CASE STUDY

Breaking down a data pipeline monolith

Introducing a “pipelines as products” approach to improve security vulnerabilities

EQUAL EXPERTS

Having a robust approach to security vulnerabilities is essential for business safety. When you're a global financial services organisation, it's vital to ensure any security issue is spotted quickly and addressed in a fast and efficient way with reduced risk to the business.

Our client sought support improving the data pipelines within its security vulnerabilities team. The current system was a monolith, with one mechanism handling all incoming and outgoing data. It made deploying changes a risky and time-consuming process which often resulted in deployments being rolled back due to errors.

Equal Experts introduced the concept of individual pipelines as products to break down the monolith and separate elements into individual pipelines which could be quickly changed with reduced risk to the rest of the system.

The initial deployment of individual data pipelines marked the start of our client's journey in breaking down the monolith. The upskilling of the client's team in frameworks, data strategies, data pipeline design and quality assurance has also seen significant improvement in code quality with the number of deployments rolled back reducing from 80% to almost zero during the engagement.

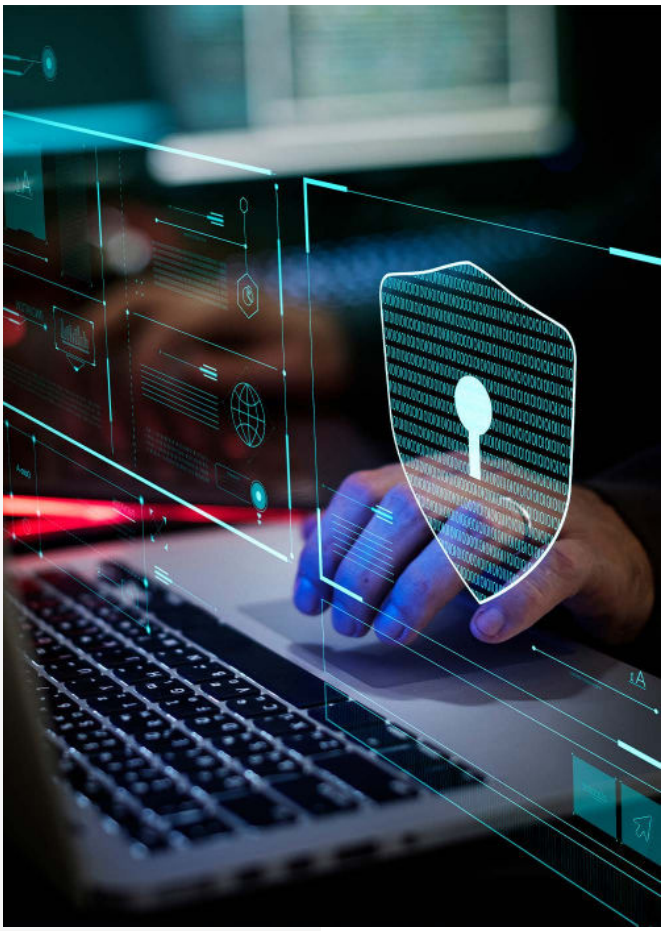
Plan for breakdown

of data monolith in place

80% to almost 0%

of deployments rolled back reduced during the first four months of engagement

About the client



Our client is a global financial services group operating in asset management, retail and business banking, wealth management, leasing and asset financing, market access, commodity trading, renewables development, specialist advisory, capital raising and principal investment. Headquartered and listed in Australia, the organisation employs more than 17,000 staff.

INDUSTRY	ORGANISATION SIZE	LOCATIONS	EQUAL EXPERT SERVICES
Financial	17,000 staff	Australia	Data

Challenge

Breaking down a security data pipelines monolith

The security vulnerabilities data pipeline is a vital piece of the organisation's security system, reporting on vulnerabilities throughout the entire infrastructure. It included scanning containers and devices with entry points to the network, providing data to numerous systems and generating reports about security vulnerabilities to be addressed.

It was also a monolith, with one central mechanism handling all the incoming data as well as the outgoing data sent to data consumers, such as Microsoft Power BI and others. This single mechanism meant that when the requirements for a data pipeline changed and updates were required, the process affected the entire monolith. Any changes were risky and took significant amounts of time to deploy, slowing down the organisation's ability to react quickly to security vulnerabilities. It also caused high risk to the organisation if it ever became unavailable, leaving the organisation unable to detect security vulnerabilities.

The complexity of the monolith meant the team struggled to deploy quality code quickly, with four in five deployments ultimately rolled back due to error. Addressing these issues and breaking down the monolith was a high priority for the client as it touched every aspect of the organisation.

Solution

Pipelines as products approach

EE proposed a “pipelines as products” approach which would treat pipelines as individual products rather than projects, allowing them to be managed individually and resulting in the breaking down of the monolith.

This would optimise the security vulnerabilities by making it simpler and quicker for changes to be deployed to individual pipelines when required. The approach would optimise security vulnerabilities by simplifying and expediting the deployment of changes to individual pipelines when necessary while minimizing the risk of impacting the entire system. It would also reduce the exposure in the event of failure, ensuring that the ability to detect and manage potential security threats remained due to limited dependencies.

To break down the monolith we began separating out pipelines and introduced Apache Airflow as an initial paved road to demonstrate the concept. This enabled us to create individual pipelines based on their specific purpose and requirements. The first Airflow DAG pipeline went into production in November 2022 with further pipelines deployed later in the engagement, helping our client to begin to reduce the size of the data pipelines monolith.

As part of the engagement, we also put fundamentals in place within the security vulnerabilities team in order to upskill them on how to design data pipelines and deliver quality deployments. This included putting in place a testing framework and quality standards that needed to be met throughout the development and delivery lifecycle. This enabled the team to work more effectively and efficiently as well instilling confidence in the quality of their code.

Results

Creation of a well-designed, extensible data platform

Utilising the pipelines as products approach and spending time upskilling the security vulnerabilities team resulted in the creation of a well-designed extensible data platform with quality, reliability and repeatability built into the design. This approach and upskilling of team members have empowered the client to continue breaking down the data pipelines monolith for themselves after our involvement ended.

The breakdown of the monolith and enhancement of the team's skills within testing frameworks and quality controls also improved the quality of deployed code. Whereas 80% of deployments previously had to be rolled back due to issues, this dropped to almost non within the first four months of the engagement. Embedding these core concepts and skills within the team will enable them to continue working in an efficient and effective way going forward. As a result, churn in the organisation also reduced, with the team more confident in their work and gaining a sense of control, instead of feeling stuck in a perpetual cycle of rolling back deployments and being unable to deliver changes quickly when required.

Want to know more?

Are you interested in this project?

Or do you have one just like it?

Get in touch. We'd love to tell you more about it.