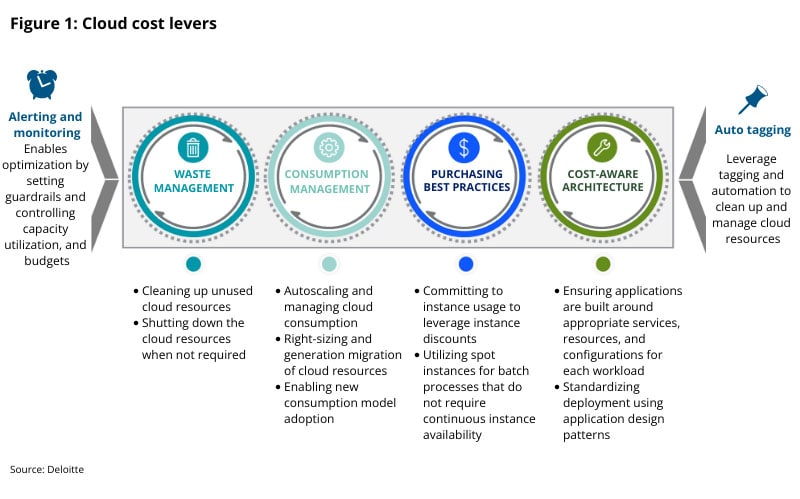
# Optimizing the value of cloud: Four key optimization levers

key optimization levers

n the previous post in this three-part series, we discussed how businesses can optimize cloud costs using four crucial levers: waste management, consumption management, purchasing leading practices, and cost aware architecture. Further, each of these levers is supported by two enablers: alerting and monitoring, and automated tagging. In this post, we explore each of these levers in-depth.

While all the cost optimization levers work in tandem with each other, there are clear benefits that can be achieved with each of these levers, as shown in figure 1.

[](https://www2.deloitte.com/content/dam/Deloitte/us/Images/blog/posts/us-cloud-key-optimization-levers-new.jpg)[Click on this image to expand]

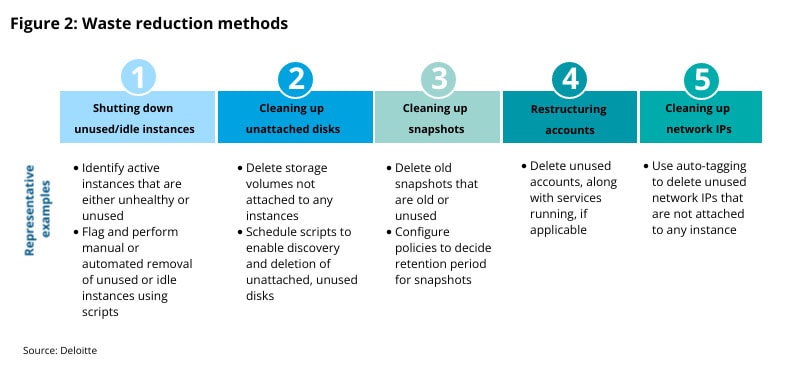
#### Cost optimization levers in depth

Let’s dive deeper into each of these levers and enablers to explore how they can help businesses optimize their cloud spend.

**Waste management**

Waste management is a critical FinOps lever that helps minimize wasteful expenses and refers to the management of resources that generate waste in terms of usage and cost. For example, an instance and its underlying storage no longer in use are potential waste, leading to avoidable expenses. Waste management helps to navigate the challenges posed by cloud environments where resources are being continually provisioned, scaled, and de-provisioned. Significant cloud spend is also wasted via instances that are over-provisioned and due to lack of optimized networks across the three major public cloud providers every year.

Common ways to reduce cloud waste are included in figure 2.

