MySQL Cluster for Location & Presence-Enabled Services

Delivering Enriched Mobile Services

A MySQL Business Whitepaper

Table of Contents

Page #

1. Executive Summary

Communications Service Providers (CSPs) are crossing an important threshold with data revenues now significantly contributing to their bottom line. New devices, such as 3G “smartphones” are capable of delivering a host of rich applications that drive additional sources of revenue, and commonly include Location and Presence Enabled Services. Traveling caffeine junkies would be mortified not to have "find the nearest coffee shop" at their 3G fingertips! Even those that traditionally have been the slowest to adopt new technologies are now showing off Facebook pictures of their travels and family from their mobile devices.

IP (Internet Protocol) based multimedia services are now recognized as the future of the mobile communications industry. To survive and thrive in today’s hyper-competitive telecommunications marketplace, operators are aggressively migrating towards an IP based services architecture, capable of supporting multimedia and service integration across a variety of devices. IMS (IP Multimedia Subsystem) is a major enabler of the transition towards IP convergence, allowing operators to cost-effectively introduce new services faster in order to satisfy rapidly changing subscriber demands.

IMS is the chief enabler for Location-Based Services and Presence Management Services. Both of these capabilities are rapidly emerging as a means to both enrich existing offerings via new “Value Added Services” (VAS) and to enable a whole new class of applications. These services offer the potential for operators to not just grow revenues from existing subscribers, but to also differentiate themselves from traditional and emerging (i.e. Internet-based) competitors in the on-going battle to attract and retain new generations of customers.

The increasing adoption of Location-Based Services and Presence Management is creating huge volumes of data within the telecommunications network which must be efficiently processed and managed to ensure rich service functionality and a seamless user experience, both of which are necessary in order to monetize such services.

> *“The database needs to act as the repository for a single list of contacts that are presence enabled, handle the multiples identities of users in the different systems and synchronize the data from phones, PDAs, or PCs...we have successfully implemented MySQL Cluster Carrier Grade Edition for our highly mission critical XDMS application which will enable the next generation of converged services ”.*
>
> **François Leygues, Software Architect, Alcatel Lucent**

The purpose of this whitepaper is to explore the different types of applications that can take advantage of location and presence awareness, and to consider the unique requirements for data management. We will then present a solution based on MySQL database technologies that allow operators to fully leverage the market opportunity presented by these new applications.

2. The Market Landscape for Location & Presence-Enabled Services

Whether offered separately or in combination, location and presence-enabled services allow CSPs to deliver a new generation of highly compelling, highly personalized mobile communications services to their subscribers. They also enable CSPs to forge new business relationships and funding models with partners such as content companies and advertisers who can use location and presence data to reach their customers in new ways.

The market opportunity for such services is growing dramatically with ABI Research estimating revenues generated by Location Based Services (LBS) reaching $12bn by 2013.

CSPs do enjoy many unique advantages that enable them to firmly entrench themselves within the value chain of service delivery. These advantages include:

- owning and / or operating the network itself
- leveraging existing close relationships with their customer base
- the amount of data collected on each subscriber through the delivery of multi-play services.

Whether the CSP wants to deliver services themselves or partner with developers or Over The Top (OTT) companies who need access to the underlying network, AAA (Authentication, Authorization & Accounting) and billing infrastructure, location and presence-enabled services represents a significant market opportunity for the CSP community.

2.1. Location Based Services

Through location awareness, wireless providers have a distinct source of value which differentiates their mobile services from applications accessed from wired internet connections, i.e. via PCs. Using mobile devices, location-based services (LBS) leverage a user's physical location to provide enhanced services and experiences.

The adoption of Location Based Services (LBS) has been driven by three dynamics:

- government mandates, initially in the United States but now extending to Europe, which require CSPs to provide location data to assist the timely response of emergency services
- the availability of ubiquitous and plentiful bandwidth which has transformed the use of subscribers' mobile devices from a simple communications tool to full lifestyle-enabler for business and personal relationships.
- Smart-phone & location-enhanced devices

The first LBS were introduced in Japan during the year 2000, and relied on triangulation from mobile cells to calculate a user's location. GPS within mobile handsets is becoming more popular with Gartner forecasting that 29% of all new handsets sold in 2009 will be GPS-enabled.[1] The adoption of Bluetooth, Wi-Fi and RFID within positioning system technologies is further increasing network coverage for LBS.

The variety of LBS applications is growing, with the most common forms including:

- ***Navigation.*** Forrester estimates that one third of North American consumers owns a navigation device[2]. The majority are standalone devices, however the mobile handset has become the fastest growing platform, with 6x greater adoption rates than other devices.
- ***Local Search.*** Very often combined with navigation services, users can search for specific services or businesses in their locality, and then receive directions to them.
- ***Social (and/or professional) networking.*** Users can find the location of their contacts (assuming appropriate permissions and opt-ins have been granted) on a map, or receive alerts when they are in close physical proximity. With two thirds of the Internet population visiting social networking or blogging sites, accounting for 10% of all Internet time, and with the growing propensity of users to access such sites from their mobile device (i.e. 23% in the UK, 19% in the US)[3], using LBS to enrich social networking via mobile devices represents a huge market opportunity for CSPs in data traffic alone.

