BIG DATA ANALYSIS WITH IBM CLOUD COMPUTING

312621104025:Santhosh S

Phase-2 Document submission

Objective:

Big Data Analytics accelerates the process by which firms transform information into insight. Making educated decisions about products, operations, marketing, and other company objectives follows from these insights. Large data sets need storage, which may be costly to maintain.

Introduction:

Big data analytics and cloud computing are two of the most important technologies in the IT industry. Big data refers to the storage, processing, and analysis of large amounts of data, while cloud computing is about providing the infrastructure to enable such processes in a cost-effective and efficient manner

Innovation:

Cloud computing and big data have led to several innovations in the field of data management. Here are some of the key points:

1. **Distributed computing systems**:
Cloud computing has enabled the use of distributed computing systems like Hadoop and Spark to store and process large amounts of data at a lower cost than traditional on-premises solutions.

2. **Cloud-based machine learning services**:
Cloud-based machine learning services like Amazon SageMaker or Google Cloud AI Platform can be used to build predictive models that can analyze big data. These models can help organizations gain insights into their data and make better business decisions.

3. **Real-time big data analytics platforms**:
Cloud computing has enabled the development of real-time big data analytics platforms like Apache Kafka or Google Cloud Pub/Sub. These platforms allow organizations to process large amounts of streaming data in real-time and make decisions based on that data.

4. **Cost-effective storage**:
Cloud storage is a cost-effective way to store large amounts of data. It allows organizations to store their data in the cloud and access it from anywhere in the world.

5. **Scalability:**

Cloud computing provides scalability, which means that organizations can easily scale up or down their computing resources based on their needs.

6. **Data security:**

Cloud computing provides a secure environment for storing and processing big data. Cloud providers offer various security measures like encryption, access control, and firewalls to protect data from unauthorized access.

7. **Cost savings:**

Cloud computing can help organizations save costs by reducing the need for expensive hardware and software. It also allows organizations to pay only for the resources they use, which can be more cost-effective than traditional on-premises solutions.

8. **Improved collaboration:**

Cloud computing enables teams to collaborate more effectively by providing a centralized platform for storing and sharing data. This can help improve productivity and reduce errors.

9. **Real-time data processing:**

Cloud computing has enabled real-time data processing, which means that organizations can make decisions based on up-to-date information.

10. **Scalability:**

Cloud computing provides scalability, which means that organizations can easily scale up or down their computing resources based on their needs.

Conclusion:

Big Data is not a new term but has gained its spotlight due to the huge amounts of data that are produced daily from different sources. From our analysis we saw that big data is increasing in a fast pace, leading to benefits but also challenges