

COMPUTING SYSTEMS HANDLING LARGE AMOUNTS OF DATA: BIG DATA APPLICATION AND ADOPTION

Ravinder Kumar*

(M.Tech., M.D.U., UIET, Rohtak)

V. & P. O.- Dubaldhan, District: Jhajjar, Haryana-124202, India

Email ID: ravinder143a@gmail.com

Accepted: 08.10.2022

Published: 01.11.2022

Keywords: Computing Systems, Big Data, Application.

Abstract

A recent boom in the usage of the internet among the masses can be felt significantly all over the world. In recent times due to the pandemic situation, the internet now has become a necessity for all. The usage of technology in varied platforms has been significantly impacting people's overall lifestyle, which in turn is defining the economical, technological, and fundamental growth of the world. With the advancement in technology, optic fiber and wireless networks have made it possible to transfer data at lighting speed around the world that leads to increase in volume and the rate at which data is created. However, this increased growth has led to many challenges, such as variety of data, portability of huge data and cyber security. The object of this study is to understand and categorize different characteristics of Big Data, that includes tendency, definitions, quick development rate, volume, the management, investigation, and security.

Paper Identification



**Corresponding Author*

Introduction

Big data means incredibly huge datasets that can be dissected to find trends, patterns, and associations, particularly connected with human way of behaving and associations. These datasets are in many cases excessively huge and complex to be handled and broken down utilizing customary data handling tools and methods. However, specific programming and calculations are expected to process and break down big data efficiently. Big data can emerge from various sources, for example, online entertainment, web search inquiries, web-based business exchanges, sensor data, and monetary records. It tends to be unorganized, unstructured, or semi-organized, and may incorporate text, pictures, sound, and video.

There are a few qualities of big data that make it not quite the same as conventional data: volume, variety, velocity, and veracity. Big data examination can be utilized in various enterprises and fields, including medical services, money, promoting, and government, to acquire experiences and go with better choices.

Handling Big Data

Handling big data relates to the procedure of storing, managing, and analyzing large and complex data sets

that can range from terabytes to petabytes in size. Some common challenges associated with handling big data include:

1. **Storage:** Storing large amounts of data can be a challenge, particularly if you need to store data in a structured format that allows for efficient querying and analysis.
2. **Processing:** Analyzing big data can require significant computing resources, including processing power, memory, and storage.
3. **Quality:** Ensuring the quality and accuracy of big data is essential, as poor quality data can lead to incorrect insights and decision-making.

There are several approaches that organizations can take to handle big data, including using specialized software and hardware tools, implementing distributed computing systems, and leveraging cloud computing resources.

Some common tools and technologies used for handling big data include Hadoop, Spark, and NoSQL databases such as MongoDB and Cassandra. These tools are designed to store and process a lot of data effectively and can be used in a variety of big data applications, including data warehousing, real-time analytics, and machine learning.

Adoption of big data technologies can be challenging for organizations due to the complexity of the technologies and the need for specialized skills to use them effectively. It is important for organizations to carefully assess their needs and determine whether the benefits of using big data technologies outweigh the costs and effort required to implement them.

Computing Systems dealing with Big Data

Computing systems are systems that perform computations, which are defined as the manipulation of data according to a set of instructions. These systems can range from simple calculators to large, complex computer systems that are used for a variety of

purposes, such as running businesses, processing scientific data, or communicating over the internet

There are several ways to handle large amounts of data in computing systems. Some common approaches include:

1. **Distributed systems:** In a distributed system, divide the data into smaller chunks and process them concurrently on multiple machines. This can be done using frameworks such as Hadoop or Spark, which allow developers to write code that can be executed on a cluster of computers.
2. **Parallel processing:** In parallel processing, multiple processors or cores are used to work on different parts of a problem simultaneously, resulting in faster processing times.
3. **Cloud computing:** Cloud computing allows you to access large amounts of computing resources and storage over the internet, rather than assembling and keeping up with your own infrastructure. This can be a cost-effective technique to handle large amounts of data.
4. **In-memory computing:** In-memory computing involves storing and processing data in the memory of a computer, rather than on a disk. This can be much faster than traditional disk-based storage, but it is limited by the amount of memory available.
5. **NoSQL databases:** NoSQL databases are designed to handle large amounts of data that is often unstructured or semi-structured. They are often used in distributed systems and are known for their scalability and flexibility.
6. **Cluster computing:** A cluster is a group of computers that work together as a single system. In a cluster computing system, data is divided up and processed by multiple

machines, which can improve the speed and scalability of data processing tasks.

7. Columnar databases: These databases store data in columns rather than rows, which can make it more efficient to retrieve and process specific columns of data, particularly for large datasets.

To take advantage of the insights that can be gleaned from big data, organizations must adopt the appropriate tools and technologies to process, store, and analyze these large datasets. This typically involves the use of specialized software and hardware platforms, such as distributed databases and cluster computing systems, to handle the scale and complexity of the data.

Example of Computing Systems

- **Cloud Computing**

Cloud computing is a model of computing in which resources, such as processing power, storage, and systems administration, are given over the web as a help. It permits clients to access and utilize these assets on request, without the need to buy, introduce, or keep up with the fundamental equipment and programming themselves.