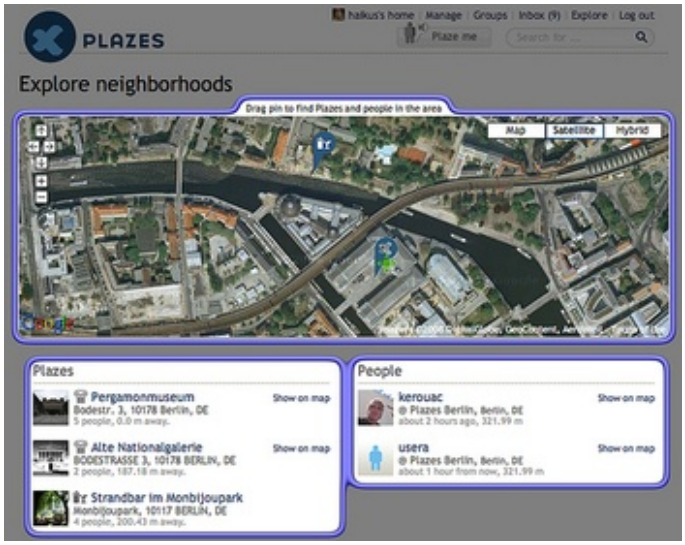


Figure 1: ***LBS Revenues are Forecast to Reach $12bn by 2013***

- ***Asset tracking.*** Extending the capabilities of RFID, LBS has witnesses major commercial adoption for the tracking of physical corporate assets, so far mainly focused on parcels and vehicles. Again, when combined with navigation capabilities and other data-driven services such as real-time analytics, applications can be delivered that allow route optimization based on local, real-time traffic updates.
- ***Mobile Advertising and Mobile Commerce.*** With the potential to enable operators to create new business models such as advertising-funded services, mobile commerce presents enormous opportunity to CSPs. Working together, operators and advertisers can create far more targeted, personalized and relevant campaigns by combining location data with demographic and historic behavioral information, leading to much higher click-through rates. Coupons can also be delivered, typically using SMS (Short Message Service), enabling users to redeem them in local businesses.
- ***Gaming.*** LBS is at the heart of new gaming experiences such as treasure hunts and pursuit games that aim to bridge the divide between the virtual and real world, blending the best of both technologies.
- ***Location-Based Media.*** Adopted increasingly by the tourist and events industries, LBS can be used to enable the delivery of on-line guides and commentary for specific points of interest or sporting events.

[1] http://www.ft.com/cms/s/3/1df00558-c193-11de-b86b-00144feab49a.html?nclick_check=1
[2] Phone Based Navigation Picks Up, Forrester, 2009
[3] Motorola Research

- ***Billing.*** Applications like Home Zone Billing – charging users differentiated rates for calls, depending on their location - is enabling more mobile operators to compete with traditional fixed-line vendors.

2.2. Presence Management

Presence is a dynamic profile of status information for people, applications or devices that is visible to others on the network. It indicates both the ability and the willingness for a user to receive communications from another contact. Presence management has been popularized by Instant Messaging (IM) and Voice over IP (VoIP) applications, and is being used more frequently in a range of services, some of which include:

- ***Unified Communications.*** Presence status can be extended across multiple devices, including a mobile handset, PC, IM client, etc. Unified Communications solutions can then intelligently route communications to the channel and device that best serves the user's state or preferences at a particular moment in time.
- ***Social networking.*** Users can communicate with friends and colleagues in real-time, from their preferred social networking platform. As discussed in the LBS section above, the accessing of social networking tools from mobile devices is growing dramatically.
- ***Collaborative workgroups.*** In commercial environments, co-workers can instantly determine who is available across the network for ad-hoc meetings or urgent questions, without the need to rely on telephony and email. This allows teams to communicate more freely and efficiently, especially those that are geographically distributed as it stimulates planned as well as spontaneous communication between remote end-users
- ***Push To Talk.*** When complemented by presence management, the full value of PTT in delivering faster and lower cost communications can finally be realized.
- ***Content Sharing.*** By adding presence management to existing content distribution and sharing applications, users can start to extend social networking platforms into real time collaborative experiences.
- ***Call Completion Solutions & Call Ring Back Services.*** Around 40% of all mobile communications are initially unsuccessful as a result of called parties being unavailable (i.e. device turned off; in a meeting; taking another call; in areas of poor network coverage). Using presence management, it is possible for callers to be notified once their contact becomes available again, or for the called party to subscribe to specific ring-back tones when they have messages waiting in their in-box. Such capabilities enable seamless mobile communications, while also ensuring operators maximize revenues by enabling a higher ratio of successful calls, and via Value-Added Service subscriptions.

2.3. Enabling Location & Presence Based Services

The examples above are by no means exhaustive, but are designed to provide an indication of some of the most common applications for LBS and Presence Management. Of course, delivering services that are location and presence-aware do not guarantee an immediate return to CSPs.

At a commercial level, privacy concerns and pricing have to be addressed. At a technologies level, there are many components from the hardware to network and software infrastructure that need to be developed and integrated to deliver a total solution. APIs (Application Programming Interfaces) providing access to both the networks and Service Delivery Platforms (SDPs) have to be developed and openly published if CSPs want to attract a vibrant eco-system of developers and content providers to release offerings for an operator's subscribers.

One of the most critical elements to the technology stack is the strategy employed for processing and managing data used and generated by these next generation applications.

3. Data Management Demands of Location & Presence-Enabled Services

For an increasing number of next generation communication services, the data - and access to it - not only enables the service itself, but also forms a major component of the service's revenue model. With the emergence of IMS-based standards, new generations of converged services are dependent on having fast and constantly available access to data. Defining and implementing the appropriate technology to satisfy the very unique data management demands of location and presence-enabled applications is therefore critical to their success.

To provide an example of how data and service logic combines to deliver a compelling experience to the user and revenue to the operator, effective data management enables a user's on-line presence and location to be updated and processed in real-time while reflecting the subscribers preferences, contacts and privacy permissions. These updates

can trigger related applications such as the receipt of marketing coupons for local restaurants, coupled with automated scheduling of lunch appointments with contacts, and then logging all of this activity in CDRs (Call Detail Records) for billing and business analysis.

