ISSN: 2321-7782 (Online)
Volume 2, Issue 3, March 2014
International Journal of Advance Research in Computer Science and Management Studies
Research Article / Paper / Case Study
Available online at: www.ijarcsms.com

Database Forensic Analysis

Karen B. Alexander
Department of Computer Science
Sipna College Of Engineering And Technology
Amravati

Abstract- Nowadays, data security policy demands auditing for the high performance databases so as to make sure data integrity and additionally to notice information change of state if any. relative databases use varied auditing capabilities, that contains the examination of knowledge and operations for accuracy, lawfulness and to report risks and to create recommendations to push sound-operating practices. information auditing is that the method to be administrated on continuous basis. This records and analyzes the information activity for coverage on some amount. However the information is tampered by choice or accidentally either by licensed or unauthorized users at any instance bypassing auditing system too. Any reasonably suspicious behavior should be inspected and analyzed more with information forensics. During this paper, a framework is projected for analyzing and reconstructing the activity of any suspicious behavior at intervals information. The aim is to spot, analyze, validate, interpret, generate rhetorical report and preserve the proof for digital investigations. To prove within the idea, the information MySQL information is studied and analyzed for this projected framework

.Keywords- MySQL , database, forensics, artifacts, logs, query cache.

I. INTRODUCTION

Today nearly all the applications like employing a high performance info for coping with large quantity of information. Therefore the info security communities area unit arising with varied techniques and approaches to keep up knowledge privacy, integrity, and handiness. However it's discovered that digital attacks area unit targeting the informations resulting in database security and threats [1, 2 and 3]. The results of these is that the present rules area unit specifying investigations and response to security breaches or policy violations. Thus Organizations should take into account their incidence response policies rigorously, that area unit a part of their overall security policies. Totally different information have alternative ways of auditing the database activities. In SQL server, auditing will be done through a SQL trace that is one among the protection policies. this can be AN interface created on the market to spot poorly running SQL statements, and to correct alternative performance issues. AN application named SQL Profiler collects the events. However solely SQL Trace isn't the one to be smitten by for watching the info. it's same that even Microsoft discourages the employment of SQL Traces on a production system [5], as a result of once enabled it will consume far more than memory locations, CPU cycles, and space. Additionally SQL Trace doesn't audit or monitor systems on continuous basis. This ancient auditing system doesn't have intelligence designed into it. The worst factor concerning it's that no modification is feasible to trace to what, when, or WHO is being audited. In alternative words, it's quite tough to trace malicious activity. it's same that SQL Trace is nice at amassing an enormous quantity of information, however is insufficient to find the "needle within the haystack" that's proof of malicious activity. Equally in MySQL 'Information_Schema table' provides access to info information. Secondly, audit systems area unit written to an area file or table which may be insecure as a result of the information isn't well protected. If a malicious activity is administered, the suspected person will tamper with audit logs or will delete records feat no traces of the activity. an equivalent can happen for a DBA WHO needed to perform actions they're not approved to hold out. however because the info systems build various redundant copies of sensitive knowledge things in table storage, indexes, logs, materialized views, and temporary relations, thus once knowledge is deleted,

it's not destroyed however it typically persists on disk [7, 8]. it's tough to predict whether or not the sensitive knowledge is lost when deletion or it's the chance of recovery. These remnants of past knowledge and activities referring to info systems will be disclosed through info rhetorical analysis. This analysis will the method of extracting data and knowledge from info internals primarily from logs, data files, tablespaces, through trigger operation, view etc. Therefore, knowing the management system structure may be a precondition for rhetorical analysis. to investigate the information, it's vital perceive|to grasp} and understand very well, however the info is made together with the understanding of forensics. With this information one will establish and may come back up to helpful rhetorical investigation report in quality time. The in depth options for MySQL info five.5 area unit highlighted during this paper. The analysis would be administered on the rhetorical copies of info that is obtained mistreatment MySQL utility programs in our projected framework. The info rhetorical investigations can reach the databases to see exactly once the attack occurred, that knowledge was compromised, from wherever the attack passed off and fairly WHO performed the attack. The goal here is to style a technique to validate hypothesis concerning past activities in a very manner that's respectable in court.

