reservations for cloud resources that provide cost savings over on-demand pricing in exchange for a commitment to a specific term. |
| Resource provisioning | Allocating and adjusting resources per workload needs to avoiding over or under provisioning. |
| Resource tagging | Applying metadata to cloud resources to categorize, track, and manage them efficiently for cost allocation and optimization. |
| Security posture | The overall security stance, readiness, and effectiveness of an organization's cloud environment in safeguarding its assets, data, applications, and services from potential threats, vulnerabilities, and attacks. |
| Cloud cost allocation | The process of attributing cloud costs to specific projects, departments, or teams, enabling better budgeting and resource optimization. |
| Auto-scaling | The process of automatically adjusting the number of compute resources based on changing workload demands. |
| Multi-cloud | Using multiple cloud service providers simultaneously to access a variety of cloud services and resources |
| Automation | Using scripts, tools, and workflows to automate various cloud management tasks, such as provisioning, scaling, and configuration management. |

*Last updated: September 2023*

©2024 Uzabase, Inc. All Rights Reserved. The information contained herein: (1) is proprietary to Uzabase Inc. and/or its content providers; (2) may not be copied or distributed; and (3) is not warranted to be accurate, complete or timely. Neither Uzabase Inc. nor its content providers are responsible for any damages or losses arising from any use of this information.