International Journal of Trend in Scientific Research and Development (IJTSRD)
Volume 5 Issue 5, July-August 2021 Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

Cloud Analytics: Ability to Design, Build, Secure, and Maintain Analytics Solutions on the Cloud

Rajan Ramvilas Saroj

Department of MCA, YMT College of Management, Kharghar, Navi Mumbai, Maharashtra, India

ABSTRACT

Cloud Analytics is another area in the IT field where different services like Software, Infrastructure, storage etc. are offered as services online. Users of cloud services are under constant fear of data loss, security threats, and availability issues. However, the major challenge in these methods is obtaining real-time and unbiased datasets. Many datasets are internal and cannot be shared due to privacy issues or may lack certain statistical characteristics. As a result of this, researchers prefer to generate datasets for training and testing purposes in simulated or closed experimental environments which may lack comprehensiveness. Advances in sensor technology, the Internet of things (IoT), social networking, wireless communications, and huge collection of data from years have all contributed to a new field of study Big Data is discussed in this paper. Through this analysis and investigation, we provide recommendations for the research public on future directions on providing data-based decisions for cloud-supported Big Data computing and analytic solutions. This paper concentrates upon the recent trends in Big Data storage and analysing, in the clouds, and also points out the security limitations.

KEYWORDS: *Data analytics, business intelligence, cloud analytics, reference architecture, SaaS BI*

How to cite this paper: Rajan Ramvilas Saroj **"Cloud Analytics: Ability to Design, Build, Secure, and Maintain Analytics Solutions on the Cloud"** Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-5 | Issue-5, August 2021, pp.317-321, URL: www.ijtsrd.com/papers/ijtsrd43728.pdf

IJTSRD43728

INTRODUCTION

Cloud computing analytics helps streamline the business intelligence process of gathering, integrating, analysing, and presenting visions to enhance business decision making. Cloud analytics allows for the simultaneous recording and processing of data regardless of proximity to local servers. ... The data stored to clouds helps make business run more efficiently and gives companies a better understanding of their customers' behaviour. As clouds become more secure, consistent, and reasonable, the use of data analytics in cloud computing will also continue to grow. When choosing which cloud storage device could best fit a business, the question becomes how much data storage is needed and what presentation demands will be placed on the cloud. Aside from its increased user-friendliness and utility, big data analysis on cloud drives also exports many IT demands, such as hosting and upholding servers, to cloud service providers. Companies can spend less money on servers and instead focus on boosting their staff and product.

Cloud computing is built around a series of hardware and software that can be remotely accessed through any web browser. usually files and software is shared and worked on by multiple users and all data is remotely unified instead of being stored on users' hard drives. Not only does analytics determine what might attract new customers, often analytics recognizes existing patterns in data to help better serve existing customers, which is typically more cost effective than establishing new business. In an ever-changing business world subject to limitless variants, analytics gives companies the edge in identifying altering environments so they can take initiate suitable action to stay good.

The data processing is done on a private or public cloud to avoid the expense and conservation of on-premises data storage and compute. Cloud-based analytics is also called a Software as a Service model or Cloud Analytics as a Service model. Some companies use a hybrid model that keeps some functions on-premise while moving others to a cloud.

The role of data analytic in Cloud computing

Cloud computing is categorized by multiagency, fast deployment, low cost, scalability, rapid provisioning, elasticity, and global network access. Cloud computing provides enormous computing resources to the user applications through the internet in the large-scale scattered computing environment.

Resources are always available on the cloud and are made available to different user domains on a mandate basis. Cloud Computing providers often utilize a "software as a service" model to allow customers to easily process data. Typically, a console that can take in specialized commands and parameters is available, but everything can also be done from the site's user interface. Some products that are usually part of this package comprise database management systems, cloud-based virtual machines and containers, identity management systems, machine learning skills, and more.

In turn, Big Data is often generated by large, network-based systems. It can be in either a standard or non-standard format. If the data is in a non-standard format, artificial intelligence from the Cloud Computing provider may be used in addition to machine learning to normalize the data.

From there, the data can be connected through the Cloud Computing platform and utilized in a variety of ways. For example, it can be searched, edited, and used for future understandings.

