



A Guide to High Availability Clustering

How MySQL[®] Supports 99.999% Availability



A MySQL[®] Business White Paper

June, 2004

Table of Contents

Executive Summary	3
Introduction to MySQL Cluster	3
Advantages of MySQL Cluster	4
Architecture of MySQL Cluster.....	4
High Availability Reduces Costly Downtime.....	5
High Performance Only a Main Memory Database Can Deliver	5
Automatic Fast Failover Without Custom Programming.....	6
Incrementally Scale Applications to Meet Capacity	6
Hot Backup and Restore without System Interruption	6
Significantly Reduce the Cost of Downtime.....	6
Reduce Hardware Costs.....	7
Lower Maintenance Costs.....	7
Easy to Use Administration	7
Services & Support.....	7
Conclusion.....	7
About MySQL	8

Executive Summary

More and more organizations are relying on online and electronic means to grow their business, improve efficiencies, and cut costs. This increased reliance on Internet and networked technologies to meet business goals is forcing organizations to deliver 'always on' availability to maximize revenues and satisfy customer expectations. Downtime for such applications can mean an immediate loss in revenues. For example,

- **Telecom providers and network operators** are required to provide non-stop, always available service to an increasing number of users to stay competitive. In addition, telecom providers must extend their services to include additional content and video.
- **Financial institutions** must continue to accommodate the increasing volume of users and time-sensitive transactions without sacrificing real-time performance requirements.
- **Public web sites and company portals** that provide self service applications over the Internet must be able to accommodate the growing online community and the exponential growth in data that needs be delivered.
- **E-commerce companies** and others that are growing their online sales channel must accommodate the growing audience of online shoppers.

For these and other organizations who cannot afford downtime, a high availability database solution is a central and critical component to their success. Without a high availability database they compromise customer satisfaction and their revenue stream. At the same time, budgetary constraints are forcing organizations to meet these demands in a more cost-effective manner than can be achieved by using a proprietary database. Proprietary databases typically require high annual software licensing costs, in addition to large up front hardware and software investments. And traditional solutions to high-availability are expensive to develop and maintain due to their complexity, making them cost-prohibitive for many organizations.

This white paper describes how MySQL Cluster is able to cost-effectively meet the need for high availability database requirements. This white paper will also describe the unique architecture and features that MySQL Cluster implements to achieve 99.999% availability.

Introduction to MySQL Cluster

MySQL is an open source database with over 4 million installations. MySQL is proven for its ease of use and reliability and has been adopted by some of the largest organizations worldwide include Sabre, Yahoo!, and Lufthansa. One of the primary reasons for MySQL's widespread adoption is its ability to reduce database TCO. On average, companies using MySQL are experiencing per-project savings of between \$250,000 and \$500,000. Larger projects, or enterprise deployments, are yielding savings of millions of dollars.

MySQL Cluster is the high-availability version of the world's most popular open source database which provides "5 nines" availability. MySQL Cluster was developed in close cooperation with one of the largest telecom equipment vendors in the world to meet the high availability, throughput, and response-time requirements of telecom providers and network operators. MySQL Cluster implements a distributed, highly fault tolerant architecture with no single point of failure. In essence, MySQL Cluster delivers mainframe reliability on commodity hardware making it the most cost-effective solution for high availability. Using MySQL Cluster organizations can deliver the scalability, reliability and performance to run the most demanding applications in the world.

Advantages of MySQL Cluster

MySQL Cluster's unique parallel fault tolerant architecture provides the following benefits:

- **99.999% Availability** – provided by a fault tolerant architecture
- **Cost Effective** – requiring less hardware, lower maintenance costs, and affordable licensing compared to proprietary databases.
- **High Performance** – that only an in-memory database can provide.
- **Linear Scalability** – to incrementally scale your system without a high initial hardware investment.
- **Easy to Administer** –reducing your need to hire additional database administrators.
- **No Single Point of Failure** – using a distributed node-based architecture with fast failover
- **Hot Backups** – to back up to system without interruption
- **Automatic Fast Failover** - enabling systems to automatically fail over in less than a second.

Architecture of MySQL Cluster

In order to achieve fault tolerance a number of distributed components to run the application, access the database, and manage the distributed environment are required. These components are typically distributed across machines and continuously share the latest changes to the database with one another. If a node or machine fails, there are other nodes that contain the same information so an application can continue to run without interruption. The system must also ensure that transactions are successfully executed and database information remains consistent throughout the recovery process.

