University Librarians Association (ULA) Sri Lanka

Conference Proceedings

ICULA 2010

6th International Conference of the University Librarians Association

14th & 15th July 2010

'University Librarianship: an academic challenge and an opportunity"

Chief Editors

Dr. Wathmanel Seneviratne
Dr. Chaminda Jayasundara
Manoja Samaradiwakera

Ceylon Continental Hotel
Colombo Sri Lanka

Proceedings of the 6th International Conference of ULA held from 14th - 15th July 2010 at Ceylon Continental Hotel Colombo, Sri Lanka.

"University Librarianship: an academic challenge and an opportunity"

ISBN 978-955-1359-03-4

Lassana Printers & Publishers
749, Avissawella Road, Wellampitiya.
Tel: 011 - 2572297/ 2572515
Fax: 011- 2547662 Email: lasprint@sltnet.lk

Assuring the Security and Accuracy of Data using a MySql Driven Database in an Automated Library Environment: case study for the library of University of Ruhuna

Hettiarachchi. N.[1] and Fernando. I.D.K.L.[2]

Abstract

After initiating the automated circulation service in the Main library of University of Ruhuna the problem of assuring the security and accuracy of data has become a major problem. These daily influxes of valuable data mainly include circulation data and library material related entities of the library. To provide a secure and accurate data service, the library need to have a good backup service. At the beginning this was managed by a schedule back-up via a routine ***cron deamon*** on a Debian (Linux) system to copy the database into several back-up media. But this result a risk of either loss or accuracy of data in the library database, since there is a time delay between backup taking times. In this paper we present the systems administration functions designed and implemented by the UoR Library by setting up a replication server which is capable of maintaining an identical copy of the original master database down to each millisecond of modification for the MySql database used by the Information System of UoR (ISURu), a derived product of Koha (University of Ruhuna, 2005). The replication server acts as a slave system which becomes a master at an event of a failure with a fully updated database which is identical to the original one. Later the system was fine-tuned by adding the service Heartbeat which is capable of switching between master and slave servers automatically in an event of a failure. The combined effects of these newly updated features provide capacity to assure the security and accuracy of data in the automated library environment of University of Ruhuna.

Key words: *Data security, replication, MySql database, Koha, library automation, automated circulation*

Introduction

The main library of the University of Ruhuna (UoR) has been rendering its service to the university since 1985 at which time the university was located at Wellamadama, Matara (University of Ruhuna, 2010). All the library functions were executed manually till 1992.

---

[1] N. Hettiarachchi, University of Ruhuna, Sri Lanka
nimal@lib.ruh.ac.lk

[2] I.D.K.L Fernando, University of Ruhuna, Sri Lanka
kusala@lib.ruh.ac.lk

In 1992, the University Grants Commission of Sri Lanka offered computers for all universities to promote library automation facilities in all Universities in Sri Lanka (Hettiarachichi, 2001). According to the annual report of the University (2005), the library has initiated work with CDS/ISIS (Computerized Documentation Service/Integrated Set of Information Systems) to manage the library related data. According to Hettiarachichi (2001) library started the OPAC (Online Public Access Catalog) in 2000. The conversion of the library software environment into the open source library management system - ISURu took place in 2005, which was completely novel and highly cost effective and economical approach that can be considered as a quantum leap in the technical development of the library functions.

The library of University of Ruhuna started automated circulation system in 2009 (Univereisty of Ruhuna, 2009) that facilitated the library requirements of nearly 10000 (University of Ruhuna, 2009) of its students and members of staff. Automating the library services will greatly benefit not only the library patrons but also the staff members of the library, who are responsible for providing different information services in particular fields (Sonker and Jayakanth, 2003). As the first step of this automation process, the Ruhuna library launched a test run of its circulation in 2008 and expanded it to a live service from 2009 onwards which compelled the library to investigate ways and means to manage a bulk volume of daily data with a super care.

Technical Background

Versatile library software which consists of a high performing database management system offers a significant support to maintain and organize a massive bulk of data with a good security in an automated library system of a University. There are various types of library management systems in the world (Hettiarachichi, 2001). With respect to the economy and above mention factors, Ruhuna University library selected an Open source library management system (koha). Koha was developed in 1999 in New Zealand by Katipo Communications Ltd with funding from the Horowhenua Library Trust (Tong, 2004). The most significant character found in the Koha Library Management System (KLMS) is that it is an open source or is providing the source code of the program (Perens, 2008) allowing modifications and even re-distribution according to requirements of the user however, under the so called GPL license (GNU general public license, internet).

