zensar

CIOs guide to **Know**, **Control**, and **Optimize** (KCO) the cloud spend

Whitepaper

zensar

Contents

Cloud adoption has rapidly increased over the last five years and will grow at approximately 30% compounded annual growth rate (CAGR). The cloud adoption provides faster time-to-market, high-velocity operations, innovation and scale. The cloud provides a paradigm shift in spend management from the capital expenditure (CapEx) to the operational expenditure (OpEx) model. The spend management for cloud services consumption requires a different type of governance approach when compared to traditional IT resources cost management when the workload is on-premises.

For a chief information officer (CIO), cloud hosting spends where most of the workload is on the cloud is expected to be 40 to over 50 percent of the overall IT spend. That is a significant portion of the CIO's budget. According to research by RightScale, "Most organizations have 30-35 percent wasted spend on cloud resources". CIOs need to establish processes for cloud spend management to avoid wasted spend. We believe cloud services cost management requires the right processes, tools, and discipline. It is recommended to consider cloud spend management as a new task rather than something done on an ad-hoc basis. The focused cloud cost management task can make cloud spends predictable and provide opportunities to optimize organizational cloud spends.

This whitepaper provides a simple framework for efficiently managing the cost of cloud services. We also share good practices to enable proactive cloud services cost management across the heterogeneous cloud. We have included a list of valuable tools from key providers for cloud cost management.

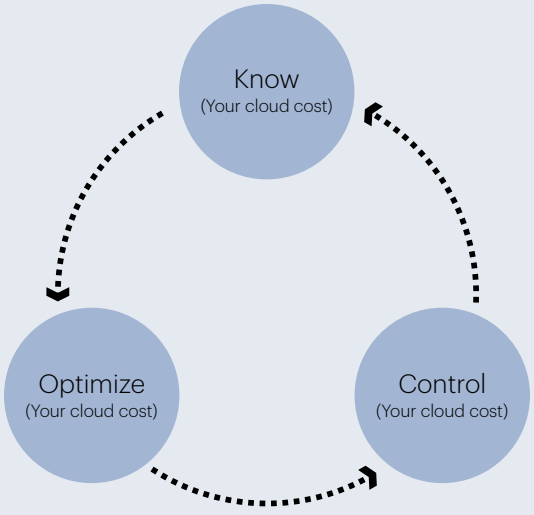
Manage cloud cost

Cloud service providers provide a wide range of services that enable customers to create innovative products using a combination of cloud services.

- Each of the major cloud hosting providers offer an average of over 150+ cloud products
- Amazon Web Services (AWS) offers 290 different virtual machine sizes

The flexibility to use many services creates a complex bill of materials (BOM) of cloud services to track the spending.

We have created a simple framework for managing cloud costs from our learnings in cloud cost management in the marketplace. The cloud cost management framework is called **KCO** (pronounced **Kay-KO**) and has three main phases — **Know, Control,** and **Optimize** (KCO). Each of these phases has key processes described in the sections below. The framework phases work in a continuous cycle as



depicted in figure-1

Figure-1: KCO Cloud Cost Management Framework

Know your cloud cost

The transformation of the compute, network, and storage spend from primarily CapEx to OpEx model requires new processes and tools. The stakeholders across finance, engineering and management must be able to see cloud OpEx as per their role for their decision making.

Peter Drucker once famously said, "You can't manage what you can't measure!". We need to see cloud OpEx spends closely from the bottom up to the desired organizational hierarchy for cloud spend management. The correct granular analysis of the OpEx will give greater visibility into cost spends across finance and engineering departments and among the senior management.

Establish the right hierarchy for the cloud resources: Stakeholders like to see spends in various ways. They want to see cloud OpEx by:

- Project
- Program (group of projects)
- Business unit
- Product (or business value stream)
- Business function (e.g., sales, marketing, HR, etc.)
- Environment (e.g., production, testing, etc.)
- Resource type (e.g., compute, network, storage, etc.)
- Region
- Others

To develop the capability to visualize cloud spends at various granular levels, you need to establish the right cloud resource hierarchy in your cloud environment. We want to share a simple view of resource hierarchies that is agnostic to cloud providers. The hierarchy shown in Figure-2 is a logical hierarchy of cloud resources that can be applicable for medium to large-size organizations. Each key cloud provider supports a similar logical hierarchy of resources. However, they may use slightly different terminologies for defining hierarchy.`

It is a good practice to group resources, as depicted in Figure-2.

