International Journal of AI & Data Science Review
Volume1, Issue1, 2025

DATA ENGINEERING GROUPS TO DEDICATE INCREASED EFFORT ON OPTIMIZING DATA CLOUD EXPENSES

Noor Akhma[1]
Department of Electrical and Information Engineering, Faculty of Engineering, Universitas Gadjah Mada, Yogyakarta, Indonesia

Article Info

Received: 28-01-2025 ***Revised:06-03-2025*** ***Accepted:17-03-2025*** ***Published:27-03-2025***

ABSTRACT

With its unmatched scalability, flexibility, and operational efficiency, cloud technologies have emerged as the foundation of many enterprises in the digital age. But its quick acceptance has also brought with it a number of difficulties, especially when it comes to controlling data processing and storage costs. There is growing need to make sure these expenditures are both financially viable and technically sound as more businesses move to the cloud. The proactive steps data engineering teams take to minimize costs are highlighted in this article, which looks deeply into the changing environment of data cloud prices. It also examines the underlying reasons for rising expenses, the challenges businesses have throughout their optimization process, and the creative fixes put in place. We want to provide a thorough manual for companies looking to achieve the ideal balance between cost-effectiveness and superior performance in the data cloud environment by looking at actual case studies and using industry expertise.

1. INTRODUCTION

Backgroundontheriseofcloudtechnologies:

The use of cloud technology has increased at an unprecedented rate in a number of sectors throughout the last ten years. When companies realized that cloud platforms might provide scalable, adaptable, and affordable alternatives to conventional on-premises infrastructure, the change got underway (Botta, De Donato, Persico, & Pescapé, 2016). A number of causes led to this change:

Scalability: One benefit of cloud platforms is that they enable companies to scale their resources according to demand, ensuring they only pay for what they use. They may so swiftly scale their resources up or down in response to demand.

- **Flexibility:** Businesses may effectively deploy apps, data storage options, and other services by using the cloud without having to make large upfront hardware expenditures..
- **Collaboration:** Regardless of physical location, cloud technologies enable smooth team

communication by sharing resources and tools.

Innovation: Because cloud providers often introduce new features, tools, and integrations, the ecosystem around cloud computing is always evolving. This makes it possible for organizations to stay up to date with the most recent advancements in technology (Ouf, Nasr, & Helmy, 2010).

The significance of cost optimization in the data cloud: As more and more companies depend on cloud platforms for data processing and storage, cost management becomes crucial. Despite the cloud's scalability and flexibility, proper administration is required to keep costs from going out of hand. The following explains the importance of cost optimization in the data cloud:

- **Budget Control:** Given the volume of data created every day, storing and processing it may be a substantial burden for an organization's IT budget. Businesses are certain to get the maximum return on their investment when cost optimization is done well.
- **Resource Efficiency:** Cost optimization involves more than simply reducing expenses; it also involves making sure resources are utilized effectively. By optimizing expenses, businesses may make sure they take every precaution to save resources.
- **Predictability:** Planning a budget might be difficult when dealing with fluctuating cloud expenses. By using cost optimization techniques, businesses may increase the predictability and consistency of their cloud spending, which will help with financial planning. (Khedr, Nasr, & Elmasry, 2019)
- **Competitive Advantage:** Effective cloud cost optimization allows businesses to refocus savings on key projects and innovation, providing them a competitive advantage over rivals who may overpay on cloud services.

As cloud technology usage continues to increase, organizations must comprehend and use cost-optimization techniques in the data cloud to make sure they can fully utilize the cloud's potential without going over budget.

2. THE CURRENT LAND SCAPE OF DATA CLOUD COSTS

The data cloud pricing environment has been changing quickly, particularly as organizations depend more and more on cloud platforms. The costs of keeping enormous volumes of data in cloud databases and the processing power needed to handle this data are referred to as data processing and storage charges (Muralidhara, 2017). The cost of data storage may vary based on the kind of data, how long it is kept, how quickly it must be accessible, and how complicated any operations are on the data. The quantity, availability, and redundancy of the data are the main factors influencing data storage expenses. For example, it will cost more to store data that must be accessed regularly and duplicated across many regions for redundancy than it costs to

store archive material that is accessed seldom. However, processing costs depend on the kind of computations being done, how long they take, and how many computing resources are used. It may be resource-intensive and, as a result, more expensive to run complicated machine learning models or real-time data analytics (Khan et al., 2022). However, these expenses are driven by more than simply the sheer amount of data or the complexity of procedures. The growing costs in the data cloud space are caused by a number of causes. More data is being absorbed into cloud systems than ever before due to the growing range of data sources, including user-generated content and Internet of Things devices. Additionally, there is a push to store and analyze even more data as firms realize the value of data-driven insights, which raises the demands on processing and storage capacity (Marr, 2015). Costs may also increase due to regulatory requirements for data protection and compliance. Maintaining audit trails, ensuring data encryption, and adhering to data residency regulations may all have a cost impact (Padhy, Patra, & Satapathy, 2011). The cost structures of various cloud providers might differ, which could result in inefficiencies if not properly handled. Another factor is vendor-specific pricing methods.