The diagram below shows both the usage of data and its requirements for efficient processing and management within location and presence-enables services.

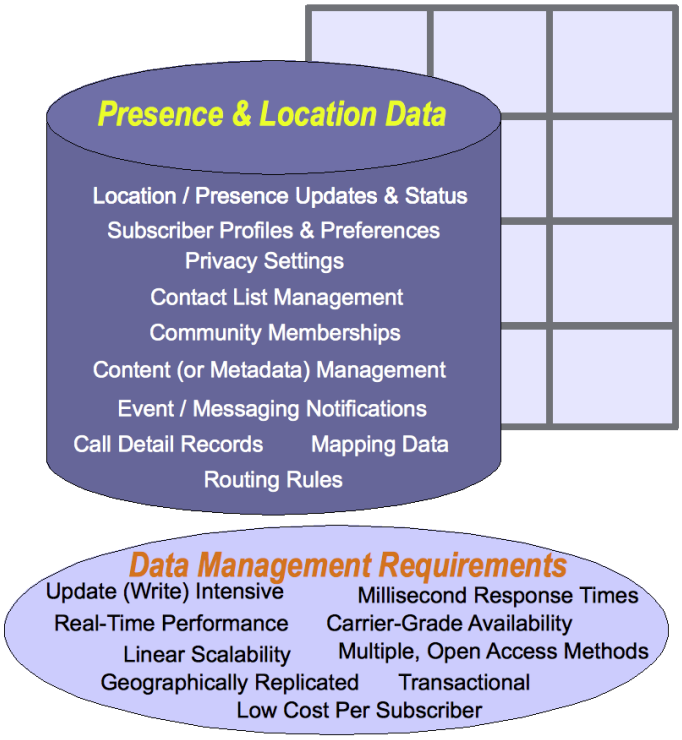


Figure 2: ***Data Management Requirements for Presence & Location-Based Services***

As illustrated above, effective data management solutions for location and presence-enabled services are characterized by a unique set of requirements. A database designed for these types of in-network applications must support a real time, live data experience. It must be able to couple the advantages of relational databases with the responsiveness and uptime demanded by communications services, all while minimizing the cost per subscriber.

3.1. Meeting Subscribers Expectations for Performance

From a user's perspective, the service they are subscribed to must be able to provide immediate responsiveness as it services requests and performs location and presence updates. To achieve this, the database must be able to do the following:

- Ensure real time responsiveness by delivering consistently low levels of latency as users interact with the service, irrespective of load on the system.
- Support constant updates (writes). Presence and location-enables services place very high update (write) demands on their underlying database.
- Support peak levels of throughput, typically measured as the volume of simultaneous users and updates per second.

By meeting the performance expectations of subscribers, customer satisfaction is achieved resulting in higher adoption rates, greater loyalty and improved ARPU (Average Revenue Per User).

3.2. Meeting Subscribers Expectations for Service Availability

Data and transactional states are usually the hardest parts of an application to make highly-available. Implementing a database that is itself highly available makes it easier for the application to become highly available as well. This approach permits delegating the complexity of data management and transactional states to the database layer. The

clear advantage of this design is that the database will always be the most competent, efficient and reliable mechanism in handling these duties when compared to other components within the system.

The essential characteristics needed to achieve high availability are discussed below.

3.2.1 Minimizing Planned and Unplanned Downtime

A system architected for High Availability has been designed specifically to minimize downtime by providing resilience in the event of system components failing or being taken off-line for maintenance. It is important to recognize that downtime can be divided into "Planned" and "Unplanned" downtime, as illustrated below:

Figure 3: ***Differentiating Planned from Unplanned Downtime***

Planned Downtime	Unplanned Downtime
Hardware Upgrades	Hardware Component Failure
Software Upgrades	Software Bugs and Crashes
Bug Fixes	Operator Error and Malicious Code
Extending Schema	Data Center Failure (i.e. physical disaster)
Back-Up Routines	

To create a true, highly available environment for communications services, it is important to select only those databases that are designed to minimize both planned and unplanned downtimes.

Combining technologies such as clustering, redundancy, on-line operations and geographic replication can enable applications to exceed 99.999% availability, equating to just five minutes of downtime per year, by minimizing both planned and unplanned downtime.

3.2.2 Fast Failover and Recovery

The goal of any technology used to maximize system uptime must be to instantly detect failures and then "failover", to clustered systems, typically in under one second, which can then continue processing transactions without any impact to the user. Multi-minute failover times can not be tolerated by services such as presence and location-enabled applications.

Database recovery is another element that is critical in delivering high availability. It is therefore important to select databases that enable either replacement or recovered systems to quickly join and re-synchonize with the active database cluster so that normal levels of system resources are re-established to meet the performance requirements of the service. Database recovery also needs to be automatic, without requiring manual intervention by administrators or complex programming logic embedded into the applications themselves.

3.3. Scaling the Service to Meet the Demands of Rapid Adoption

It is impossible to predict how successful (or not!) a new communications service can become. What may start out as service appealing to several hundred users can quickly spiral to attract millions of mobile devotees through the increasingly viral nature of communications enabled by Web 2.0 and IMS technologies. Similarly, mobile marketing and commerce campaigns may start by offering just a few promotions to a limited set of subscribers, but rapid expansions via partner adoption and user opt-ins can result in the infrastructure needing to scale rapidly to support increasing update and transaction volumes.

The web and telecommunications industries have long recognized the cost and flexibility advantages of "scale-out" computing over the more traditional enterprise models of "scale-up" computing. Architectures such as ATCA (Advanced Telecommunications Computing Architecture) are becoming more common place, where the processing and database load is deployed and scaled over low cost, commodity server blades built to comply with NEBS (Network Equipment Building Standards).