MySQL Cluster is a flexible distributed database management system that consists of three kinds of nodes which replicate data, access the database, as well as manage the system.

- **Database Nodes** are the main nodes of the system. All data is replicated and stored on the database nodes and the database handles all database transactions. If a database node fails, then there is always another database node storing the same information.
- **Application Nodes** are the applications connecting to the database. If a database node fails then there is always another database node storing the same information allowing the application node to continue to run.
- **Management Server Nodes** handle the system configuration. The management server nodes are used only at startup and for system reconfiguration. Management server nodes can be stopped and restarted without affecting the ongoing execution of the database and application nodes.

This architecture ensures there is no single point of failure. Applications continue to run and data remains consistent, even if any one of the database, application, or management server nodes fail.

In addition, nodes and computers can be distributed across geographies. This is important to recover a database or an entire cluster in case of a physical disaster in a particular location.



Plus, MySQL Cluster delivers a flexible distributed architecture and is highly configurable. For example, organizations can choose how many database nodes they want to run and how they want to distribute them across computers and geographic locations. This gives them the control to implement the proper level of performance, reliability, scalability and fault tolerance to match their application requirements.

High Availability Reduces Costly Downtime

A recent IDC study found that downtime is the largest contributor to database TCO¹. In order to eliminate costly downtime, database management systems must provide a distributed fault tolerant architecture that replicates data to other machines or across regions so they can continue to run without interruption in the event of hardware, network or application failure.

A recent IDC study found that downtime is the largest contributor to database TCO¹

IDC
2002

A unique parallel server architecture enables MySQL cluster to achieve 99.999% availability. This translates into less than 5 minutes downtime per year including regularly scheduled maintenance operations. In this architecture, each node has its own memory and storage disk and the database is distributed across several machines within the cluster. Data is synchronously replicated across nodes without having to use an externalized shared disk. If one or more database nodes fail the application can fail-over to another node, in order to complete the transaction successfully.

Should all nodes fail due to hardware faults for example, MySQL Cluster ensures an entire system can be safely recovered in a consistent state. This is achieved by replaying the log to get the database up-to-date. The log is stored on the files system for each database node and stores all database operations (inserts, deletes, updates, etc). A local checkpoint protocol is used to create a transaction consistent snapshot of all node data to disk.

Furthermore, in case of disaster (fire, earthquake, flood), MySQL Cluster enables entire clusters to be replicated across geographies, to other sites to ensure system availability. To recover from system failures, MySQL Cluster uses a checkpointing protocol to commit all transactions to disk ensuring success.

High Performance Only a Main Memory Database Can Deliver

In addition to high availability, high-performance is a critical requirement to accommodate the high volumes of database requests and transactions, particularly in the financial services and telecom industries. Telecom/IP database performance requirements are in the range of 1000's to 10,000's of requests per user. Plus, these performance requirements typically need to be met while maintaining high availability.

MySQL Cluster is a main-memory database designed to meet these performance requirements. MySQL Cluster is implemented using a Shared Nothing Architecture, so each database node has its own disk and memory storage rather than requiring a shared disk which adds to the expense and can limit performance in proprietary solutions. MySQL Cluster keeps all data in memory for transaction and fast failover purposes and limits Input/Output bottlenecks by asynchronously writing transactions logs to disk. Therefore, typical response times for MySQL Cluster are in the range of a few milliseconds. Plus, this architecture enables MySQL to handle tens of thousands of distributed transactions per second that are also replicated across the database nodes.

¹ IDC, Maximizing the Business Value of Enterprise Database Applications on a Unix Platform. 2002.

Automatic Fast Failover without Custom Programming

MySQL Cluster provides failure detection and extremely fast failover time with sub-second response to ensure applications remain highly available in the event of a failure. This is critical in banking and airline reservation systems that are extremely time sensitive. MySQL Cluster uses synchronous replication to propagate transaction information to all the appropriate database nodes so applications can automatically fail over to another node extremely quickly. When failed nodes are detected they are automatically restarted using a node recovery protocol which provides the restarting nodes with the necessary data from the surviving nodes.