In summary, while the cloud has unmatched benefits in terms of scalability and flexibility, it also presents difficulties for cost management and optimization, particularly when it comes to data processing and storage.

3. CHALLENGES IN DATA CLOUD COST OPTIMIZATION

Identifying Hidden Costs:

The most costly hidden costs in cloud computing are those that become obvious right away when firms first move to the cloud or expand their activities there. In 2016, Alkhanak, Lee, Rezaei, and Parizi

- **Data Egress Fees:** Although it is usually cheap to post data to the cloud (ingress), egressing data from the cloud (egress) may be quite expensive, particularly for large-scale operations.
- **License and Software Costs:** Additional software licenses may be needed for some cloud services, which might raise the total cost.
- **Premium Support and Service Costs:** Specialized services or higher support levels may come with extra costs.
- **Costs of Unused Resources:** Waste may result from resources that are provided but not used, such as over-provisioned storage or inactive virtual machines.

Balancing Performance and Expenses:

Finding the ideal balance between cost and performance is a delicate task.

- **Over-Provisioning:** Performance may be guaranteed by making sure applications have more than adequate resources, but

if those resources are not used to their full potential, it may also result in needless expenses.

- **Under-Provisioning:** However, reducing resources in an attempt to save money might have a negative impact on application performance, resulting in subpar user experiences or even outages.
- **Optimal Resource Selection:** Cost and performance may be greatly impacted by selecting the appropriate database, compute instance, or storage type.

Evolving Cloud Pricing Models: Cloud service providers' pricing structures are often changing, which makes it more difficult for companies looking to save expenses.

- **On-Demand vs. Reserved Instances:** Reserved instances may save money for workloads that are predictable, while on-demand pricing gives flexibility. On the other hand, overspending may result from committing to reserved instances without enough foresight.
- **Spot Instances:** These are often less expensive, but they may be discontinued if resources are required elsewhere, which might cause problems.
- **Savings Plans:** Cost optimization tactics get much more difficult when considering the savings programs that many cloud providers provide, which give reductions based on consumption commitments.
- **Hybrid and Multi-Cloud Deployments:** Cost-tracking and optimization issues arise when using services from different cloud providers or when integrating on-premises infrastructure with the cloud. Petcu (2013)

Although there are many advantages to the cloud, navigating its financial environment may be difficult. Companies need to be proactive, using tactics and technologies to save expenses without sacrificing functionality or performance.

4. STRATEGIES EMPLOYED BY DATA ENGINEERING TEAMS

Teams of data engineers are always looking for methods to reduce expenses while maintaining effective operations. Proactive monitoring and notifications are one of their main tactics. Teams may handle any cost overruns before they worsen by putting in place real-time monitoring tools, which quickly notify them of any abnormalities or spikes in consumption. This strategy has the advantages of cost management and effective resource use. Choosing affordable storage options is another crucial tactic. Selecting the best cloud storage option that fits the demands of the company and its budget is crucial since there are so many alternatives available, ranging from object and block storage to file systems and databases. To ascertain which storage options provide the highest performance at the most competitive prices, data engineering teams often carry out extensive assessments and benchmarking. (Mujeeb, Zhang, Huang, Madonski, and Ahmad, 2022). In order to

optimize savings, a lot of teams are also using cloud providers' savings programs and reserved instances. With these alternatives, companies may get cheaper prices in return for committing to certain consumption levels. Data engineering teams may make well-informed judgments about which reserved instances or savings plans to buy by examining their past use trends and projecting future requirements. This will ultimately result in significant cost savings. To put it simply, data engineering teams are at the forefront of minimizing cloud expenses, ensuring organizations get the most out of their cloud investments, via careful monitoring, wise storage decisions, and exploiting available pricing models.

5. CASE STUDY:A SUCCESSFUL COST OPTIMIZATION JOURNEY

Case Study-Deloitte's Cloud Cost Optimization (Deloitte, n. d.):

Background of the Organization:

Leading global consulting company Deloitte has been leading the way in helping organizations with their digital transformation initiatives. They have seen that while cost savings and predictability are often mentioned by IT managers and company executives as the main benefits of moving to the cloud, many firms need assistance in order to reap these advantages.