It is therefore necessary to select databases that can be distributed across these types of architectures while being able to handle the rapid increases in capacity and performance demanded by next generation communications services.

Coupling scalability with high availability is especially critical for communications services. If a service has to be restarted to recognize and use new servers deployed into an infrastructure, then the service will experience "planned" downtime. Therefore databases must be selected that enable the infrastructure to be scaled on-line, with no

interruption to either the application or the users. Such on-line scaling ensures the service can accommodate the most unpredictable workloads typical in mobile communications applications, without disrupting application availability.

3.4. Integrating New Services with Existing Infrastructure

CSPs are moving away from application-specific data stores towards centralized databases that serve the needs of multiple services. Such an approach enables CSPs to accelerate time to market by re-using existing database assets, while also reducing operational cost and complexity, as there are fewer data stores to administer.

The reality of such data management initiatives is to blend both federated databases and consolidated databases to provide the required data management capabilities to services in the network. Many applications require access to subscriber data, from OSS (Operational Support Systems) such as AAA and provisioning via LDAP, through to web interfaces for converged services into BSS (Business Support Systems) such as billing and business analysis.

As a result the database for presence and location-enabled services must be easily integrated into existing telecoms infrastructure via multiple industry-standard access methods, including SQL (Structured Query Language), LDAP, Java and HTTP along with native data access via common APIs and programming languages such as C++.

3.5. Summary of Data Management Requirements

To meet the data management demands of Presence and Location-Based Services, the underlying database must be able to address the following requirements:

- Achieve high levels of performance, specifically low latency response times and high update rates (writes), even under extreme loads;
- Deliver carrier-grade levels of availability with fault resilience and sub-second failover in order to minimize both planned and unplanned downtime;
- Extreme levels of dynamic and on-demand scalability;
- Provide multi methods to access data in order to meet diverse service requirements from a single, centralized database.

4. Introducing The MySQL Cluster Carrier Grade Edition Database

MySQL Cluster Carrier Grade Edition (CGE) is a real-time, distributed database combining the flexibility of a highly available relational database with the low TCO of open source.

Featuring a “shared-nothing” distributed architecture with no single point of failure, MySQL Cluster CGE is designed to deliver 99.999% availability demanded by next generation, converged communications services such as presence and location-enabled applications.

MySQL Cluster CGE's real-time design delivers predictable, millisecond response times with the ability to service tens of thousands of transactions per second. Support for in-memory and disk based data, automatic data partitioning with load balancing and the ability to add nodes to a running cluster with zero downtime allows linear database scalability to handle the most unpredictable and dynamic mobile communications workloads.

MySQL Cluster CGE is already proven in the toughest telecommunications environments delivering higher database throughput and faster response times at 10x lower cost than proprietary clustered shared-disk databases[4], with the added benefit of running on commodity hardware and operating systems. Customers include Alcatel Lucent, Cisco, Deutsche Telekom, Ericsson, Nokia Siemens Networks, Telenor and UTStarcom.

[4] http://miscalculation/why-mysql/case-studies/mysql-cs-alcatel.php

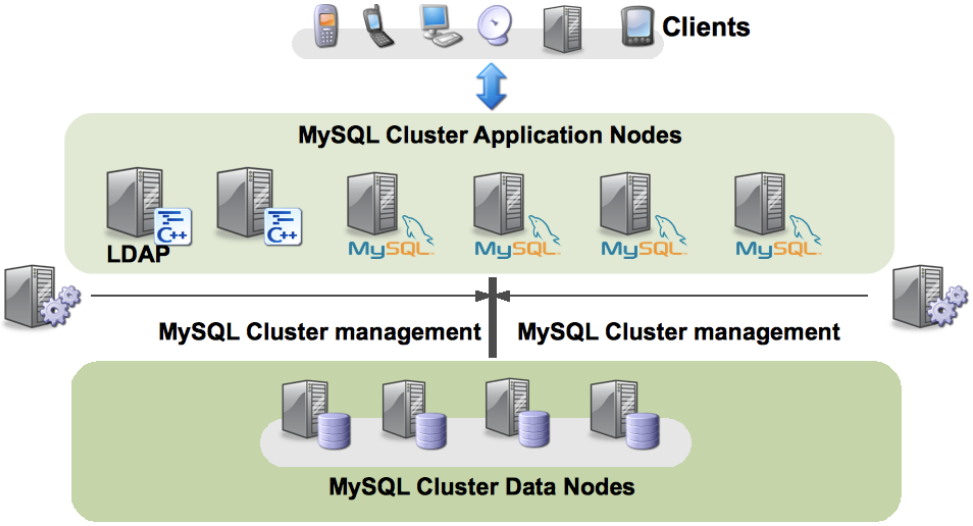


Figure 4: ***The MySQL Cluster Architecture Eliminates any Single Point of Failure***

To learn more about the MySQL Cluster architecture, refer to the MySQL Cluster Architecture and New Features whitepaper posted in the Resources section as follows:
http://www.mysql.com/products/database/cluster/resources.html

4.1. Advantages of MySQL Cluster CGE for Location & Presence Enabled Services

The following section of the whitepaper explores the capabilities of MySQL Cluster CGE in serving location and presence-enabled applications.

4.1.1 Achieving Real Time Performance & Linear Scalability with MySQL Cluster CGE

Real-time performance with high throughput and linear scalability are critical requirements to accommodate the massive volumes of database updates and transactions associated with location and presence-enabled services. Database performance requirements are typically tens of thousands of operations per second, with a consistent response time latency of just a few milliseconds.