According to the official website of Koha, it is the first open-source Integrated Library System (ILS) which is being used in libraries all over the world and which is built using library standards and protocols that ensure interoperability between Koha and other systems and technologies, while supporting existing workflows and tools. Koha source code is written in Perl, using a MySql database to store its data, and is accessed via an Apache web server (Weir, 2005). Since it is a fully

functional library management system, it is equipped with modules for circulation, cataloging, acquisitions, serials, reserves, patron management, branch relationships, and also with support for MARC21 and UNIMARC, E-mail and/or paper based patron's overdues and other notices, Print libraries' own barcodes, web based OPAC system, report generating etc. (Koha, online; Tong, 2004). It also enables support for various operating systems such as Linux, UNIX, Windows and MacOS without any indecision. Another significant feature that provide by koha is that it uses a dual database design that utilizes the strengths of the two major industry-standard database types (text-based and relational database management system - RDBMS). This design feature ensures that Koha is scalable (capable of being easily expanded or upgraded on demand) enough to meet the transaction load of any library system irrespective of its size (koha, internet).

As well as the organization and management of library environment, security of library data and information is also a significant matter of consideration (Newby, 2002). When data security is addressed, it may be from the point of view of corporate information security management, rather than library environments (Davies, 1992). A library should be capable enough of keeping data without lost and in a way that they can be access whenever needed. According to the Newby (2002), data security includes personnel security, privacy, policy and computer security of both users and library. Since the data in an automated library environment are stored in computer based electronic equipment, they are mainly depending on electricity. Frequent failures of electricity can be a serious threat to data that stored in a library. Apart from that software bugs routinely corrupt data files, users accidently delete their whole work-stuff, hackers and disgruntled employees erase disks or even a natural disaster can destroy the entire database without a pre-warring of the danger (Nemeth, et al., 2003). According to Newby (2002), a library would include personnel security and policies when the time they taken steps for effective backups and should assure the physical integrity of computing facilities for its clients.

At the beginning, the Ruhuna university library managed a physical backup system to secure the library database. System manipulated to get backups in scheduled hours within a day through a cron demon (Welsh & Kaufman, 1995), since the Ruhuna university library supported by Debian operating system (Debianhelp online, 2010). However, it was a temporary solution for assuring the security of the data that stored in the library database. The cron daemon is a long running process that executes commands at specific dates and times (e.g. hourly, daily or weekly) (Nemeth, et al., 2003). It was a time consuming physical backing up process that copying the whole database of the Ruhuna library system. Since the library database was a MySql driven one it was easy to modify the system by using the specific backing up features that are available in MySql to obtain a more advanced service for data security.

MySql is an open-source database management system for relational databases (Ullman, 2006) and it is a diligent database management system that can handle a

large number of data. According to the manual of Global Backup & Recovery (internet), MySql offers a variety of backup strategies from which users can choose the methods that best suit with the requirements of their own. MySql replication is a very efficient and safe way to protect databases that are used for critical systems and it does not effect on the data operations (MySql replication, internet). This can use to create a warm standby server that can continue the services which are provided by the primary database server in a case of failure. Replication enables data from one MySql database server (the master) to be replicated to one or more MySql database servers (the slaves). Since the replication is where one MySql server mimics the behavior of another, at each millisecond the slave server is updating with fresh data simultaneously with fresh modifications at master database (Ullman, 2006).

Managing the Switching between these two (master and slave) servers need to be managed by the system administrator in an event of a failure and this has become a major demerit that found in a replicated system and a factor of risk when providing a continuous service for the library users. So library system of Ruhuna was needed to be updated in to an automated condition of switching to workable server in case of a system failure. The complication was solved by MySql Heartbeat. Heartbeat is a software solution for Linux (Using Heartbeat with MySql and DRBD, Internet). It is not a data replication or synchronization solution, but a solution for monitoring servers and switching active MySql servers automatically in the event of failure. Heartbeat needs to be configured to manage the switching between two servers in such a failure. The resource configuration defines the individual services that should be brought up (or taken down) in the event of a failure (High Availability and Scalability, online). Heartbeat needs to be combined with MySql Replication or DRBD (Distributed Replicated Block Device) to provide automatic failover. DRBD is a solution supported by Linux. DRBD creates a virtual block device (which is associated with an underlying physical block device) that can be replicated from the primary server to a secondary server (High Availability and Scalability, online).

