**International Journal of Computer Engineering and Technology (IJCET)**
Volume 15, Issue 5, Sep-Oct 2024, pp. 475-484, Article ID: IJCET_15_05_044
Available online at https://iaeme.com/Home/issue/IJCET?Volume=15&Issue=5
ISSN Print: 0976-6367 and ISSN Online: 0976-6375
Impact Factor (2024): 18.59 (Based on Google Scholar Citation)
DOI: https://doi.org/10.5281/zenodo.13845994



# CLOUD COMPUTING FOR EFFICIENT IMPLEMENTATION OF BUSINESS ANALYTICS

**Ajinkya P. Chatur**

**Shiragi S. Pattalwar**

**ABSTRACT**

*With exponential growth in data collection of different variety, volatility and with high velocity, we are ending up with huge data. Managing, storing, processing and understanding the hidden details in that data, has become very important for businesses and given rise to the field of business analytics. In cloud computing with the help of cloud services, we can setup resources efficiently, easily and in very less cost. Today, we have 3 large cloud providers Amazon Web Services (AWS), Microsoft Azure, Google Cloud Platform (GCP). This paper pro- vides a comprehensive comparison of different services provided by these cloud providers for different stages and use cases of business analytics. It delves into core services, cost, performance and implementation aspects. It also explores some common use cases of business analytics and try to provide a comparison between these services with a goal to provide clarity to users about service selection. Key- words: business analytics, cloud computing, cloud service providers, Amazon Web Services (AWS), Microsoft Azure, Google Cloud Platform (GCP).*

**Keywords:** Business Analytics, Cloud Computing, Cloud Service Providers, Amazon Web Services (AWS), Microsoft Azure

**Cite this Article:** Ajinkya P. Chatur and Shiragi S. Pattalwar, Cloud Computing for Efficient Implementation of Business Analytics, *International Journal of Computer Engineering and Technology (IJCET)*, 15(5), 2024, pp. 475-484.
https://iaeme.com/MasterAdmin/Journal_uploads/IJCET/VOLUME_15_ISSUE_5/IJCET_15_05_044.pdf

## 1. INTRODUCTION

### 1.1. Purpose of the paper

This paper aims to evaluate and compare the major cloud platforms—AWS, Azure, and GCP—interms of their capabilities and offerings for business analytics. By analyzing various services pro-vided by these platforms, the paper seeks to guidebusinesses in choosing the most suitable cloud provider based on their specific analytics needs.

### 1.2. Scope

The focus is on different products provided by AWS, Azure and GCP to fulfill the needs of different phases of business analytics. This includesdata warehousing, business intelligence tools, and data integration services. The paper evaluates these services in terms of cost, performance, scalability, and ease of implementation.

### 1.3. Significance

As businesses increasingly rely on data-driven decisions, selecting the right cloud platform for analytics is crucial as there are many options. This paper provides an in-depth analysis and comparison to help stakeholders make informed decisionsbased on their business requirements.

## 2. OVERVIEW OF CLOUD PROVIDERS

### 2.1. Amazon Web Services (AWS)

#### *2.1.1. History and Evolution*

AWS, launched by Amazon in 2006, has become a leading cloud service provider with current market cap of 31%. Over the years, AWS has expanded its offerings significantly, including a comprehensivesuite of services for business analytics [1].

#### *2.1.2. Core Business Analytics Services*

**Amazon Redshift**: A fully managed data ware- house service known for its high performance andscalability [2].

**Amazon QuickSight**: A cloud-based business intelligence service that provides interactive dash- boards, visualizations, and reporting [3].

**AWS Glue**: A managed ETL (extract, transform,load) service that simplifies the process of data preparation and loading [4].

**AWS Data Pipeline**: A web service that helpsin processing and moving data between different AWS compute and storage services [5].

### 2.2. Microsoft Azure

#### *2.2.1. History and Evolution*

Launched in 2010, Microsoft Azure has grown tobecome a major player in the cloud space with cur- rent market cap of 24%. Azure's analytics services are designed to support a wide range of data processing and business intelligence needs [6].

### 2.2.2. Core Business Analytics Services

**Azure Synapse Analytics**: Previously known as Azure SQL Data Warehouse, it combines big dataand data warehousing capabilities [7].

**Power BI**: A suite of business analytics tools thatprovides interactive visualizations and business intelligence capabilities [8].