A framework projected during this paper for the sake of rhetorical analysis works in 2 stages. within the initial stage, careful multiple logs of MySQL area unit created on the market that is employed for analysis. These logs grow into substantial sizes that hold vital data. However, together with the helpful data it additionally contains routine primarily based general operational knowledge which cannot be needed at the time of research. Thus extracting the helpful data from logs that area unit required for the target analysis may be a crucial task. To do so, MySQL utility programs area unit used for creating rhetorical copies of the databases and its multiple log files (text files, binary log files etc.) for analysis. These files of info area unit then parsed and browse mistreatment Perl script written to provide the relevant data as per the condition set by the investigator. The expected results of this script is that the information having traces of actions from the multiple files that helps to predict the identification and activities of unauthorized events administered. Then, the supported illation rules set down mistreatment knowledgeable information the choices area unit taken from the hold on data within the derived information to urge the relevant and filtered data for analysis. this can be achieved by matching similar patterns and behavior of the system all over from information. A pre-final log analysis report is generated at this stage. Within the second stage, we tend to establish and reconstruct the activity from the varied MySQL server objects known and picked up as delineate in section five for additional analysis. The activity that is then reconstructed should be valid against the antecedently generated log analysis report back to provides a final rhetorical report.

II. Related Work

Database rhetorical is an important space that should want awareness from analysis perspective. the dearth of analysis during this space is thanks to the quality of databases that aren't totally understood in an exceedingly rhetorical context nonetheless. The paper InnoDB info Forensics shows however the MySQL tables within the .frm files area unit designed and the way vital info is saved. The aim was to spot and name the bytes and to interpret them. There with information, it had been attainable to observe inconsistencies within the info. however there was no information discovered for the multiple log files and cache for additional analysis [11]. David Litchfield projected a LogMiner tool that permits Associate in Nursing Oracle DBA Associate in Nursingd rhetorical analyst to spot the actions taken on an Oracle info although the auditing options are turned off. LogMiner may be a utility that may be accustomed Associate in Nursingalyze the redo log files that area unit created by an Oracle info [8, 9, and 12]. Dragoon (Database rhetorical Analysis safeGuard Of arizONa) by Pavlou and Snodgrass may be a image auditing system for change of state detection and info rhetorical analysis. A survey study was administered antecedently on "Database Security Threats and Challenges in info Forensic" that highlights the work done by numerous researchers within the space of info forensics.

III. MySQL's Internal Structure

MySQL is one among the foremost fashionable Open relational database management systems. Figure 1 below shows main parts of MySQL.

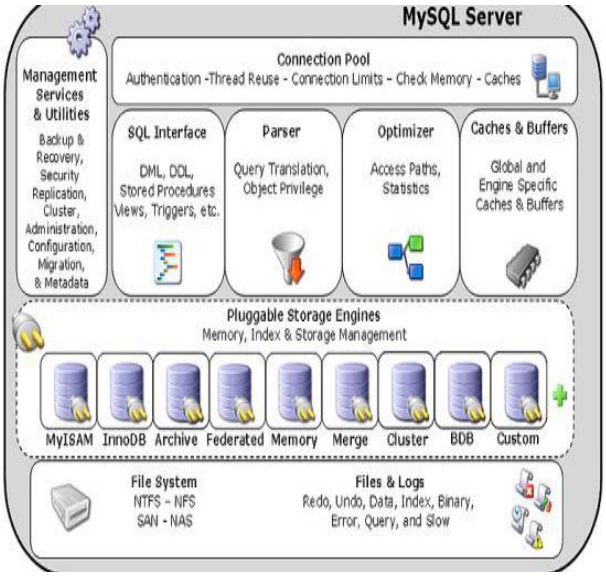

Fig. 1) MySQL Architecture

From the above figure we can see that, the default storage engine for newly created tables in MySQL database 5.5 is InnoDB [5, 10]. MySQL contains the data directory that stores all kinds of information managed by the MySQL server. It also stores all the databases, status files and log files. It includes features for transactions, stored procedures, views, and triggers, etc. The storage management of these engines forms the mines for database forensic investigation. All of the database components should be well known for investigations.

