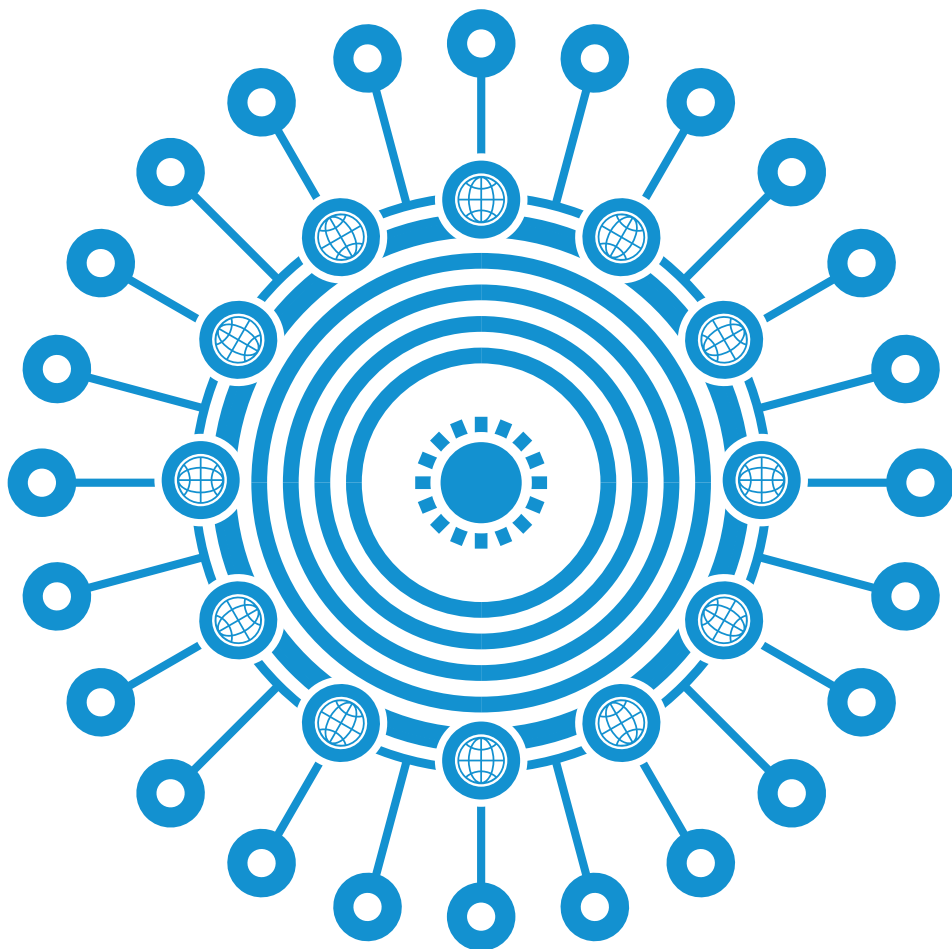


A Guide to Efficient Database Infrastructure Operations



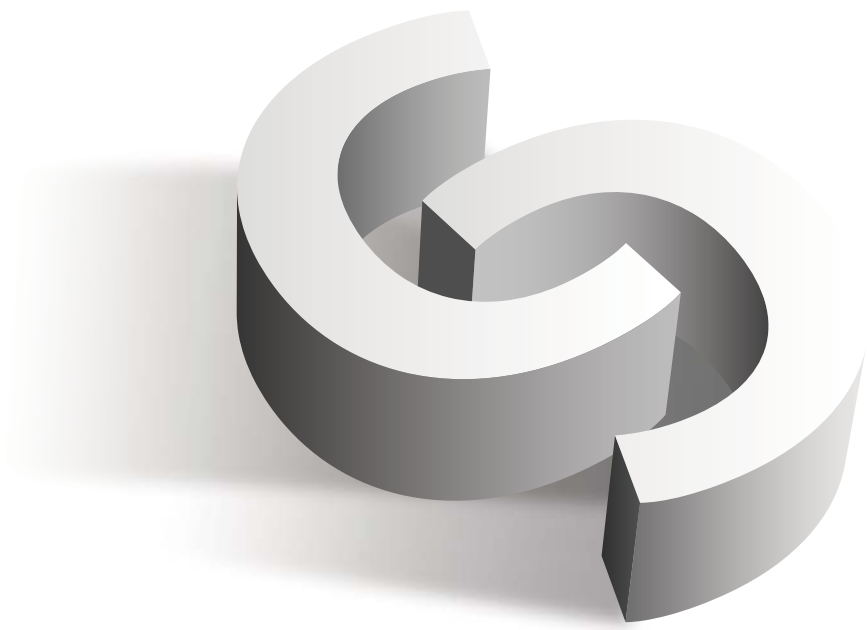




Table of Contents

1. Database Infrastructure Operations in the Modern Enterprise	4
2. The Impact of New Trends on Database Infrastructure	6
3. Why are Organisations Struggling with Their Database Operations?	7
4. Driving Down the Cost of Operations	9
5. Taking Control of Their Data – Every Company’s Number One Job	13
6. About Severalnines	17
7. Related Resources from Severalnines	18

Database infrastructure operations in the modern enterprise

The management and operation of database infrastructures is an area that companies often find themselves spending more on than they expect. At the moment, many database management tools and processes are inefficient, do not offer adequate functionalities and consume a large amount of resources. This manifests into escalating operating costs, project delays and database outages.

For the typical firm, database operations encompass a number of functions. These include the initial deployment of a solution as well as numerous management options, including configuration management, performance monitoring, SLA management, backups, patches, version upgrades and scaling. Development and test teams also rely on operational staff to clone production environments, either for testing, benchmarking, troubleshooting or migration purposes.

All of these are essential if enterprises are to have control over their database infrastructure and minimise the risk of falling victim to unnecessary downtime. However, it is sometimes the case that companies without significant expertise in database administration will only focus on some of these areas and take shortcuts with others, which can lead to problems.