1. Resource grouping by resource groups
2. Resource groups can be grouped by projects or application or product
3. Projects can be grouped by program or group of projects or group of products
4. Programs can be grouped by portfolio or organization responsible for overall spends in the hierarchy

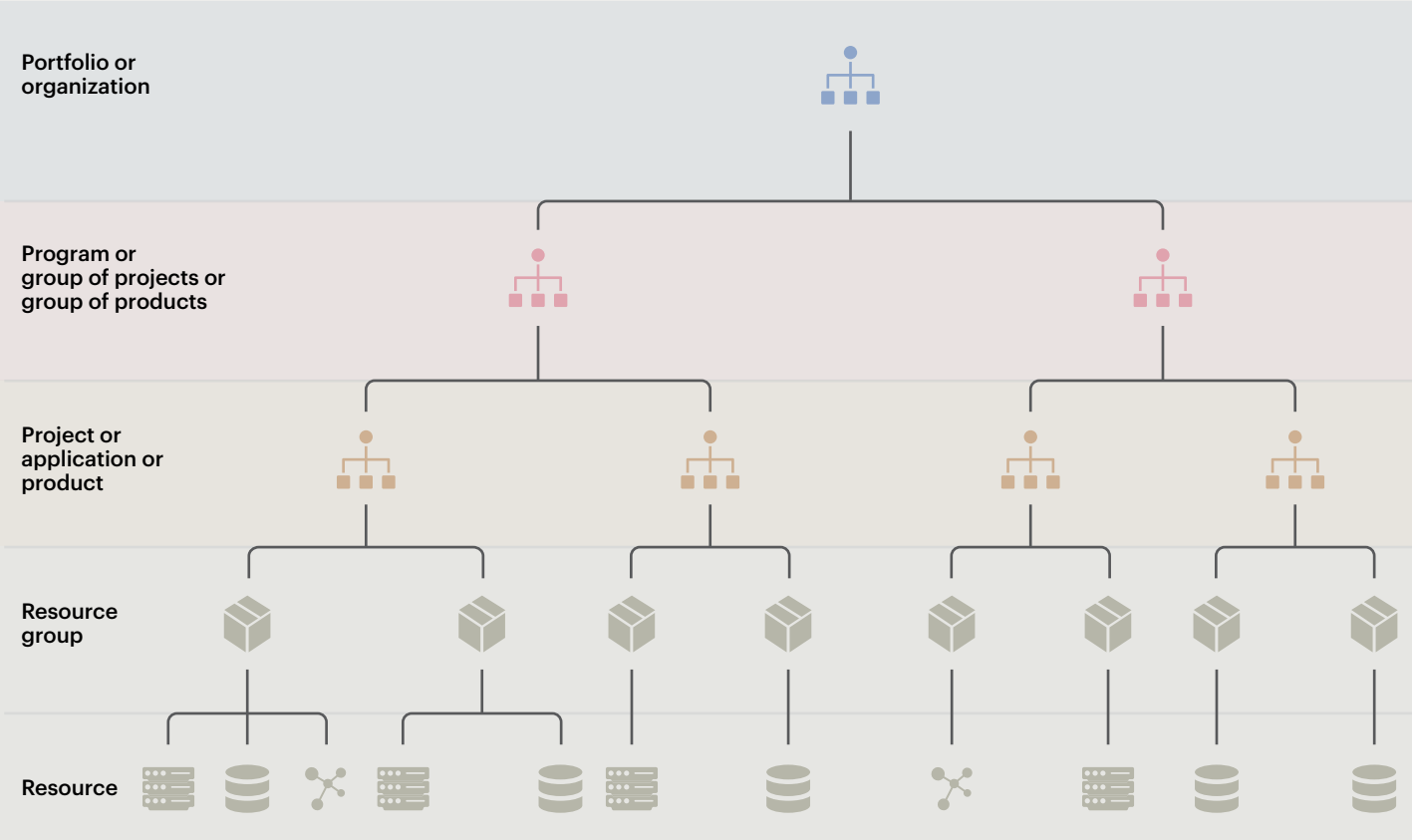


Figure-2: Logical resource hierarchy in cloud

In addition to the resource hierarchy, organizations should leverage labels or tags for cloud resources and the layers of the hierarchy. For example, the environment label can be used to label all resources in the testing environment. Labels add the convenience of categorical filtering of cloud spends.

The combination of hierarchy and labels gives the flexibility to create cloud spend reports from the granular level of individual resources to the project, program, & organizational levels.