[](https://www2.deloitte.com/content/dam/Deloitte/us/Images/blog/posts/us-cloud-waste-reduction-methods.jpg)[Click on this image to expand]

**Consumption management**

Consumption management refers to the management of resources that are currently in use but are over-provisioned. For example, an instance in use with low usage metrics might be an over-provisioned resource. With the ease that cloud resources can be provisioned spur of the moment, it becomes imperative to manage consumption patterns to avoid excess costs. Some common consumption management practices are:

* **Replacing expensive storage:** Evaluate storage requirements against parameters such as ideal storage type, redundancy options, and access frequency to select the optimal storage and prevent overspending.
* **Migrating to lower-cost instance types or regions:** Switch costly instance series to cheaper, older instance series where possible, and select regions with lower costs.
* **Rightsizing instances:** Analyze resource consumption of virtual machines to determine the optimal instance size, and use cost-monitoring tools.
* **Leveraging a reserved and spot instance strategy:** Analyze current instance usage and future demand levels to identify a reserved instance strategy, and utilize spot instances for batch processes that do not require continuous instance availability.
* **Automating optimization:** Use automation to create alerts that trigger automated actions to resolve inefficiencies, and use autoscaling systems to utilize resources effectively.
* **Tagging analysis:** Apply organization-wide tagging policies for better tracking and visibility into account-level spending patterns, and allocate costs to appropriate cost centers and business units.
* **Reducing data transfer costs:** Analyze data transfer across one-way and two-way traffic and between instances on different availability zones, virtual private clouds, or regions. Identify savings opportunities by placing workloads where data transfer costs can be optimized.

**Purchasing leading practices**

Purchasing leading practices refer to tactics to utilize all available procurement levers to manage the complexities of cloud provider pricing models, optimize rates, and drive unit cost efficiency. Over long periods, businesses can significantly reduce their cloud spend by leveraging purchasing leading practices while provisioning cloud resources. Following these leading practices can result in the following benefits:

* **Direct cost savings by:**
  + Helping achieve the exact portfolio of resources at a discounted price
  + Creating opportunities to continuously increase the resource-to-dollar ratio
* **Gaining operational efficiency by:**
  + Creating customized instance configurations that best meet requirements
  + Scaling up or down by integrating reserved resources with on‑demand resources
  + Providing flexibility to modify, convert, exchange, resell and cancel reserved resource

Implementing cloud purchasing leading practices is not a one-time activity, but a gradual process that evolves with time. Businesses can take three key steps to help ensure that they can purchase cloud resources on favorable terms

**1. Data gathering and filtering**

* Ensure that all instances are properly “tagged” to identify consumption trends
* Filter out instances with low utilization (e.g., instances with uptime of less than 65%) to check if convertible RIs or spot instances can replace them.

**2. Data aggregation**

* Aggregate hourly data by various attributes, such as instance type, environment, application type, operating system, tenancy, and availability zone.
* Review the most expensive groups to identify right-sizing opportunities.

**3. Scenario identification**

* Consider savings plans and spot instances along with convertible reserved instances to achieve maximum discounts, along with automating the purchasing activity.
* Identify regions/zones to buy resources from without compromising on latency.

**Cost-aware architecture**

Traditional cost optimization efforts focus on optimizing applications already running on public cloud infrastructure through the IaaS-focused levers—consumption management, waste management, and purchasing tactics. However, very few organizations have evaluated the potential of addressing the application architecture and revisiting architectural decisions to drive even higher levels of optimization.

A cost-aware architecture can help optimize costs by ensuring applications are built around appropriate services, resources, and configurations for each workload. Application design patterns are adopted by business units to standardize deployment and manage workloads. Additionally, these patterns can help set preventive and detective controls (or guardrails) to enable future cost tracking.