Proprietary clustered database systems typically provide failover capabilities through a shared disk architecture. In order for a failed node to recover, it must recreate the database by reading and replaying the log files. This is a time consuming process that typically takes several seconds. Plus, the shared disk architecture is cost prohibitive and adds a single point of failure. A shared disk typically adds a minimum of \$30k to the cost of a clustered database solution. The advantage of synchronous replication is that it eliminates the time consuming operation of recreating and replaying log files for applications to fail over successfully. Plus, MySQL Cluster database nodes are able to automatically restart, recover, and dynamically reconfigure themselves automatically in less than a second without having to write the recovery logic into the application.

Incrementally Scale Applications to Meet Capacity

The ability to scale in a near linear fashion is necessary to cost-effectively grow your system. MySQL's parallel server architecture enables MySQL Cluster to scale in a near linear fashion. This gives organizations the flexibility to start small and make incremental investments to increase capacity as their needs grow. This eliminates the need for a large initial hardware and software investment in what tends to be an over powered configuration. In addition, adding machines (nodes) to a system can be done while in operation, without any downtime.

Hot Backup and Restore without System Interruption

MySQL Cluster enables organizations to make online backups and restore backups without having to interrupt or take down the system saving administration time and resources. For commercial applications such as ecommerce or customer service, this means you can provide a higher level of service and prevent catastrophic failure from interrupting business.

Significantly Reduce the Cost of Downtime

MySQL Cluster not only lowers up-front license costs with affordable commercial licensing under a dual licensing mechanism, but also significantly reduces system downtime. Furthermore, a highly portable standards-based environment allows you to cost-effectively distribute your applications across new or existing low cost hardware and operating system infrastructures, including Linux and Windows.

Reduce Hardware Costs

By being able to efficiently use low-cost commodity hardware, MySQL Cluster requires significantly less investments in hardware than proprietary clustering solutions. Other clustering solutions must rely on expensive multi-way machines.

In addition, by implementing a parallel server architecture processing power is distributed and shared by all nodes within a cluster. No standby hardware is required, so you get full advantage of all the horsepower you have invested in.

Also, the shared nothing architecture of MySQL Cluster eliminates the cost of a SAN (Storage Area Network), a centralized data storage location which requires dedicated hardware and introduces a single point of failure. Furthermore, to eliminate the single point of failure, this shared disk architecture requires even more hardware for the SAN to effectively failover.

Lower Maintenance Costs

MySQL Cluster is designed to be largely self-governing so very few system parameters need fine-tuning, further reducing the risk of costly errors. As a result, there are typically fewer conflicts with other software and hardware, and less need for manual intervention. This also means that MySQL Cluster will have a much lower maintenance costs, since fewer database administrators are required.

Easy to Use Administration

An easy-to-use management command-line interface makes it easy to monitor and maintain the database. It is easy to configure the server locally or remotely using management clients that connect to the management servers. Plus, a management API makes it possible to write programs that automatically monitor and maintain the database.

Services & Support

MySQL AB provides extensive consulting, training and technical support services to ensure the success of any organizations database application project. Our services can be tailored to meet your specific needs and requirements so you can implement cost-effective enterprise database applications with confidence. MySQL has a proven track record gained through thousands of successful customer deployments that can lower your risk and maximize return on investment.

Conclusion

High availability is a core requirement for all organizations that depend on real-time online enterprise systems to sell their products, service their customers, and do business with partners. For them, system downtime immediately translates into lost revenue and increased costs such as lost orders and increase call center load.



Until recently, commercial solutions for high availability have been too and too expensive for many companies. Using MySQL Cluster, organizations can make their core IT systems highly available in a cost-effective manner. The combination of affordable pricing and the ability to run on commodity hardware enables companies to withstand hardware, network, application, and system failures so they can continue to deliver their business critical services.

About MySQL

MySQL AB develops and markets a family of high performance, affordable database servers and tools. The company's flagship product is MySQL, the world's most popular open source database with more than 5 million active installations. Many of the world's largest organizations, including Yahoo!, Sabre Holdings, Cox Communications, The Associated Press and NASA, are realizing significant cost savings by using MySQL to power Web sites, business-critical enterprise applications and packaged software. MySQL AB is a second generation open source company, with dual licensing that supports open source values and methodology in a profitable, sustainable business. For more information about MySQL, please go to www.mysql.com.