

LogAuditor Enterprise: Integrated Management System for Various Log Data

Authors: Mitsunori Kori* and Takashi Fujimura**

1. Introduction

LogAuditor Enterprise provides integrated management of various logs generated by information systems for a company's internal control and security management. Mitsubishi Electric has used its high-speed processing technologies to achieve integration, high-speed accumulation and searching of logs of different formats, while reducing the storage capacity required, which were hard to realize in the past. Analysis templates for outputting audit reports are also available.

2. Problems Related with Log Management

Amid the increasing interest in internal control and security management among corporations, the logs generated by various information systems need to be stored as evidence. However, the volume of such logs may reach tens of terabytes per year. And whereas logs used to be managed for each information system, today integrated management is required for reducing management costs and increasing the efficiency of problem analysis.

In conventional log management, a general-purpose RDB (Relational Database) was often used. However, since RDBs were developed for applications based on OLTP (On-Line Transaction Processing), they have various formats and are not suitable for efficiently processing logs that contain huge volumes of data¹. As the types and volumes of logs have increased, log management using RDB involves the following problems.

- The data formats must be unified beforehand and logs of formats not specified in advance are difficult to handle.
- The time required for processing related to log accumulation or log search is too long.
- The cost of long-term storage is very high.

3. LogAuditor Enterprise

LogAuditor Enterprise solves these problems and provides integrated management of various large logs. LogAuditor Enterprise generally consists of LogAuditor/PSF (Power Staging Facility) which imports logs, LogAuditor/LDB (Log Database) which stores and

monitors logs, and LogAuditor/AQL (Analytical Query Language) which is a log analysis engine. A Microsoft Excel add-in is provided as a front end for analysis. Table 1 shows the operating environment of LogAuditor Enterprise.

Table 1 Operating environment of LogAuditor Enterprise

Server	Microsoft Windows Server 2003
Client	Microsoft Windows XP Professional Microsoft Windows 2000 Professional

LogAuditor/PSF collects and processes various types of a company's internal log data, and has the following features:

- Fine and detailed data processing and edit functions
- High productivity and maintainability
- Major RDBs and CSVs as data sources can be applied.

LogAuditor/LDB² is a new type of database that can accumulate given logs, and has the following features:

- Logs, regardless of type, can be gathered and stored in a manner that allows the original logs to be restored completely. Particular log types need not be specified in advance.
- High-speed accumulation of terabyte-size logs and high-speed searching by regular expression specification
- Storage volume is reduced by data compression. Time-series management such as back-up or deletion of logs based on ranges such as daily or the like (patent pending)

LogAuditor/AQL is a database management system suitable for data compilation and analysis, and has the following features:

- Logs are held as structured data suitable for compilation and analysis
- High-speed data search and compilation
- Reduced required storage volume by data compression
- Conformance to standard SQL (Structured Query Language)

The analysis front end is an add-in tool that directly

¹ LogAuditor is a trademark owned by Mitsubishi Electric Information Technology Corporation.

² Microsoft, Excel, Windows, Windows Server 2003, Windows XP, and Windows 2000 are trademarks owned by Microsoft Corporation, U.S.A.

produces analysis reports by Microsoft Excel, and has the following features:

- Preparation and use of analysis templates
- Flexible atypical analysis and easy-to-use wizard type operation method
- Seamless operation from Excel and automatic generation of summary sheets
- Drill-through function to move from summary values to breakdown analysis data
- Linking of primary log data

4. High-Speed Processing Technology of LogAuditor Enterprise

LogAuditor/LDB and LogAuditor/AQL offer high-speed processing of large logs thanks to Mitsubishi's unique large-scale data high-speed processing architecture SISA (Scalable Intelligent Storage Architecture).

In both LogAuditor/LDB and LogAuditor/AQL, logs are automatically compressed to reduce the required storage volume to about 1/10 or less of the standard. In addition, storage input/output is reduced by data compression, thus increasing the speed of accumulation and searching. Figure 1 shows an example of the reduction of data volume when a PC operation log is stored in LogAuditor/LDB. Compared to the RDB, the storage volume is reduced to about 1/23.