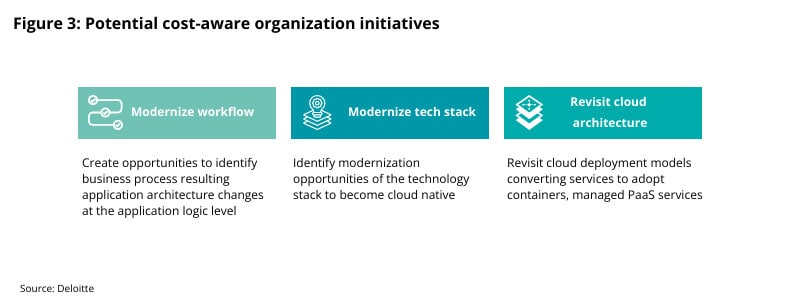
Cost-aware architectures focus on three main goals:

1. **Be better, faster, and more cost effective.**Apply an application lens to cost optimization by looking under the hood of applications and evaluating all aspects of the architecture for greater effectiveness, efficiency, and agility.
2. **Drive cost-aware culture.**Drill down into the application layer to understand cloud consumption that product development teams understand, and re-enforce best practices at the organization level.
3. **Enable cloud-native capabilities.**Set the foundational structure to effectively respond to the increasing velocity, complexity, and volume of business needs.

Key measures in achieving cost-aware architecture are listed below:

* **Address spend for large monolithic applications** where significant cloud consumption occurs and right-size accordingly.
* Develop **architectural recommendations** to improve the cost profile for your applications.
* Evaluate architectural decisions with a cost lens (individually or as a group) to determine the **ROI of executing application changes**.
* Build a **cost road map** alongside a technical road map to provide a clear picture of **benefits realization**.
* **Standardize architecture patterns** for better governance across applications.
* **Leverage cloud-native capabilities**, such as containers and serverless technologies, to manage costs.
* Build processes aimed at **eliminating technical debt**.
* Drive cultural norms toward **enterprise-wide cost** **awareness** through change management efforts.

Cost-aware architecture-based optimization efforts could lead to a more robust architecture and can be categorized into following initiatives (figure 3):

[](https://www2.deloitte.com/content/dam/Deloitte/us/Images/blog/posts/us-cloud-potential-cost-aware-org.jpg)[Click on this image to expand]

#### Enablers for cloud cost optimization

While each of the above levers is crucial to optimize cloud costs, they are collectively driven by two enablers: auto-tagging, and alerting and monitoring.

**Auto-tagging:**

Auto-tagging refers to mapping cloud resources and services to organizational and operational metadata by leveraging scripts and automations. This can help eliminate the need to manually tag resources, enabling more efficient control of cloud resources. Effective auto-tagging can help enterprises enhance their operational efficiency, improve governance, and add conditional tags based on business context.

**Alerting and monitoring:**

Alerting and monitoring is using alerts driven by capacity, utilization, or budget thresholds for effective control of cloud resources and services. This enables holistic and sustainable cloud financial management. Alerting and monitoring helps to optimize current cloud costs and sets guardrails for future deployments and cost tracking.

Alerting and monitoring help businesses by enabling real-time usage and cost monitoring, improving real-time data analysis, reducing incident response time, driving operational effectiveness, enhancing spend visibility, and helping right-size resources.

#### Wrapping it up

To recap, cloud cost optimization involves a combination of improving visibility and accountability of cloud spend, continuous efforts to reduce known and unattributed costs, and enhancing planning and forecasting. For effective cloud cost optimization, businesses should start with assessing their current-state maturity levels, cloud costs, governance mechanisms, FinOps capabilities, and operating models.

Depending on their maturity levels, businesses should identify gaps against the target state they want to achieve. Next, enterprises should align their cloud cost optimization efforts across four key levers: waste management, consumption management, purchasing best practices, and cost-aware architecture.

Each of these levers, supported with enablers such as alerting and monitoring and auto-tagging, help drive operational and financial improvements, with direct impacts on the bottom line and productivity. These efforts should be aided with cost optimization tools provided by cloud service providers and third-party market leaders.