**Azure Data Factory**: A cloud-based data integration service that allows for the creation, scheduling, and orchestration of data pipelines [9].

**Azure Databricks**: An Apache Spark-based analytics platform optimized for Azure [10].

## 2.3. Google Cloud Platform (GCP)

### *2.3.1. History and Evolution*

GCP, launched by Google in 2008, focuses on pro- viding high-performance, scalable cloud solutionswith current market cap of 11%. It offers a range of services tailored for business analytics and dataprocessing [11].

### *2.3.2. Core Business Analytics Services*

**BigQuery**: A fully managed, serverless data warehouse service that enables super-fast SQL queries and interactive analysis [12].

**Data Studio**: A free tool for creating interactive dashboards and reports [13].

**Dataflow**: A fully managed service for stream and batch data processing, based on Apache Beam[14].

**Dataproc**: A managed Spark and Hadoop servicethat simplifies the process of running big data processing frameworks on Google Cloud [15].

# 3. COMPARATIVE ANALYSIS

## 3.1. Core Services for Business Analytics

### *3.1.1. Data Warehousing*

- **AWS Redshift:**

– **Features**: Columnar storage, parallel query execution, automatic backups, data encryption, and integrations with various BI tools [2].

– **Performance**: Redshift offers high performance with its massively parallel processing (MPP) architecture. It sup- ports complex queries and large-scale data analytics [16].

– **Scalability**: Easily scalable by adding nodes or resizing clusters. Supports on-demand and reserved instances [17].

- **Azure Synapse Analytics:**

– **Features**: Unified analytics experience, serverless SQL pool, dedicated SQL pool, integration with Apache Spark,and real-time data processing [7].

– **Performance**: Combines big data and data warehousing capabilities, optimized for both data lake and data ware-house scenarios [18].

– **Scalability**: Scales dynamically based on workload requirements. Supports various sizes of SQL pools and integrates with on-demand and provisionedresources [19].

- **GCP BigQuery:**

– **Features**: Serverless architecture, real- time data analysis, built-in machine learning capabilities, and automaticscaling [12].

– **Performance**: High-speed query execution and processing capabilities. Optimized for large-scale data analytics withlow latency [20].

– **Scalability**: Automatically scales basedon query and storage requirements. Provides a fully managed experience with-out needing to handle infrastructure [21].

### *3.1.2. Business Intelligence*

- **AWS QuickSight:**

– **Features**: Interactive dashboards, visualizations, machine learning insights, and integration with AWS data sources[3].
– **Ease of Use**: User-friendly interfacewith drag-and-drop functionality. Integration with other AWS services simplifies data analysis [22].
– **Integration**: Supports integration with various AWS data sources, including Redshift, RDS, and S3 [23].

- **Azure Power BI:**

– **Features**: Rich visualizations, interactive reports, data modeling, and integration with Azure and other data sources [8].
– **Ease of Use**: Intuitive interface with ex-tensive data connectivity options. Customizable dashboards and reports [24].
– **Integration**: Seamless integration with Azure Synapse, Azure SQL Database, and various external data sources [25].

- **GCP Data Studio:**

– **Features**: Customizable dashboards, interactive reports, integration with Google services, and collaborative features [13].
– **Ease of Use**: Simple interface with easy-to-create visualizations. Integration with Google data sources and third-party connectors [26].
– **Integration**: Connects with BigQuery, Google Sheets, and other data sources, allowing for flexible reporting [27].

### *3.1.3. Data Integration*

- **AWS Glue:**

– **Capabilities**: Data cataloging, ETL processes, schema discovery, and integration with various AWS services [4].
– **Ease of Use**: Managed ETL servicewith automated data preparation and transformation features. Simple jobscheduling and monitoring [28].

- **Azure Data Factory:**

– **Capabilities**: Data movement, data transformation, pipeline orchestration, and integration with various on-premises and cloud data sources [9].
– **Ease of Use**: Visual interface for de- signing data workflows. Integration with Azure services and extensive connectivity options [29].

- **GCP Dataflow:**

– **Capabilities**: Stream and batch dataprocessing, flexible data pipelines, and integration with Apache Beam [14].
– **Ease of Use**: Managed service for creating and running data pipelines with minimal infrastructure management [30].

## 3.2. Cost Analysis

### *3.2.1. Pricing Models*

- **AWS Pricing:**

– **Redshift**: On-demand pricing starts at $0.25 per hour for a dc2.large node. Re-served instances offer lower rates [2].

