The Benefits of Adopting Compliant Database DevOps at PASS

PASS

redgate

Executive Summary

In early 2018, PASS was in the middle of a journey that was long, demanding and sometimes confusing. Like many organizations, they were trying to plan and put into place the processes that would enable them to comply with the General Data Protection Regulation (GDPR), due to come into force on May 25.

Even though the GDPR is a European directive, it affects any organization which collects, processes, and stores the data of European citizens. Personally identifiable information has to be protected, access to it restricted, and organizations have to demonstrate they have made every effort to comply.

It has also become the catalyst which has prompted other, similar legislation to be proposed. The Stop Hacks and Improve Electronic Data Security (SHIELD) Act is scheduled to come into force in New York, and it will be joined by the California Consumer Privacy Act (CaCPA) in 2020.

This presented a challenge for PASS. A community of over 250,000 data professionals with almost 300 Local Groups around the world and 24 Virtual Groups, PASS has a huge database that is constantly being accessed and updated through their website. The IT team were already looking at how they could improve their process for deploying changes to the database, and meeting the new regulatory demands was adding another burden to their workload.

At the same time, as an association dedicated to data professionals, this meant that the solution had to be a benchmark that PASS members could follow.

This paper talks about PASS' journey with Redgate and shows that applying DevOps practices to the database like automation and smaller, more frequent releases also introduces, by default, compliance with the data privacy regulations many organizations now face.
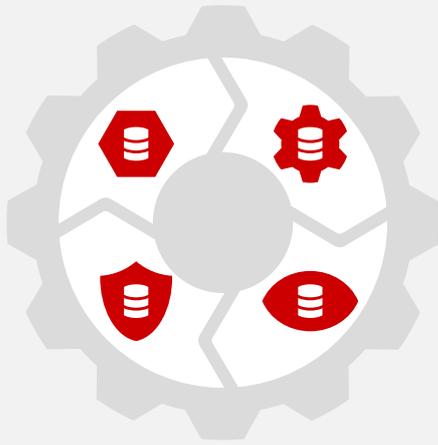
The Challenge

An independent, not-for-profit, user-run association, PASS was co-founded in 1999 by CA Technologies and Microsoft to promote and educate SQL Server users around the world. Since then, PASS has expanded and diversified their membership and are now a global community for data professionals who use the Microsoft data platform.

Their mission is to empower members to connect, share, and learn through networking, knowledge sharing, and peer-based learning. They encourage the exchange of information through local and virtual groups, online events, local and regional events, and international conferences; delivering high-quality technical content for in-depth education and professional development.

Central to these efforts is the PASS website which members use to access webinars, recordings and resources, book conferences and events, and to get involved with the community. The website is constantly evolving and a team of developers work behind the scenes to support it, add new features, and improve the experience for users.

Alongside the website itself, they are also responsible for maintaining and updating the SQL Server database which collects, processes, and stores the constant stream of data generated by member activity.

While the team were confident in their ability to enhance and improve the events, content, and learning opportunities PASS offers, there were two big concerns with the database.

Firstly, making changes to the database was a cumbersome process. Small ad hoc changes to items like stored procedures were tested locally, tested again in the user acceptance testing (UAT), and staging environments, and then reviewed one last time before deploying to production. At each step, hand-rolled scripts were used, which posed a risk that breaking changes could make it through if there was a mistake.

Secondly, the enforcement of the GDPR was fast approaching. The team were already familiar with meeting the requirements of legislation like Canada's Anti-Spam Law, which prevents the distribution of unsolicited or misleading emails, social media posts, and text messages. The reach of the GDPR, however, is far wider and deeper, and requires personal data to be protected in development environments with measures like pseudonymization.

Many development teams, however – including the one at PASS HQ – like to use up-to-date copies of the production database to test proposed changes against. Using a copy with a limited, anonymous dataset is possible, but inevitably means changes are tested with a database that is neither realistic, nor of a size where the impact on performance can be assessed. This results in the increased likelihood of breaking changes hitting production, which would further exacerbate the problems the team already faced.

Discussions on how to resolve both issues led to the conclusion that the database development process had to change.

The team needed to move away from the shared development model, with every developer accessing and updating the databases in the UAT and staging environments. Instead, each developer required their own copy of the production database to test changes against, with the ability to merge the changes later. Every copy, however, had to be sanitized of all personal data and any financial details. The hand-rolled method of creating deployment scripts also had to be replaced with a process that would protect data while minimizing potential errors.

The Planning

The IT team at PASS HQ wanted to go further than simply improving the database development process. They wanted to incorporate a DevOps approach by introducing an automated pipeline which would handle the merging of changes and generate the deployment scripts required, while maintaining compliance with regulations like the GDPR, SHIELD, and CaCPA.

Redgate's SQL Provision solution supports the database DevOps approach the team was looking for, while keeping compliance central to the process. It enables clones of databases, with a tiny footprint, to be spun up quickly, and can also mask those databases to bring production-like data into development and test, while ensuring sensitive data isn't exposed.

Working with Redgate, the team decided to equip every developer with Redgate's SQL Toolbelt and introduce SQL Provision so that they could all have their own fully masked copy of the production database. That way, they would have the tools they needed to bring DevOps to the database, while the personal data of PASS members would be protected in a demonstrable way throughout the process.