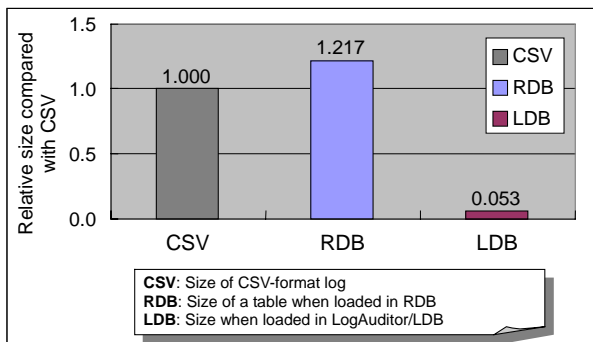


Fig. 1 Reduction of data by data compression (LogAuditor/LDB)

Both LogAuditor/LDB and LogAuditor/AQL execute compression, extension, and searching by parallel processing using multiple processors, distribute the data to multiple storage devices, and perform input/output operations in parallel. As a result, the system is highly scalable in accordance with the log volume. Figure 2 shows an example of the full search performance of LogAuditor/LDB with PC operation logs.

LogAuditor/LDB can extract logs by judging the log types upon log accumulation, without having to specify

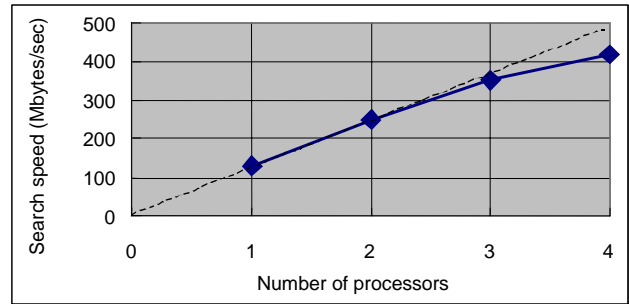


Fig. 2 Log full search performance (LogAuditor/LDB)

the log types prior to log accumulation. In the conventional character-string pattern matching method, the complicated pattern matching process was too slow for judging the type of log. But with Mitsubishi's own sDFA (size-reduced Deterministic Finite Automaton)³ (patent pending) technology, a high speed of approximately 100 million characters/sec. is attained regardless of the search condition, thus solving the speed-related problem (see Fig. 3). Indexing is often used for boosting the speed of database searching, but indexing is unsuitable for log management because it decreases the accumulation speed and increases the storage volume. With LogAuditor/LDB, all requirements regarding accumulation speed, storage volume, and search speed are satisfied by high-speed character-string pattern matching technology.

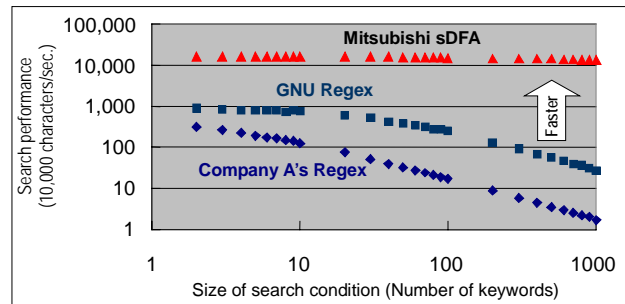


Fig. 3 Comparison of text search performance

5. Application Examples of Integrated Log Management Solution

As a solution making use of LogAuditor Enterprise, we provide an "Analysis Template" which outputs audit reports for internal control based on the corporate business operation flow execution log, PC operation log, file server access log, and the like. Analysis Template outputs Microsoft Excel type audit reports in accordance with the definitions of the structure of the integrated log DB, structure of the data mart for log analysis, and the log import style.

It is difficult to intuitively grasp a huge volume of information contained in a log. For example, an increase in the number of accesses to confidential files or its relationship with the execution status of business operation flow cannot easily be identified by viewing the logs of the file server and business applications. The audit reports are designed to present tables and graphs as reports by visualizing the intangible log data (see Fig. 5). Access to confidential files and business operation status with respect to users are clearly recognized and can be used for verifying or reviewing security management measures. Furthermore, causes can be analyzed in detail from a required position in the audit

report by using the "Search back in log breakdown" function.

Integrated Log Management Solution is useful for information security management, internal control, information infrastructure management, information system implementation management, and many other fields.

6. Outlook

We plan to expand the scalability and improve the Analysis Template to handle larger logs and diversified log types.

