Using data virtualisation to detect an insider breach


George Smyth

George Smyth, Rocket Software

The latest figures from Lloyd's of London indicate that a worldwide cyber-attack could result in losses of $53bn, with potential consequences akin to that of a natural disaster[1]. Some eye-watering sums have been racked up by recent crimes – for example the WannaCry attack cost $8bn globally while NotPetya caused $850m in damages.

When faced with scenarios like these, the response from global organisations is to invest significant effort and money in making data security walls thicker. For banks and insurance firms, the need for rigorous data security is particularly strong, given the serious ramifications of losing customers' sensitive financial information. However, an exclusive focus on the dangers menacing an organisation's data from the outside can divert attention from another area of significant risk – the insider threat.

Since 2013, more than seven billion data records have been exposed as a result of a breach.[2] It is estimated that 9% of these were at the hands of malicious insiders looking to steal data for financial gain or simply for revenge. A 2017 survey of 4,000 office workers in the UK, Germany, France and Italy confirmed the dangers that come from within a business – 29% of respondents reported having intentionally sent unauthorised data to third parties, while 14% admitted that they would consider selling their IT log-ins to a third party.[3]

The everyday threat

An insider threat can come in many forms, from a current or former employee to a contractor or third-party supplier. Essentially, anyone who has had access to confidential company information represents a risk and has the potential to be a perpetrator of a range of malicious crimes, including data theft; identity theft; monetary theft; or data corruption or deletion. Often, these insiders are policed by reductive or ineffective security measures (think password-only access), rather than the big guns that are deployed to fend off cyber-criminals.

"A PwC report highlights that inadvertent actions by employees were the number one cause of breaches in 2015. This often comes down to a lack of data protection training and unclear processes"

Insiders can have a level of knowledge of internal IT systems and data security processes that make them a far more serious threat than people consider them to be. There is also the case of accidental insider breaches, caused by employees as a result of human error. A PwC report highlights that inadvertent actions by employees were the number one cause of breaches in 2015.[4] This often comes down to a lack of data protection training and unclear processes in relation to new technologies, such as the cloud or BYOD programmes.

Recent cases demonstrate the damage that insiders can inflict on a financial services business, not just in terms of the bottom line but also on company reputation and customer satisfaction. Bupa is the latest insurance firm to find itself the victim of a malicious employee, with the worker copying and transferring information related to 547,000 international health insurance plan customers.[5] With the Ponemon Institute reporting that 54% of companies believe that it can take at least 10 months to rebuild the reputation of a business after a data breach, Bupa has its work cut out to restore customer trust.[6]

The problem with big data

Banks and insurers, like companies in most other sectors, are continuing to grapple with the data governance issues that have developed thanks to the rise of big data. Businesses operate using a wide array of data sources and it is simply not feasible to attempt to consolidate these in a physical location. Previously, data was once siloed across traditional analytics databases and enterprise data warehouses (EDWs). However, nowadays, businesses are turning to more modern solutions – such as Apache Hadoop and Spark, plus NoSQL – to support the processing, storage and retrieval of relational and non-relational data that is held in disparate computing environments.

With the task of managing big data becoming increasingly complex, it is also becoming increasingly difficult to detect when this data is being altered or accessed without authorisation. This is leaving the door wide open for insiders to exploit the complicated data environment for their own personal gain. In fact, Verizon's '2017 Data Breach Investigations Report' highlights that in 60% of all insider attacks, the person in question was planning on stealing the data with the intention of converting it to cash.[7]

Improve data access

The reliance on traditional operating protocols and technology is making it even easier for insiders to take advantage of lapses in data security. For example, system audit logs are used to record data access and data downloads, but they are typically not reviewed until the following day, or even week. This creates a large window of time for data to be stolen or compromised. In the case of an insider data breach, by the time the logs have been examined, the damage has already been done.

Solutions such as Extract, Transform, Load (ETL) are also adding to the difficulty of protecting against insider threats. Used on its own, the technology cannot deliver the accuracy and speed needed by financial services institutions to detect when their own employees are using or accessing data without authorisation. Part of this stems from the fact that ETL makes multiple copies of the data, with these copies ending up spread across the company's IT estate, with each copy being another potential point of data loss. At this point, it becomes almost impossible to discover if a copy of this data has been compromised. This means that any instances of tampering or unauthorised access after this point cannot be picked up by the infosecurity team.

The issues with using ETL also become apparent when viewed in relation to typical data security measures. Some organisations instigate the principle of least privilege (PoLP) to combat against an infiltration of their networks by a cybercriminal, but also to stop rogue insiders gaining access to and exploiting sensitive customer information. But this security protocol does nothing against the threat of an insider who has been given the 'privilege' to access certain data sets. This individual still has the means and opportunity to take advantage of his or her position for dishonest means. And, in reality, many businesses take the path of least resistance and grant employees full access to an entire database, rather than restricting their contact with sensitive information.

With ETL creating multiple copies of data sets, there are more points at which sensitive data can be exposed to those with the necessary privilege to access it; this data trail becomes incredibly difficult to keep track of. When you have multiple copies of multiple pieces of information, it becomes too easy to put data security measures aside and simply forget where this data is stored. Infosecurity teams may not even realise that there are storage systems where sensitive data is mixed in with the harmless data that is being used as part of everyday tasks.

Don't wait to view data

One solution for reducing the threat of insider attacks lies in data virtualisation. The technology removes the need to create copies of the data and ensures the access controls for the data are enforced and can be monitored. For companies using mainframes, which includes the majority of banks and insurance companies, data virtualisation can also allow live monitoring of audit logs (SMF records) to help prevent inappropriate activity within seconds rather than days. For businesses, the popularity of data virtualisation technology is growing. Forrester predicts the data virtualisation industry will be worth $6.7bn by 2021, which is a substantial increase from the $3.3bn valuation in 2015.[8]

The need for the technology is increasing, in large part, because of the reliance of banking and financial organisations on increasingly complex data management ecosystems, including the solutions mentioned previously – Hadoop, Spark and NoSQL – and the integration of these with existing on-premise databases and cloud-hosted systems. Instead of trying to group and store data together in one physical location – for example, a giant information warehouse – virtualised data is distributed across various disparate servers. Using data virtualisation, these information streams can all be brought together, avoiding the time-consuming processes inherent in ETL and providing a single point of data access across the company's IT estate.

Essentially, the technology eliminates the need to move data from one storage silo to another, allowing staff to view the information 'virtually', without creating multiple copies of data sets. Data virtualisation also means that information can be viewed in real time. Having the ability to scrutinise information on a second-by-second basis and validate data across various storage systems means that any anomalies can be identified almost instantly. Most importantly, the solution can build an accurate picture of who has access to data and when, so banks and insurance firms can recognise when insiders are acting maliciously.

To win the fight against insider threats, companies also rely heavily on analytics solutions. Data virtualisation supports more efficient, accurate analytics by ensuring that only the most relevant information to the task at hand is extracted from the IT system. This leaves behind the unnecessary data that would do nothing but slow down the analysis and allows banks and insurers to allocate more resources to scrutinising the data that will help them detect insider fraud.

Poor data access

With the reputational and financial damage that comes with suffering a breach by someone from within the organisation, banks and insurers must look more closely at their data storage infrastructure. Knowing who has access to sensitive information and when is a critical part of a full data security programme.

The temptation for personal gain can be too much to resist for some employees, while others may see the opportunity to steal data as a justified response after being fired. Whatever the case for the malicious activity, the excuse that an information security professional didn't spot the attack amidst the mountains of business data will not stack up in the face of investor and public backlash.