

Simplifying Data Management and Security in the Cloud

Data and machine learning can transform business, but it takes the flexibility, scalability and security of an object-based cloud storage solution.

Oil and gas companies are sitting on data which could, McKinsey suggests, realize up to \$1bn in cost savings or production increases.

Data is both the biggest opportunity and greatest challenge facing CIOs and IT leaders. On the one hand, data is key to game-changing improvements in productivity, efficiency and performance, powered by machine learning and advanced analytics. On the other, the sheer quantities of data pose puzzling questions for management and security. How can organisations integrate and optimise it for these new applications? How can they manage and secure it in a way that safeguards the business and its customers – not to mention complies with incoming regulations like the EU’s General Data Protection Regulation (GDPR)?

The answer lies in the cloud and in object storage; a combination that provides the flexibility and scalability Machine Learning applications need, plus the resilience and security required to handle the most sensitive personal and business data.

Extracting value

The potential value of data has become apparent across all industries and sectors. In the oil and gas industry, firms are turning to their data assets to optimise their investments and identify which drilling sites to prioritise. Companies are also looking for ways to maximise the potential of existing facilities. Yet according to a 2015 report by McKinsey, of 30,000 data points collected from a typical oil rig, less than 1% are utilised. Oil and gas companies are sitting on data which could, McKinsey suggests, realize up to \$1bn in cost savings or production increases.

It’s a similar story in capital markets. The US Market Regulatory Organization, FINRA, monitors up to 75 billion market transactions every day, while Bloomberg sees more than 60 billion market ticks passing through its service daily. Combine that with unstructured data – social media posts, political activity, new scientific or medical research – and there’s a mass of data that investment firms could be using to achieve competitive advantage, yet only a few companies have the expertise and means to do so.

In healthcare, data can transform patient outcomes. It’s growing faster and more affordable to sequence a human genome – hours and a thousand dollars rather than the days and ten million of ten years ago. Analysis is a different story. The quantities of data are vast, with the New York Genome Center generating somewhere between 10 and 30 terabytes per day. It’s a challenge for storage, let alone analysis. The same goes for advances in digital medical imaging or the data captured through a patient’s life. MRI scans, X-rays, vaccinations, blood samples, vital signs and medical histories all contain information that could improve diagnosis and treatment, if only its value could be realised.

Machine Learning is the key that could unlock it, with Gartner predicting that AI will be a top investment priority for more than 30 percent of CIOs by 2020. In a Harvard Science Review article, Machine Learning: The Future of Healthcare, Puneet Gupta talks of how machine learning could be used to predict mortality and length of life remaining, or more

Data gravity is rapidly shifting to the cloud, with IoT, data providers and cloud-native applications leading the way.

significantly 'help physicians diagnose patients, especially in cases involving relatively rare diseases or when outcomes are difficult to predict.' It could be used to optimise medication or find points for concern in MRI scans. Applied to the human genome, it could assist the developing field of Precision Medicine, identifying patterns within high volume genetic data sets, then translating these into computer models that could predict an individual's probability of developing a certain disease.

In the oil and gas industries, McKinsey has reported on a North Sea Operator that piloted advanced analytics on a mature, semi-submersible production platform, using three years' worth of data collected from 5,000 sensors totalling hundreds of gigabytes. The work resulted in an expected production increase of up to 0.5%, with potential for additional 1 to 2% gains through identifying and mitigating further bottlenecks. Machine Learning could also play a vital role in enabling firms to accurately identify, purchase and optimise new on-shore and off-shore assets. In Capital Markets, Machine Learning could help with everything from fraud protection to anti-money laundering to market timing strategy.

What's holding companies back? Partly, it's a lack of in-house data science expertise; more than half the respondents in Gartner's 2017 AI development strategies survey indicated that a lack of necessary staff skills was the number one challenge in adopting AI. Much, however, comes down to the costs and challenges of storing and managing data. Simply working out what to store, where to store it and how long to store it for can be an obstacle for some enterprises. The costs can be astronomical, while sorting, cataloguing and structuring this data is far from a trivial task.

What's more, the growth of data protection regulation is a major concern for many businesses, and even more so in regulated industries looking after sensitive personal or financial data. Companies need to balance accessibility with resilience and security. There's no room for data siloes lying unmonitored.

Analytics meets Object Storage in the Cloud

The cloud can help with all these concerns. Machine Learning and AI applications are finding their natural home there, as platforms like the IBM Cloud Platform and Watson Data services enable developers and business users to pull together cognitive, predictive modelling and analytics applications on-demand, without high levels of in-house data science expertise.

Cloud also solves the storage problem, as platforms like IBM's Cloud Object Storage simplify data management and security. COS cleverly divides data objects into slices, with those slices dispersed across multiple devices and locations for resiliency and security. Slices can be stored across regions to maximise availability, or a single region – even a single datacentre – to suit regulatory or performance needs. Each slice is encrypted, and erasure coded, ensuring that the slices can only be reassembled and decrypted at the client's primary data centre. What's more, it's secured in transit by carrier grade TLS/SSL. All This happens transparently but ensures availability and security.

Dynamic, scalable and simple to manage, COS gives companies the capacity to store their data and the flexibility to make full use of it, with close integration with IBM Cloud and Watson data services plus support for REST APIs and SDKs.

It's a natural fit. Cloud provides the capacity, flexibility and scalability to store new data sources, plus the high-performance, scalable platform for machine language and analytics. This also matches current ideas of Data Gravity, where the data that needs to be correlated for analysis and the systems running that analysis inevitably move together. In the words of a Gartner report from October 2016, 'Data gravity is rapidly shifting to the cloud, with IoT, data providers and cloud-native applications leading the way. It is no longer a question of "if" for using cloud for data and analytics; it's "how." Together, COS, IBM Cloud and Watson data services has the answer.