Visibility and accountability: Various stakeholders from the finance, engineering, & management departments will participate in cloud OpEx governance. We are required to define three types of roles.

- **Administrators:** Can provide access to others, manage resources, and manage billing
- **Contributor:** Can read and manage billing
- **Reader:** Can view billing reports

The stakeholder can be assigned these roles at a particular stage in the resource hierarchy. The stakeholder with the part of a contributor can read billing data up to the hierarchy assigned to them. For example, a stakeholder tasked as a contributor for project-1 cloud spend cannot see billing data for project-2.

For the analysis of cloud spends, stakeholders should also know about the various aspects of cloud costs.

1. Resource unit price for your organization
2. Type of resource pricing, such as reserved instance pricing, spot pricing, and pay-as-you-go pricing
3. Overview of resources in their scope as per the hierarchy

The stakeholders are recommended to analyze the following reports daily.

- Spend by project, program, and portfolio level
- Actual spend vs. budget vs. forecast
- Spend by resource type (e.g., compute, storage, network, etc.)

Control your cloud cost

Organizations require a disciplined approach to monitor and control cloud spends. Many organizations have the role of cloud spend analysts responsible for regularly monitoring, analyzing, and optimizing cloud spends. They can also alert others on the spending. The cloud spends control function is generally modeled as a hybrid of centralized and decentralized. The portfolio-level cloud cost control function is centralized, and the centralized team will focus on the portfolio, or organization-level cloud spends. The stakeholders from finance and engineering teams can participate. The project and program level cloud spend control are decentralized across program teams. The program-level teams will regularly look at cloud spends at a program and project level.

Setup for success: For effective control of cloud spends, organizations need an effective resource hierarchy setup. Role-based access control (RBAC) allows stakeholders across finance, engineering, and management teams to track spends without affecting the setup. They can use labels or tags to track costs at a more granular level. The recommended resource hierarchy is depicted in Figure-2. The RBAC-based access is provided in the visibility and accountability section. The process for label creation and using labels for tracking cloud spends at a granular level is specific to organizations. The general categorization of labels are:

1. Labels for the environment (e.g., development, testing, performance, production, etc.)
2. Labels for departments for shared resources
3. Labels for business processes and/or value streams. This can help track cloud costs by process and value stream (e.g., order to cash (OTC), procure to pay (PTP), order fulfillment, etc.)

We recommend the centralized team maintain the label strategy for consistency and traceability.

Manage cloud cost: Managing cloud spends requires budgeting, forecasting, actual spending, using business events data to create more accurate budgets, and understanding cloud spends.

Budgeting and forecasting cloud spend: Organizations generally do CapEx and OpEx budgeting and forecasting using tools like Workday Adaptive and Oracle Hyperion tools. Organizations should start budgeting for cloud OpEx as part of the budgeting cycle. The cloud OpEx budgeting can leverage historical cloud spend data and future business events to predict cloud spend budget. The major cloud vendors also provide tools to budget cloud spend for the next cycle. The cloud spends budget should be entered into the cloud management platform. The budget configuration should have multiple thresholds levels from the baseline budget to raise alerts to stakeholders if spending is crossing the thresholds. This will enable proactive tracking of actual cloud spend as compared to the budget and forecast.

Analyze actual spend vs. budget and forecast: The cloud spend analysis should be done by stakeholders daily for their scope of resources. The analysis should track actual spending to the budget and forecast. In case there is an increase in spending above the set thresholds, stakeholders should be alerted. The stakeholders should act to:

1. Understand the rise in cloud spends. In most cases, it is due to business reasons (e.g., eCommerce cloud spend may increase due to a successful marketing campaign launch)
2. Plan to make changes in the budgeted funds for upcoming days within the given cycle
3. Remove inefficient cloud resources that are causing unwanted cloud spends

Other activities: Stakeholders should export the cloud billing data to analyze historical trends outside the cloud provider's billing portals. For example, stakeholders can export cloud billing data to BigQuery or Redshift, or any other data lake platform. The historical data can create more insightful reports using tools like Looker, Tableau, and Power BI.

There are a few recommended reports that should be created using historical data. These reports have to be set up and are not available by default from the cloud providers.

