

Is your Big Data properly protected?

In the past few years there has been much hype around Big Data and why organisations need to harness and analyse it in order to deliver business value and overall success. Big Data initiatives are on the rise as enterprises strive to increase efficiency, improve business decision making and attain that important competitive edge.

 By [Johan Scheepers](#) 28 Dec 2016



©Jorgen Ekstrand via [123RF](#)

Despite the importance of Big Data to business, however, it is often the case that systems put into place to protect it are completely inadequate. This is due to the complexity of designing and developing solutions to support such large data sets at a cost that fits in with IT budgets.

However, inadequate data protection on what is fast becoming a mission critical asset introduces unacceptable levels of risk. Organisations need to adopt data protection and management that is Big Data aware, so that automated disaster recovery and enhanced visibility can be applied unilaterally across all data sources and platforms.

In addition, solid protection and recovery solutions help businesses to fully realise the power of Big Data initiatives, further enabling increased efficiency and enhanced decision-making through advanced analytics and improved innovation. Here are five key considerations for businesses to ensure that their Big Data enjoys the same level of protection as all of their other enterprise data volumes.

1. Unstructured data is growing exponentially

Unstructured data from a variety of sources, including social media and videos, can contain valuable customer intelligence if it is effectively managed and protected. This data is growing at an exponential rate and analyst firm IDC estimates that by 2020 there will be as much as 44.1 zettabytes of unstructured data. Despite this growth and the understood importance of Big Data, multiple surveys show that the majority of organisations still place greater value on protecting their structured data sources, and many do not even have unstructured data on the radar.

In order to achieve effective information governance, it is essential that data management solutions can incorporate Big Data. This means that they need to easily accommodate large volumes of unstructured data from many sources to ensure the comprehensive data protection and recovery business requires.

2. Consider your integration options

Obviously the most economical approach to Big Data protection is to integrate it into existing recovery infrastructure. However, many systems are not capable of providing the required levels of visibility into leading Big Data tools such as Hadoop, Greenplum and GPFS. Effective Big Data awareness means that data protection solutions are able to help organisations map Big Data implementations and architectures. This in turn is essential in delivering the required insight to ensure protection and recovery as a whole, or across selected nodes, components, and/or data sets.

3. Big Data can be a security and compliance risk

Unstructured data is not just social media, but includes textual data as well. It comes from a plethora of different sources, including email, PowerPoint presentations, Word documents, collaboration software, instant messages and more. When these vast and ever-growing data sets are added to the volumes of structured and analytical data used to derive business insight, the complexity increases. The result is frequently a multi-structured mess that introduces risk across security and compliance.

In order to address this challenge, organisations need a sophisticated converged backup and archive process. This will assure Big Data recoverability as well as discovery, without adding complexity or heavy resource overheads.

4. Disaster recovery needs to be automated

Human error that results from manual disaster recovery processes introduces risk. This risk is multiplied when it comes to Big Data, thanks to the volumes of data and complexity involved as well as the many sources and repositories of Big Data and other organisational data. In order to ensure adequate protection organisations need an intelligent approach to protecting the infrastructure of Big Data initiatives.

This requires the ability to automate disaster recovery for these multi-node systems. In addition, organisations need to look for solutions that do not permit all users to have all access. A single user interface for setting rules and policies across any combination of physical, virtual, cloud and Big Data environments is essential. This ensures more efficient and reliable data management and disaster recovery.

5. Data portability is key

Big Data can originate from and reside in multiple infrastructure options, including cloud, on premises, in virtualised solutions or traditional data storage – or any combination of these. For the most effective and useful backup and recovery, data needs to be portable across infrastructure, regardless of its origin. This delivers the required levels of agility to avoid vendor lock in – essential to ensure organisations have the agility they need to leverage new innovations and developments with ease.

Big Data needs to be top of the data protection and recovery agenda

Big Data is immensely valuable for business intelligence, advanced analytics and technological innovation and its value will only continue to grow. Organisations need to ensure that their IT protection, compliance,

security and recovery reflects this value. However, Big Data does not override regular, structured business data. Organisations still need to ensure that their traditional disaster recovery plans and solutions are tested and effective.

Dealing with Big Data requires business to carefully examine overall data protection and recovery infrastructure to ensure readiness for Big Data as well as protection of regular data assets.

ABOUT JOHAN SCHEEPERS

Johan Scheepers is Commvault systems engineering director for MESAT
[View my profile and articles...](#)

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