– **QuickSight**: Pricing based on the num-ber of users and the amount of data pro-cessed. Starting at $9 per user per month [3].

– **Glue**: Charged based on the data pro- cessed and the number of Data Process-ing Units (DPUs) used. Pricing varies by region [4].

- **Azure Pricing:**

– **Synapse Analytics**: On-demand andprovisioned pricing. Starting at $0.25 per hour for data movement and distribution nodes [7].

– **Power BI**: Free tier available, with Proand Premium options starting at $9.99 per user per month and $20 per user permonth, respectively [8].

– **Data Factory**: Pricing based on thenumber of pipeline activities and data movement. Data movement costs start at $0.25 per GB [9].

- **GCP Pricing:**

– **BigQuery**: On-demand pricing is $5 per TB of data processed. Flat-rate pricing available for high query volumes [12].

– **Data Studio**: Free of charge, but data sources may incur costs [13].

– **Dataflow**: Pricing based on the numberof workers and the duration of process-ing. Starts at $0.01 per GB of processeddata [14].

### *3.2.2. Cost Comparison*

- **Example Pricing**:
  - **Scenario**: Analyze a 1TB dataset with 50 queries per day.
    - **AWS Redshift**: Estimated cost of $0.25 per hour x 24 hours x 30 days = $180 per month [2].
    - **Azure Synapse**: Estimated cost of $0.25 per hour x 24 hours x 30 days = $180 per month [7].
    - **GCP BigQuery**: 1TB x $5 per TB = $5 per query x 50 queries = $250 per month [12].

- **Cost Efficiency:**

– **AWS and Azure**: Both offer reserved instance options that reduce costs for long-term commitments [2][7].

– **GCP**: BigQuery's on-demand pricing can be cost-effective for infrequent queries but might become expensive with high query volumes [12].

## 3.3. Performance and Efficiency

### *3.3.1. Benchmarking*

- **Query Performance:**

– **AWS Redshift**: Benchmarked for com-plex queries with high performance dueto its columnar storage and parallel pro-cessing [16].

– **Azure Synapse**: Optimized for bothlarge-scale data and interactive queries,with high performance across various workloads [18].

– **GCP BigQuery**: Known for its fastquery execution, even on large datasets,due to its distributed architecture [20].

### *3.3.2. Scalability*

- **AWS Redshift**: Scales by adding nodes or resizing clusters. Supports both on-demand and reserved instances for cost and performance optimization [17].
- **Azure Synapse**: Provides dynamic scaling of data pools based on workload requirements. Supports both on-demand and provisioned resources [19].
- **GCP BigQuery**: Automatically scales with-out requiring user intervention. Serverless architecture supports high scalability [21].

### *3.3.3. Resource Management*

- **AWS**: Offers detailed monitoring and management through AWS CloudWatch. Re-source allocation and scaling are configurable based on workload demands [22].
- **Azure**: Provides monitoring and management through Azure Monitor and Azure Ad-visor. Automatic scaling options help man- age resources efficiently [24].
- **GCP**: Uses Stackdriver for monitoring and management. Resource scaling is handled automatically based on usage patterns [27].

## 3.4. Implementation Details

### *3.4.1. Ease of Setup*

- **AWS Redshift**: Simple setup with a variety of instance types. Provides automated back-ups and scaling options. Requires some con-figuration for optimal performance [2].
- **Azure Synapse**: Easy integration with ex- isting Azure resources. Provides a unified workspace for data and analytics. Requires setup of dedicated SQL pools and Spark pools [7].
- **GCP BigQuery**: Serverless architecture sim- plifies setup. No infrastructure management required, making it easy to start analyzing data quickly [12].

### *3.4.2. Integration Capabilities*

- **AWS**: Integrates seamlessly with other AWSservices like S3, RDS, and DynamoDB. Sup-ports third-party BI tools and data connectors[22].
- **Azure**: Provides robust integration with Azure services like Azure Data Factory and Azure Machine Learning. Supports a wide range of data sources and external tools [24].
- **GCP**: Integrates well with Google services and supports various third-party data sources and tools. Provides APIs and connectors for flexibility [27].