*To learn more:*

###### *A blog post by*[*Nikhil Roychowdhury*](https://www2.deloitte.com/us/en/profiles/nikhil-roychowdhury.html)*, principal, Deloitte Consulting LLP;*[*Nik Jethi*](https://www.linkedin.com/in/nikjethi/)*, senior manager, Deloitte Consulting LLP;*[*Farhan Akram,*](https://www.linkedin.com/in/farhanakram91/)*manager, Deloitte Consulting LLP;*[*Rishabh Kochhar*](https://in.linkedin.com/in/rishabhkochhar)*, senior consultant, Deloitte Consulting LLP.*

This post is the first in a series of three posts that highlight how businesses can get started with optimizing their cloud costs. While in the first post we look at why cloud cost optimization is a vital business need, and describe a practical framework that businesses can adopt, the subsequent posts will dive deeper into each lever of the framework and describe how businesses can derive value from them.

**Growing need of cloud cost optimization**

The increasing adoption of cloud computing has helped transform and streamline the way businesses operate, but enterprises find themselves facing a new challenge that has a direct impact on their bottom line—inefficient cloud spend. With Infrastructure as a Service (IaaS) and Platform as a Service (PaaS), businesses incur operating expenses (OpEx) as they are billed based on ongoing consumption, instead of capital expenses (CapEx) associated with capacity in a data center. Consequently, businesses often find themselves trying to create complex cost estimates of their consumption on cloud, which affects their financial management—ultimately having an impact on businesses’ bottom line.

There are six key factors that drive complexity of cloud cost optimization:

* **Complex and multifaceted pricing:** Cloud billing models and pricing structures have options and combinations that are hard to understand, and selecting the best options can be challenging.
* **Granularity of cloud bills:**Cloud bills often run into thousands of line items despite low consumption, making it difficult to attribute costs and consequently charge back.
* **Click provisioning of cloud services**: Easy access to cloud services through web consoles and APIs can lead to an increase in the number of cloud services being used, driving up expenses.
* **Availability of multiple architectures:** Applications can be built using multiple architectures and services, making it hard for businesses to identify the most cost-effective development strategy.
* **Constant check in cloud offerings**: Cloud services, features, instance types, and pricing models undergo frequent updates, and businesses can find it hard to understand their financial impact.
* **Lack of standardization:** Most cloud platforms do not provide any standardization of billing models, billing formats, APIs, or services.

**Preparing for cloud cost optimization**

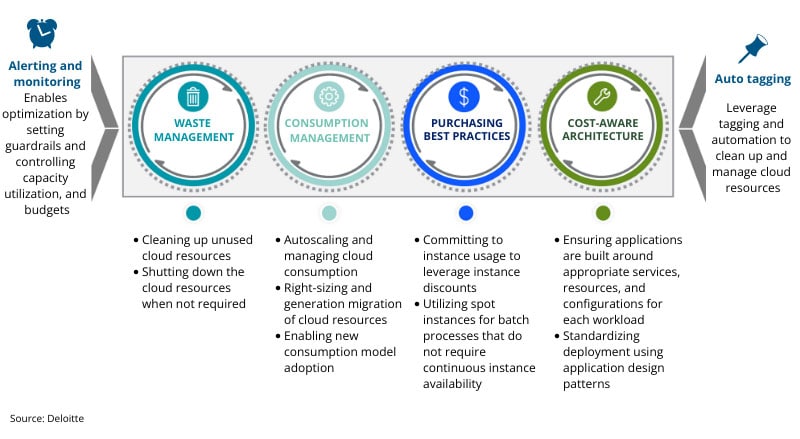
Given the complexities involved with optimizing cloud costs, businesses should begin with examining their current spend and existing practices across the FinOps capabilities. This can help identify existing gaps, thereby helping define a strategy and road map for optimization execution.

