



Amazon Athena

- *Rajesh Shirke*

Amazon Athena

- ❖ **Interactive query service** to analyze data in Amazon S3 using standard SQL and get results in seconds
- ❖ Athena is serverless - no infrastructure to set up or manage
- ❖ Pay only for the queries you run
- ❖ Athena scales automatically (executing queries in parallel)
- ❖ Fast results even with large datasets & complex queries

Amazon Athena

- ❖ Analyze unstructured, semi-structured, and structured data
Example - CSV, JSON, or columnar data formats – Apache Parquet and Apache ORC)
- ❖ Athena integrates with Amazon QuickSight and AWS Glue Data Catalog (offers a persistent metadata store for data in S3)
- ❖ AWS Service Integration with Athena

<https://docs.aws.amazon.com/athena/latest/ug/athena-aws-service-integrations.html>

<https://docs.aws.amazon.com/quicksight/latest/user/welcome.html>

<https://docs.aws.amazon.com/glue/latest/dg/what-is-glue.html>

Demo

- Upload CloudWatch log to S3
- Attached policies to the user account
- Create bucket policy
- Athena - add table/columns and select data format
- Run SQL queries
- Pricing calculator <https://aws.amazon.com/athena/pricing/>

QuickSight Integration with Athena



<https://aws.amazon.com/blogs/big-data/derive-insights-from-iot-in-minutes-using-aws-iot-amazon-kinesis-firehose-amazon-athena-and-amazon-quicksight/>

Amazon Athena

AWS IoT

- Easily and securely connects devices
- The IoT Rules Engine continuously processes incoming messages, enabling devices to interact with other AWS services

Amazon Kinesis Firehose

- Capture and automatically load streaming data into Kinesis Analytics, S3, Redshift, Splunk, and Elasticsearch Service
- Enables near-real-time business intelligence and the use of dashboards

Amazon QuickSight

- Fast, cloud-powered business analytics service to build visualizations, analysis, and get business insights from your data.
- Run SQL queries using Athena on data in S3 to build business dashboards

References

- ❖ <https://docs.aws.amazon.com/AmazonCloudWatch/latest/logs/S3ExportTasksConsole.html>
- ❖ <https://docs.aws.amazon.com/athena/latest/ug/querying-AWS-service-logs.html>
- ❖ <https://docs.aws.amazon.com/AmazonCloudWatch/latest/logs/WhatIsCloudWatchLogs.html>

Further Reading

- ❖ <https://aws.amazon.com/blogs/big-data/visualize-over-200-years-of-global-climate-data-using-amazon-athena-and-amazon-quicksight/>
- ❖ <https://aws.amazon.com/blogs/big-data/querying-openstreetmap-with-amazon-athena/>
- ❖ <https://blog.panoply.io/an-amazonian-battle-comparing-athena-and-redshift>
- ❖ <https://blog.panoply.io/the-spectrum-of-redshift-athena-and-s3>

Amazon Athena

