# SYSFORE TECHNOLOGIES

SUCCEED WITH BIG DATA ON CLOUD

The current term that's making the rounds is **Big Data**. But what is big data, and how do we use it? Simply put, big data is data that, by virtue of its velocity, volume, or variety (the three Vs), cannot be easily stored or analyzed with traditional methods.

The term covers each and every piece of data your organization has stored till now. It includes all the data stored both on-premises or in the cloud. It could be papers, digital, structured and non-structured data within your company.

To put it another way, big data is data that doesn't fit well into a familiar analytic paradigm. It won't fit into the rows and columns of an Excel spreadsheet. It can't be analyzed with conventional analytical tools, and it is too big to fit on your normal computer's hard drive.

So, What exactly is Big Data?

Gartner defines Big Data as "high volume, velocity and variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making".
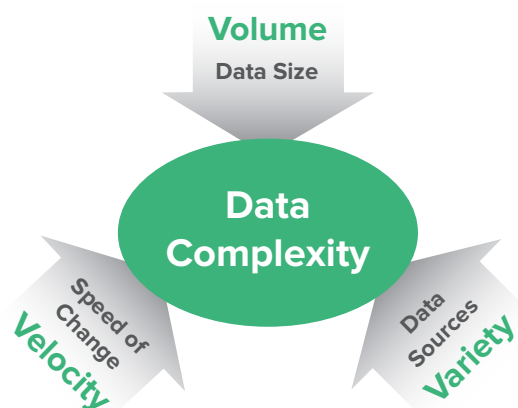
Big Data is the data – whether or not categorized – that is present in your servers. It can be used to analyze and predict different types of results using the same data. It is not needed to use all of this Big Data for all analysis. Different analysis uses different parts of the BIG DATA to produce the results and predictions necessary.

Volume

It refers to management of data storage. The tremendous amount of data that is generated every single minute is too big to handle in your computers. It is just too big to be stored using the conventional methods. This confirms the general observation that physical capacity and performance of computers double about every two years.

Velocity

It refers to the speed of data processing. There are two types of data that is used for data processing. The static data which once entered, doesn't change. The streaming data where there is a constant influx of data. This presents a special challenge for the analysis. It combines readily usable data with different analysis tools.

Variety

The datasets in the Big Data is present in both the structured and unstructured format. It includes the formatted data set in a spread sheet, hierarchical structures and documents. The unstructured data includes texts, blog posts, books, comments, video, images, and audio. It is estimated that 80 percent of enterprise data may be unstructured. It is this data variety that is pushing companies towards Big Data solutions.

## Using Big Data and Hadoop for analysis

As the name itself suggests, the enormous data cannot be manipulated and analyzed using the normal spreadsheets and regular tools of database management. They need special analysis tools like Hadoop, so that all the data can be analyzed at once. It may include multiple iterations of analysis.

Hadoop is an open source platform for consolidating, combining and understanding large-scale data in order to make better business decisions. Hadoop is the technology powering many (but not all) big data analytics infrastructures.

It formats or structures the unformatted data to make it suitable for data mining and subsequent analysis. Hadoop uses a distributed computing architecture, consisting of multiple servers using commodity hardware, making it relatively inexpensive to scale and support extremely large data stores.

There are 2 key parts to Hadoop:

➤ HDFS (Hadoop distributed file system) which lets you store data across multiple nodes.

➤ MapReduce which lets you process data in parallel across multiple nodes.
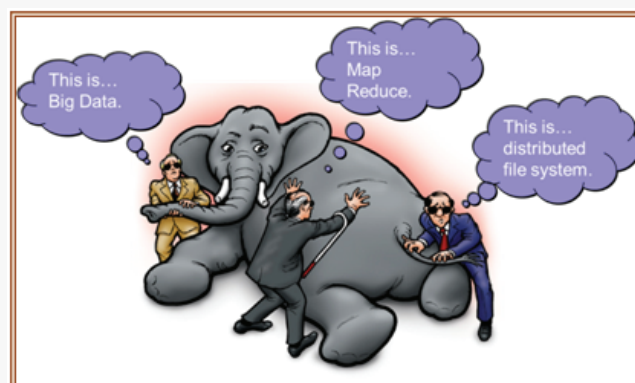
The idea behind MapReduce is that Hadoop can first map a large data set, and then perform a reduction on that content for specific results. A reduce function can be thought of as a kind of filter for raw data. The HDFS system then acts to distribute data across a network or migrate it as necessary.

The important characteristic is that you're able to **draw insights from large quantities of data,** independent of specific technologies.

How is Big Data being used?

Today, with Big Data, organizations can harness all their data, both inside and outside the enterprise, to enable smarter, more precise action and planning than ever before. The impact can be huge for revenues, customer satisfaction and brand awareness.

The Big Data is best understand in the context of where it is being used. It could be used for the consumers, businesses, scholarships and for research. It can be used for predictive marketing analysis for businesses.

This is when big data is used to help decide who the audience would be for something before they actually get there. This is trying to predict any number of events that are often associated with a whole series of commercial transactions.

About Sysfore:

Sysfore is a systems integrator (SI) specialized in building computing systems for enterprise clients using the best of cloud, mobile, and responsive web technologies. We serve a global client base, offering consulting, technology and managed services.

We are a team of about 100 people, passionate about using the best technology and tools to achieve results for our customers. We call Bangalore, India as Home. We have worked with customers from across the globe, successfully delivering quality work products and making a positive contribution to their businesses. Since the beginning, we've always been committed to a customer's success and go that extra mile towards achieving delight.

Sysfore is a recognized Microsoft Gold Partner and Amazon Web Services Consulting Partner delivering results for customers, through comprehensive consulting, deployment and cloud managed solutions.

We take pride in having over 70 global clients, done over a 100+ cloud consulting engagements, 80+ cloud migrations, 50+ POCs and deployed over 50 Storage and DR Solutions across multiple industries.

For more details or information, connect with us:

Email: info@sysfore.com

Call us : +91-80- 4110-5555

Website: www.sysfore.com