#### *3.4.3. Security and Compliance*

- **AWS**: Offers extensive security features, including encryption, IAM (Identity and Access Management), and compliance certifications (e.g., GDPR, HIPAA) [22].
- **Azure**: Provides robust security and compliance features, including Azure Security Center, encryption, and compliance with variousstandards (e.g., ISO, SOC) [24].
- **GCP**: Known for its security-first approach with features like encryption, identity management, and compliance with global standards (e.g., GDPR, HIPAA) [27].

## 4. USE CASE EXAMPLES

### 4.1. Use Case 1: Retail Sales AnalyticsScenario

A retail chain wants to analyze sales data to im-prove inventory management and sales strategies.

- **AWS Solution:**
  - **Data Ingestion**: Use Amazon Kinesis to stream sales data in real-time [2].
  - **Data Storage**: Store raw data in Amazon S3 [2].
  - **Data Warehousing**: Process and analyze data using Amazon Redshift [2].
  - **Analytics and Reporting**: Create inter-active dashboards with Amazon Quick-Sight [3].
- **Azure Solution:**
  - **Data Ingestion**: Process real-time data using Azure Stream Analytics [7].
  - **Data Storage**: Store data in Azure BlobStorage [7].
  - **Data Warehousing**: Analyze data withAzure Synapse Analytics [7].
  - **Analytics and Reporting**: Use Power BI to create visualizations and reports [8].
- **GCP Solution:**
  - **Data Ingestion**: Collect sales data usingGoogle Cloud Pub/Sub [12].
  - **Data Storage**: Store data in GoogleCloud Storage [12].
  - **Data Warehousing**: Analyze data withBigQuery [12].
  - **Analytics and Reporting**: Generate interactive reports using Data Studio [13].

### 4.2. Use Case 2: Customer Behavior AnalysisScenario

A company wants to track and analyze customer behavior on its website to enhance user experienceand marketing strategies.

- **AWS Solution:**
  - **Data Ingestion**: Use Amazon Kinesis to capture real-time customer interactions [2].
  - **Data Storage**: Store data in Amazon S3[2].
  - **Data Warehousing**: Analyze data withAmazon Redshift [2].
  - **Analytics and Reporting**: Use Amazon QuickSight for detailed visualizations [3].

- **Azure Solution:**
- **Data Ingestion**: Use Azure Stream An-alytics for real-time processing of customer data [7].
- **Data Storage**: Store data in Azure BlobStorage [7].
- **Data Warehousing**: Analyze data using Azure Synapse Analytics [7].
- **Analytics and Reporting**: Visualizedata with Power BI [8].

- **GCP Solution:**
- **Data Ingestion**: Use Google CloudPub/Sub for capturing data [12].
- **Data Storage**: Store data in GoogleCloud Storage [12].
- **Data Warehousing**: Process data withBigQuery [12].
- **Analytics and Reporting**: Create dash-boards with Data Studio [13].

### 4.3. Use Case 3: Financial ReportingScenario

An enterprise needs to generate and analyze financial reports for monthly and quarterly evaluations.

- **AWS Solution:**
- **Data Ingestion**: Use AWS DataPipeline to automate data workflows [5].
- **Data Storage**: Store financial data in Amazon S3 [2].
- **Data Warehousing**: Analyze data withAmazon Redshift [2].
- **Analytics and Reporting**: Use Amazon QuickSight for reporting and analysis [3].

- **Azure Solution:**
- **Data Ingestion**: Use Azure Data Factory to orchestrate data workflows [9].
- **Data Storage**: Store data in Azure BlobStorage [7].
- **Data Warehousing**: Analyze data withAzure Synapse Analytics [7].
- **Analytics and Reporting**: Analytics and Reporting: Use Power BI for generating financial reports [8].

- **GCP Solution:**
- **Data Ingestion**: Use Dataflow for pro-cessing financial data [14].
- **Data Storage**: Store data in GoogleCloud Storage [12].
- **Data Warehousing**: Analyze data withBigQuery [12].
- **Analytics and Reporting**: Use DataStudio to create detailed reports [13].

## 5. CONCLUSION

### 5.1. Summary of Findings

- **AWS**: Provides a robust set of tools with strong integration capabilities and high performance but may be more complex in termsof setup and cost management.
- **Azure**: Offers a unified analytics experiencewith good integration and ease of use but may have higher costs for certain services.

- **GCP**: Known for its serverless architecture and high performance with competitive pricing, though it may lack some integrationdepth compared to AWS and Azure.