As a distributed database, MySQL Cluster CGE employs a parallel server architecture with multiple active master nodes. This ensures updates and transactions can be load balanced and automatically scaled across multiple data nodes simultaneously, with each SQL node able to access and update data across any node in the cluster. Such an architecture ensures MySQL Cluster is extremely effective at scaling a very high volume of writes – which are typical in location and presence-enabled services.

MySQL Cluster CGE also offers a flexible data storage architecture with the ability to store both indexes and data in-memory or data on disk. As a result of this in-memory characteristic, MySQL Cluster is able to limit disk-based I/O bottlenecks by asynchronously writing transaction logs to disk, therefore maintaining real-time performance.

To further ensure real-time processing of updates and requests, users can configure MySQL Cluster to lock threads, such as the execution and maintenance threads, to specific CPUs, and set real-time scheduling parameters. This approach ensures all MySQL Cluster CGE processes remain active, avoiding them entering sleep or wait modes.

MySQL Cluster was recently benchmarked using the DBT2 test suite and achieved 251,000 transactions per minute with just four data nodes[5]. Each transaction involved around 30 database operations, and so MySQL Cluster was able to sustain around 125,000 operations per second, with an average response time of just 1.5 milliseconds. This performance increases represented a 4x improvement in scalability over previous versions of MySQL Cluster.

[5]http://blogs.sun.com/hasham/entry/mysql_cluster_7_performance_benchmark

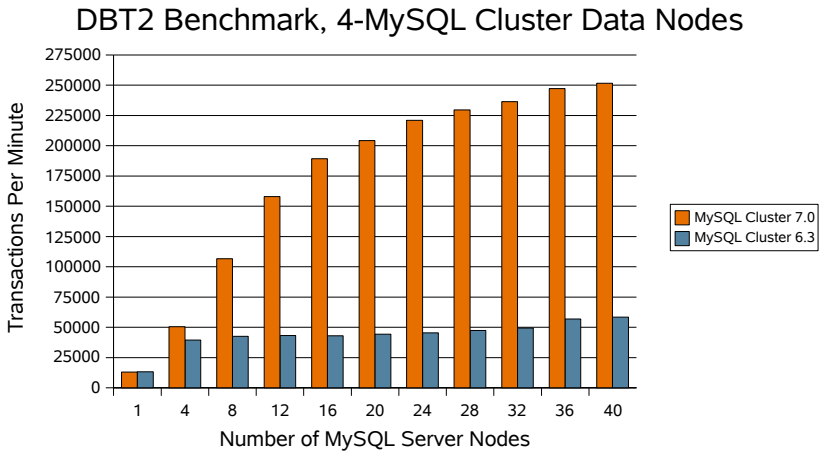


Figure 5: ***MySQL Cluster achieves over 250k TPM, or 125k Operations per Second with Average Latency of just 1.5ms***

Note: *More than one MySQL Server node was installed on each physical server, with a number of the server instances used as load generators to the MySQL Cluster database. An alternative deployment could use multiple connections from each MySQL Server, which would have resulted in fewer MySQL Server nodes being used.*

In addition to scaling performance, MySQL Cluster CGE can be deployed into a range of scaling scenarios. The distributed architecture means it can be easily scaled across multiple, low cost servers. Using the latest multi-threaded data nodes, MySQL Cluster can be used in server architectures equipped with multiple cores and threads. MySQL Cluster also supports the addition of nodes to a running cluster without interruption to the applications or users, enabling it to handle the most unpredictable and volatile workloads demands of converged communications services.

With support for up to 255 nodes, organizations have the flexibility to start small and make incremental investments to increase capacity as service adoption revenue streams grow. This eliminates the need for a large initial hardware and software investment in what tends to be over powered and costly configurations.

Through the combination of a distributed, parallel server architecture coupled with in-memory index and data support and choice of scaling options, MySQL Cluster CGE is highly suited to demanding, real-time and update-intensive requirements of presence and location-based services.

4.1.2 Achieving High Availability with MySQL Cluster CGE

The architecture of MySQL Cluster is designed to deliver 99.999% availability, which includes both regularly scheduled maintenance operations, as well as systems failures (i.e. planned and unplanned downtime).

This level of availability is achieved via a distributed, shared-nothing architecture and by synchronous replication of data within the cluster, which automatically propagates transaction information to all appropriate Data Nodes, before committing the transaction. If one or more database nodes fail during a transaction, the application simply retries the transaction and the remaining Data Nodes will successfully satisfy the request.

The advantage of synchronous replication is that it eliminates the time consumed by shared disk architectures in having to recreate and replay log files in order for the application to fail over. MySQL Cluster CGE detects any failures instantly and control is automatically failed over to other active nodes in the cluster, without interrupting service to the clients. In the event of a failure, the MySQL Cluster database nodes are able to automatically restart, recover, and dynamically reconfigure themselves, all of which is completely transparent to the application.

Should all nodes fail due to a complete power failure or multi-system hardware failure, MySQL Cluster CGE ensures that an entire system can be safely recovered to a consistent state. This is achieved by maintaining a series of checkpoints on the local disks of each Data Node, giving durability to the database.