System administrators are increasingly being asked to manage databases, especially in smaller businesses. But while a system administrator can get a database cluster up and running, database administration is a distinct role and skill-set. If the data is critical, the company management should make sure operational staff are trained and equipped with the appropriate tools. The average cost of downtime for a small to medium-sized business is calculated to be \$74,000 per hour¹ and with human error a common reason for outages, it's clearly an area businesses cannot afford to cut corners with.

Patching of databases is another area which is overlooked. McAfee has estimated more than 40 per cent of firms will not apply essential patches for at least three months after release, while 16 per cent are not aware of their patching frequency². The use of open source databases had historically been associated with small, non mission-critical use cases. However, recent evidence suggests that enterprises are now turning to open source databases to reduce database costs and avoid supplier lock-in.

McAfee has estimated more than 40 per cent of firms will not apply essential patches for at least three months after release.

The harsh reality is that while projects very often only think of the upfront costs of

¹ Quorum - Disaster Recovery Report Q1 2013

² McAfee - Overcome the Challenges of Database Patching in Production Environments

deploying a database, such as hardware and the necessary software licences, 80 per cent of total database costs is spent on ongoing operational expenditure³.



Many of the largest costs of running a database may not be immediately obvious on the surface

The impact of new trends on database infrastructure

Mobility, social tools and the rise of big data in particular mean enterprises are having to cope with more information than ever⁴. There are many reasons for this, such as increased penetration of superfast internet capabilities, greater adoption of mobile devices and more apps that automatically produce potentially useful data. It has been estimated that by 2016, the internet will be four times bigger than in 2012, with annual global IP traffic expected to hit 1.3 trillion gigabytes by that year⁵.