A. Structure of the Data Directory

The data directory stores its databases forensic related files. For rhetorical investigations, it's necessary to grasp the structure and contents of the info directory in order that the investigator is aware of however the server uses the filing system to represent databases and tables, additionally as wherever the server logs square measure set and what they contain. underneath Windows, the default information directory location usually is C:\ Program Files \ MySQL \ MySQL Server 5.0\ data

All the databases managed by the server are contained within the MySQL data directory. These square measure organized as shown in Figure 2 below.

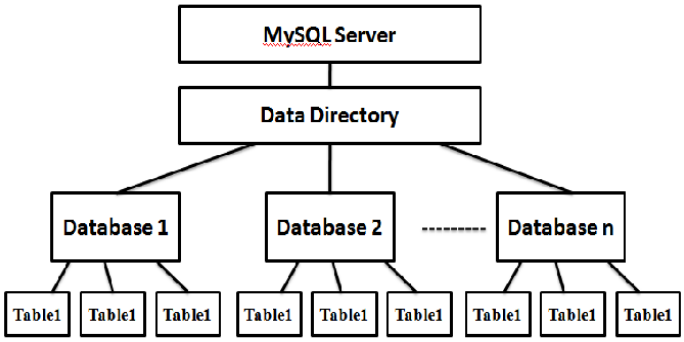

Fig. 2) MySQL Data Directory Structure

From the higher than figure that depicts the MySQL information Directory Structure we have a tendency to see that:

Each and each info is found below the information directory itself.

Tables, views, likewise because the triggers inside a info correspond to files within the info directory.

The data directory conjointly could contain alternative files like, server's method ID (PID) file, that is written by the server itself on the beginning.

Server generated standing and log files also are an area of knowledge directory. These files provide the directors some necessary info concerning the server's operation. It's a standard approach for these directors to use the information directory so as to find these files.

B. Information Schema

Information Schema is a table that provides an access to the database metadata, all kinds of information about the MySQL server such as the database name or table name, data type of columns, or even access privileges. All sort of information about all the databases in MySQL server are stored in this table with only a read privilege.

C. MySQL Status and Log Files

MySQL data directory contains a number of status and log files. MySQL Server maintains logical logs and storage engine maintains physical or logical logs (additionally). Data directory of the server is the default location for each file with its default name (HOSTNAME). The table lists only the server-level status and log files. Individual storage engines may have their own logs or other files. The InnoDB engine of MySQL server has two types of logs : undo log and a redo log. The undo log rolls back transactions. It also displays the data's older versions. The redo log stores the information which is utilized at the time of crash recovery process. It gives permission to the recovery process so that the transactions are re-executed that could or could not have completed before the crash occurred. Once it again executes those transactions, the database goes into a consistent state. Log files are most important data facts for investigation, since they contain those statements which include sensitive information such as passwords.

D. MySQL Utility Programs For Forensic Analysis

Some MySQL utility programs are studied and analyzed which can further be used at the time of forensic analysis.

1. mysqldump (A Database Backup Program)

The mysqldump shopper may be a quite a backup program which will be accustomed dump a info or a set of databases for a backup purpose or transfer to a different SQL server. This dump generally contains all the SQL statements to make the table, populate it, or both. This program may also be accustomed generate files in XML format..

2. mysqlaccess (Client for Checking Access Privileges)

The mysqlaccess may be a diagnostic tool for the MySQL distribution that checks access privileges. These privileges' ar outlined for a bunch name, user name, and additionally numerous info combinations. It additionally checks access exploitation solely the user, database, and host tables. It doesn't check tables, columns, or routine privileges.

3. myisamlog (Display MyISAM Log File Contents)

The contents of a MyISAM log file are processed by the program myisamlog. it's the choices to perform a recovery operation, specify record position file and record position, perform associate update operation, show version data etc.

