ALTIMETRIK®

The Altimetrik
Cloud Hygiene Advantage

From Data Cleanup to ROI Boost

Authored by: K V N Prithvi, Venkat Reddy Yenugu & Ritesh Srivastava

WHITEPAPER

Agenda

Executive Summary

In the fast-paced realm of modern business, striking a crucial balance between operational excellence and financial prudence has become a necessity, not merely a choice. The widespread adoption of cloud services has ushered in unparalleled flexibility and scalability, yet it has not been without its inherent challenges, particularly in the arena of cost management.

This whitepaper meticulously delves into the intricacies of how the industry is affected by cloud cost inefficiencies linked to cloud utilization and provides some tips and best practices to help you optimize your cloud spending.

Introduction

The emergence of cloud computing has ignited a transformational wave across diverse industries, fundamentally reshaping the landscape within which businesses operate, expand, and innovate. The enticing attributes of cloud computing, including cost-effectiveness, scalability, accessibility, security, and global connectivity, have drawn enterprises toward its virtual domains.

However, amid these undeniable advantages, the shift from on-premises systems to the cloud introduces a spectrum of challenges, including security, data privacy, cost management, data governance, skill gaps, and vendor lock-in.

Industry Trends

Cloud computing has become a widely embraced and efficient method for delivering IT services and resources via the internet. Nonetheless, managing the expenses associated with cloud computing presents a complex and daunting challenge, particularly for businesses spanning various industries and sectors.

A significant number of companies have found that their cloud spending has correlated with an increase in cloud waste, which refers to the inefficient or unnecessary utilization of cloud computing resources. According to a survey conducted by StormForge, involving 131 IT professionals, the extent of cloud waste could amount to as much as 47% of a cloud budget. Flexera's data indicates that, in 2021, cloud waste constituted an average of 30% of companies' cloud budgets. In 2022, this figure escalated to 32%.

Notably, Gartner's statistics reveal that the collective cloud expenditure reached nearly $500 billion in 2022, with approximately $120 billion of this sum representing an exorbitant amount of money wasted due to cloud inefficiency. These statistics underscore the concerning trends associated with cloud waste.

Key statistics pertaining to cloud waste are as follows:

- Six out of ten organizations find their cloud costs exceeding initial expectations.
- Nearly half (49%) of businesses relying on cloud services struggle to maintain control over their cloud-related expenses.
- A significant 75% of organizations report a notable surge in cloud waste.
- Approximately 32% of a typical cloud budget is squandered due to inefficiencies.
- An alarming 42% of CIOs and CTOs identify cloud waste as their foremost challenge.
- More than half of enterprises are grappling with the challenge of realizing a return on investment (ROI) in the realm of cloud computing.

The Importance of Cloud Hygiene: **Navigating Challenges for Success**

While we readily acknowledge the complexities that accompany the adoption of cloud solutions, it's imperative to recognize that their advantages far outweigh these challenges. Embracing prevailing industry trends is essential, and effectively surmounting these obstacles demands the implementation of strategic approaches.

The central objective of this white paper is to address issues related to managing cloud costs and ensuring proper data governance. The recommendations provided for maintaining cloud hygiene are specifically tailored to meet these challenges head-on.

The intricacies of cloud cost management and governance arise from various factors, including:

1. **Complexity in Pricing Models:** Cloud service providers offer a multitude of services, each with their own intricate pricing structure. Diverse services entail distinct cost frameworks, making it a formidable challenge for entities to accurately predict and manage their cloud-related expenditures.

2. **Dynamic Workloads:** Cloud resources' dynamic nature, allowing them to scale based on demand, is a fundamental aspect of cloud computing. However, this dynamism can lead to unpredictable expenses if resource utilization isn't vigilantly monitored and controlled.

3. **Over Provisioning:** In the pursuit of achieving optimal performance and uninterrupted availability, organizations might inadvertently allocate excessive cloud resources, consequently leading to escalated costs. Unnoticed instances of unused or underutilized resources can result in wasteful financial outlays.

4. **Limited Visibility:** Effectively managing cloud-related expenses becomes a challenge when entities lack a comprehensive view of their cloud consumption patterns and associated spending. Without a detailed understanding of resource utilization, identifying areas for cost optimization becomes a formidable task, and maintaining an accurate record of cloud assets can be challenging.