What this means for businesses is a greater need for large-scale database clusters that can react quickly to shifting demands and handle ever increasing volumes of data. For instance, retailers can experience large fluctuations in demand depending on current trends. In the peak selling period between Black Friday and Christmas, they may see their transactions increased up to ten-fold. As a result, database solutions that can quickly scale to deal with this elasticity are vital.

New technologies such as the cloud are also having a major impact on the way in which companies scale their database infrastructure. In the past, if firms needed to scale their databases, they would move to larger and more powerful hardware. However, scale up (or vertical scaling) is hard to do in cloud environments, because of the storage IO bottleneck that comes with large scale virtualization. Database clustering provides a means to horizontally scale out on a number of smaller cloud instances.

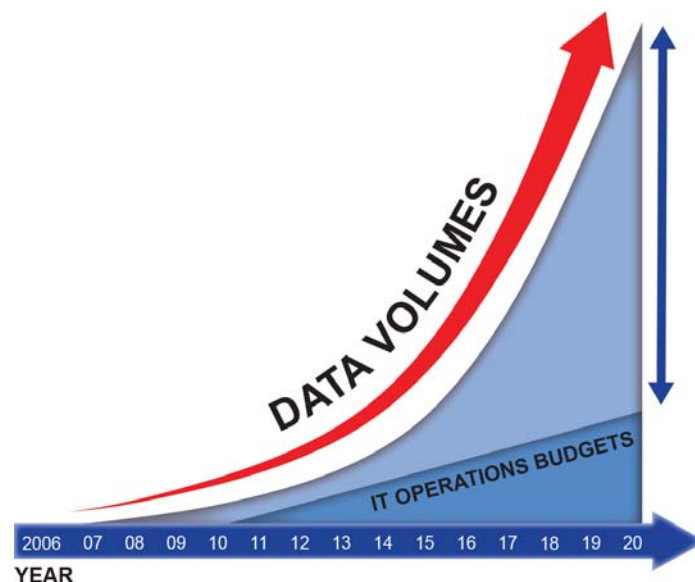
Another key trend is the increasingly globalised world, with businesses now expected to operate 24/7. This means traditional practices of planned overnight downtime is no longer adequate. Database clusters are required to offer a high level of redundancy and are expected to operate despite server failures or even if a whole data centre goes down.

⁴ **Gartner** - Gartner Says Actionable Analytics Will Be Driven by Mobile, Social and Big Data Forces in 2013 and Beyond

⁵ **Cisco** - Cisco's VNI Forecast Projects the Internet Will Be Four Times as Large in Four Years

Why are organisations struggling with their database operations?

IT Operations consumes a major piece of the operational budget of organisations, with Gartner estimating this at around 70 per cent for a typical company⁶. As data management becomes more pivotal to the overall IT infrastructure, it has also become important that organisations optimise the availability and performance of their database infrastructure. However, database systems are complex, complicated and difficult to manage. Maintaining and operating a database infrastructure requires skilled resources and investment, but the availability of these resources at an affordable price is an increasing problem.



