several9s

Database automation for Europe's largest research agency

CNRS

Industry: Research
Technologies: Galera, MariaDB, MySQL
Hosting: Cloud
Datacenters: Private
Products: ClusterControl

USE CASE

Database Automation for LAMP applications

WHY SEVERALNINES

Easy to use management tools,competence in database clustering

BACKGROUND

The CNRS – Centre National de la Recherche Scientifique (National Center for Scientific Research) is a public organization under the responsibility of the French Ministry of Higher Education and Research. As the largest fundamental research organization in Europe, CNRS carries out research in all fields of knowledge.

The CNRS currently counts over 33,000 employees of which 24,747 are CNRS tenured employees. It is organized as 10 institutes, 18 regional offices, ensuring decentralized direct management of laboratories as well as 1,100 research units (96 % are joint research laboratories with universities and industry).

The CNRS – DSI (Directorate of Information Systems) is the organization that defines and implements information systems for the running and management of the institution's various activities within scientific research as well as support processes including payroll, accounting, finance, and budgeting.

The DSI is responsible for all systems used by administrative and scientific directorates as well as laboratories of the CNRS.

CHALLENGE

CNRS – DSI operates over 100 LAMP applications and because of the simplicity and ease of use of MySQL, no dedicated DBA resources were needed to manage these database instances. They do have ORACLE databases as well in addition to MySQL, but those are managed separately.

MySQL is provided out-of-the-box by their Red Hat installation processes. However, over time, new needs became

apparent, which their infrastructure would struggle to support.

Historically, the LAMP applications supported relatively modest business requirements, that is to say, 8 am through 6 pm availability during the 5 days of a typical working week. The databases were relatively small in size, with only a few hundred simultaneous users.

SOLUTION

The team at CNRS – DSI decided to conduct a study to find out how they could:

- Painlessly migrate existing applications to a new database solution with the following constraints:
    - No application code modification and either null or most minimal database schema structure modification
- Provide a robust and highly available database solution
    - Service available even if a database server fails and no service shutdown while backup or patch operations
- Provide a heavy-load capable solution
    - Increase capacity in the background without stopping or even disturbing the actual service
- Provide a transparently scalable solution
    - No need for dedicated DBAs
- Provide a solution easy to manage and monitor

These technical criteria led the team to a choice between the following database cluster solutions… MariaDB Galera Cluster, Oracle MySQL Cluster, or ScaleDB.

Oracle MySQL Cluster was disregarded because of its licensing costs and because the master/slave setups were not evaluated as a robust high-availability solution. ScaleDB
was not considered because of its perceived lack of maturity. MariaDB Galera Cluster was the team's final choice because:

- Painless import existing of MySQL databases into MariaDB Cluster
- The high availability and scalability tests were successful
- MariaDB Cluster is open source (CNRS is a governmental research agency and the organization is highly sensitive to this aspect)

While a Galera Cluster is more complex to install and manage than a single MySQL node, the team found in Severalnines ClusterControl a comprehensive and powerful tool to manage and operate the cluster. Prior to going live, the tool was used to easily deploy test and QA environments, as well as to detect application incompatibilities with Galera. One functionality that was of particular interest was the ability to transparently manage load balancers through ClusterControl.

The team went live after the evaluation. The actual migration was straightforward and required no schema changes. At the time of writing (October 2014), the system has been running live for over a year without issues, which has brought the peace of mind that The CNRS/DSI now has a data layer that is highly available. No traffic is dropped due to a database limitation, even when subjected to extreme traffic conditions.

"We decided to invest into the Enterprise of ClusterControl and it's proven to have a very good ROI value. ClusterControl will be our companion in our future database technology roadmap," said Olivier Lenormand, Technical Manager, CNRS/DSI.

> *Things are changing at CNRS, as we start to offer more critical LAMP-based services to our customers": 24/7 applications with large databases, and tens of thousands of simultaneous users. It soon became clear that we needed a more robust and scalable solution than what*

Olivier Lenormand, Technical Manager CNRS

OUTCOME

With the positive experience of the first initial applications, the CNRS-DSI management decided that all LAMP applications must be compatible with Galera/MariaDB; all future LAMP applications will be running in this environment. Deployment automation, management, and monitoring framework will be provided by Severalnines.

> *In Severalnines we found a partner that is much more than a perfect database management tool provider: we have a partner that helps us define the architectures of our LAMP projects and leverage the capabilities of MariaDB Cluster.*

Olivier Lenormand, Technical Manager CNRS

Beside these classical LAMP applications, CNRS/DSI is deploying a cloud storage solution for thousands of its users (https://owncloud.org/). For obvious performance and availability reasons, MariaDB/Galera has been chosen as the database component in place of the classical standalone MySQL; and Severalnines ClusterControl has been naturally chosen as the management tool for this critical service.

The CNRS – DSI team could not have gone down the MariaDB Galera Cluster path without the proper tools to manage the cluster infrastructure. Additional manpower would not be available, so it was important to keep the operational overhead to a minimum and stay agile.