5. **Charges for Data Transfer and Egress:** Transferring substantial data volumes to and from the cloud can result in additional charges. Organizations must also consider data transfer expenses when designing their cloud architecture.

6. **Unplanned Growth:** As organizations embrace cloud services, their usage may expand more rapidly than expected, resulting in unforeseen financial increments. This situation can arise when cloud resource adoption lacks proper planning and diligent oversight.

7. **Multi-Cloud Complexities:** Entities that utilize multiple cloud providers may face supplementary complexities in managing costs. Each cloud platform operates with its own billing and management interface, making it challenging to obtain a unified view of expenditures. Ensuring consistent governance policies and maintaining uniformity across diverse environments can also pose challenges.

8. **Optimizing Resource Utilization:** Striking a balance between resource availability, optimal performance, and cost-efficiency can be intricate. Allocating excessive resources can result in unnecessary expenses, while insufficient allocation may lead to performance bottlenecks.

9. **Enforcing Governance Policies:** Maintaining consistent governance policy enforcement across diverse cloud services, geographical regions, and teams presents a multi-faceted challenge. Overlooking this crucial aspect could potentially expose organizations to security vulnerabilities and non-compliance risks.

10. **Organizational Alignment:** Achieving alignment across various departments, teams, and stakeholders in adhering to cloud governance policies and recommended practices can be demanding. Effective collaboration and transparent communication play a pivotal role in addressing this challenge.

In navigating these challenges and embracing cloud hygiene, organizations can unlock the full potential of cloud technology while ensuring fiscal responsibility and compliance. This white paper serves as a guide, providing actionable insights and strategies to optimize cloud operations and drive business success.

Drivers of Cloud Computing & Cost Management

One of the main drivers of cloud computing cost management is the need to align the cloud strategy with the business goals and objectives. Cloud computing can offer many benefits, such as scalability, flexibility, innovation, and security, but it also requires careful planning and governance to ensure that the cloud services and resources are aligned with the business needs and expectations. For example, choosing the right cloud service model (such as SaaS, PaaS, or IaaS), the right cloud provider (such as AWS, Azure, or Google Cloud), and the right cloud deployment model (such as public, private, or hybrid) can have a significant impact on the cloud costs and performance.

Optimizing Cloud Efficiency: Strategies for Effective Cloud Hygiene

In our pursuit of cost-effective and efficient cloud management, we introduce the concept of **Cloud Hygiene.** This encompasses a set of cohesive strategies and guidelines that address the complexities of cloud expenses and resource management. Cloud Hygiene represents an ongoing commitment to refinement and readiness, encompassing diverse facets from adept cloud management to prudent cost control.

Outlined below are the recommended strategies within the framework of Cloud Hygiene:

- **Right-Sizing Resources:** Consistently assess the usage of resources and select appropriately fitting instances for tasks. Adapting instance sizes to match real usage patterns can result in notable cost reductions. Diverse cloud providers offer an array of monitoring and logging tools that facilitate the collection of data about CPU, memory, network, and storage utilization. This data can be employed to pinpoint application workloads, including times of peak usage, seasonal spikes, and traffic fluctuations. The choice of the ideal instance type and size should be guided by specific use case requirements. Across major cloud platforms, a variety of instance types are available, including General Purpose, Compute Optimized, and Memory Optimized, among others.

  - **Scenario:** During non-peak hours, if application instances are only operating at 20% capacity. The instances are underutilized, and they're being billed for more power than the application requires.

  - **Solution:** Analyzed the usage patterns and scaled down the instances with lower specifications during off-peak times. This way, we optimized costs without sacrificing performance.

- **Usage of Spot Instances:** Leverage the potential of Spot Instances, which are surplus cloud resources accessible at a discounted rate. Leading cloud service providers grant access to untapped computing capacity, often at a notably reduced expense compared to on-demand computing power. These instances are subject to price fluctuations determined by supply and demand, reminiscent of an auction model. Spot Instances operate when the prevailing spot price falls below the specified maximum price. Instances could be terminated if the price surpasses the set limit. While cloud providers retain the right to terminate Spot Instances at any point, they prove exceptionally economical for applications designed to handle interruptions gracefully through techniques such as checkpoints or stateless architectures. With a primary focus on tasks like data analysis, batch processing, rendering, and scientific simulations, Spot Instances emerge as a cost-effective solution for non-time-critical workloads, particularly within lower-priority environments.