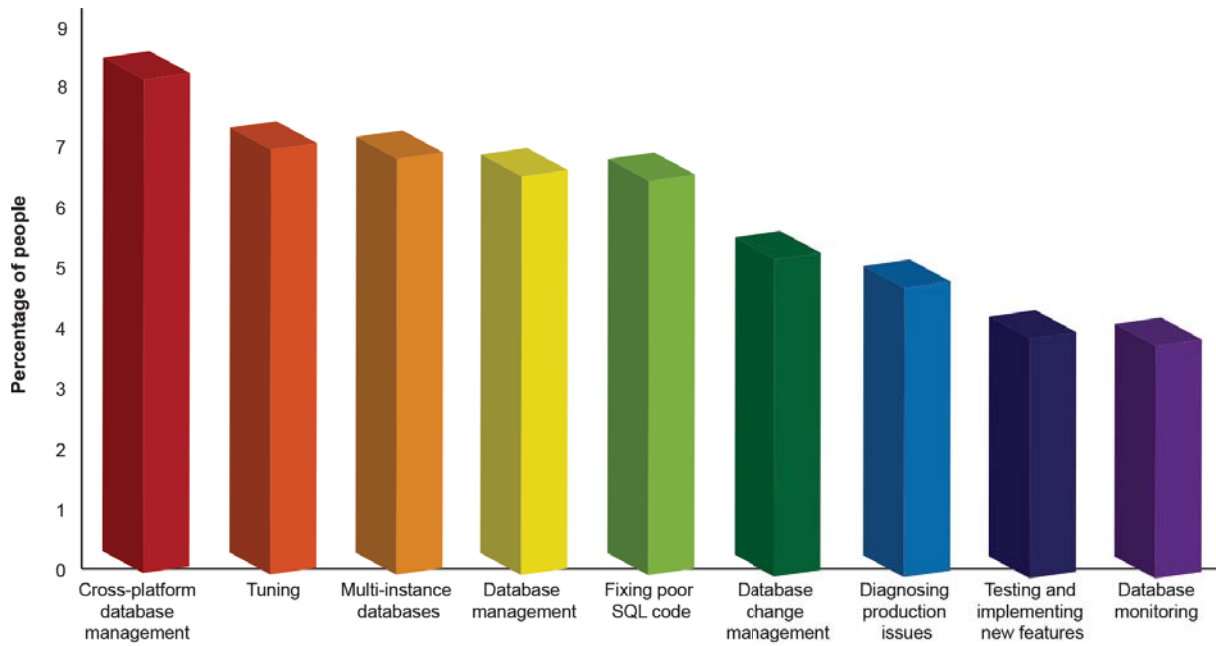
Even though data volumes are increasing exponentially, IT Operations budgets are remaining largely static, requiring firms to do more with less resources^{7,8}

As businesses continue to add database capacity at record rates and start adopting scale out database clustering, they are having to re-evaluate management solutions to identify opportunities to increase operational efficiency and cost savings.

⁶ Gartner - IT spending - How do you stack up?

⁷ IDC via the Economist - The Leaky Corporation (02/11)

⁸ Gartner - Gartner Worldwide IT Spending Forecast, Q4 2012



Challenges faced by firms

Database administrators have identified some of the key challenges they face when managing databases⁹.

⁹ Embarcadero - Database Trends Survey Report

Driving down the cost of operations

Tech-savvy businesses, regardless of size, are leading the data management revolution by demanding operational solutions that address the following requirements:

- **Automated Deployment**

If deployments are not automated, then they have to be done by hand. There are varying levels of automation, attempting to solve some or all of the problems with manual deployments. Deployments are time consuming, and a number of steps have to be repeated for development, test, user acceptance, staging and production environments.

Typing in a lot of manual commands can be error prone. Also, cloning a production database is done by the few people having access to the production cluster. The same few people are probably the only ones to know how any home-grown deployment scripts work. If these people are unavailable, the project may need to delay a release. In case of problems, it might be down to one person to apply the necessary fixes. Finally, with the number of databases in potentially different locations, it becomes hard to keep track of versions. An automated deployment tool can solve these problems.

- **A Complete Monitoring System**

In terms of monitoring infrastructure, there are numerous expensive enterprise monitoring solutions available, as well as a myriad of open source solutions – each addressing specific pain points. For instance, in the open-source space: Graphite and Cacti provide trending, Nagios provides alerting and Statsd and Collectd gather raw metrics. Integrating these systems is a daunting task. Identifying issues in production is one of the most tedious manual tasks DBAs can perform, with over a third wanting ways to automate this¹⁰.

This is an area where system admins spend time developing custom solutions. Often, such solutions morph into their own significant code base that must be maintained and improved on an ongoing basis. A complete monitoring system would provide all the required pieces to create a holistic view of the database cluster. Other features include real time data to know what is happening now, higher resolution of data for better accuracy and pattern recognition and Adaptive Baseline Alerting to report on emerging problems as deviations from normal behaviour.