1. Cloud spend as a percentage of the unit cost of goods sold (COGS)
2. Cloud spend by value stream or process (e.g., order to cash value stream, per order fulfillment, etc.)
3. Change in cloud spend by projects/products

Govern: The cloud spend reports are emailed daily to the stakeholders, and cloud spend alerts are raised in real-time to make decisions. The finance, engineering, and management teams should meet regularly for governing cloud spends. We recommend bi-weekly meetings to review cloud spends and identify action items for governing and optimizing cloud spends. The cloud OpEx should also be a part of the organizational level OpEx spend review generally done by the finance department.

Optimize your cloud cost

According to research by Flexera in 2020, 73 percent of organizations want to optimize their cloud spends. This section will cover the areas of wasted spends and how organizations can optimize spends better.

Figure-3 depicts various options for cloud spends optimization. These options are depicted on the scale of the potential of savings and the complexity of doing suggested changes.

Reserved instance plan or committed use discount: Many major cloud providers offer heavy committed use discounts (CUD) and reserved instances (RI) on resources. Analyze the resource usage pattern by resource type (e.g., compute, storage, network, etc.) over a period and identify a threshold unit count of resources that is the bare minimum in your environment. You can commit to using those resources upfront to choose CUD or RI.

Spot instance pricing: Major cloud providers allow bids for the resource unit cost from the unused capacity. The spot instance price is significantly lower than the regular price. For non-time-sensitive workloads, you can plan to use spot instance pricing, e.g., data processing and batch workload.

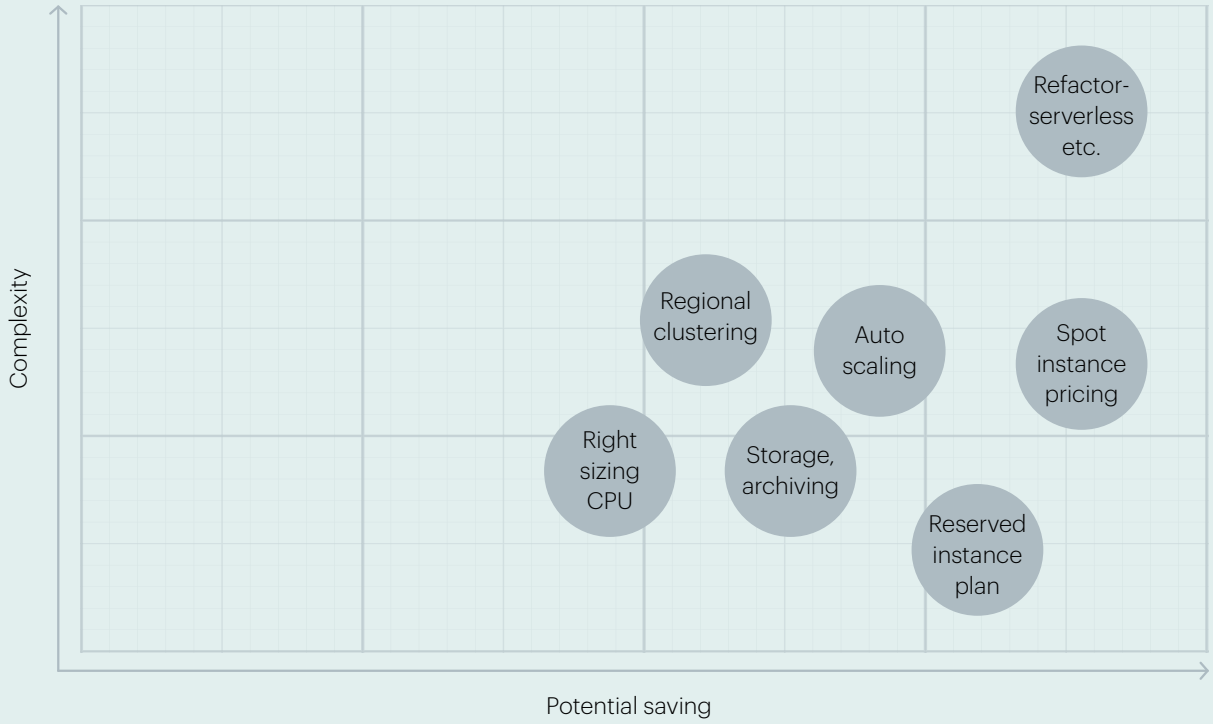


Figure-3: Cloud spend cost optimization

Right-sizing CPU: We tend to over-estimate the resource requirements, causing overprovisioning of resources in the cloud. It is common to find overprovisioned resources in a cloud environment. Approximately 60 percent of the cloud spend is to compute resources. Organizations should look at the average CPU usage of cloud resources and identify areas where CPU configuration can be optimized for cost savings.