### 5.2. Recommendation

- **For Large-Scale Data Warehousing**: AWSRedshift and Azure Synapse Analytics are strong contenders. GCP BigQuery is also a viable option for its serverless architecture.
- **For Business Intelligence**: Power BI offers extensive features and integration, while QuickSight and Data Studio provide more streamlined solutions.
- **For Cost Efficiency**: GCP's BigQuery offers competitive pricing, especially for variable workloads. AWS and Azure have cost- effective options with reserved instances.

### 5.3. Future Outlook

**Emerging Trends**: The evolution of AI and machine learning integration, serverless computing, and advancements in data processingtechnologies will continue to shape the land-scape of cloud analytics.

## REFERENCES

[1] Amazon Web Services, "Amazon Web Services (AWS) Overview," AWS, 2024. [Online]. Available: http://aws.amazon.com/about-aws/

[2] Amazon Web Services, "Amazon Red- shift," AWS, 2024. [Online]. Available: https://aws.amazon.com/redshift/

[3] Amazon Web Services, "Amazon Quick- Sight," AWS, 2024. [Online]. Available: https://aws.amazon.com/quicksight/

[4] Amazon Web Services, "AWS Glue," AWS, 2024. [On- line]. Available: https://aws.amazon.com/glue/

[5] Amazon Web Services, "AWS Data Pipeline," AWS, 2024. [Online]. Available: https://aws.amazon.com/datapipeline/

[6] Microsoft Azure, "Microsoft Azure Overview," Microsoft, 2024. [Online]. Available: https://azure.microsoft.com/en-us/

[7] Microsoft Azure, "Azure Synapse Analytics," Microsoft, 2024. [Online]. Available: https://azure.microsoft.com/en-us/services/synapse-analytics/

[8] Microsoft Azure, "Power BI," Microsoft, 2024. [On- line]. Available: https://powerbi.microsoft.com/

[9] Microsoft Azure, "Azure Data Factory," Microsoft, 2024. [Online]. Available: https://azure.microsoft.com/en-us/services/data- factory/

[10] Microsoft Azure, "Azure Databricks," Microsoft, 2024. [Online]. Available: https://azure.microsoft.com/en-us/services/databricks/

[11] Google Cloud Platform, "Google Cloud Platform Overview," Google, 2024. [Online]. Available: https://cloud.google.com/

[12] Google Cloud Platform, "BigQuery," Google, 2024. [Online]. Available: https://cloud.google.com/bigquery/

[13] Google Cloud Platform, "Data Studio," Google, 2024. [Online]. Available: https://datastudio.google.com/

[14] Google Cloud Platform, "Dataflow," Google, 2024. [Online]. Available: https://cloud.google.com/dataflow/

[15] Google Cloud Platform, "Dataproc," Google, 2024. [Online]. Available: https://cloud.google.com/dataproc/

[16] D. Naous, J. Schwarz, C. Legner 2017. Analytics as A Service: Cloud Computing andThe Transformation of Business An Alytics Business Models And Ecosys Tems *25th European Conference on Information Systems (ECIS), Guimarães, Portugal, June 5-10, 2017*, pp. 487–501. ISBN 978-0-9915567-0-0

[17] J. Ereth, H. Baars 2015. Cloud-Based Business Intelligence and Analytics Applications – Business Valueand Feasibility. *PACIS 2015 Proceedings.*, Paper 36.

[18] L. Barthelus 2010. Adopting cloud computing within the healthcare industry: opportunity or risk? *OnlineJournal of Applied Knowledge Management*, 4 (1) (2010).

[19] I. Neaga, Y. Hao 2014. A Holistic Analysis of Cloud Based Big Data Mining. *International Journal of Knowledge, Innovation and Entrepreneurship*, 2 (2)(2014), pp. 56-64

**Citation:** Ajinkya P. Chatur and Shiragi S. Pattalwar, Cloud Computing for Efficient Implementation of Business Analytics, International Journal of Computer Engineering and Technology (IJCET), 15(5), 2024, pp. 475-484

**Abstract Link:** https://iaeme.com/Home/article_id/IJCET_15_05_044

**Article Link:**
https://iaeme.com/MasterAdmin/Journal_uploads/IJCET/VOLUME_15_ISSUE_5/IJCET_15_05_044.pdf



✉ **editor@iaeme.com**