Our experience has shown that the following five steps can help businesses get started with reviewing their existing processes and strategy:

[](https://www2.deloitte.com/content/dam/Deloitte/us/Images/blog/posts/us-cloud-preparing-for-cloud-optimization.jpg)[Click on this image to expand]

**A framework for optimizing cloud costs**

A thorough assessment of an enterprise’s current-state capabilities and governance models will help organizations identify cloud costs optimization opportunities. These opportunities can be systematically evaluated leveraging a framework that is built upon **four**crucial levers: waste management, consumption management, purchasing best practices, and cost-aware architecture. Each of these levers is supported by **two**enablers: alerting and monitoring, and automated tagging.

[](https://www2.deloitte.com/content/dam/Deloitte/us/Images/blog/posts/us-cloud-key-optimization-levers-wo-title.jpg)[Click on this image to expand]

While all cost optimization levers work in tandem with each other, there are clear benefits that can be achieved with each of these levers. The levers, mapped to their benefits, are:

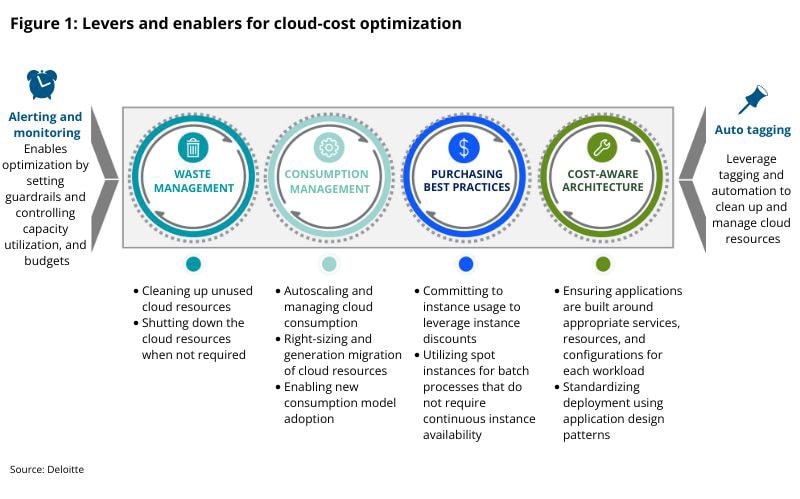
* **Waste management**
  + Shut down unused instances: Reduce cost of unused instances; shutting down instances enables run features for different environments.
  + Unattached disks cleanup: Promote reduction of waste and enable IT teams to focus holistically on core capabilities, improving time to market.
  + Snapshots cleanup: Identify areas to delete snapshots that are old or unused; decide retention for snapshots.
  + Account restructuring: Reduce costs by deleting unused accounts along with services running, if applicable.
  + Network IP cleanup: Delete unattached network IPs to reduce unused resources.
* **Consumption management**
  + Expensive storage replacement: Reduce costs by evaluating the storage requirements on parameters such as storage type required (SSD vs. HDD).
  + Instance rightsizing: Remove resource consumption of your VMs, determine correct instance size for optimum cost and performance and to reduce overprovisioning.
  + Automating optimization: Autoscale systems to utilize resources effectively, which results in continuous optimization.
  + Tagging analysis: Improve asset costing, which helps manage expenses and derive better cost insights.
  + Instance discounting: Provide spin transparency across the enterprise and in understanding the performance of assets.
* **Purchasing best practices**
  + Migration to lower-cost instance types or regions: Reduce technical debt and improve asset planning driven through efficient budgeting and forecasting.
  + Leveraging reserved and spot instance strategy: Promote efficient budget planning and tighter spend control through improvements in FinOps.
  + Enterprise agreement: Create opportunities to obtain better deals and discounts on cloud cost, depending on how accounts are set up.
* **Cost-aware architecture**
  + Modernize workflow: Create opportunities to identify business process resulting in application architecture changes at the application logic level.
  + Modernize tech stack: Identify modernization opportunities for the technology stack to be cloud native.
  + Refactor architecture: Revisit cloud deployment models converting services to adopt containers and managed PaaS.