4. myisamchk (MyISAM Table-Maintenance Utility)

The myisamchk utility gets data concerning your info tables or checks, repairs, or optimizes them. The myisamchk utility works with MyISAM tables (tables that have .MYD and .MYI files for storing knowledge and indexes).

IV. Logs And File Analysis

Log files ar usually terribly massive and from time to time have complicated structure. whereas the method of generating log files is sort of easy and easy through the utility programs of the information, however its analysis can be an incredible task that needs monumental process resources, lasting and complicated procedures. The log files unendingly grow into large sizes that hold in most cases a bulk of the knowledge} with traditional operational data that is of less importance in associate analysis. Extracting the helpful info from logs required for the target analysis is one amongst the difficult and a tough task. rhetorical analysis procedures usually demand extracting info from the utmost log files and correlating them to possess a broader understanding of the case.

A framework is meant and planned for this analysis work. The system design for the logs and file analysis is shown below.

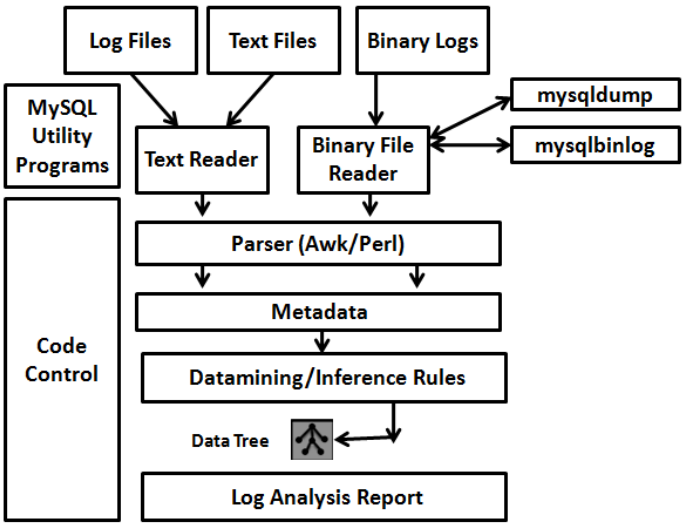

Fig. 3.) System Architecture for database forensic analysis: Stage I

Tasks to be allotted for information rhetorical analysis would be as:

1. Establish and collect the databases, log files, binary logs and text files at information server mistreatment MySQL Utility programs.

2. The MySQL utility programs ar accustomed dump the information to create a backup copy for analysis. The elaborate info like user access, timestamp, date etc. is to be copied. The parsers (Awk/Perl) are going to be accustomed scan large and multiple log files and text files. The essential operate of awk are going to be to go looking files for lines (or alternative units of text) that contain bound patterns. Once a line matches one amongst the patterns, awk performs specific actions on it line. It keeps process input lines during this means till it reaches the top of the input files.

3. Binary log info is 1st retrieved mistreatment MySQL utility programs like mysqlbinlog and is then given to the binary reader.

4. The information is collected to make data. It contains the gathering of facts from the extracted log knowledge/ system data.

5. data processing techniques is then accustomed build illation rules supported knowledgeable information to create the choices for extracting most relevant info from the designed up data [22]. Failure to keep inferences might delay the analysis in best case and can hamper it from generating any answer to the matter in worst.

6. Finally a close log analysis report is generated.

7. The planned framework would promote apply from its design to facilitate knowledgeable information dissemination. It once fashioned can provide economical information transfe rmechanism. this may facilitate in automation.

8. The script (control code) is written for this whole method.

V. Artifact Collection In MYSQL Server

A rhetorical methodology may be a logical and well-thought-out order of operations that's dead throughout a digital investigation. Rhetorical methodologies facilitate guarantee investigations square measure documented, repeatable, and dead in a very manner that's court friendly, ought to the collected information got to be submitted as proof in a very court of law [18]. MySQL Server artifacts reside among OS files and square measureas of memory that are expressly reserved for SQL Server use. These information facts will exist among giant, core MySQL Server files, like information information or dealings log files, or among smaller, less visible files. These artifacts kind the prime assortment of information that may be used for information investigation.