Regional cluster: Analyze the egress charges across regions. If there are unnecessary data transfers between regions, plan for regional clusters so that transaction fulfillment does not require communication with other regions.

Refactor - serverless, PaaS: The pricing of the serverless resources (and PaaS resources) is cheaper when compared to dedicated cloud resources. For new workload or workload going through significant changes, organizations can consider leveraging serverless or PaaS services for optimizing the cloud spend.

There are many success stories of cloud cost optimization. Organizations can optimize cloud spends significantly by taking a disciplined approach, e.g., a leading specialist eCommerce company was able to optimize cloud spends by 30 percent, and another leading digital marketing organization was able to optimize cloud spends by 45 percent.

Cloud spend management best practices and tools

Cloud spend optimization is a function that requires processes, tools, and training for people to manage cloud spends effectively.

Figure-4 illustrates cloud spending management good practices — **Know** your cloud spend, **Control** your cloud spend, and Optimize your cloud spend functions of the **KCO (Kay-KO)** framework. Major cloud providers have essential tools for cloud spend measurement and reporting. These tools are also listed in the illustration.

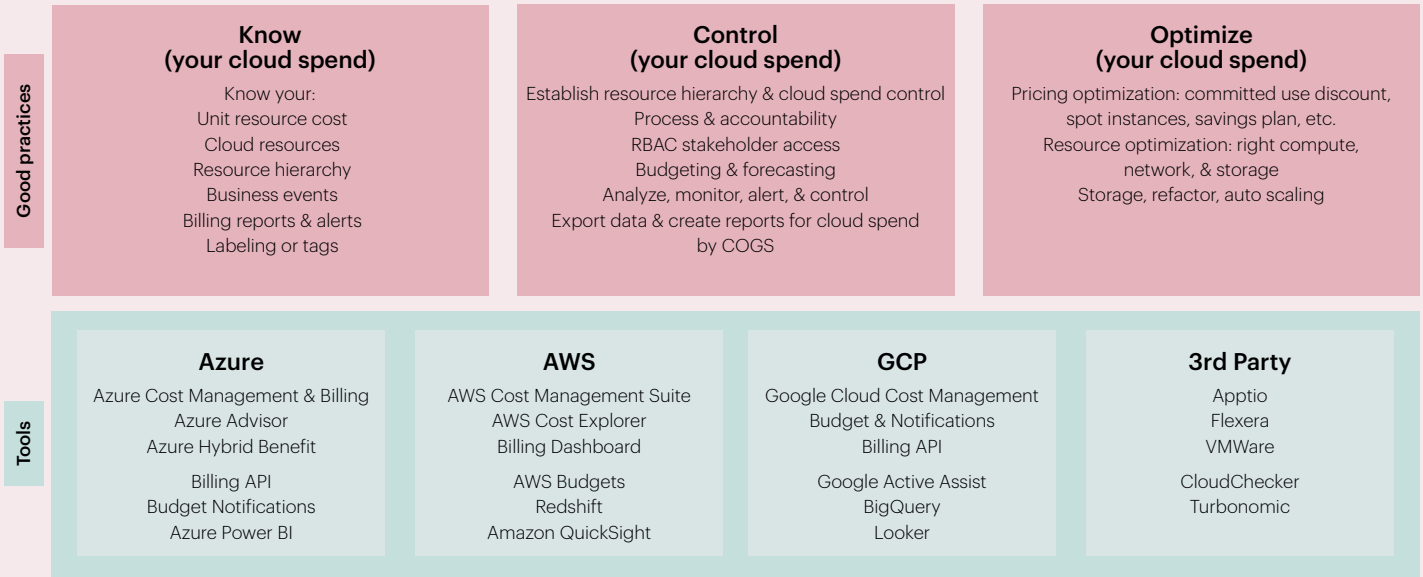