**Wrapping it up**

Optimizing cloud costs is important for every business. In today’s hypercompetitive environment, it is very easy to lose track of cloud services and resources being used, as businesses scale and develop new products to meet market needs and maintain a competitive edge. This tends to have a direct impact on an enterprise’s bottom line. Without a well-thought-out strategy to optimize cloud costs, the cloud becomes yet another cost center like traditional IT.

Cloud cost optimization involves a combination of improving visibility and accountability of cloud spend; continuous efforts to reduce known and unattributed costs; and enhancing planning and forecasting. For effective cloud cost optimization, businesses should start with assessing their current-state maturity levels, cloud costs, governance mechanisms, FinOps capabilities, and operating models. Depending on their maturity levels, businesses should identify gaps compared to the target state they want to achieve.

n the previous posts in this three-part series, we discussed [how businesses can optimize their cloud](https://www2.deloitte.com/us/en/blog/deloitte-on-cloud-blog/2023/a-practical-guide-to-optimizing-the-value-of-cloud.html) using a framework composed of four levers: waste management, consumption management, purchasing best practices, and cost-aware architecture. Each of these levers is supported by two enablers: alerting and monitoring, and automated tagging.

[](https://www2.deloitte.com/content/dam/Deloitte/us/Images/blog/posts/us-cloud-key-optimization-levers-enablers.jpg)[Click on this image to expand]

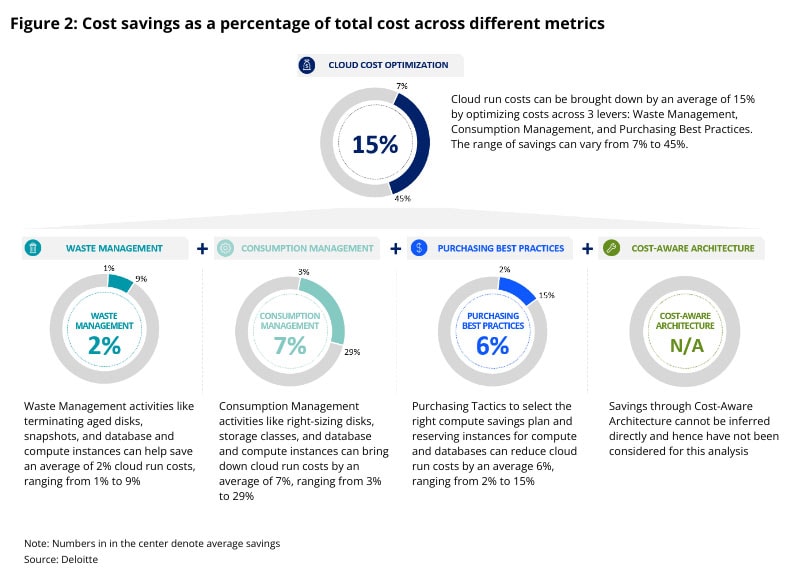
Then [we dove deeper into each of the optimization levers](https://www2.deloitte.com/us/en/blog/deloitte-on-cloud-blog/2023/four-key-optimization-levers-for-cloud.html) and enablers to get a better understanding of how businesses can approach cloud spend optimization. In this final post in the series, we explore the impact businesses can expect from cloud cost optimization initiatives.

#### Expected savings from cloud-cost optimization initiatives

Deloitte recently analyzed anonymized data from several of our recent FinOps engagements to assess the value of the savings that businesses can achieve by driving efforts toward cloud-cost optimization. Early results of our analysis suggest that businesses can save up to 45% (15% on average) on their cloud costs by optimizing those costs across waste management, consumption management, and purchasing best practices levers.