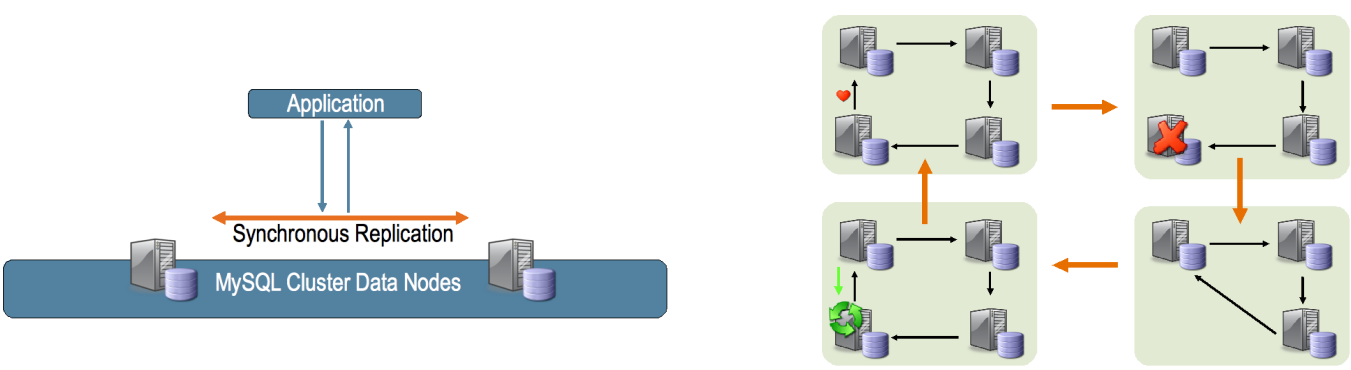


Figure 6: ***With Synchronous Replication and Sub-Second Failover, MySQL Cluster CGE Minimizes Downtime***

> *"MySQL Cluster delivers the high-availability that enables us to guarantee continuous broadband internet access and VoIP services to our subscribers. This has had an immediate impact in significantly improving customer satisfaction, and has reduced the cost of operating our network."*
>
> **Lars-Ake Norling**
> **B2, Telenor**

To further support continuous operation, MySQL Cluster CGE allows on-line updates to a live database schema, in addition to upgrades and maintenance to the underlying hardware & software infrastructure, as well as on-line back-ups. A new feature recently introduced to MySQL Cluster CGE is the ability to add nodes on-line to a running cluster. This is achieved through MySQL Cluster's ability to automatically re-partition data as new nodes are added, ensuring the cluster maintains continuous operation and application connectivity.

The ability to withstand site failures by replicating clusters across multiple remote locations is an important capability for many global on-line communications services. Geographic Replication is an optional feature of MySQL Cluster CGE, commonly implemented to:

- Achieve higher availability within the data center or across a geographic WAN
- Provide lower latency data access in different geographies.
- Replicating real-time data to other database tables for complex data analysis, without impacting the performance of the main, real-time production database

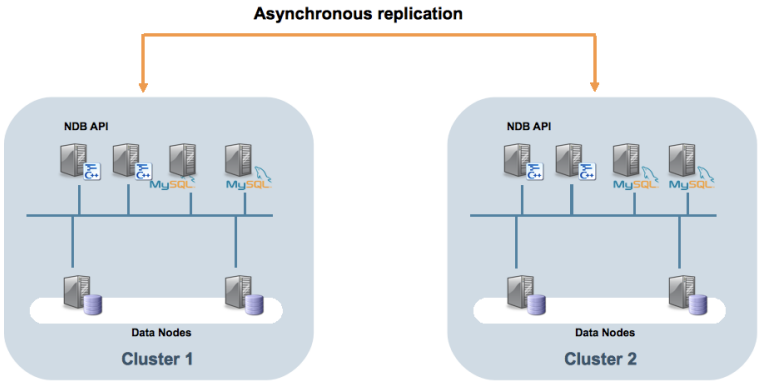


Figure 7: ***Geographic Replication provides multi-site protection for critical on-line applications***

Through the capabilities described above, MySQL Cluster CGE is able to eliminate both planned maintenance and unplanned downtime in order to deliver 99.999% availability demanded by location and presence enabled applications.

Requirement of On-Line Application	MySQL Cluster Capability
High Reliability & Availability	Support for ACID transactions Distributed Shared-Nothing Architecture Synchronous Data Replication Automated Sub-Second Failover Automatic Recovery and Data Synchronization Local & Global Checkpoints to Disk for Durability Geographic Replication Between Clusters On-Line Upgrades On-Line Schema Updates On-Line Add Nodes On-Line Back-Up High Availability Consulting 24x7 Technical Support Professional Training and Certification
Performance & Scalability	Hybrid In-Memory & Disk Based Storage Parallel Server Architecture for High Read/Write Performance Automatic Data Partitioning Enabling Load Balancing Distributed Database for Scaling Out Multi-Threaded Data Nodes for Scaling Up On-Line Add Nodes for On-Demand Scaling Multi-Master Clustesr with Conflict Detection & Resolution Asynchronous Replication for Scaling across Geographies
Interoperability	World's Most Popular Open Source Database Support for Multiple Hardware Platforms and Operating Systems SQL, Java, C, C++, LDAP, Web Services Data Access Support for Full Range of MySQL Connectors
Low TCO	Open Source Freedom, Standards and Economics Low Cost Service Offerings Runs on Commodity Hardware, including Rack Mounts and Blades Eliminates Costly Shared Storage Minimized Design Complexity and Maintenance

Figure 8: ***MySQL Cluster CGE meets the toughest challenges of next generation communications services***

4.1.3 Flexible Data Access with MySQL Cluster

Application developers can easily integrate new and legacy applications using their preferred database independent method. MySQL Cluster Carrier Grade Edition provides multiple data access methods that work together. These include SQL, a low-level C++ API, Java, LDAP and Web Services. This allows developers to select the data access method that best fits their development and application requirements. It also allows a single instance of MySQL Cluster to service a range of applications that previously would have required their own local database.

4.1.4 Minimizing Cost Per Subscriber

As a means to both accelerate time to market and reduce costs, the telecommunications industry is rapidly migrating towards commodity hardware, especially that which is based on ATCA standards. MySQL Cluster is specifically designed for commodity hardware, including rack and blade servers equipped with Intel, AMD and SPARC processors, running a range of industry standard operating systems including Linux and Sun Solaris in production environments, as well as Mac OSX and Windows in development environments.

