severalnines

Managed Hoster gains competitive edge through new cross-region multi-datacenter architecture

Customer Case Study: Sphinx Information Technology and ComplyWorks

December 2012

1. Sphinx Information Technology and ComplyWorks

Sphinx Information Technology is an integrated IT services and solutions provider helping clients working in over 90 countries worldwide realize the business value and efficiency of incorporating IT in their daily operations. The company provides a variety of services under two main business hubs:

- GuardianHELP
- GuardianDEV

One of the company's clients, ComplyWorks, is a global leader in Compliance Management Solutions for the Oil & Gas, Construction and Mining industries. Their solutions automate the collection and monitoring of relevant compliance data for job sites and regions, and provide that data to decision makers as actionable intelligence. Over 10,000 companies rely on their systems to securely exchange private business compliance data.

ComplyWorks At a Glance

Industry: Oil & Gas, Construction and Mining
Location: Calgary, Canada
Hosting Provider: Sphinx Information Technology, Canada
Use Case: Galera Cluster for MySQL
Why Severalnines: Deploying and managing database clusters across multi-datacenters

2. Challenge

ComplyWorks' experienced rapid growth of their customer base, which has doubled every year for the past 3 years. With over 10,000 companies using the system, the scale of the application was increasing dramatically. The datacenter was having to handle increasingly large data volumes. With users spread across different geographies, it was important to find a way to move operations closer to the user and thereby minimize network latency.

Another important aspect was to protect against the eventuality of a datacenter going down. Despite all precautions one might take, there are enough studies out there to show that the dream of a zero downtime datacenter is probably still a dream. ComplyWorks needed a belt and suspenders approach to system availability to eliminate all failure points. The application needed a cross-region multi datacenter setup.

'The database was the hardest problem – we needed it to be able to read and write from all datacenters, and handle any consistency problems. Maintaining consistency in the face of the distances and failures was not going to be trivial', says Ibrahim Hamouda, CEO of Sphinx Information Technology.

"The database was the hardest problem – we needed it to be able to read and write from all datacenters, and handle any consistency problems. Maintaining consistency in the face of the distances and failures was not going to be trivial"
- Ibrahim Hamouda, CEO of Sphinx Information Technology.

The operations team had experience with different MySQL Replication topologies, including master-slave replication with one updateable master, or circular replication with multiple updateable masters.

Although circular replication was a potential solution for a multi-datacenter setup, it was very fragile. Any problems with any of the nodes in the ring would cause the whole replication scheme to break. Also, since replication was asynchronous, serving writes from multiple locations might lead to conflicts (eventual consistency). Since data integrity could not be guaranteed, circular replication was not a realistic solution.

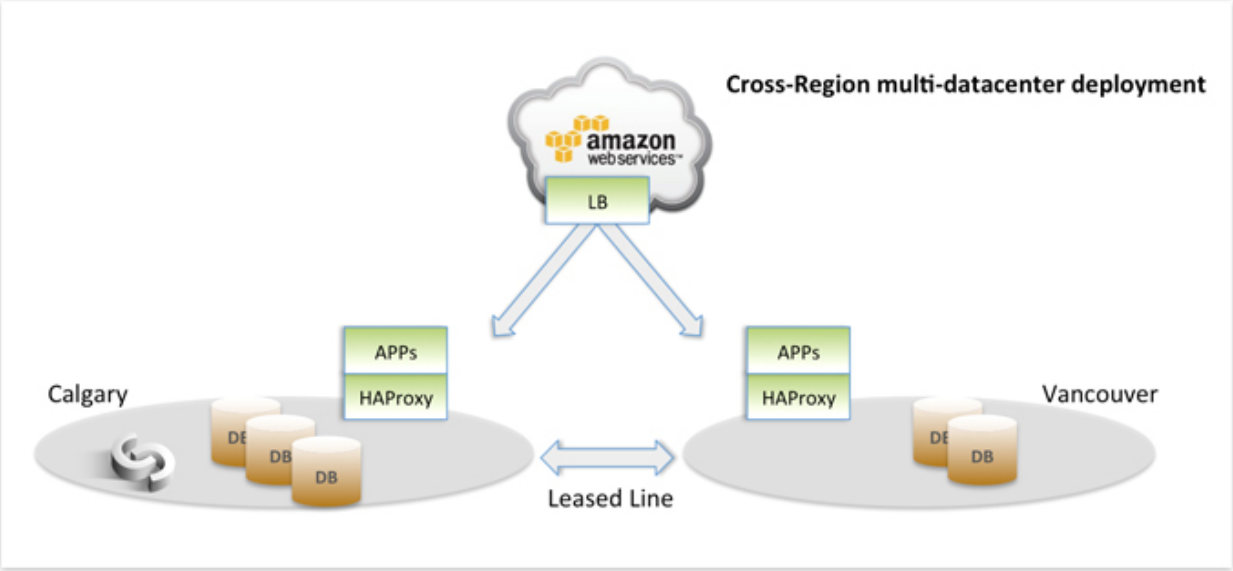
3. The Solution

The team ran evaluations on two multi-master solutions for MySQL:

- MySQL Cluster
- Galera Cluster for MySQL

Each solution were evaluated for its ability to span across at least 2 data centers and provide strong consistency of data. The datacenters were in Calgary and Vancouver, i.e., about 1000 km apart with a latency of approximately 20 milliseconds. This ensured different power and network backbones on the grid. A third datacenter located in Africa will be added during 2013.

Galera was chosen as it performed better in a cross-region multi-datacenter setup. In this case the certification-based replication was better adapted to WAN latencies than MySQL Cluster's two-phase commit protocol. Also, since Galera uses the InnoDB storage engine transparently, there was no need to spend time to optimize application performance by migrating their existing InnoDB tables and tune queries for the MySQL Cluster's NDB storage engine.



Using the Severalnines Configurator and ClusterControl, the operations team were quickly able to set up the different clusters. The team also leveraged a whole toolset from loading data, benchmarking and cluster management to speed up the evaluation.

"It would have taken us much longer without these tools, and thankfully, our ops team could instead concentrate on finding the right clustering solution for our requirements", says Ibrahim Hamouda. "A cross-region multi-datacenter database infrastructure translates into a powerful competitive advantage for our global clients who require geolocality and cannot afford any downtime. This is a game changer for us."

"A cross-region multi-datacenter database infrastructure translates into a powerful competitive advantage for our global clients who require geolocality and cannot afford any downtime. This is a game changer for us."
- Ibrahim Hamouda