

The alternative to this problem is the big data approach. Many IT tools are available for big data projects. The storage requirements of big data are taken care of by Hadoop cluster. Apache Spark is capable

of stream processing (e.g., advertisement data). When used, these tools can dramatically reduce the time-to-value – in most cases from more than 2 years to less than 4 months. The benefit is that many speculative projects can be approved or abandoned based on the result.

Organizations whose data workloads are constant and predictable are better served by the traditional database, whereas organizations challenged by increasing data demands will want to take advantage of the Hadoop's scalable infrastructure. Scalability allows servers to be added on demand to accommodate the growing workloads.

There are hybrid systems, which integrate Hadoop platforms with traditional (relational) databases, that are gaining popularity as the cost-effective systems for organizations to leverage the benefits of both the platforms.

1.4.3 Advantage of “Big Data” Analytics

Big Data Analytics is advantageous in the following two ways when compared to the traditional analytical model:

1. Big Data Analytics uses a simple model that can be applied to volumes of data that would be too large for the traditional analytical environment. Researchers say that a simple algorithm with a large volume of data produces more accurate results than a sophisticated algorithm with a small volume of data. The algorithm by itself is not too great but its ability to apply it to huge amounts of data without compromising on performance gives the competitive edge.
2. Big Data Analytics has sophisticated model developed for it. Present day's database management system (DBMS) vendors provide analysis algorithms directly. Most companies go well beyond this and develop newer, more sophisticated statistical analysis models.

Advantages of using Hadoop over traditional storage systems such as relational DBMS (RDBMS) are as follows:

1. Scalability – nodes can be added to scale the system with little administration.
2. Unlike traditional RDBMS, no pre-processing is required before storing.
3. Any unstructured data such as text, images and videos can be stored.

4. There is no limit to how much data needs to be stored and for how long.
5. Protection against hardware failure – in case of any node failure, it is redirected to other nodes.

Multiple copies of the data are automatically stored.