Data Security Threats of Log Aggregation

Ruiyang Liu*, Senior Engineer
Guangzhou Huali Science and Technology Vocational College, 7 Huali Road, Guangzhou Huali Science Park, Zengcheng District, Guangzhou City, Guangdong Province, CN 511300;

Abstract: The development of cloud computing and distributed technologies has shifted log management operations to an aggregated form. Diverse information devices as well as information systems perform data collection and status analysis through a unified log aggregation management system. However, with the increasingly diverse business forms, the threats confronted by log aggregation systems are also mounting, among which data leakage, data tampering, data loss, and security vulnerabilities dominate the main positions. Data leakage imposes a severe challenge to the data privacy of enterprises and users, data tampering misleads the information system of enterprises and even adversely affects the decision-making of enterprises, data loss devastates the past data of enterprises and impinges on the operation of enterprises, and security vulnerabilities inflict potential safety hazards to the information systems accessed to log aggregation.

1. INTRODUCTION

In line with the continuous development of information technology, cloud computing has gradually occupied the market share of traditional servers [1]. Traditional local log systems of the server have various defects, which is difficult to meet the daily management needs of cloud computing and distributed technology. In addition, the amount of data generated by computer systems is increasing, and the amount of data and business forms are becoming more and more diversified. This new information technology development makes the requirement for efficient and secure log management become an important issue [2]. In order to adapt to cloud computing and distributed processing business forms, the log management business has changed from a localized management form to a unified management form, namely, a log aggregation system [3]. Log aggregation works by collecting and analyzing logs from various sources and extracting meaningful information from them for various troubleshooting, security monitoring, and compliance auditing.

Log aggregation systems can not only adapt to cloud computing and distributed technology but also usher in a great many new functions and benefits. However, the emergence of new functions also means various security risks. This paper focuses on security issues related to log aggregation, including data privacy, integrity, availability, and confidentiality. The purpose of this paper is to make a comprehensive analysis of these problems in order to ensure the safe and reliable operation of the log aggregation system.

The rest of this paper is described below. Section 2 outlines the importance of log aggregation in modern computing environments. Section 3 discusses the security challenges associated with log aggregation, including data privacy, integrity, availability, and confidentiality. Finally, this paper provides a summary of this research and overviews future research directions.

2. IMPORTANCE OF LOG AGGREGATION

Log aggregation collects and analyzes logs from different information devices, and extracts the status information of information devices for fault diagnosis, safety monitoring, and compliance audit [4]. Log aggregation is becoming more and more important in cloud computing. In the distributed environment of cloud computing, distributed computer information systems will produce a large amount of data, and the log data of information equipment contain important information, which can provide powerful support for information security and business decision-making of enterprises.

Real-time analysis of log aggregation data allows enterprises to monitor the real-time running status of information systems and identify potential failures and security threats [5]. In addition, log aggregation can also improve the efficiency of enterprise operation and maintenance, and simplify the process of log management and analysis by analyzing the status of information equipment and reasonably scheduling the software and hardware of the information system.

The importance of log summary is not only embodied in the internal of the enterprise, but also in the...
relationship between the enterprise and the outside world. For example, regulators require companies to report periodically on the operations of their enterprise information systems, and these reports must be based on a large amount of data generated by the systems. In addition, in financial, medical and public service industries, information security required by laws and regulations also needs to use a log aggregation system to provide data support.

However, the log aggregation system itself entails some security problems. Since logs contain sensitive business information from different information systems, data privacy and confidentiality are important issues. At the same time, the log aggregation system also needs to ensure the integrity and availability of data to ensure that the data is not tampered with or lost, which may affect the security of enterprises and business operations.

3. DATA SECURITY FLAWS OF LOG AGGREGATION

The importance of log aggregation systems in the information environment has been discussed earlier because log aggregation systems provide in-depth analysis and monitoring of computing systems and network activity. Information collection and data management of log aggregation systems are also facing challenges in data security. Data security vulnerabilities of log aggregation systems include data leakage, data tampering, security vulnerabilities, and data loss. Data leakage will lead to the disclosure of confidential information, privacy violation, identity theft, etc. Data tampering will damage the integrity and reliability of data, security vulnerabilities will lead to unauthorized access and attacks. Data loss will lead to unrecoverable data and damage data integrity. Data security flaws in log aggregation systems need to be taken seriously and solved so that data and system security in computing environments can be better protected.

3.1 Data leakage

When an enterprise prepares to launch a log aggregation system, it needs to meticulously contemplate the data leakage problem faced by the log aggregation system. A log aggregation system is one of the important data sources of a company. It contains a large amount of sensitive data, such as company information and user information. If these data are accessed by unauthorized third parties, substantial harm can be wreaked on enterprises.