Finally, a good tool recognises that a cluster system consists of multiple nodes - potentially of different node types - and automatically builds and updates a logical topology of the database cluster. This helps operational staff to monitor a cluster as a single entity, rather than individual servers, while also supporting drill-downs to individual nodes.

¹⁰ Embarcadero - Database Trends Survey Report

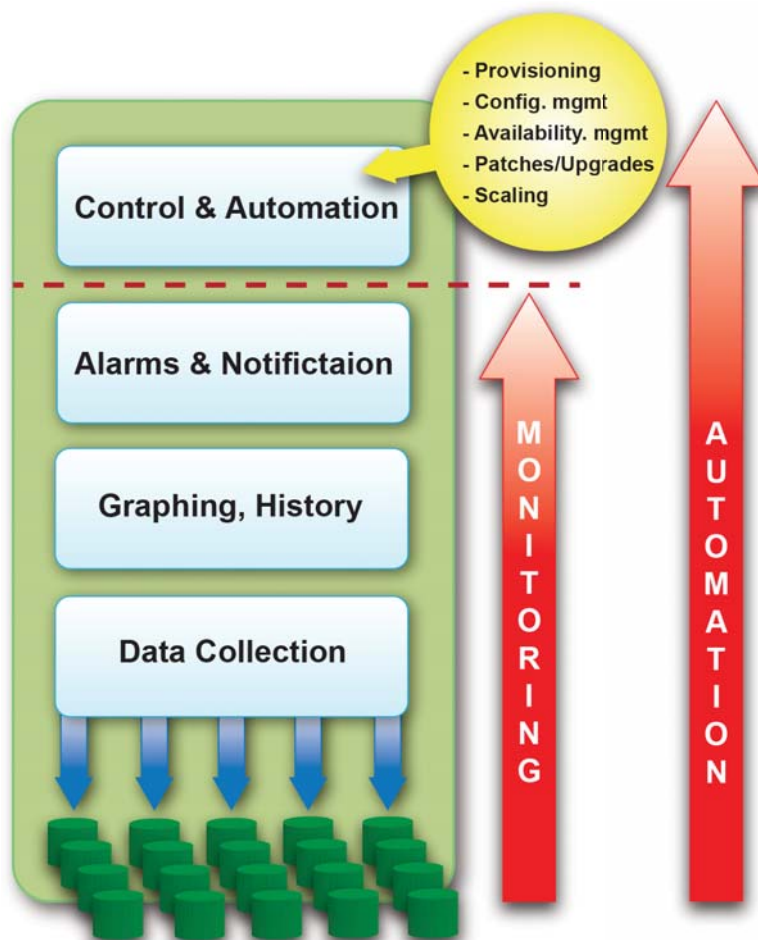
• Performance Management

Dealing with performance problems is usually one of the biggest post-deployment nightmares faced by database administrators. For instance, DBAs spend an inordinate amount of time combing through cumbersome SQL logs to find problems.

A web interface would provide status and transaction/query information of all nodes, as well as all resources (CPU, RAM, Disk, Network). A single, consolidated view of all the nodes is fundamental. It gives a quick overview of the overall health of multiple clusters, and tells which database nodes require the closest attention. This helps identify the location and source of issues and bottlenecks, and facilitates the user workflows that drive problem resolution.

• Integrated management and automation

Cluster architectures, by their inherent complexity, require more management and administration. Management operations include rolling out configuration changes across a cluster, patching, version upgrade or downgrade, backups, adding/removing nodes with subsequent reconfiguration of a running cluster. These procedures tend to be very manual. For instance, a configuration change requires an admin to manually edit configuration files and distribute them to all other cluster nodes, and to determine if rolling restarts are required. A rolling restart is a sensitive procedure in itself, and in case of problems during the procedure, the admin needs to know how to abort and rollback the changes.



Specialised automation tools offer a range of improvements over a basic monitoring solution.

