

The right software can ensure 'Big Data' doesn't overwhelm your business

 By [Modeen Malick](#)

3 Dec 2018

Everyone is talking about 'Big Data', a critical element in the shift toward digitalisation. But what exactly is it, what does it mean for business and how can software assist organisations to overcome Big Data challenges so that they do not become overwhelmed by its sheer scale and complexity?



Modeen Malick is senior systems engineer for MESAT at Commvault

Big Data for business

As the business embarks on or evolves its digital transformation efforts, more and more technology come into play. The Internet of Things (IoT) with its sensors and smart devices, cloud-based applications, robotics, automation, and other emerging technologies, are all creating massive amounts of data over and above a business's traditional data.

Data increases and when analysed, it delivers invaluable business insights at a granular level. These findings help the business to define its next operational or sales models, transform processes, direct target markets and develop products and solutions which cater specifically to a market need. In other words, Big Data helps businesses to quantify their digital efforts.

How is it different from 'regular' data?

Big Data differs from traditional data in that it is based on a distributed cluster which brings supercomputing levels of computer power to the data and is resilient to local failures. It uncovers the ever-changing relationships between data from multiple individual data sources. Due to the compute power and resilience required to analyse and interpret Big Data, it is not very storage efficient.

Big Data sources include operational data coming from transactional systems, streaming data and sensors; 'dark data' that businesses already own but aren't actively using; commercial data; social media data; and public data which is available from government and independent open data initiatives on a multitude of topics which may be relevant to the business.

When Big Data projects go into production, they become similar to conventional applications in that they need service levels. This is where software comes into play.

Big Data challenges

The most obvious challenge businesses face with their Big Data projects is the sheer volume of data. Traditional data protection solutions that struggle to scan and process regular data fast enough will struggle even more to manage the quantities of Big Data at a fast-enough rate to be effective. The scale, complexity and completely varied nature of Big Data makes it tricky to manage.

Big Data is also relatively new somewhat immature. As a result, many native data tools require manual customisation and scripting, which can be a counterproductively long and tedious process.

While Big Data comes with standard resilience built in, it is unable to protect itself against logical and user errors. Hardware resilience does not automatically allow data to revert to an earlier point in time where it was logically consistent, therefore the potential for failure exists.

Finally, it can be complicated to manage Big Data in terms of capacity. Big Data arrives in big surges which businesses need to be prepared for and have capacity for.

Simplifying Big Data

Software can simplify Big Data management for businesses and eradicate most of these challenges. It can apply policies to newly created data that can run globally across a business's entire infrastructure, detecting any changes to the environment, making data management much simpler.

Key functions of software include indexing, collection and recovery which can integrate directly with an application via an Application Programming Interface (API), simplifying the process as it engages the Big Data name node to identify and understand each new data set.

Big Data management software also helps to manage performance. Backups take full advantage of the distributed architecture using load balancing. Once multiple media agents have been deployed, the software will intelligently use the fastest nodes and switch over to another one if one fails.

Intelligent Big Data

It's important to choose intelligent software which understands data structure and reacts accordingly, enabling fast data restoration when necessary. A fast and intelligent way to do restores with very large data sets is to provide native access to the data. This delivers immediate availability, avoiding the need to perform a restore. It also means limited workloads can be run immediately.

Software which integrates seamlessly with the cloud allows for the provisioning of both data and compute. This enables cloud-based data recovery and cloud-based development and testing operational use cases with pre-built workflows.

ABOUT MODEEN MALICK

Senior Systems Engineer at Commvault

- Containerisation tech is vital for businesses needing to do more with less - 6 Apr 2021
- Don't let your data management strategy get stuck in the past - 13 Nov 2019
- How the World Wide Web gave birth to IoT - 24 May 2019
- Securing your data across hybrid and multi-cloud environments - 12 Apr 2019
- The right software can ensure 'Big Data' doesn't overwhelm your business - 3 Dec 2018

[View my profile and articles...](#)

For more, visit: <https://www.bizcommunity.com>