VI. Artifact Analysis

A good understanding of MySQL Server artifacts is needed to satisfy the target of associate degree investigation The analysis starts with a collection of log analysis report as shown in system design in Figure 3.

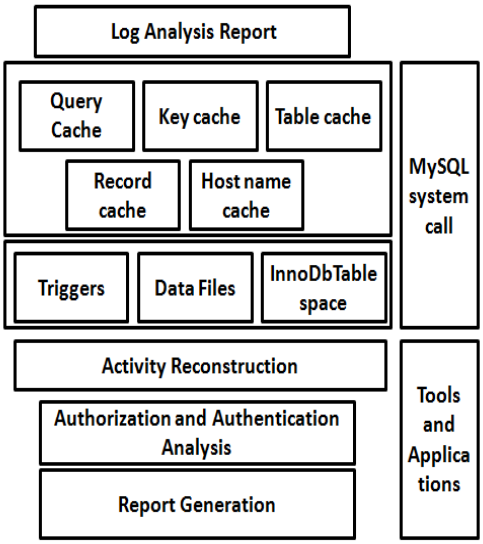

Fig. 4) System Architecture for database forensic analysis: Stage 2

A script would be written to scan and extract information from every log files. The output yields a tree containing the information extracted from the corresponding log file supported the abstract thought rules which might be later on applied to urge an ensuing tree that holds the facts that may be employed in more analysis or creating choices. The selections area unit then processed with the assorted artifacts known and picked up as shown in Figure five Stage two.

The artifacts (multiple cache, triggers, information files and InnoDb Tablespace etc.) area unit known, collected and valid against the Log analysis report. A rhetorical methodology is meted out for more investigations for Activity Reconstruction and analyzing the assorted attested and approved users. Finally a rhetorical report is generated with the tool and application developed

VII. Conclusion

The study concludes that there is no difference in buying preference in respect of Peraiyur taluk and Vadipatti taluk consumerskk.

In this paper, therefore we've given a framework for the aim of analyzing along with the reconstruction of the activity of any suspicious behavior occurring inside the information itself. The aim of the on top of paper description is to spot the specific drawback, analyze it, validate further as interpret and generate a rhetorical report thereby conserving the proof for the sake of digital investigations.

References

1. Sohail Imran, Dr. Irfan Hyder. (2009) "Security Issues in Databases", Second International Conference on Future Information Technology and Management Engineering, IEEE.
2. John oltsik. (2009), "Database security and Compliance Risk", ESG Market research study, Application Security, Inc.
3. U.S. health insurance portability and accountability act (HIPAA). Available at www.hhs.gov/ocr/hipaa
4. "MySQL 5.5 Reference Manual", www.dev.mysql.com/doc/refman/5.5/
5. Nina Godbole and Sunit Belapure. (2011) "Cyber Security, Understanding Computer Forensics and Legal Perspectives", Wiley-India. ISBN: 978-81-265-2179-1.
6. Jasmin Azemovic and Denis Music.(2010). "Methods for Efficient Digital Evidences Collecting of Business Processes and Users Activity in eLearning Environments", IEEE, 2010 International Conference on e-Education, e-Business, e-Management and e-Learning.
7. Article by David Litchfield (2011, August), www.darkreading.com/database-
8. security/167901020/security/attacks-breaches/231300307/database-forensics-still-in-dark-ages.html.
9. Article by David Litchfield (2011, August),
10. http://www.computerweekly.com/Articles/2007/08/03/225987/New-database-forensics-tool-could- aid-data-breach-cases.htm
11. Peter Frühwirt, Markus Huber and Martin Mulazzani, Edgar R. Weippl (2010), "InnoDB Database Forensics", 24th IEEE International Conference on Advanced Information Networking and Applications.
12. Paul M. Wright, (2005) "Oracle Database Forensics using LogMiner", June 2004 Conference, SANS Institute 2005