Patches and upgrades is another area that require constant attention, as these have to be installed on a regular basis to protect systems from known stability problems or security holes. However, many firms are failing to do this. Evidence suggests companies are actually getting worse at patching databases, with the number of users applying patches in a timely fashion down by ten percentage points in 2012 compared with the previous year¹¹.

If you have thousands or even hundreds of databases and Oracle releases a patch every quarter, then it's almost impossible for you to go each and every quarter, for each and every database that you have and retest all your applications, actually apply the patch and absorb that downtime.

Slavik Markovich, vice-president and CTO of database security for McAfee¹².

A good management tool can boost DBA productivity and help ensure that operational staff are not taking short-cuts on the management of their production clusters.

• Easy Management

A well designed web-based interface can greatly increase usability. Manual processes increase the potential risk for human errors, whereas automated procedures are accurate, consistent and repeatable. A good tool also provides an integrated approach to operations management, otherwise operational staff will end up using different tools that only solve specific pain points. In addition, operational tasks on database clusters usually need to be done in an online fashion so as to not interrupt clients or applications accessing the database.

• Flexible and Scalable

The management platform must grow with the organization to gracefully manage different types of database clusters, potentially deployed on-premise or on different public clouds.



One pane of glass to manage all databases, regardless of location

^{11, 12} Independent Oracle Users Group survey 2012

- **Lower downtime**

Outages attributable to database infrastructure generally fall into one of three categories: planned maintenance, system errors, or human factors. Companies typically focus on the first two, but the last one, human factors, is also a major contributor, with individual errors being cited by database professionals as a root cause in over 50 per cent of cases¹¹. A tool helps reduce the exposure of the database infrastructure to human intervention and errors.

- **Lower cost of ownership**

Low cost applies not only to the initial purchase price but also the costs of installation, configuration, integration and ongoing management and administration. Operating a database costs four times the purchase price, and a tool that simplifies ongoing operations can lower labour and operational costs, resulting in significant savings for the business.

Taking control of their data – every company's number one job

Some of the customers that have benefited from Severalnines' ClusterControl product include:

- AVG Technologies
- Mail and Guardian Online
- Ping Identity
- SCI Solutions
- Sphinx Information Technologies
- Technicolor
- Telekom Malaysia Research & Development
- Vertical Systems

Sphinx Information Technologies

Sphinx Information Technologies is a managed hosting provider focused on the oil and gas industry. The company is headquartered in Calgary, Canada's energy capital. Sphinx acts as a full-fledged business partner to its customers to create global state of the art IT infrastructures in an industry that is increasingly becoming data-driven.