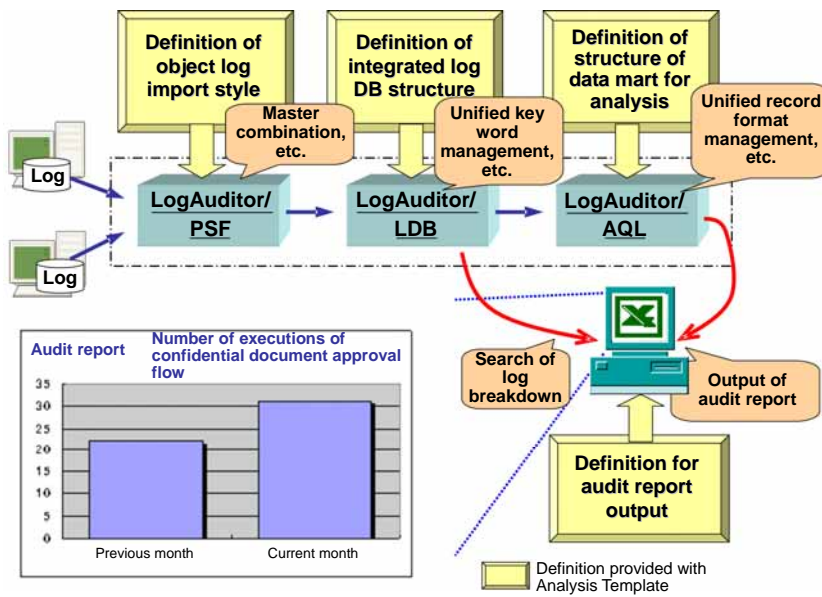


Fig. 4 Structure of analysis template

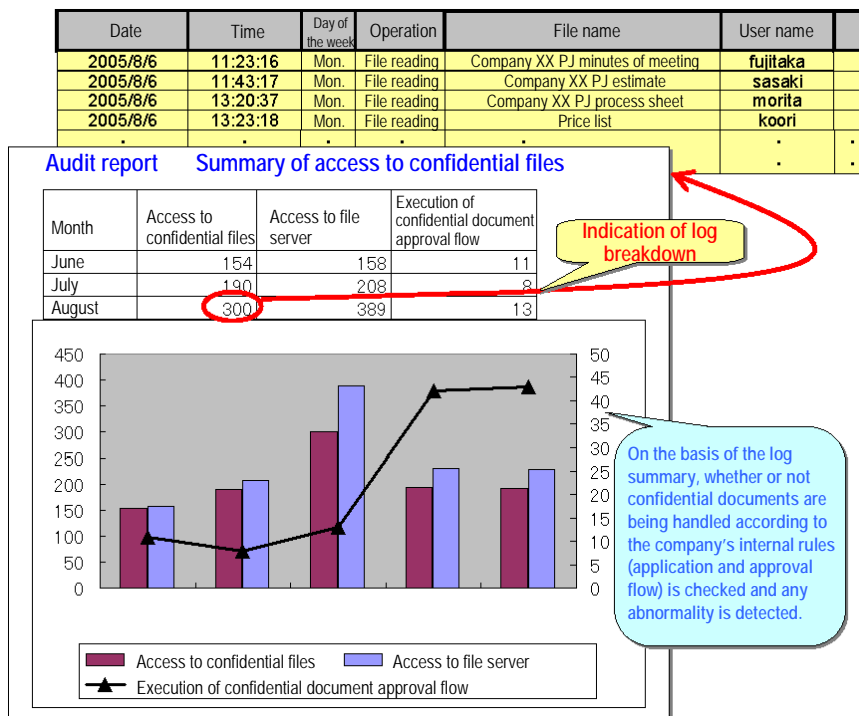


Fig. 5 Example of audit report

References:

- (1) Sah, A.: A New Architecture for Managing Enterprise Log Data, Proc. of LISA 2002, 121 to 132 (2002)
- (2) T. Nakamura, et al.: Realization of Large-Scale Log Database, 68th Information Processing Society National Conference, ID-2 (2006)
- (3) T. Nakamura, et al.: High-Speed Pattern Matching Method for Large-Scale Regular Expressions, 67th Information Processing Society National Conference, 4F-5 (2005)
- (4) T. Fujimura, et al.: Compliance Promotion Solution to Support Information Risk Management and Internal Control, Mitsubishi Electric Corporation Technical Report, 80, No. 4, 281 to 284 (2006)