"MySQL Cluster delivers carrier-grade levels of availability and performance with linear scalability on commodity hardware. It is a fraction of the cost of proprietary alternatives, allowing us to compete aggressively, and enabling operators to maximize their ARPU"

Jan Martens, Managing Director
SPEECH DESIGN Carrier Systems GmbH

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

MySQL Cluster is designed to be largely self-governing. There are very few system parameters that require fine-tuning, further reducing the risk of costly errors. As a result, there are typically fewer conflicts with other software and hardware

components, and less need for manual intervention. This also means that MySQL Cluster will have much lower maintenance costs, since fewer database administrators are required.

Beyond hardware and administration, MySQL Cluster is an open source high availability database solution that allows developers to download and start building their next generation services without the usual costs and time-to-market delays associated with adopting proprietary databases. 24x7 technical support services, consulting services and affordable licensing for MySQL Cluster are all available at a fraction of the cost of proprietary solutions.

> *"MySQL Cluster Carrier Grade Edition is a product of high quality, extremely robust and meets our demands in terms of performance and high availability. We evaluated shared-disk clustered databases but the cost would have been at least 10x more."*
>
> **Francois Leygues**
> **Alcatel-Lucent**

Running MySQL Cluster CGE on commodity hardware systems has enabled the telecommunications industry to radically transform both their development cycles and cost models to accelerate the delivery of new services while improving ARPU with increased margins.

5. MySQL Cluster Case Studies in Location & Presence Enabled Services

5.1. Alcatel Lucent: XDMS Application

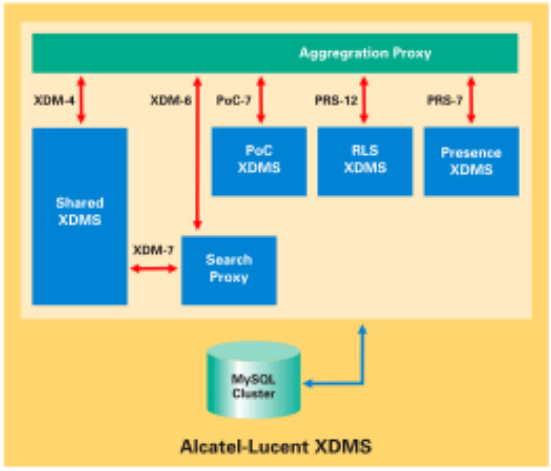
Alcatel Lucent originally deployed MySQL Cluster Carrier Grade Edition to power subscriber databases used in their HLR / HSS (Home Location Register / Home Subscriber Server) and Subscriber Data Management solutions.

> *"We have successfully implemented MySQL Cluster Carrier Grade Edition for our highly mission critical XDMS application which will enable the next generation of converged services"*
>
> **Francois Leygues**
> **Alcatel-Lucent**

The proven ability of MySQL Cluster to meet their requirements for real-time performance, continuous availability and linear scalability led Alcatel Lucent to evaluate the database as they began architecting the XDMS solution.

At the heart of ALU's next generation IMS services, XDMS was designed as a building block for managing security, subscriber authentication, contact lists and multiple user identities across multiple communications services and devices. A key feature of the XDMS solution was presence management which served as an enabler for IM, VoIP, PTT and on-line gaming.

MySQL Cluster CGE was selected to act as the single application repository for presence-enabled contact lists, identity management and data synchronization across different devices.



MySQL Cluster CGE enables Alcatel Lucent's XDMS system to achieve all of the following with a low TCO:

- Millisecond replication capabilities so that subscriber data is available in real-time across all nodes in the cluster, and is resilient to failures
- Low latency of transactions to support a high quality user experience
- Ability to scale-out incrementally, adding capacity via low cost nodes on an as-need basis, and scaling to handle millions of users. The XDMS system is designed to support up to 50 million concurrent subscribers generating up to 1,000 transactions per second.
- Geographic Redundancy to ensure data is replicated across different data centers

Learn more about Alcatel Lucent's experiences with MySQL Cluster here:
http://www.mysql.com/why-mysql/case-studies/mysql-cs-alcatel.php

5.2. Italtel: The Service Box (TSB)

Italtel is an acknowledged leader in next generation network solutions designed to create and grow new services for CSPs and ISPs (Internet Service Providers).

TSB has been developed by Italtel to integrate social networking sites with mobile telephony, allowing users to interact via web, voice, instant messaging and SMS from any device. The platform also provides location and presence based services.

> *"We decided to use MySQL for our products because we found that it had wide-spread, proven deployments – and met our stringent reliability and scalability requirements for the communications industry".*
>
> **Franco Serio, CTO**
> **Italtel**

MySQL Cluster CGE was selected by Italtel to power both the subscriber database and presence management services of the TSB platform. Italtel needed a data management solution that combined the flexibility of a Relational Database Management System (RDBMS) with unprecedented levels of reliability and scalability on low cost, commodity hardware in order to keep costs low.

Learn more about Italtel's experiences with MySQL Cluster here:
http://www.mysql.com/why-mysql/case-studies/mysql_cs_italtel.php

5.3. SPEECH DESIGN: Call Completion Services

SPEECH DESIGN is a leading European provider of messaging and mobility solutions for the enterprise and CSP marketplace. More than 130,000 companies and millions of subscribers benefit daily from SPEECH DESIGN's Value Added Services (VAS).

To enrich mobile communications, SPEECH DESIGN has developed the innovative Thor Application Platform. Layered on top of the OpenCall Media Platform, HP's media gateway solution, the Thor Application Platform offers Call Completion and Next-Generation Messaging services.