Here is an illustration of Python code that shows how to involve the AWS SDK for Python (Boto3) to transfer a record to a S3 can (a kind of storage services given by AWS):

```
import boto3
# Create an S3 client
s3 = boto3.client('s3')
# Set the name of the bucket and the file to
upload
bucket_name = 'my-bucket'
file_name = 'path/to/local/file.txt'
# Upload the file to the bucket
s3.upload_file(file_name, bucket_name,
file_name)
```

This code creates an S3 client using the 'boto3' library, and then uses the upload_file method to upload the specified file to the specified bucket.

Other cloud providers have their own SDKs and APIs for interacting with their cloud services. For example, Azure has the Azure SDK for Python, and GCP has the Google Cloud Client Library for Python.

- **Distributed Computing Systems**

Distributed computing systems involve using multiple computers to perform a task. There are many ways to implement distributed systems. A common approach is the Client-server model.

In this model, there is a central server that manages the tasks and a group of clients that perform the tasks. The clients send requests to the server, which coordinates the distribution of tasks and returns the results to the clients.

Here is an example of a simple distributed system using the client-server model in Python:

```
import socket
# Set up the server socket
server_socket =
socket.socket(socket.AF_INET,
socket.SOCK_STREAM)
server_socket.bind(('localhost', 8080))
server_socket.listen()
# Set up the client socket
client_socket =
socket.socket(socket.AF_INET,
socket.SOCK_STREAM)
client_socket.connect(('localhost', 8080))
# Send a message from the client to the server
client_socket.sendall(b'Hello, server!')
# Receive the message from the server
data = server_socket.recv(1024)
```

```
print(repr(data))
# Close the sockets
client_socket.close()
server_socket.close()
```

This example creates a server socket that listens for incoming connections on port 8080 and a client socket that connects to the server. The client sends a message to the server, and the server receives and prints the message.

Adoption and Application

The adoption of big data technologies and practices has become increasingly common in recent years as organizations seek to gain insights and make data-driven decisions. There are several factors that have contributed to the adoption of big data, including:

1. **The growth of data:** The proliferation of digital devices and the Internet has led to a massive increase in the amount of data being generated.
2. **The need for real-time analysis:** Businesses need to be able to analyze data in real-time in order to make timely and informed decisions.
3. **The availability of tools and technologies:** There are now a wide variety of tools and technologies available to help organizations manage and analyze large volumes of data.
4. **The desire to gain competitive advantage:** Many businesses see the use of big data as a way to gain a competitive advantage over their rivals by using data to make better decisions and optimize their operations.

In order to successfully adopt big data, organizations need to have a clear understanding of their data needs and goals, as well as the tools and technologies available to them. They should also ensure that they have the necessary infrastructure, such as hardware and software, and that they have the necessary skills and expertise to effectively manage and analyze their data.

There are many applications of big data in various industries, including healthcare, finance, retail, and manufacturing. Some examples of how big data can be used include:

1. **Enhancing financial services:** Big data can be used to analyze customer data and identify trends that can help financial institutions make better lending and investment decisions.
2. **Fraud detection:** Big data can be used to identify patterns and anomalies that may indicate fraudulent activity, helping organizations to prevent losses and protect their assets.
3. **Customer analytics:** This is done through the use of data analytics tools and techniques, which allow businesses to identify patterns and trends in the data that can be used to inform their decision-making processes.
4. **Supply chain optimization:** Big data can help organizations to optimize their supply chains by analyzing data on demand, inventory, and logistics, and using this information to make more informed decisions about production, distribution, and fulfillment.
5. **Predictive maintenance:** By analyzing data from sensors and other monitoring systems, organizations can predict when equipment is likely to fail, allowing them to schedule maintenance and repairs before problems occur, reducing downtime and improving efficiency.
6. **Healthcare:** Big data can be used to improve patient outcomes by identifying trends and patterns in healthcare data that may indicate potential health issues, and by providing doctors with access to a patient's complete medical history.

Conclusion

Big Data is definitely not a new term yet has acquired its spotlight because of the immense measures of

information that are generated everyday from various sources. From examining various sources the conclusion can be drawn that Big Data is expanding at a high speed, prompting benefits and additional difficulties. After all the research it is safe to say that computing systems make it possible for us to manipulate and access all the information that the huge amount of data around us have to give. The variety in computing systems provide us the flexibility to adapt and apply Big Data to use in various fields and in numerous ways. Overall, computing systems play a critical role in handling big data by providing the necessary hardware and software infrastructure to process, store, and analyze large volumes of data.

10. *BBC News* :

<https://www.bbc.com/news/technology>

References

1. Hillbert M, Lopez P (2011) *The world's technological capacity to store, communicate and compute information. Science III:62–65*
2. Statista, "Statista," 2020. Available: <https://www.statista.com/statistics/871513/worldwide-data-created/>. Accessed 21 Jan 2021
3. Reinsel D, Gantz J, Rydning J (2017) *Data age 2025: the evolution of data to-life critical. International Data Corporation, Framingham*
4. Kaisler S, Armour F, Espinosa J (2013) *Big data: issues and challenges moving forward, Wailea, Maui, HI, s.n, pp 995–1004*
5. Bakshi, K.: *Considerations for Big Data(2012): Architecture and Approaches. In: Proceedings of the IEEE Aerospace Conference.*
6. Elgendy, N.(2013): *Big Data Analytics in Support of the Decision Making Process. MSc Thesis, German University in Cairo.*
7. EMC(2012): *Data Science and Big Data Analytics. In: EMC Education Services.*
8. Wikipedia :<https://www.wikipedia.org/>
9. *The Guardian* : <https://g.co/kgs/crmDBS>