Using the new data discovery and classification tool which was introduced in SQL Server Management Studio 17.5, the columns in the PASS database which contained sensitive data were identified, classified, and labelled. This information was then used to build a masking script in SQL Provision, which was applied to the PASS database before an image was generated from which the clones are created. These clones can be spun up in seconds so that, even if developers test a potential change which breaks their own copy, another copy can be provisioned instantly.

With the data privacy concern addressed, the IT team then turned to the deployment process.

Firstly, all database changes were put into version control using SQL Source Control, which plugs into and integrates with Team Foundation Server (TFS) – Microsoft's Application Lifecycle Management (ALM) system – which was already being used to version control the application.

Secondly, the developers were encouraged to use SQL Compare and SQL Data Compare to save time finding the differences between database versions, fix errors, and push changes through the UAT and staging environments.

And thirdly, SQL Change Automation was used to generate the migration scripts to safely automate the deployment of changes to production.

Beyond these changes in how the database is now developed, big releases have been broken down into small steps, with updates and bug fixes now scheduled into releases every two weeks.
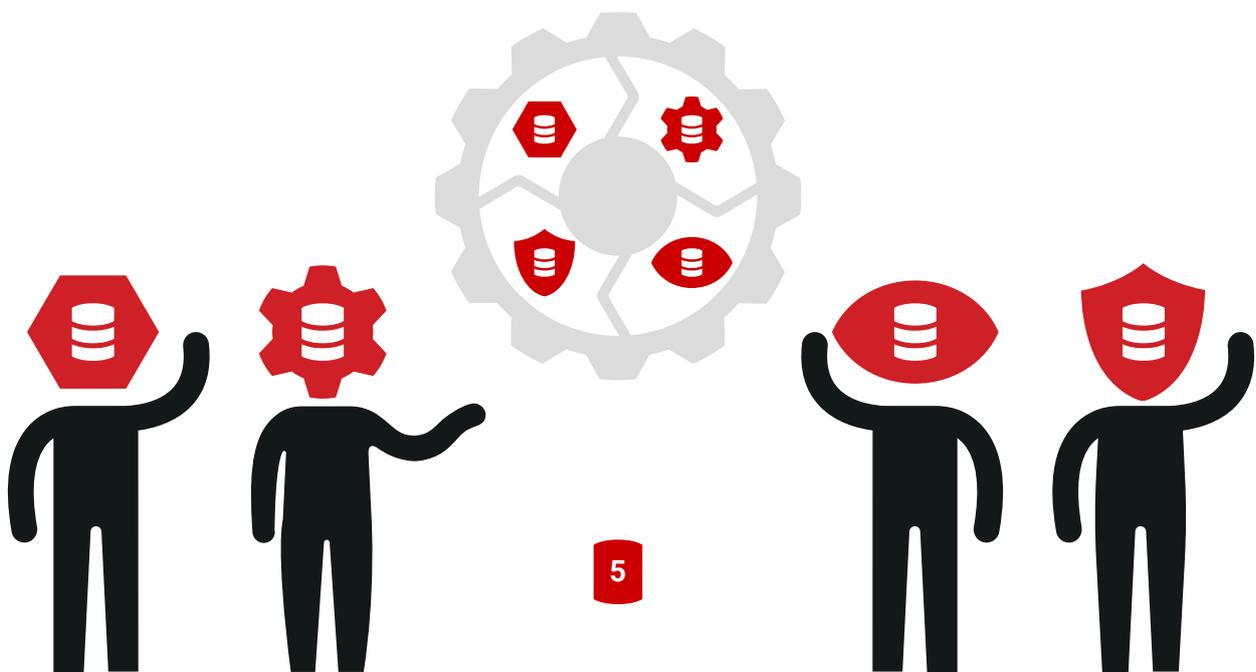
The Outcome

What started as a conversation about meeting data privacy concerns has resulted in a new database development and deployment process at PASS that is streamlined, trouble-free and, importantly, compliant with the GDPR and other legislation.

Developers can now test their changes against their own copies of the production database which, while masked, are fully representative of the real database and can be provisioned to them in seconds.

Version controlling the changes has stopped the merge conflicts that were happening previously, and the automated generation of migration scripts has transformed deployments from long, worrying evenings to a smooth, error-free process.

This new way of developing the database has inspired the team to look into how they can apply similar processes to the way they develop and deploy applications, in order to gain the same advantages. At PASS, rather fittingly, it's the database that's taking the lead when it comes to applying modern development processes.
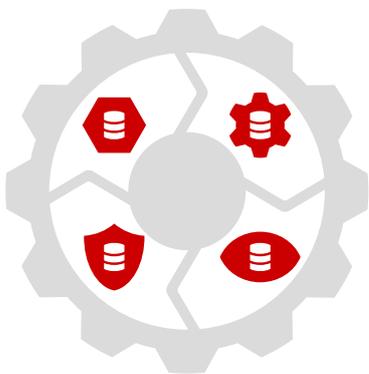
Compliant
Database DevOps

To find out more about the benefits of Compliant Database DevOps go to www.red-gate.com/DevOps