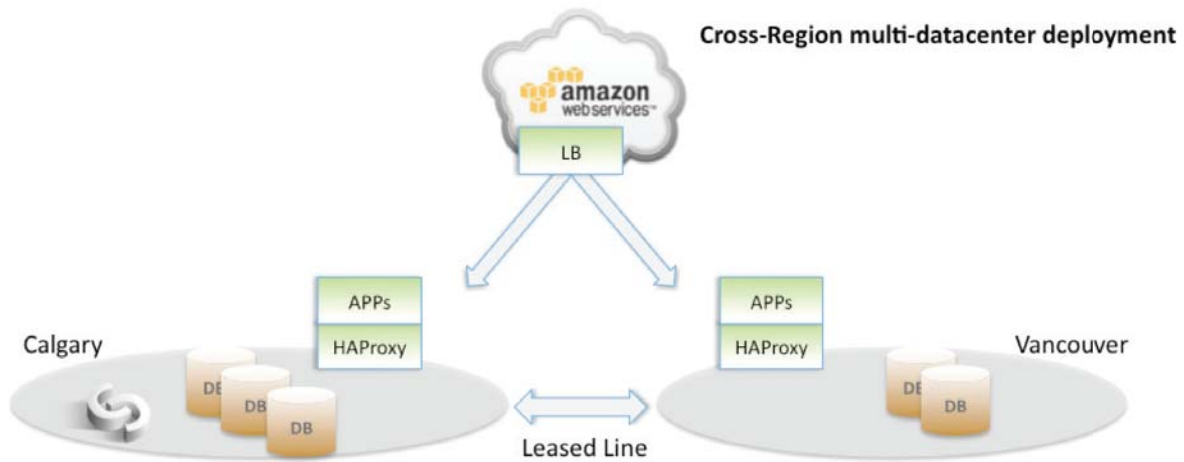


One of the company's clients, ComplyWorks, a global leader in Compliance Management Solutions for the oil and gas, construction and mining industries, was experiencing rapid growth of their customer base. With over 10,000 companies using their Software-as-a-Service (SaaS) system, the scale of the application was increasing dramatically.

Sphinx therefore implemented a cross-region multi-data centre architecture to scale the system across multiple data centres while ensuring business continuity in the eventuality of a data centre failure.

Severalnines allowed us to deploy our database cluster in 15 minutes, across two data centres with two nodes in each facility. We were able to use the management and monitoring software very easily to manage the cluster. A scaling operation to add a node to increase capacity takes us less than five minutes.

Ibrahim Hamouda, chief executive officer at Sphinx Information Technologies



Ping Identity

For Ping Identity, a company providing cloud identity and access management to 45 of the Fortune 100 companies, downtime is not an option. The company, headquartered in Denver, Colorado, runs their SaaS-based PingOne single sign on service in a hybrid cloud architecture.



Ping deployed a clustered MySQL setup across two internal data centres and one public cloud and their services were unaffected despite losing a data centre to hurricane Sandy¹³.

ClusterControl presents value to our operations team in several different ways, one of which is performance analysis. When first implementing Galera clusters, we found a number of performance bottlenecks which we quickly found and resolved as a direct result of ClusterControl. One of those being 'Health Report' which provides insight into cache hit ratios, percentage of max connections, open file limits, table lock contention along with other valuable metrics. Other benefits include disaster recovery (backups), cluster configuration management, central logging and node management.

Michael Ward, site reliability engineer specialising
in high availability solutions and system architecture, PingOne

Mail & Guardian

Mail & Guardian, one of South Africa's major news publishers, was looking for a way to reduce the cost of operations.

Managing the number of databases was already a major pain for the IT operations team and could potentially inhibit the business from growing.



¹³ Ping Identity - Why PingOne's heart doesn't skip a beat: MySQL Galera

The team moved to a clustered architecture, and used ClusterControl as the backbone of their database operations. By using Rackspace Cloud services, they could also quickly add capacity to their database cluster to meet sudden increases in traffic.

Nearly all of our tedious database maintenance work has been automated or made redundant. This has freed up hundreds of hours a month that can be used for development, other vital work and research. It's also meant we can now serve twice as much traffic without breaking a sweat, and that we are able to monitor our cluster with 100 per cent confidence that what we're seeing reflects reality.

Alistair Fairweather, general manager of digital, Mail & Guardian

Vertical Systems

Vertical Systems, one of the longest established and largest suppliers of technology to the UK travel industry, was also looking at open source databases to reduce database costs.

Historically a Microsoft SQL Server house, they had been slowly switching over to MySQL. The company had no MySQL DBA, but could quickly set up a MySQL Cluster using Severalnines.



The other benefits with the Severalnines solution is the continued monitoring it provides. You have a dashboard where you can get an overview of the health of your MySQL systems. A quick glance can show you query monitoring and all that. It is a detailed bit of software for monitoring your databases 24/7 to find any potential issues. It is like having your very own DBA [database administrator] 24/7.

Peter Bentley, network manager at Vertical Systems

AVG Technologies

AVG (NYSE: AVG) is a global leader in the security software market, with the mission is to simplify, optimise and secure the internet experience. The company has grown its user base to 143 million active users as of September 30th, 2012 and offers a product portfolio that targets the consumer and small business markets and includes Internet security, PC performance optimization, online backup, mobile security, identity protection and family safety software.