Challenges Faced:

Over 20 assets from over 50 accounts were moved to the cloud by a Deloitte service group customer. They saw a notable rise in their cloud hosting expenses after the transfer. Among the difficulties were:• A 40% rise in monthly expenditures above pre-migration costs.
• Insufficient understanding of the primary cost drivers. • Tools to track cost and use patterns are required. • There are no aggressive steps in place to curb these rising expenses.

Solutions Implemented and Results Achieved:

In order to identify areas for improvement, Deloitte worked closely with the client to create a clear roadmap that included measures for measuring data in real time. The following were some of the project's main recommendations:
• Investigating and recording fast victories for instant savings. Using the resources at hand, do a thorough study of cloud utilization and expenses. Finding architectural upgrades to take use of cloud-native services. • Detecting underused or unused cloud services and resources via automation. • Using asset team inputs and data analysis insights to provide customized suggestions.

The Results:

• Increased responsibility across teams leads to

improvements in operational efficiency.

• Without affecting normal company operations, monthly cloud hosting expenses were lowered by 15% for the first three months, 20% for the next six, 40% for the next twelve, and 50% for the next sixteen.
• Other high-priority projects were funded with the savings.
• It enhanced budgeting and forecasting accuracy.
• Greater openness on the expenses of all assets, environments, and accounts.

CaseStudy-Cloud Cost Optimization Reveals Fin Ops Massive Benefits (GlobalDots, n.d.) :

Background of the Organization:

GlobalDots is a company that focuses on security and performance solutions for the cloud. Having served companies of all sizes for more than 17 years, they have been at the forefront of technological innovation. They have experience in a number of areas, such as cloud optimization, DDoS prevention, and content delivery networks (CDNs).

Challenges Faced:

1. **Rising Cloud Costs:** Businesses faced rising expenses as they grew and depended increasingly on cloud infrastructures. These expenses were unpredictable, which made financial planning and budgeting difficult.
2. **Complexity in Cloud Management:** Managing and optimizing many cloud services got complex. This intricacy often results in waste and inefficiency.
3. **Lack of Visibility:** Companies need assistance in properly understanding their cloud use and expenses. Their capacity to make well-informed judgments was hampered by this lack of openness.

Solutions Implemented and Result Achieved:

1. **FinOps Implementation:** FinOps, a blend of technology, best practices, and culture, was established by GlobalDots to provide financial responsibility to the cloud's variable expenditure model. This strategy made sure organizations were getting the most out of every dollar they spent on the cloud.
2. **Continuous Monitoring and Optimization:** Through ongoing cloud use and cost monitoring, GlobalDotsco could quickly discover and fix inefficiencies. They were able to make the most use of their resources because to this proactive attitude.
3. **Enhanced Visibility:** With the correct tools and procedures, companies were able to see how much cloud they were using. They were able to more correctly and effectively anticipate expenses and manage resources because to this openness.

The Results:

4. **Significant Cost Savings:** By optimizing cloud resources and removing waste, businesses experienced significant reductions in their cloud bills.
5. **Improved Efficiency:**

Businesses might more efficiently manage resources, improving performance and cutting expenses, if they had a better grasp of cloud utilization.

6. **Empowered Decision Making:** Businesses were better able to make well-informed choices about their cloud strategy because to increased visibility and analytics.

6. FUTURE TRENDS IN DATA CLOUD COST OPTIMIZATION

Predictive Cost Management:

Predictive cost management forecasts future cloud costs using cutting-edge technology and approaches. The examination of past data and present use trends forms the basis of this strategy.

- **Historical Data Utilization:** By looking at historical cloud use and related expenses, businesses may find trends and patterns. Predictive models are based on this historical viewpoint.
- **Real-time Monitoring:** Predictive models get a dynamic input from ongoing monitoring of actual cloud use, guaranteeing that forecasts are based on the most recent facts.
- **Budgeting and Planning:** Businesses might create more precise budgets with a predictive view on cloud spending. This proactive approach to budgeting guarantees effective resource allocation and helps prevent unforeseen expenses.

Businesses may be proactive instead than reactive with predictive cost management. They may make strategic choices to optimize spending, such scaling up or down resources, by projecting future costs.

Integration of AI and ML for Smarter Optimization:

Cloud cost optimization is changing as a result of AI and ML. Large datasets may be processed and analyzed by these technologies at previously unheard-of rates, yielding insights that might result in considerable cost savings.

- **Data-driven Insights:** Massive volumes of cloud consumption data may be sorted through by AI and ML models to find inefficiencies, underutilized resources, or places where expenses can be cut.
- **Automated Recommendations:** In addition to offering insights, these models might suggest certain courses of action, such redistributing resources or ending instances that aren't being utilized.
- **Proactive Cost-saving Measures:** Without the need for human participation, certain sophisticated AI-driven solutions may even take proactive measures in response to their analysis, guaranteeing optimum resource usage. (Prakash & Deepika, 2020)

Businesses may improve their automation in cloud cost management by combining AI and ML. This reduces the requirement for a human effort in cost optimization and leads to greater resource use.