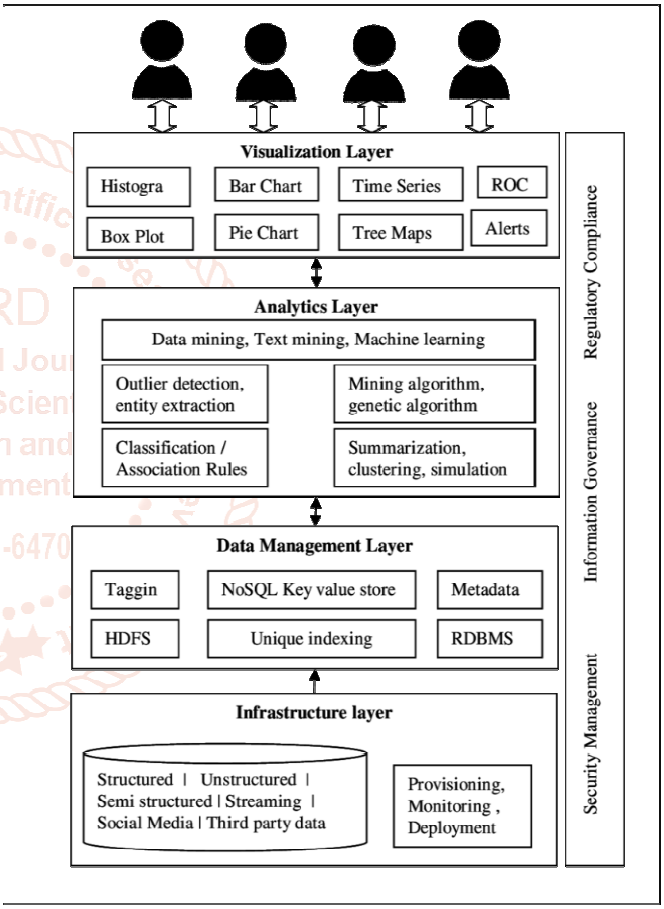
This cloud structure allows for real-time processing of Big Data. It can take huge "explosions" of data from intensive systems and interpret it in real-time. Another common relationship between Big Data and Cloud Computing is that the power of the cloud allows Big Data analytics to occur in a segment of the time it used to.

CLOUD ANALYTICS ARCHITECTURE:

Cloud analytics, which refers to the application of cloud services for a single or all components of an analytic solution, enables organizations to control social data, Internet data, and third-party data by using pay per use model. Integration of these data with the enterprise data will provide more insight into customer preferences and demands. The need for a new interactive database model such as NoSQL, which efficiently stores and access a huge volume of data from the cloud evolved as the analytics applications utilize datasets with more reads than writes. The compute-intensive module of the analytics is the primary candidates of the analytics application to be moved on to the cloud for the advantage of fast query responses and reduced decision-making time.

Infrastructure Layer:

This layer is responsible for storing and managing data. The layer consists of structured and unstructured data assembled from various data sources. The machine/sensor data, social data, video streaming data, and third party. data are available on pay per use basis for easy integration with enterprise data. The entire infrastructure management activities such as dynamic provisioning, monitoring, and automatic deployments are handled by this layer. The requirement spikes and the rapid growth of the database are handled by the horizontal and vertical scalability feature of cloud organization. This is also the foundation layer of the cloud analytics architecture that enables organizations to leverage fast and low-risk infrastructure deployments.



Data Management Layer:

It is a data storage repository layer which is also termed as "data lakes". A huge amount of raw data in structured, unstructured, and semi-structured format is stored as flat files for analytics purposes. The schema and data format requirements are not defined until a query is raised. Each data element of this layer is attached with a unique identifier and is tagged using metadata tags . This layer is maintained using Hadoop-oriented object storage, where the queries are applied to retrieve intelligence . It differs from the Enterprise Data Warehouses (EDW) as it employs schema-on-read processing, possesses greater agility and reusability along with a low cost of storage. Node

failure and data movements between the nodes are routine operations in data management using cloud setup. This is done to maintain the fault tolerance feature of cloud analytics using automatic replication of data across nodes in the cloud cluster of this layer.

