**Sustainability on AWS Cloud**

By Krunal Shah

# Introduction

Innovation is key to achieving sustainability goals—challenges such as decarbonization of operations to water conservation are addressed through technologies that drive sustainable transformation. Customers can build sustainable solutions ranging from carbon tracking to energy conservation to waste reduction, using AWS services to ingest, analyse, and manage sustainability data using AI and ML technologies offered by AWS.

[**Moving to AWS Can Help Significantly Lower Carbon**](https://sustainability.aboutamazon.com/carbon-reduction-aws.pdf)

Studies by 451 Research have shown that AWS’ infrastructure is 3.6 times more energy efficient than the median of U.S. enterprise data centres surveyed and up to five times more energy efficient than the average in Europe. 451 Research also found that AWS can lower customers’ workload carbon footprints by nearly 80% compared to surveyed enterprise data centres, and up to 96% once AWS is powered with 100% renewable energy—a target we’re on path to meet by 2025.

# Optimizing your AWS Infrastructure for Sustainability

As many organizations align their business with sustainable practices, it is important to review all functional areas for sustainable growth. As an AWS Engineer if you’re building, deploying, and maintaining an IT stack, improving its environmental impact requires informed decision making.

This can be done majorly in 3 ways for which details are as below:

## **Optimizing the Compute layer of your AWS infrastructure**

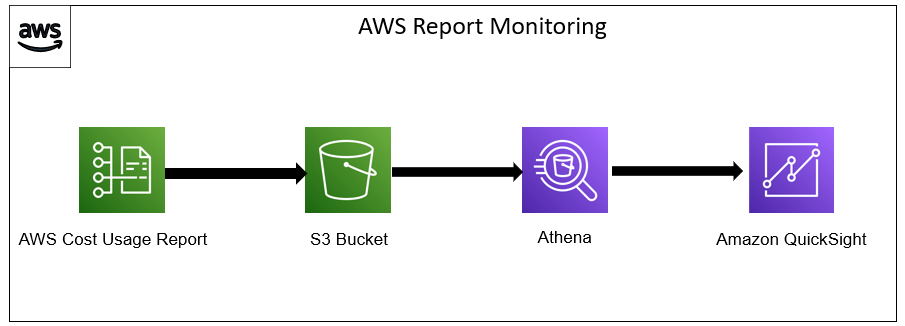
Compute services make up the foundation of many customers’ workloads, which brings great potential for optimization. Improving the efficiency of your workload by using the fewest number of compute resources and achieving a high utilization could help you to use resources more efficiently and save costs.

The following table shows you the most widely used compute service and the compute metric to measure:

|  |  |
| --- | --- |
| **Service** | **Measuring Metric** |
| Amazon Elastic Compute Cloud (Amazon EC2) | CPU Utilization |
| Total Number of vCPUs |
| Amazon Elastic Container Service (Amazon ECS) | CPU Utilization |
| Amazon EMR | IsIdle |

You can also make use of AWS Cost and Usage report to understand more about your resource utilization. Below is a simple graphical representation of how the solution looks like.

Use Athena to query the reports for AWS CUR stored in S3 bucket, you can also use quick sight for creating custom dashboards for graphical overview about the usage.



In addition to this below are some of the additional ways to reduce idle resources and maximize utilization.

* Integrate Amazon EC2 Auto Scaling.
* Right-size resources with AWS Cost Explorer and AWS Graviton2.
* Adopt a serverless, event-driven architecture.
* Use Amazon EC2 Spot Instances.

## **Optimizing the Storage layer of your AWS infrastructure**

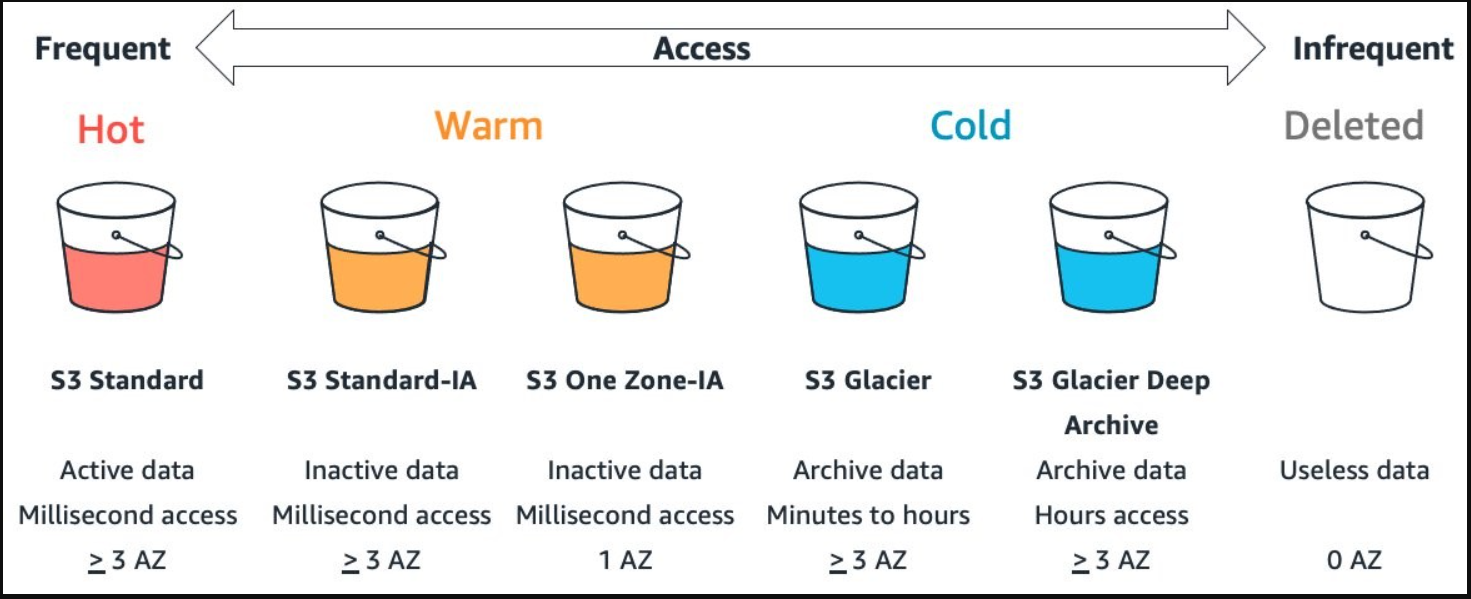
|  |  |
| --- | --- |
| **Service** | **Measuring Metric** |
| Amazon Simple Storage Service (Amazon S3) | BucketSizeBytes |
| S3 Object Access |
| Amazon Elastic Block Store (Amazon EBS) | VolumeIdleTime |
| Amazon Elastic File System (Amazon EFS) | Storage Bytes |
| Amazon FSx for Lustre | FreeDataStorageCapacity |
| Amazon FSx for Windows File Server | FreeStorageCapacity |