Problem Statement

Automated Library system of University of Ruhuna is holding valuable information other than the library items that placed in the library. This valuable information is the data about library materials and circulation records of users. Traditionally, the most sensitive data that libraries collect are circulation records (Newby, 2002). By necessity, these are linked to identifying information for individual patrons, who borrow books or other materials. These data must be kept in a good security to provide a continuous library service without any failure in the automated library system.

After initiated the automated circulation process the library administration of University of Ruhuna face a problem of maintaining a huge daily bulk of data with a high security. The tar command that programmed through a cron job was not

diligent and efficient enough to supply a perfect security to Ruhuna library database. On the other hand, the corn job can only be updated in hours and therefore, there is a risk of maintaining updated database to its last minute.

Objectives

The objective was to find out the means to assure the accuracy and security of the data, which associate with the library system by cutting down the intervention of a middle man to maintain backups and automated switching between replicated servers in an event of a failure.

Method

This case study explains the practical and applicable approaches that were implemented by the library of University of Ruhuna to mitigate the problem of maintaining the security and accuracy of library data in an automated library environment by applying following methods and technical solutions respectively.

Managing the Data Backups by using 'CronDemon'

A cron- job was scheduled by including a tar command in to the crontab file found in Linux (Debian/Leney) operating system to get backups of the library database in specific times of a day, i.e. the tar command will copying the MySql database of KLMS at 06:00am, 12:00noon and 05:00pm on each day in to a tape cartridge and USB hard disk.

Managing the Data Backups by using Standard MySql Replication

To facilitate a high efficient backup service, the master server which running the MySql data base was replicated with another secondary master (slave) server which running with the same version of MySql server (version 5). Since this service is capable of keeping backups to its last minute of updating of the master server while the pervious cron-job manage the usual scheduling in several time slots at night. There by the cron at 12:00noon was terminated and new cron was programmed to get **–tar** backups at 12:00 midnight by ensuring a physical backup of the koha database. The replication system was established according to following steps.

1. Setting the Replication Master by editing the **my.cnf** file and modifying it by adding the following options.

```
[MySqld]
log-bin=MySql-bin
server-id=1
```

2. Setting the Replication Slave by editing the **my.cnf** file and modifying it by adding the following options.

```
[MySqld]
server-id=2
```

3. A new user account was created by granting privileges required for replication in the master server.
4. Create a snapshot of the data in the master database by using the *MySqldump* tool.
5. Import the dump file in to the slave and create the database.
6. To set up the slave to communicate with the master for replication, following statement the on the slave was executed.

```
CHANGE MASTER TO  MASTER_HOST='master_host_name',
MASTER_USER= 'replication_user_name',
MASTER_PASSWORD='replication_password',
MASTER_LOG_FILE= 'recorded_log_file_name',
MASTER_LOG_POS=recorded_log_position;
```

7. Start the replication by executing following command.

```
START SLAVE;
```

A schematized diagram of the MySql replication architecture is showing in figure 01.

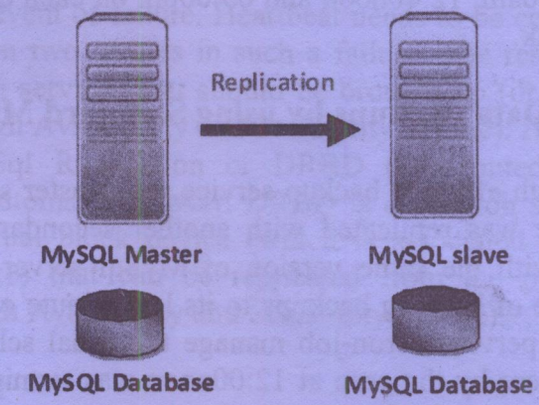

Figure: 01. MySql replication architecture

System updated with the automated failover with MySql with Heartbeat

Previously modified MySql replication system was then combined with MySql Heartbeat to provide the continuous monitoring of servers and switching the service to an active another server automatically which is in the replicated system in an event of a failure. These systems make a fully automated security service for the Ruhuna library that alive through out the whole day long. Heartbeat was

configured as figure 02 bellow and Steps that taken to establish the heartbeat are mention bellow.