3.1.1 Technical hazards of data leakage

First, the data leakage of the aggregated log system is likely to incur the disclosure of user privacy. Log aggregation systems may contain sensitive information about users, such as user names, passwords, personal information, and so on. If this information is leaked, users' privacy will be violated, which may lead to users' dissatisfaction and resistance. Especially for some businesses involving users' personal property and safety, such as e-commerce, banking, medical care, finance and other industries, the disclosure of users' privacy may bring losses to users and damage the reputation of the enterprise, thus affecting its business and market position.

Second, the data leakage in the aggregated log system may also lead to the disclosure of enterprise secrets. Data in an aggregated logging system may contain sensitive enterprise information, such as business processes, customer information, and trade secrets. If this information is leaked, corporate secrets could be unveiled, causing irreparable losses. For some sensitive enterprise information, such as financial data, R&D projects, core technologies, etc., information leakage may cause a huge blow to the company's commercial value and market competitive advantage.

Thirdly, the leakage of log data collected by aggregation log system may also cause interference to enterprise business, which makes enterprise business unable to run normally, thus affecting enterprise productivity and economic interests. An enterprise information business system is usually composed of multiple components, some of which are supported by collected log data. If these data are leaked, the information business system will fail to operate normally, thereby affecting the productivity and economic interests of enterprises.

3.1.2 Business risks of data leakage

The consequences of a data leakage go beyond damaging a company's reputation; a breach of user data can also lead to a loss of trust among customers, who may switch to a competitor's service, putting a damper on the company's business and revenue [6].

First, data leakage not only damages the reputation and finance of the company, but also harms users themselves due to information disclosure. For example, the compromised data may contain personal identification information, such as name, address, date of birth, national insurance number, etc., which can be used for identity theft and fraud. If the data contains sensitive personal physical health information or financial information, such as medical records or credit card information, the user's health and financial security will also be threatened.

Second, data leakage may also expose companies to legal liability. In some cases, the leaked data may contain private personal information, such as gender, sexual orientation, political beliefs, etc., which is considered to be protected by law in some countries and regions. If the leaked data is related to this information, the company may face huge fines. For example, according to the new data protection regulations of the European Union, called GDPR, the company may face fines as high as several million yuan, or the company may even confront legal proceedings and investigations [7].
3.2 Data tampering

3.2.1 Technical hazards of data tampering

With the continuous development of information technology, the technical structure of information systems is becoming more and more diversified, and various log systems are extensively used to adapt to various newly developed information system structures. An aggregation log system is one of the common log management methods. Because of the working mode of the aggregation log, it is necessary to collect a large amount of data from information business systems, and then analyze and manage it uniformly. Therefore, the data of aggregated log systems face the risk of being tampered with, and it is a security threat to the information system that cannot be ignored, which will bring serious consequences to enterprises and organizations.

First, if the data in the aggregated log system is tampered with, three is a high possibility to be stuck in information leakage and privacy leakage. Hackers can use tampered log data to obtain sensitive information, such as login credentials, customer names, and addresses. The disclosure of these pieces of information will bring serious economic losses and reputation risks to enterprises or organizations.

Second, tampering with the data in the aggregated logging system can lead to misleading and incorrect decisions by the enterprise. When decision-makers of the enterprise make decisions based on incorrect data, organizations may suffer heavy losses [8]. For example, decisions based on tampered log data deviate from the true state of the information system, so decisions will lead to financial losses, security vulnerabilities, and confidentiality issues.

Third, tampering with the data in the aggregated log system will lead to business interruption and system crashes of other information systems, thus disrupting the availability of business processes and business or organizational services [9]. In some cases, hackers can use tampered log data to disturb the information system. When the information system is misled by the tampered data and then makes wrong decisions according to the wrong data, it is possible to cause the collapse of the whole system, resulting in business interruption and economic losses.

3.2.2 Business risks of data tampering

On the one hand, tampering with aggregated log system data can also lead to compliance problems. Many organizations must comply with certain regulations and industry standards, such as Health Insurance Portability and Accountability Act (HIPAA), Payment Card Industry Data Security Standard (PCI DSS), and other specifications. Each industry and region in which the organization is located has its own regulatory standards, and enterprises must meet these provisions. If hackers use tampered log data to steal personal identification information, credit card information, or other sensitive data protected by laws or regulations, organizations may be saddled with costly fines and legal proceedings [10].

On the other hand, due to the particularity of the aggregate log system, a large number of information system devices are needed to connect with the aggregate log system. In consequence, tampering with aggregate log data can also lead to an overall system attack. For instance, an attacker can spoof a malicious request and inject it into the aggregate log to make the system administrator think it is a normal user request. When the administrator makes a decision based on this fake data, this wrong decision will lead to system failure or data leakage. By the same token, attackers can tamper with key data such as login information and permission information in logs, thus realizing malicious behaviors such as system access overwrite and data theft. These behaviors will cause the information system of the company to fail to operate normally, and then affect the profitability of the daily operation and commercial operation of the company.