- **Scenario:** Development and testing environments don't require continuous availability. Application teams need resources to test code and deploy applications, but these environments can be spun up and down as needed.

- **Solution:** recommended Spot Instances for testing and development, taking advantage of cost savings. Developers can request instances when needed and release them when the work is done.

➢ **Commit to Reserved Instances (RIs):** Commit to long-term usage by purchasing reserved instances. RIs offer substantial cost savings, especially for steady-state workloads with predictable usage, often cutting costs by up to 75% or more compared to on-demand instances. The savings depend on factors such as the reservation term, payment option, and instance type.

- **Scenario:** If Applications have predictable level of traffic and infrastructure requirements are relatively stable

- **Solution:** Analysed historical usage patterns and committed to reserved instances for a specific type and size that matches application baseline workload.

➢ **Optimize Storage Costs:** Tailor storage choices to match data access frequency and performance needs. Cloud providers offer various storage classes, including Standard, Premium, Cold, and Archive storage. Reserve higher-performance storage for data requiring swift responsiveness. Leverage lifecycle policies that automate data migration to lower-cost storage tiers as access frequency decreases or explore tiered storage solutions that intelligently relocate data to the most cost-effective tier based on usage patterns. For virtual machines and databases, prioritize block storage that balances performance and capacity appropriately. Maintain flexibility by resizing volumes to meet real-time demands.

- **Scenario:** Application teams have a large dataset that is critical for compliance reasons, but it is infrequently accessed. This could be historical transaction data, backups, or logs that need to be retained for regulatory purposes.

- **Solution:** In such scenarios, instead of keeping this data on high-performance, expensive storage solutions, we moved it to a lower-cost storage tier designed for archival purposes.

➢ **Leverage Serverless Services:** Explore serverless computing to pay only for compute time during function execution. However, it's important to note that serverless might not be the best fit for every use case. Long-running processes, applications with specific resource requirements, and certain types of latency-sensitive workloads might still benefit from more traditional compute models. Serverless services automatically scale and manage resources based on demand. Serverless services are particularly well-suited for tasks like microservices, APIs, data processing, real-time file processing, etc. This can be more cost-effective than maintaining continuously running servers.

- **Scenario:** Consider there is a system where certain tasks need to be executed in response to events, such as file uploads, database changes, or incoming messages.

- **Solution:** All the major cloud providers provide serverless functions to execute code in response to these events. This way, application teams don't need to maintain a constantly running server, and you only incur costs when the function is triggered.

➢ **Autoscaling:** Implement autoscaling to automatically adjust the number of instances based on workload demands. Scaling down during periods of low usage can save costs, while scaling up ensures performance during peak times. Instead of opting for larger instances, consider distributing your workload across multiple smaller instances. All the major cloud providers facilitate autoscaling to dynamically adjusting the number of instances based on traffic patterns which ensures only running the necessary number of instances at any given time.

➢ **Idle Resource Management & Consolidate Resources:** Identify and decommission idle or unused resources. Terminating resources that are not actively contributing to operations can prevent unnecessary costs. Identify areas where multiple small instances or databases can be consolidated into larger ones to eliminate redundant resources.

- **Scenario:** At certain times the application traffic significantly increases, the current infrastructure will not be sufficient and will struggles to handle the load

- **Solution:** enabled auto scaling in the application. Set up policies that automatically add more compute resources (such as virtual machines or containers) when the demand increases and remove them during periods of lower traffic.

➢ **Monitor and Optimize Data Transfer:** Minimize data transfer between cloud regions or services to reduce associated costs. Consider leveraging content delivery networks (CDNs) for frequently accessed content to reduce data transfer charges. Utilize caching solutions provided by the cloud service providers for in-memory caching to reduce the load on application and database servers.

- **Scenario:** Application development and testing environments are active only during specific working hours, but they remain idle during nights and weekends.

- **Solution:** Implemented a schedule or automation to automatically shut down or scale down development and testing instances during periods of inactivity. This prevents you from paying for resources that are not being used.