Our research indicated that, while waste management activities, such as terminating unused resources, are relatively straightforward to implement, they will also result in only limited savings (potentially from 1% to 9%). The majority of the savings typically come from consumption management and purchasing tactics levers. We also found that, based on the FinOps maturity of an organization, consumption management activities like right-sizing can bring savings up to 29% (7% on average). The research also revealed that purchasing best practices, such as selecting the right compute savings plan and reserve instances for computer and databases, can bring savings up to 15% (6% on average).

Our analysis did not consider cost savings due to cost-aware architecture because of the limited scope of this lever on the engagements we analyzed. However, the savings are expected to further improve with cost-aware architecture-focused optimizations.

[](https://www2.deloitte.com/content/dam/Deloitte/us/Images/blog/posts/us-cloud-key-optimization-cost-savings.jpg)[Click on this image to expand]

#### Value savings and realization

In addition to reporting savings from cost optimization, it’s essential for businesses to quantify the additional value delivered. Cost optimization benefits are typically quantified in terms of lower costs per business outcome. The benefits from cost optimization, however, go above and beyond cost reduction or avoidance and encompass improvements in efficiency and business value.

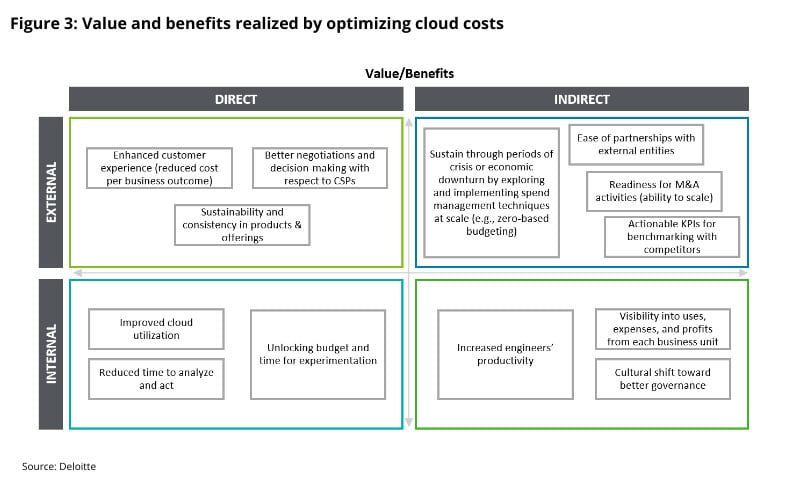
**Benefits in business outcomes through unit economics**

Tracking the business outcomes of an effective cost optimization strategy on cloud is best done through the lens of cloud unit economics—a system that fosters profit maximization by tracking unit cost metrics specific to the delivery and development of cloud-based software or tooling against marginal revenue. The difference between marginal cost and marginal revenue helps in making informed decisions by presenting facts, such as the break-even point for the service, business-unit performance, etc. Additionally, the impact of end-user behavior on cloud costs can be measured and used to drive more cost-conscious behavior.

Unit economics can also highlight areas where usage may need to be shaped/governed. The amount of business value achieved per unit of cloud spend or metrics like cost per weekly active user or cost per million transactions can help leaders focus on key products/offerings and double down on efforts for products that they know can serve customers at the lowest cost and beat the competition. The connection between current business demand and cloud costs at levels such as per-customer, per-feature, per-product, etc. are key to driving optimization efforts.

The end goal of cloud-cost optimization efforts is to maximize business outcome per dollar spent on cloud. Often, however, certain kinds of benefits escape our purview. It is recommended that businesses quantify the additional value delivered—those that go above and beyond cost reduction or avoidance and encompass improvements in efficiency and business value.