Analytics Layer:
The core layer of cloud analytics as it holds the business intelligence applications. It consists of pre-analytic or filtering tools and analytics tools. Pre-analytic tools are used to cleanse, choose and organize data that were retrieved from the data management layer, and analytics tools are used to identify the patterns and retrieve actionable insights from the data. Various analytics tools and software used by this layer are R, SAS, Mathematica, MapReduce, Pig, Hive, and ETL tools, etc.

Visualization Layer:
The layer provides a user interface for accepting user queries and displaying analytics results. This layer helps users to customize the data analytics visualizations. This layer enables subject matter experts (SME) to explore the data without any assistance from the IT experts. e elimination of assistance to extract insights enhances the quality of information retrieval as the SME directly works on the data without the interference which rejects the requirement specification degradation.

CLOUD ANALYTICS TOOLS:
The cloud analytics tool is a platform based in the cloud that allows the user to view analytics data. Through one pane of glass can monitor data sources, data models, processing applications, computing resources, analytics models, and storage resources. Diverse model, based execution and customization requirements of BI applications make it difficult to implement on a global scale that suits all customer supplies. The core benefit of cloud analytics tools is that you can monitor lots of disparate services and datasets from one location.

Cloud analytics tools:

1. **App Optics Custom Metrics and Analytics**: A data gathering service that gives businesses analytical insights into the performance of applications and IT infrastructure.
2. **IBM Cognos Analytics**: Uses AI techniques to uncover and identify patterns combined with great visuals. Has plans to suit businesses at any scale.
3. **Microsoft Power BI:** Great data visualization, create dashboards and easily share and collaborate with colleagues. Integrates machine learning.
4. **Zoho Analytics**: Available on-premises and in the cloud with easy-to-use drag-and-drop customizable dashboards.
5. **Board**: Decades of business performance management distilled into a cloud-based offering, with powerful pre-configured automated predictive modeling.
6. **TIBCO Spotfire**: AI-powered advanced analytics tool designed for enterprise users with strong search capabilities.
7. **Domo**: Pulls data from external service providers such as Microsoft Excel, Xero, Facebook, Salesforce, AWS, MySQL, and more.

IMPLEMENTATION CHALLENGES:
Cloud analytics provides an effective solution for the out-of-date BI issues but also includes various challenges in its execution. Organizations often fail to devise a collaborative plan involving different departments, before the implementation of cloud analytics solutions.

The challenges also arise due to geographically distributed storage, involvement of various third-party for the provision of intermediate services, cross-border compliance issues, and network latency.

- Security:
- Security is a significant factor when it comes to cloud analytics and computing challenges. ...
- Implementing a Functioning Cloud Analytics Application. ...
- Maintainability & Governance. ...
- Managing the Cloud Costs. ...
- Assess Total Ownership Cost.

Marketing Hype:
All companies that offer one or more components of the analytics solution are tagged as cloud analytics companies. This is one of the main challenges as organizations often get carried away by the marketing claims. Lack of clear understanding of the scope required and the scope of the cloud analytic solution will result in aincorrect selection of the product that may result in performance degradation of the Decision-making process.

Suggested Solution: Six elements of the analytics solutions and their scope of operations are to be clearly understood by the person responsible for the selection of the analytic product. Depending on the nature of data and process sensitivity either SaaS BI or analytics as a service has to be selected.

Integration:
At its most basic level, cloud integration means bringing multiple cloud environments together — either in a hybrid deployment or as multiple public clouds — so that they can operate as a single, cohesive IT infrastructure for an enterprise. The format compatibility of the data that are to be used

between the on-premise and cloud analytics solutions may lead to inconsistent decision making. Data originating from the cloud such as social media data or the SaaS application data are comparatively easier to handle than on-premise data.

Suggested Solution: Maintaining the data in the open data format or the standard data format is the optimal solution for the integration issues. The usage of data format conversion modules for integration must be minimized to reduce the software development and maintenance cost.

Security:

The 4 essential pillars of cloud security

- Visibility and compliance.
- Compute-based **security**.
- Network protections.
- Identity **security**.

There is an explosion of devices on the private network, and more workloads are being migrated to the public cloud. Meanwhile, security practitioners are inundated with security alerts to the point of unmanageability.