At AVG.com, we achieved a dramatic improvement in the database infrastructure for our online users repository. With the help of Severalnines products, we implemented a database cluster in record time, and could easily increase capacity by 20x and performance by 10x. We have been very impressed by the ease of use of the Severalnines products as well as the smart people in their support team. They are a big part of our operational peace of mind.

Václav Adamec, Systems Engineering, AVG Technologies

Read more about the benefits seen by some of Severalnines' customers on our website <http://severalnines.com/customers>

About Severalnines

Severalnines provides automation and management software for database clusters. We help companies deploy their databases in any environment, and manage all operational aspects to achieve high-scale availability.

Severalnines' products are used by developers and administrators of all skills levels to provide the full 'deploy, manage, monitor, scale' database cycle, thus freeing them from the complexity and learning curves that are typically associated with highly available database clusters. The company has enabled over 8,000 deployments to date via its popular ClusterControl solution. Currently counting BT, Orange, Cisco, CNRS, Technicolour, AVG, Ping Identity and Paytrail as customers. Severalnines is a private company headquartered in Stockholm, Sweden with offices in Singapore and Tokyo, Japan. To see who is using Severalnines today visit, <http://severalnines.com/customers>.



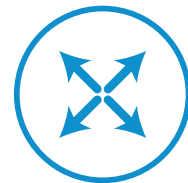
Deploy



Manage



Monitor



Scale

Related Resources from Severalnines

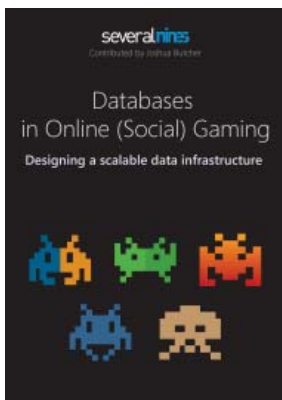
Whitepapers



MySQL Replication for High Availability

This tutorial covers information about MySQL Replication, with information about the latest features introduced in 5.6 and 5.7. There is also a more hands-on, practical section on how to quickly deploy and manage a replication setup using ClusterControl.

[Download here](#)



Databases in Online (Social) Gaming

This paper discusses the importance of databases for the gaming industry, what its requirements are in terms of database technology as well as a discussion on why MySQL is or should be the database of choice for anyone wanting to develop online social games that are reliable and stable in all their aspects.

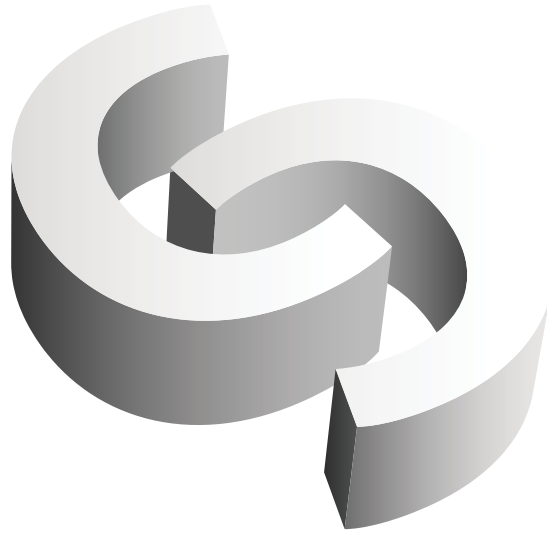
[Download here](#)



Database Sharding with MySQL Fabric

Why do we shard? How does sharding work? What are the different ways I can shard my database? This whitepaper goes through some of the theory behind sharding. It also discusses three different tools which are designed to help users shard their MySQL databases. And last but not least, it shows you how to set up a sharded MySQL setup based on MySQL Fabric and ProxySQL.

[Download here](#)



Deploy



Manage



Monitor



Scale