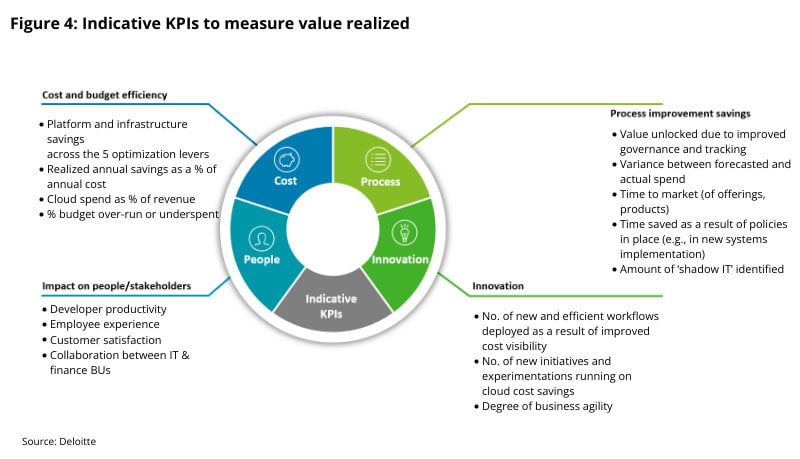
The value delivered can be seen as either internal or external to the company. For example, a robust cloud-cost optimization strategy can help add new features without driving additional costs too much. This would also count as a direct consequence of our optimization efforts. Similarly, developer productivity is indirectly enhanced through automated cost optimization practices.

[](https://www2.deloitte.com/content/dam/Deloitte/us/Images/blog/posts/us-cloud-key-optimization-value-benefits.jpg)[Click on this image to expand]

#### Tracking value realized through cost optimization

Apart from some standard KPIs highlighted above, it’s important to have a set of indicators to capture the benefits and value realization because of efforts spent on cost optimization.

The framework below suggests various KPIs that businesses can consider to track the value realized through cost optimization. Please note that the KPIs below are only representative of KPIs used by many organizations for cloud-cost optimization. Individual organizations should assess their needs and tweak as per their requirements.

[](https://www2.deloitte.com/content/dam/Deloitte/us/Images/blog/posts/us-cloud-key-optimization-kpis.jpg)[Click on this image to expand]

#### Moving forward with cloud-cost optimization

Optimizing cloud costs is important for every business. In today’s hypercompetitive environment, it is very easy to lose track of cloud services and resources being used as businesses scale and develop new products to meet market needs and maintain a competitive edge. This tends to have a direct impact on an enterprise’s bottom line. Without a well-thought-out strategy to optimize cloud costs, the cloud becomes yet another cost center like traditional IT.

Cloud-cost optimization involves a combination of improving visibility and accountability of cloud spend, engaging in continuous efforts to reduce known and unattributed costs, and enhancing planning and forecasting. For effective cloud-cost optimization, businesses should start with assessing their current-state maturity levels, cloud costs, governance mechanisms, FinOps capabilities, and operating models. Depending on their maturity levels, businesses should identify gaps vs. the target state they want to achieve.

Next, enterprises should align their cloud-cost optimization efforts across four key levers (i.e., waste management, consumption management, purchasing best practices, and cost-aware architecture). Each of these levers, supported with enablers such as alerting and monitoring and auto-tagging, helps drive operational and financial improvements, with direct impacts on the bottom line and productivity. This should be aided with cost-optimization tools provided by cloud service providers and third-party market leaders to optimize.

#### A final word

Cloud-cost optimization is not a one-time activity; instead, it is an ongoing activity that evolves with time and grows as a business matures. A well-defined cloud optimization strategy, complemented by a robust governance structure, can help businesses improve their bottom lines, enhance business agility, and become more resilient in the face of stiff competition.

*To learn more:*

Read the first blog in this series, [Optimizing the value of cloud: A practical guide to getting started](https://www2.deloitte.com/us/en/blog/deloitte-on-cloud-blog/2023/a-practical-guide-to-optimizing-the-value-of-cloud.html).

Read the second blog in this series, [Optimizing the value of cloud: Four crucial levers to optimize cloud costs](https://www2.deloitte.com/us/en/blog/deloitte-on-cloud-blog/2023/four-key-optimization-levers-for-cloud.html).