The Role of Open-Source Tools in Cost Management:

Because of their versatility and adaptability, open-source tools—which are often created and maintained by a community of experts—are quickly becoming a mainstay in cloud cost management.

- **Community-driven Development:** Open-source technologies benefit from a sizable

developer and expert community that constantly improves and updates the tool in response to real-world demands and difficulties.

- **Customization:** Proprietary tools do not provide customization to meet the specific demands of a firm, in contrast to open-source alternatives. This guarantees that the business's cloud management needs are appropriately met.
- **Cost-effective:** Because open-source technologies don't have the hefty prices that come with commercial ones, many firms are attracted to them. This enables businesses to optimize expenditure and save money without incurring new software costs.

You may save a lot of money on software purchase and cloud charges by using open-source solutions for cloud cost management. These tools' community-driven design guarantees that they stay current and applicable to the constantly changing cloud world.

7. CONCLUSION

Businesses must strategically optimize costs, particularly in cloud settings. By avoiding needless expenditures and encouraging effective resource use, it guarantees financial stability. This gives organizations a competitive edge and immediate financial gains, enabling them to reinvest savings in expansion and innovation. Data engineering teams have a direct impact on cloud expenses because of their critical role in managing and processing large amounts of data. From tool selection to data pipeline structure, their choices might have a big effect on costs. As a result, these teams need to take a proactive approach to cost management by foreseeing future cost increases and putting preventative measures in place. Additionally, since the cloud ecosystem is always changing, new tools, pricing schemes, and best practices are always appearing. To maintain ongoing cost effectiveness, data engineering teams must place a high priority on keeping abreast of these developments and modifying their approaches. In summary, the importance of cost optimization and the part data engineering teams play in this process cannot be overstated if companies want to get the most out of their cloud investments.

REFERENCES

[1] Botta, A., De Donato, W., Persico, V., & Pescapé, A. (2016). Integration of cloud computing and internet of things: a survey. Future generation computer systems, 56, 684-700.

[2] Ouf,S.,Nasr,M.,&Helmy,Y.(2010, December).Anenhancede-learningecosystembasedon an integration between cloud computing and Web2. 0. In The 10th IEEE International Symposium on Signal Processing and Information Technology (pp. 48-55). IEEE.

[3] Elmasry, H.M., Khedr, A.E., & Nasr,M.M.(2019).An adaptivetechnique forcost reduction in cloud data centre

environment. International Journal of Grid and Utility Computing, 10(5), 448-464.

[4] Muralidhara, P. (2017). The evolution of cloud computing security: addressing emerging threats. International journal of computer science and technology, 1(4), 1-33.

[5] Khan,T.,Tian,W.,Zhou,G.,Ilager,S.,Gong,M.,&Buyya,R.(2022).Machinelearning(ML)- centric resource management in cloud computing: A review and future directions. Journal of Network and Computer Applications, 204, 103405.

[6] Marr, B. (2015).Big Data: Using SMART big data, analytics and metrics to make better decisions and improve performance. John Wiley & Sons.

[7] Padhy, R. P., Patra, M. R., & Satapathy, S. C. (2011). Cloud computing: security issues and researchchallenges.InternationalJournalofComputerScienceandInformationTechnology& Security (IJCSITS), 1(2), 136-146.

[8] Alkhanak, E. N., Lee, S. P., Rezaei, R., & Parizi, R. M. (2016). Cost optimization approaches for scientific workflow scheduling in cloud and grid computing: A review, classifications, and open issues. Journal of Systems and Software, 113, 1-26.

[9] Petcu,D.(2013,April).Multi-cloud:expectationsandcurrentapproaches.In Proceedingsofthe 2013 international workshop on Multi-cloud applications and federated clouds (pp. 1-6).

[10] Ahmad,T.,Madonski,R.,Zhang,D., Huang,C.,&Mujeeb,A.(2022).Data-drivenprobabilistic machine learning in sustainable smart energy/smart energy systems: Key developments, challenges,andfutureresearchopportunitiesinthecontext of smartgridparadigm.Renewable and Sustainable Energy Reviews, 160, 112128.

[11] Deloitte-Available: https://www2.deloitte.com/content/dam/Deloitte/pt/Documents/ofertas-consultoria/cloud_cost_optimization_case_study_ecs.pdf

[12] GlobalDots– Available:https://www.globaldots.com/resources/case-studies/cloud-cost-optimization-case-study-reveals-finops-massive-benefits/

[13] Deepika, T., & Prakash, P. (2020). Power consumption prediction in cloud data center usingmachine learning. Int. J. Electr. Comput. Eng. (IJECE), 10(2), 1524-1532