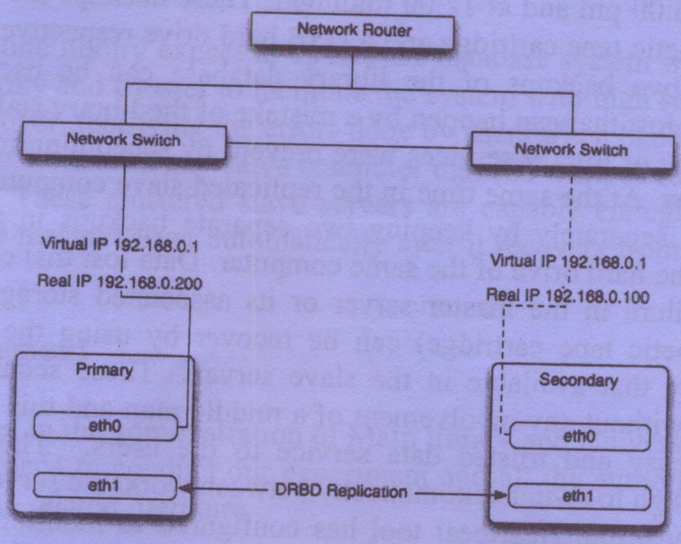

Figure: 02. Heartbeat Architecture

1. The master and slave servers were configured by adding two Ethernet cards (eth0 and eth1).

2. Those servers were configured by giving same virtual and different real IP addresses.

3. Both master and slave servers were configured as DRBD servers with unique host names.

4. These two servers were connected to the network through two different network switches.

5. **"ha.cf, haresources, authkeys"** files were configured according to the configurations of heartbeat.

Results

Having ISURu as the library management software that embedded with a MySql database management system, Ruhuna library is capable to reach the objective of providing a secure data service by maintaining one identical copy and four other backup of the library database in a single day. As shown in figure 03, the replication server keeps an identical copy of the original database till its last millisecond of update. By this the previous problem of keeping an updated library database was completely solved. In the same time four physical backups of the library database are saved in four different locations in specific time daily. In the master server there is a cron job to –tar the database folder of the original database

at 06:00 am, 06:00 pm and at 12:00 midnight. These backups are programmed to store in a magnetic tape cartridge and a USB hard drive respectively in the master server. These two backups of the library database can be use to recover an accidental data lost that can happen by a mistake of the library staff. The network traffic and delay of library services were prevent by scheduling to get backups at off library hours. At the same time in the replicated slave computer the same cron job is running separately by keeping two separate backups in a magnetic tape cartridge and the hard drive of the same computer. Data lost that can be happen by a hardware failure in the master server or its associated storage devices (USB drive or magnetic tape cartridge) can be recover by using the backups or the replicated copy that available in the slave server. These security services are running daily without any involvement of a middle man and this make the library management easy and trusted data service to the users. The system is now intelligent enough to switch automatically between workable servers in an event of a failure, because the Heartbeat tool has configured to monitor the behavior of servers throughout the day. Server system is also supported with UPSs to prevent the electric failures that can happen suddenly.

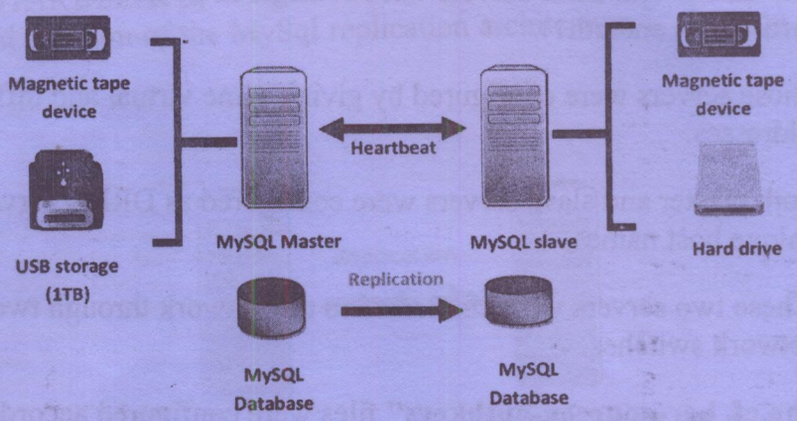

Figure: 03. Overview of data security strategy of the Ruhuna library system

Conclusion and Discussion

Library system of the University of Ruhuna competent to supply a high security assurance to users on their personal, circulation and library related data with in a fully automated library environment. Finally the Ruhuna library system is maintaining five different backups of the libraries database without a human intervention on any step of the backing up process.