Cloud security posture management:

Secure Cloud Analytics begins checking your cloud resources for risky configurations and changes upon deployment. You can also create your watch lists to be alerted to the activity of interest and to ensure cloud resources are adhering to your internal policy.

Secure Cloud Analytics

Secure Cloud Analytics provides perceptibility and risk detection in Amazon Web Services (AWS), Google Cloud Platform, and Microsoft Azure infrastructures. It is a cloud-delivered, SaaS-based key that can be deployed easily and quickly.

The solution can be deployed without software agents, instead of relying on native sources of telemetry such as its Virtual Private Cloud (VPC) flow logs. Secure Cloud Analytics models all IP traffic generated by an organization's resources and functions whether they are inside the VPC, between VPCs, or to external IP addresses. It integrates with additional Cloud Service Provider APIs like Cloud Trail, Cloud Watch, Config, Inspector, Identity and Access Management (IAM), Lambda, and many more.

Fundamental Cloud Computing Skills

- Programming. ...
- Platform Expertise. ...
- Selecting the Right Services. ...
- Managing an Integrated Environment. ...
- Maintaining Databases. ...
- Managing a Network. ...
- Securing the **Cloud** Environment. ...
- Adapting to New Roles and Technologies.

Lack of skill set:

Recruiting, training, and employing data scientists is a stimulating task. As this is an emerging discipline getting expertise from a diverse field is very difficult. The choice of the best analytic tool to assist the organization in making the decision is very essential as it is the base for generating accurate business predictions and submissions.

Regulatory Compliance:

This is an inevitable part of all cloud services due to its working. Despite the regulation variation, the environmental distribution of data is essential to maintain fault acceptance.

Suggested Solution:

Product provisions are to be screened properly to check for compliance certifications. The providers of cloud analytics services have to acquire compliance credentials such as HIPAA, PCI-DSS, US-EU Safe depending on the domain of the serving customers. These certificates are to be changed periodically.

Customization:

SaaS BI on the other hand is established for a wide range of customers and hence customization as per the business needs is not possible. Establishments are either forced to analytics needs as per the SaaS BI specifications or have to tolerate the presence of unwanted modules.

Suggested Solution:

choosing ansuitable SaaS BI similarapproximately to their condition will provide a explanation to the challenge. If the organization is very particular about its BI needs, then the data storage can be done in the cloud and analytic solutions can be developed as an on-premise application. This will eliminate the tolerance of unwanted modules.

Latency:

Features that depend on inactivity

Some capabilities in the Adobe Experience Cloud come with an innate amount of latency on top of standard processing time. Analytics for Target (A4T) requires an additional 5-10 minutes of latency to allow collected data from both platforms to be stored in the same hit. Latency greatly affects how usable and enjoyable devices and communications are. ... Most tools for measuring latency, like trace route and pings work based on ICMP packets, which are not generally used. This is a longstanding challenge due to its lack of proper responsibility. Data localization can be opted by the organization to reduce the latency but it will growth the cost of storage.

CONCLUSION:

Association of the departments in outlining the BI needs of the organization will assist in determining the choice of the right mix of cloud analytics solutions. The hybrid or mix-cloud approach is the optimal solution to address security and availability issues. Cloud computing has been a boon to big data processing. The ability to capture, store, and process data without having to concern about scaling servers and databases has democratized and advanced big data capabilities like never before. Multi-cloud data analytics is the upcoming of business- uniting the resiliency and flexibility of multi-cloud policies with the power of data analytics.

References:

[1] https://www.xplenty.com/blog/multi-cloud-data-analytics

[2] https://www.cisco.com/c/en/us/products/collateral/security/stealthwatch/datasheet-c78-739619.html

[3] https://www.ijert.org/research/a-study-of-big-data-analytics-in-clouds-with-a-security-perspective

[4] http://sersc.org/journals/index.php/IJAST/article/view/11301

[5] https://www.computer.org/publications/tech-news/trends/big-data-and-cloud-computing

[6] https://ieeexplore.ieee.org/stamp/stamp.jsp

[7] https://www.researchgate.net/publication

[8] https://www.xplenty.com/blog/multi-cloud-data-analytics/