➢ **Implement Governance and Policies:** Establish cloud governance policies to enforce best practices, cost controls, and usage guidelines across the organization. A well-defined cloud governance framework is essential to manage expenses effectively and prevent unexpected costs.

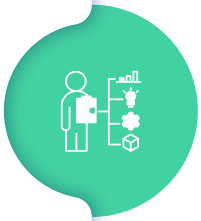
- **Scenario:** Application uses multiple cloud providers, and the data transfer costs between these providers are escalating.
- **Solution:** Utilized CDNs to cache and serve static assets, reducing the need for cross-cloud data transfer. CDNs distribute content closer to end-users, improving performance and lowering costs.

By embracing these strategies within the framework of Cloud Hygiene, organizations can optimize their cloud operations, enhance cost efficiency, and ensure the responsible utilization of cloud resources.

Governance Framework for Cloud Cost Control

Set Budgets

Establish clear budgets for teams, projects, departments to control spending.



Define Cost Allocation & Accountability

Assign ownership and accountability for managing cloud costs.

Use Cost Tags

Categorize and allocate costs accurately with resource tagging.

Monitor & Track Usage

Implement monitoring to track resource usage closely.

Cost Reporting & Visualization

Cost reports and dashboards that showcase spending patterns.

Regular Review & Optimization

Continuously review and optimize cloud cost structure.

Case Study for A Leading Pharmaceutical Manufacturer

Transforming a Small Cloud Cost Optimization MVP into a Significant Cost-Saving Opportunity

Highlights

- Established clear standard operating procedures (SOPs) for new applications, ensuring the correct infrastructure configurations are consistently set up.
- Created a standardized checklist tailored to specific use cases, tech stack ensuring comprehensive and consistent assessments.
- Prescribed a well-structured governance model for application teams, enhancing overall management and accountability.
- Successfully reduced cloud costs by 17%, resulting in substantial savings.
- Efficiently streamlined resource utilization, achieving a 50% improvement in efficiency.

Business Challenge

In the fast-paced technology sector, large businesses often encounter significant obstacles. This was the case for one of the world's leading pharmaceutical companies. Their rising cloud costs prompted them to start a Cloud Hygiene project to regain control over expenses. The client, eager to confirm the potential benefits of this effort, began a Proof of Value (PoV) journey.

This initiative revolved around a **two-fold** mission:

first, to unearth the root causes of the cost surge,

and **subsequently**, to expedite the implementation of cost-saving measures.

Solution

Our initial task was to identify the root causes. The primary issue was the persistent escalation of cloud costs, which we identified as stemming from three key factors:

i. **Inadequate Monitoring and Alert Mechanisms:** The absence of efficient monitoring and alert systems hindered our ability to conduct real-time resource assessments, ultimately resulting in sub-optimal resource utilization.
ii. **Active Lower Environments:** Lower-level environments remained active even after project completion, incurring unnecessary expenses.
iii. **Stakeholder Accountability:** Stakeholders prioritized application delivery at the expense of resource and cloud cost optimization, resulting in a noticeable lack of accountability.

Approach

To address these challenges, we implemented a **systematic strategy** that encompassed the following key steps:

Checklist Creation: Checklists were developed to address the core technology stacks contributing to higher cloud costs, specifically targeting AWS Cloud Compute and Storage, Databricks Jobs and Cluster Sizing, as well as Snowflake Design, Configuration, and Utilization.

Component	Checklist item	Action item
Memory and CPU sizing	Analyze CPU and Memory utilization for last 30 days.	If utilization is below 50 % right size the instance type to appropriate Instance type
Consolidate servers	Analyze if any two servers can be consolidated	Any application using 2 different servers for same purpose they need to be consolidated to 1
Leverage Spot Instance	Analyze the workloads and leverage spot instances where ever applicable	Once the development is completed in lower environments. spot instances can be leveraged in place of reserved and on demand instance
optimize storage cost	Analyze storage tiers and archiving mechanism	Consider using appropriate S3 storage classes for different data layers Configure Lifecycle policies to move data between classes and define retention periods

I**n-Depth Analysis and Standardization:** We conducted a thorough analysis of consumption costs within various use cases, pinpointing areas of concern. This process culminated in the formulation of standardized configuration checklists for pivotal technologies such as Snowflake, AWS, and Databricks.