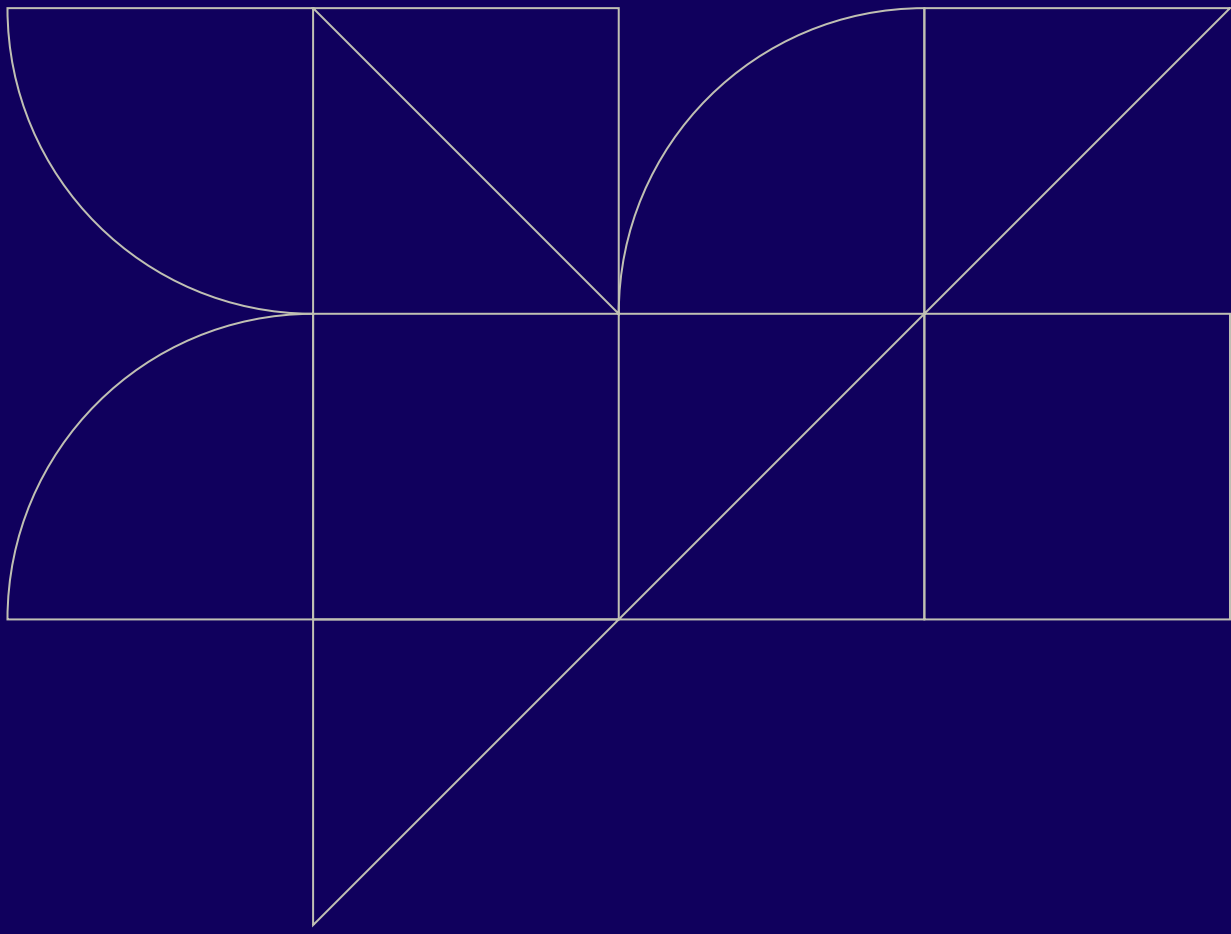


Figure-4: Cloud spend management good practices and tools

It is great to start leveraging cloud spend management from the beginning of a cloud journey, e.g., onboarding resources to the cloud. It is crucial to allocate and track cloud spends based on the hierarchy of resources and cluster resources to operate at high velocity and scale faster. Zensar implements the best practices from the beginning to kickstart the cloud onboarding journey and ensure your cloud spends are optimized to meet velocity and scaling requirements. If you have any questions or require help in workload migration/modernization to the cloud and cloud cost optimization, please reach out to us at connect@zensar.com.

Author:

Rahul Tyagi
Market Leader (North America), Advanced Engineering Services, Zensar.
rahul.tyagi@zensar.com



zensar
An RPG Company

We conceptualize, build, and manage digital products through experience design, data engineering, and advanced analytics for over 130 leading companies. Our solutions leverage industry-leading platforms to help our clients be competitive, agile, and disruptive while moving with velocity through change and opportunity.

With headquarters in Pune, India, our 10,000+ associates work across 33 locations, including San Jose, Seattle, Princeton, Cape Town, London, Singapore, and Mexico City.

For more information please contact: velocity@zensar.com | www.zensar.com