Cloud Economics 360CA

Background

Cloud economics, often referred to as cloud cost management or cloud cost optimization, is a critical discipline in the realm of cloud computing. It involves the strategies, practices, and tools used by organizations to effectively manage and control the costs associated with using cloud services, such as those offered by major cloud providers like Amazon Web Services (AWS). Cloud economics, in the context of Amazon Web Services (AWS), refers to the discipline of managing and optimizing the costs associated with using AWS cloud services. AWS offers a wide array of cloud computing services, and understanding how to maximize the value while minimizing costs is essential for organizations of all sizes. AWS cloud economics encompasses various strategies, practices, and tools to achieve this balance.

The advent of cloud computing has transformed the way businesses deploy and operate their IT infrastructure. Rather than investing in physical hardware and data centers, organizations are increasingly turning to cloud services to host their applications, store data, and scale their computing resources. This transition to the cloud provides tremendous flexibility, scalability, and agility, but it also introduces a new set of financial challenges. Cloud services are typically billed on a pay-as-you-go or subscription basis, which can result in unpredictable costs. As organizations migrate their workloads to the cloud, they must strike a balance between reaping the benefits of cloud computing and managing their cloud spending effectively.

The purpose of cost optimization

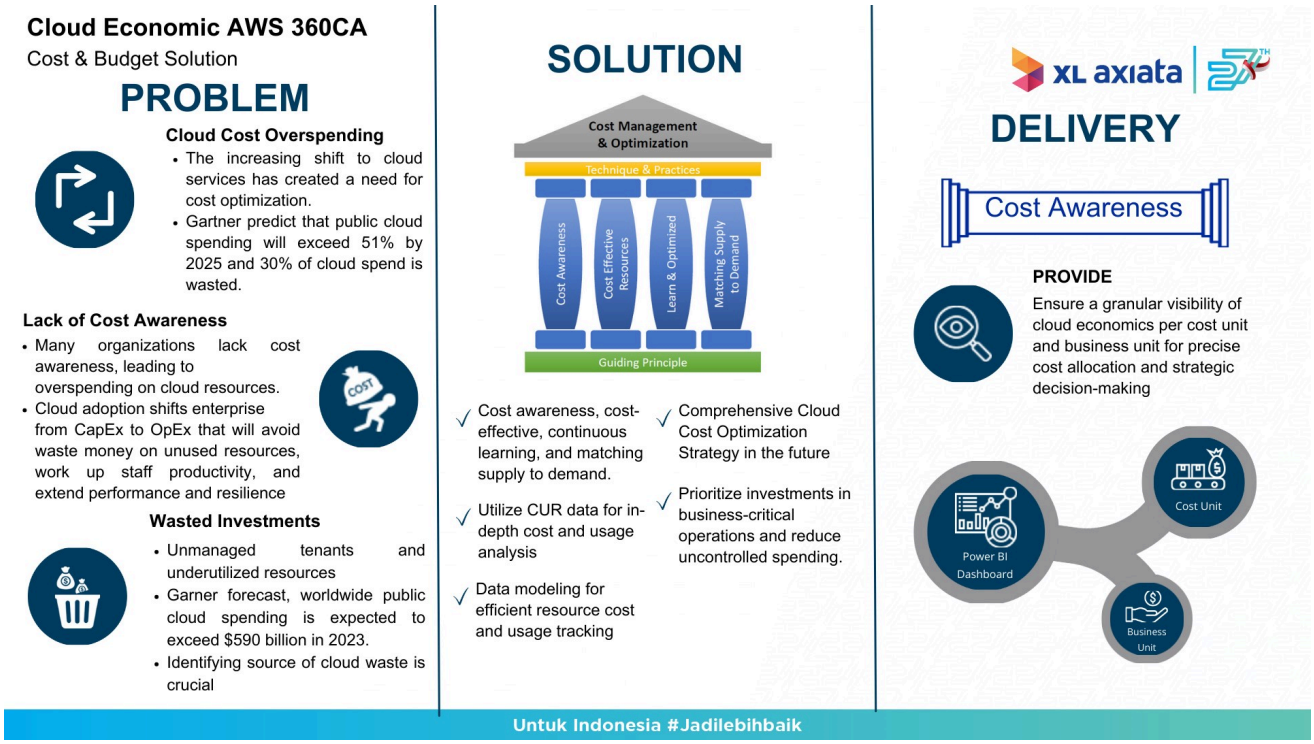
Cost optimization on AWS cloud aims to achieve efficient resource utilization, enhance performance, and maintain or improve service levels, all within a cost-effective framework. Here's the purpose of cost optimization:

1. Efficient Resource Utilization: The primary purpose of cost optimization is to ensure efficient use of cloud resources, avoiding over-provisioning and underutilization.
2. Minimize Unnecessary Expenses: Cost optimization aims to minimize unnecessary expenses by identifying and eliminating redundant or unused resources in the cloud infrastructure.
3. Enhance Performance: Optimization strategies seek to improve performance and responsiveness of cloud services without incurring additional costs, thus maximizing the value of the investment.
4. Maintain or Improve Service Levels: The goal is to optimize costs while maintaining or even improving service levels, ensuring that cost-cutting measures do not compromise the quality of services.

5. Strategic Financial Efficiency: Cost optimization contributes to strategic financial efficiency, allowing organizations to allocate resources effectively, reduce waste, and achieve better return on investment (ROI) from their cloud deployments.
6. Scalability and Agility: By optimizing costs, organizations can scale resources as needed, respond quickly to changing demands, and maintain agility in adapting to evolving business requirements in a cost-effective manner.
7. Continuous Monitoring and Adjustment: The process involves continuous monitoring, analysis, and adjustment of cloud resources, ensuring ongoing alignment with actual business needs and preventing unnecessary expenditures.

Integration of Cost Management and Visibility:

- **Cost Transparency:** Foster a culture of transparency regarding costs throughout the organization. This helps employees understand how their actions impact costs and encourages responsible spending.
- **Technology Solutions:** Implement cost management and visibility tools and software that provide comprehensive insights into financial data. Cloud-based solutions can offer real-time access from anywhere.
- **Regular Analysis:** Conduct regular analyses of cost data and performance metrics to identify areas where costs can be optimized. This may involve collaboration between finance, operations, and other relevant departments.
- **Continuous Improvement:** Use the insights gained through visibility to drive continuous improvement initiatives. This could involve process optimization, renegotiating contracts, or adopting new technologies to enhance efficiency.



What are we doing?

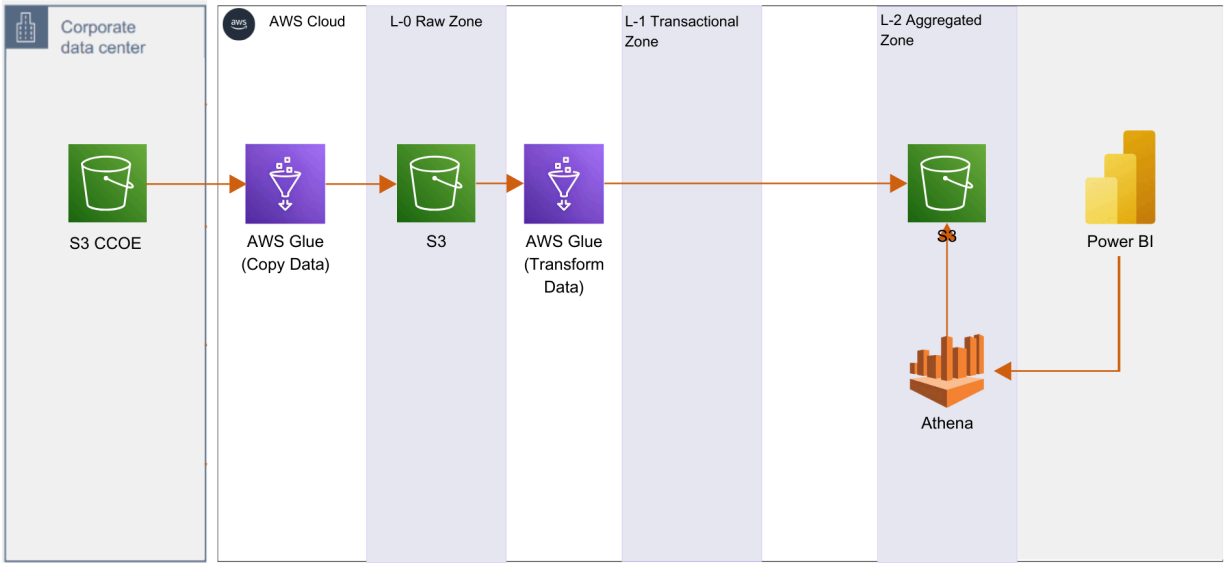
Latest status

The latest status of the project is in the design phase, where the team is working on adjusting the column naming standards of the data model to comply with the rules in EDM. They are also identifying which columns require parsing of their values. The project is currently at the design stage for both the model and pipeline, and has progressed to Athena. However, since the design is not yet finalized, further adjustments are still needed to align with the design.