Precise Execution and Recommendations: Each use case underwent the systematic execution of these checklists, enabling us to compile valuable recommendations. These recommendations were then meticulously communicated to the respective use case teams.

Implementation and Continuous Monitoring: On receiving approval from the use case teams, the recommended changes were promptly implemented. Our diligent vigilance extended beyond implementation, with continuous monitoring of cloud costs ensuring that the improvements were sustained over time.

Implementation of Cloud Hygiene

The goal was to assess the business ROI of cloud hygiene, leading to Altimetrik's involvement in establishing a Proof of Value (PoV). Consequently, the engagement was divided into two distinct phases.

Phase 1 placed a primary emphasis on root cause identification and PoV development.

- **Initial Assessment:** As part of the initial assessment, Altimetrik identified 3 major tech stacks and 4 applications in data management analytics platform as an outlier which were major contributors to the incurring cloud costs.

  To validate the approach, the checklists for each tech stack (AWS, Databricks, Snowflake) were piloted against the top 4 applications that contributed the most to cloud costs. This involved analyzing current trends for each checklist item and providing optimal recommendations based on these trends.

- **Initial Outcome:** The results from Phase 1 were encouraging! Resource utilization improved by an impressive 50%. Monthly cloud costs were reduced by a substantial 30% for the top 4 use cases.

Phase 2 shifted focus towards the swift implementation of cost-saving measures identified during Phase 1, targeting a select number of applications.

- **Accelerated Implementation:** Encouraged by the initial success, the customer decided to extend the engagement. Since a PoV had already been established, Phase 2 focused on accelerating the implementation of cloud hygiene practices. The goal was to run through the checklist and implement the recommendations across 15 applications in data management and analytics platform in a span of 3 months.

- **Accomplishments:** Phase 2 delivered significant business benefits! Among the 15 applications, resource optimization reached 50%, ensuring efficient utilization. The overall cloud cost was reduced by an impressive 17%, making substantial cost savings across a broader range of use cases.

Key Take-aways

The Cloud Hygiene engagement not only identified the root causes of high cloud costs but also demonstrated substantial business impact for the client. Through meticulous planning, checklist implementation, and continuous monitoring, the client organization achieved remarkable resource optimization and significant cost reduction, thereby ensuring a more cost-efficient and streamlined cloud environment.

This comprehensive approach by Altimetrik underscores the significance of adopting proactive, data-driven cost optimization strategies in the contemporary, cloud-centric business landscape.

Conclusion

To sum up, maintaining good practices in Cloud Hygiene is crucial for organizations aiming to manage costs effectively. To succeed, organizations need to use the right framework and governance model, making sure that those using the cloud take responsibility for their spending. This helps organizations keep their cloud expenses in line with their business goals and see how their cloud investments pay off.

Altimetrik, known for its expertise in Cloud Hygiene, understands the power of using built-in cloud tools, which are often better than third-party options. According to a recent Pepper Data survey, 40% of companies struggle with managing their cloud costs. This shows that taking steps to save money on cloud services not only helps 67% of businesses grow their revenue but also boosts profits for 64% of them. So, while cloud services are essential, having a strong governance model and smart cost-saving strategies are just as important for success.

About Altimetrik

Altimetrik is a pure play digital business company focused on delivering business outcomes with an agile, product-oriented approach. Our digital business methodology provides a blueprint to develop, scale, and launch new products to market faster. Our team of 5,000+ employees with software, data, cloud engineering skills help create a culture of innovation and agility that optimizes team performance, modernizes technology, and builds new business models. As a strategic partner and catalyst, Altimetrik quickly delivers results without disruption to the business.

Our unique Digital Business Methodology centers on business led ownership aligned to company goals. It is comprised of three pillars: experienced team of practitioners, an incremental approach, and an end-to-end self-service digital business platform. These combine to facilitate collaboration and agility between business and engineering teams to co-create products and solutions faster without disruption to the business. This is powered by a single source of truth and a culture of innovation that brings unlimited growth within reach.

We cater to companies of all sizes from Fortune 100 to digital disruptors and start-ups. We are a people-centric organization and talent is one of the central pillars of our business model and success. Employee engagement, diversity & inclusion, well-being and empowerment are central themes to our ethos. This project has been instrumental in taking our digital culture to the next level and brings our globally spread employee base on a unified platform.

www.altimetrik.com