The Call Completion solution is designed to send SMS messages to users who have either missed a call from another party, or who have voice mail waiting in their inbox. Using presence management capabilities, the party that placed the call can decide to be automatically contacted by the application, when the called party becomes available again, or the called party will receive an SMS containing a "call back request". This solution is designed to accelerate user communications, as well as recover lost revenue for the operator. It can also reduce network load as the need for users to retrieve messages via IVR (Integrated Voice Response) is mitigated.

The database layer is a key component of SPEECH DESIGN's product suite. SPEECH DESIGN is using MySQL Cluster CGE to power the THOR application platform with the database responsible for core functions and services. With carrier grade availability and its high throughput architecture, MySQL Cluster CGE ensures time-critical operations for message-transactions and notification events. The database also stores global service configurations and messaging rule sets, as well as logging all user activity via CDRs (Call Detail Records).

Learn more about SPEECH DESIGN's experiences with MySQL Cluster here:
http://www.mysql.com/why-mysql/case-studies/mysql_cs_speechdesign.php

5.4. Tema Networks: Location Based Billing

Tema Networks develops billing and Value Added Services (VAS) solutions for mobile network operators around the world.

An increasing number of CSPs are promoting the mobile device as a complete replacement for fixed lines in the home or office, however, one inhibitor to such replacement is the perception of higher mobile telephony costs.
Tema Networks has developed their Home Zone Billing (HZB) solution to address this perception. HZB is location aware, allowing for variable tariffs based on the location of the subscriber at the time they make a call. This enables the operator to offer competitive pricing of mobile communications when used as a substitute for fixed lines in the home, while maintaining revenue and margin when real mobility outside of the home is required.

Tema Network's Home Zone Billing solution has recently been adopted by one of the leading mobile operators in the European and Asian markets, serving 90 million subscribers across two continents.

Because the HZB system provides real time tariffing for both pre and post-paid subscribers, Tema Networks needed a database system capable of maintaining the most stringent throughput and availability requirements, while delivering a consistent response time of just a few milliseconds. For their new customer, Tema Networks had to deploy the HZB system over five sites, up to 6,500 kilometers apart, separated by remote terrain and spanning seven different time zones. The database is distributed across these sites and must maintain data consistency and integrity at all times while also achieving high availability.

Note: The database layer was implemented with MySQL Embedded Server Pro 5.1.

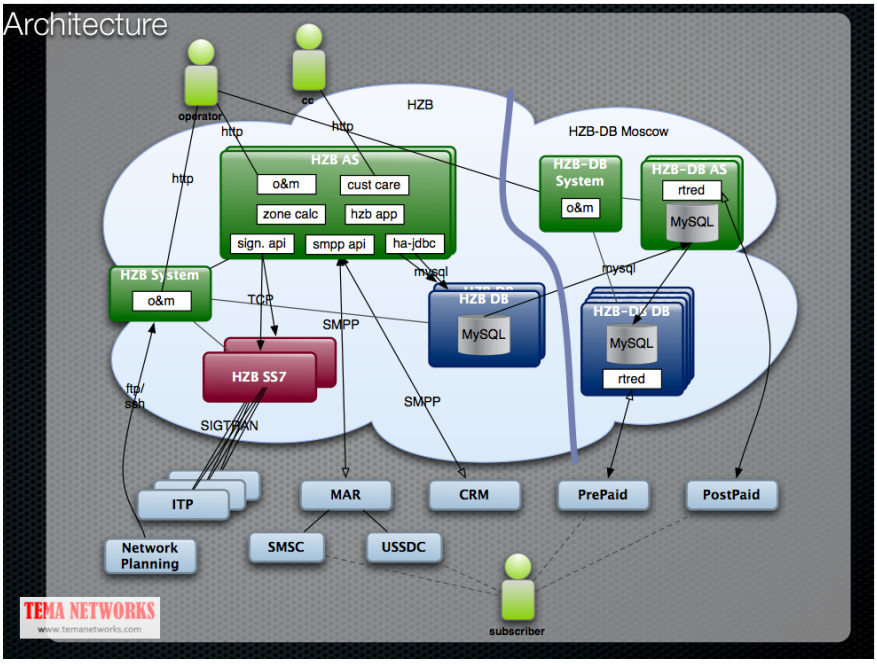


Figure 9: Architecture of Tema Network's Home Zone Billing Solution

Learn more about Tema Network's experiences with MySQL here:
http://www.mysql.com/why-mysql/case-studies/mysql_cs_temanetworks.php

6. Conclusion

In this paper we have explored how location and presence-enabled services are enabling CSPs to substitute declining voice revenues with highly compelling, highly personalized next generation converged services, thereby generating new revenue streams and acquiring new subscribers.

MySQL Cluster provides the ideal foundation for data management for these next generation services. By delivering real-time responsiveness, carrier-grade availability and linear scalability, CSPs can cost-effectively deliver services that meet user requirements and provide competitive differentiation.

With successful deployments in a host of presence and location-based services, coupled with a proven track record in areas such as Subscriber Databases (HLR / HSS), Telecoms Application Servers, AAA infrastructure, etc, it is easy for any telecommunications company to get started with MySQL Cluster today by downloading the product from the MySQL website[6] and using the resources below to start an evaluation.

[6]http://dev.mysql.com/downloads/cluster/7.0.html

7. Additional Resources

MySQL Cluster on the web: www.mysql.com/cluster

MySQL Cluster Customer Case Studies: http://www.mysql.com/customers/cluster/

Getting Started with MySQL Cluster: http://www.mysql.com/products/database/cluster/get-started.html

MySQL Cluster Architecture and New Features Whitepaper:
http://www.mysql.com/why-mysql/white-papers/mysql_wp_cluster7_architecture.php

MySQL Cluster Evaluation Guide: http://www.mysql.com/why-mysql/white-papers/mysql_cluster_eval_guide.php

---