A Library system must ensure the security and accuracy of data (Newby, 2003) to provide an effective and efficient service to its clients. This is being a significant issue in an automated library environment rather than non – automated condition.

Automated library system in University of Ruhuna performs well by providing an unbreakable service to the library users.

In future, Ruhuna library expect to update the database system with new software like MySql Proxy and Cluster to facilitate the system with high security, accuracy, automated failover, rectifying the errors done by users when handling the database and enabling to work with any salve computer that can work as server in an failure of the master. These clustered slave servers are capable enough to upgrade the database of the master server automatically after it recovers from the failure at no time.

Acknowledgement

We are indebted to the administration of Main library and University of Ruhuna to serve enough space to conduct our experiment and all the authors who supported by providing the related literature.

References

Bhojaraju, G. (2002). Automation of A.T.E.C Central Library: A Case Study.

Davies, J. E. (1992). Computer Misuse. In Chapman & Zwicky, op cit.

Hettiarachichi, N. (2001). A Comparative study on usage of information technology in University and special libraries in Sri Lanka. Dissertation for Master of Social Science: University of Kelaniya. 20-21.

Nemeth, E. Snyder, G. Hein, Trent R. (2003). Linux Administration Handbook. New Delhi: Prentice-Hall of India Private Limited.

Newby, G. B. (2002). Information security for libraries. School of Information and Library Science. University of North Carolina : Chapel Hill.

Sonker, S. K., Jayakanth, F. (2003). Koha - An Open Source Integrated Library Automation System. National Center for Science Information. Indian Institute of science: Bangalore.

Suehring, S. (2002). MySql Bible with PHP, Perl and JDBC/ODBC. New Delhi: Wiley Dreamtech India (P) Ltd. 348-351.

Ullman, L. (2006). Visual quick start guide MySql. -2nd ed.- Pearson Education. Inc. and Dorling Kindersley publishing Inc.

University of Ruhuna. (2005). Annual Report. Matara: University of Ruhuna.

University of Ruhuna. (2009).Annual Report. Matara: University of Ruhuna.

Welsh, Matt., Kaufman, Lar. (1995). Running Linux. United States of America: O'Reilly & Associates, Inc. Ch.4.

Backup and Recovery Types (1997). Retrieved May 12, 2010 from http://dev.MySql.com/doc/refman/5.5/en/backup-types.html.

Debianhelp. Retrieved 12, May 2010 from http://www.debianhelp.co.uk/schedulejobs.htm.

Global Backup & Recovery for MySql. Retrieved 12, May 2010 from http://MySql.com/products/backup/.

GNU general public license. Retrieved May 12, 2010 from http://www.gnu.org/licenses/gpl.html .

High Availability and Scalability. Retrieved May 12, 2010 from http://dev.MySql.com/doc/refman/5.0/en/ha-overview.html.

Koha Retrieved May 12, 2010 from http://koha.org/.

MySql 5.1 Reference Manual. Retrieved May 13, 2010 from http://dev.MySql.com/doc/refman/5.0/en/MySql-cluster-overview.html

MySql Replication. Retrieved May 12, 2010 from http://dev.MySql.com/doc/refman/5.1/en/replication.html.

Perens, B. (2008). The Opensource definition. [e –book] Retrieved May 10, 2010 from http://perens.com/OSD.html.

University of Ruhuna. Historical Sketch of the Library, University of Ruhuna. Retrieved May 17, 2010 from http://www.lib.ruh.ac.lk/historicalSketch.php.

Tong, T. W. (2004). Free/Open Source Software Education. [e-book] New Delhi: Reed Elsevier India Private Limited Retrieved May 10, 2010 from http://akgul.bilkent.edu.tr/iosn/foss-edu.pdf.

Using Heartbeat with MySql. and DRBD Retrieved May 12, 2010 from http://dev.MySql.com/doc/refman/5.5/en/ha-heartbeat-drbd.html.

Weir, B. (2005). Koha on Windows. Retrieved May 12, 2010 from http://www.ossrc.org.in/downloads/koha-resources-cd/koha-on-windows/Koha-on-Windows.pdf.