Data architecture

The data architecture described involves the movement and transformation of data using various AWS services. The process begins with data being copied from S3 to an AWS Glue data catalog, then transformed using AWS Glue, and finally stored in different S3 buckets (bronze and gold) before being queried using Athena.

1. **S3 to AWS Glue (Copy data)**: Data is initially copied from S3 to an AWS Glue data catalog, where it is organized and made available for further processing.
2. **AWS Glue (Transform data)**: The data is then transformed using AWS Glue, which provides built-in transforms for ETL operations. This transformation step involves processes such as data format, quality checks, deduplication, and other transformations.
3. **S3 Bronze and S3 Gold**: The transformed data is stored in different S3 buckets, such as S3 bronze and S3 gold, as part of the data lake architecture.
4. **Athena**: Once the data is in S3, it can be queried using Athena, which allows for easy analysis of the data stored in S3 using standard SQL.



PowerBI Dashboard

This is the dashboard that show data of AWS 360CA CUR Data specifically:

1. **Total cost $31.6K with 49 services from AWS**
2. **Average used CPU 6.6 Hour/Month**: This metric indicates the average amount of time the CPU is in use during a month. A higher value means that the CPU is being utilized more frequently, which can lead to better performance. However, high CPU usage can also cause system slowdowns if the CPU is overworked
3. **Average Stored Data 20.32 GB/Month**: This metric represents the average amount of data stored in the system per month. It helps to understand the data storage requirements and can be useful for managing storage resources and costs
4. **Average Processed Data 8.52 GB/Hour**: This metric indicates the average amount of data processed per hour. It can help to evaluate the efficiency of the system in handling data and can be useful for optimizing data processing pipelines.
5. **Average Daily Running Use Case 3 occurence/Month**: This metric indicates that the system is used for running a specific use case three times a month on average. It helps to understand the frequency of use of the system and can be useful for capacity planning and resource allocation.
6. **Average running instance 14.74 instance used/Month**: This metric represents the average number of instances used per month. It helps to understand the resource utilization of the system and can be useful for optimizing resource allocation and cost management.

Monitoring these metrics can help identify potential issues with system performance, resource allocation, and data processing. By tracking CPU usage, stored data, and proceed data, administrators can make informed decisions about adjusting system resources, optimizing data processing, and ensuring optimal system performance.

Dimension Cost

With a total of $31.6K as in time dimension and volume dimension. Dimension cost refers to the categorization of costs based on various dimensions, such as time, resource, or service type. AWS Cost Categories allow you to define cost categories rules across dimensions, providing more flexibility in grouping AWS resources and accounts. Dimension cost specifically:

1. **Time Dimension**, refers to the categorization of costs based on time periods, such as daily, weekly, or monthly. By categorizing costs based on time, you can better understand how your costs are changing over time and identify trends or anomalies. This can help you to optimize your resource usage and reduce costs. In this dashboard the expenditure for time cost amounts to $23.7K with storage cost $7.84K, process cost $15.84K, and discount cost $444.77. The top of storage cost is from Amazon DocumentDB (with MongoDB compatibility) with $59.15. The top of process cost is from Amazon Managed Streaming for Apache Kafka with $44.10

2. **Volume Dimension**, refers to the categorization of costs based on the amount of resource usage, such as the number of instances, storage capacity, or bandwidth. This dimension is used to group costs based on the volume of resources being used, which can help you identify trends and patterns in your cost data and make informed decisions about optimizing your resource usage. In this dashboard the expenditure for volume cost amounts to $7.2K with storage cost $108.29, process cost $7.07K, and discount cost $-0.65. The top of storage cost is from AWS X-Ray with $11.75. The top of process cost is from AWS Config with $5.44