Storing and accessing data efficiently, in addition to reducing idle storage resources results in a more efficient and sustainable architecture. Amazon CloudWatch offers storage metrics that can be used to assess storage improvements, as listed in the following table.

In the following sections, we present four concepts to reduce idle resources and maximize utilization for your AWS storage layer.

Choosing the right storage tier after analysing data access patterns gives you more sustainable storage options in the cloud.

* By storing less volatile data on technologies designed for efficient long-term storage, you will optimize your storage footprint. More specifically, you’ll reduce the impact you have on the lifetime of storage resources by storing slow-changing or unchanging data on magnetic storage, as opposed to solid state memory. For archiving data or storing slow-changing data, consider using Amazon EFS Infrequent Access, Amazon EBS Cold HDD volumes, and Amazon S3 Glacier.
* To store your data efficiently throughout its lifetime, create an [Amazon S3 Lifecycle configuration](https://docs.aws.amazon.com/AmazonS3/latest/userguide/object-lifecycle-mgmt.html) that automatically transfers objects to a different storage class based on your pre-defined rules. The Expiring Amazon S3 Objects Based on Last Accessed Date to Decrease Costs blog post shows you how to create custom object expiry rules for Amazon S3 based on the last accessed date of the object.
* For data with unknown or changing access patterns, use Amazon S3 Intelligent-Tiering to monitor access patterns and move objects among tiers automatically. In general, you must make a trade-off between resource efficiency, access latency, and reliability when considering these storage mechanisms. Figure 2 shows an overview of data access patterns for Amazon S3 and the resulting storage tier. For example, in S3 One Zone-IA, energy and server capacity are reduced, because data is stored only within one Availability Zone.



## **Optimizing the Networking layer of your AWS infrastructure**

This section focuses on the network layer of your AWS infrastructure and proposes concepts to optimize your network utilization.

When you make your applications available to more customers, the packets that travel across the network will increase. Similarly, the larger the size of data, as well as the more distance a packet has to travel, the more resources are required to transmit it. It becomes important on optimize the network components on AWS to reduce the costs.

Reducing the data sent over the network and optimizing the path a packet takes will result in a more efficient data transfer. The following table provides metrics related to some AWS services that can help you find potential network optimization opportunities.

|  |  |
| --- | --- |
| **Service** | **Metric/Check** |
| Amazon CloudFront | Cache hit rate |
|  |
| Amazon Simple Storage Service (Amazon S3) | Data transferred in/out of a bucket |  |
|  |
| Amazon Elastic Compute Cloud (Amazon EC2) | NetworkPacketsIn/NetworkPacketsOut |  |
| CloudFront Content Delivery Optimization |  |

The following strategies allow users to read the data from the source closest to them; thus, fewer requests travel longer distances.

**Use a content delivery network**

Content delivery networks (CDNs) bring your data closer to the end user. When requested, they cache static content from the original server and deliver it to the user. This shortens the distance each packet must travel.

**Optimize CloudFront cache hit ratio**

Edge computing brings data storage and computation closer to users. By implementing this approach, you can perform data pre-processing or run machine learning algorithms on the edge.

CloudFront Functions can run compute on edge locations and Lambda@Edge can generate regional edge caches.

Sogeti is a leading provider of technology and engineering services. Sogeti delivers solutions that enable digital transformation and offers cutting-edge expertise in Cloud, Cybersecurity, Digital Manufacturing, Digital Assurance & Testing, and emerging technologies. Sogeti combines agility and speed of implementation with strong technology supplier partnerships, world class methodologies and its global delivery model, Rightshore®. Sogeti brings together more than 25,000 professionals in 15 countries, based in over 100 locations in Europe, USA and India. Sogeti is a wholly-owned subsidiary of Capgemini SE, listed on the Paris Stock Exchange.

About Sogeti

Learn more about us at

www.sogeti.com