To put it in a nutshell, tampering with the data of an aggregated logging system is a serious hazard that can lead to reduced system security, data leakage, unauthorized access, and other problems. In order to protect the security and stability of the system, it bears great significance to strengthen the attention and security management of log data, continuously improve the awareness of data security and technical level, and ensure the safe and stable operation of the system.

3.3 Security vulnerability

3.3.1 Technical hazards of security vulnerability

With the increasingly widespread application of aggregated logs and the growing number of information systems connected to aggregated log systems, the importance of aggregated log security issues looms larger and larger. In the running process of a log aggregation system, there may be various security vulnerabilities on account of the system and human factors. Once these vulnerabilities are exploited by attackers, they will pose a great threat to the stability and security of the system [11].

First, serving as an independent information service system, an aggregate log system needs to be built on the general processor, operating system, and other software and hardware, all of which have potential undiscovered zero-day vulnerabilities. These undiscovered zero-day vulnerabilities make the security of the summary recording system greatly threatened.

Second, owing to the particularity of the aggregate log system, the aggregate log system needs to interact with a large number of information systems. In the process of system docking, since software and hardware architectures of different business systems vary a lot, the aggregated log system needs to use a variety of software and hardware interfaces to collect data from different business systems, and complex and diverse data interfaces can easily become the weak link of the whole information system [12]. The information between the
aggregate log system and other business systems is exposed to various attack threats.

Thirdly, the aggregation log system collects and analyzes the status of each business system, and provides data support for information system administrators to maintain information systems. An attack inflicted on the aggregate log system may mislead administrators to make wrong judgments and eventually paralyze the business system. Because aggregate log system usually runs in a distributed environment, there are various vulnerabilities, such as DDoS attack, cache overflow, and so on. It is available for attackers to use these vulnerabilities to destroy the stability of the system, thus causing system paralysis and affecting the normal use of users.

By and large, the security vulnerability of aggregation logs may cause the risk of enterprise information system intrusion and business halt. During log aggregation, sensitive information, such as business rules and processes, can be leaked to people who shouldn't know it.

3.3.2 Commercial risks of security vulnerabilities

First, the security vulnerabilities in the aggregate log system will bring huge economic losses to the company. When attackers have unauthorized access to sensitive data stored in aggregated logs, they can use this information to steal money or valuable assets. For example, if an attacker gains access to a company's financial logs, they can use that information to steal money or engage in fraudulent activities.

Second, aggregate logging systems often contain sensitive information about users, customers, and employees. If this information is exposed, the reputation of the company may be seriously damaged. Customers may lose trust in the company, while employees may become demoralized or even resign. In addition, if companies fail to comply with data privacy regulations, they may face legal penalties.

Third, the security vulnerabilities of aggregated logs may damage the integrity and availability of company information systems, and then affect the profitability of enterprises. It is possible for an attacker to exploit the data in the aggregated logs to launch further attacks on the company's information system, disrupting their services and causing significant downtime.

Finally, the security vulnerabilities in aggregation logs will have a wider social impact. If an attacker obtains a log containing sensitive information, they can use it for espionage or other malicious activities. The security vulnerabilities of aggregated logs pose a serious threat to companies, individuals, and society as a whole.

3.4 Data loss

3.4.1 Technical hazards of data loss

First, data storage depends on information hardware and software, which need to be stored in specific places. Different locations will encounter different natural disasters, such as floods, earthquakes, hurricanes, etc. These disasters will destroy physical infrastructure, including servers and data centers, resulting in data loss. These sudden natural disasters are generally accompanied by data loss and damage in data center computer rooms [13]. Hence, data loss in natural disasters is a huge loss for enterprises, and it is very difficult to recover information systems.

Second, in the running process of hardware equipment, there is a small probability of failure. Occasional hardware failures, such as hard disk failure, storage device failure, array card failure, etc., may cause data loss. When the information system is intact, the data loss in the information system may affect the daily operation of the information system. When information systems do not rely on past data, data loss may affect access to historical data. When information systems rely on past data, the loss of data may affect the current operation of information systems.

Third, information systems are often attacked by network attacks, such as malicious software, phishing, and ransomware. It is difficult for information systems to be completely immune to frequent network attacks, so network attacks are also one of the important reasons for data loss. In the face of data loss caused by cyber attacks, the common danger is data kidnapping and extortion. Most well-known enterprises need to worry about business operating costs, corporate reputation, and other issues, and are objectively willing to pay large sums of money.

Fourth, human factors are another important cause of data loss, because employees may accidentally delete or overwrite data, resulting in unrecoverable data loss. Even in large groups with good rules and regulations, the business processes caused by human factors are not standardized, which makes data security still full of uncertainty [14]. Especially those employees who do not operate according to common sense or have ulterior motives may bypass the original rules and regulations, and damage the data in the information system, resulting in data loss. In the case of artificially caused data loss, if it is unintentional misoperation, data corruption is determined by the design of the data protection framework. In the case of artificial data corruption, data corruption is driven more by the operator's comprehension of the data protection framework.

3.4.2 Business risks of data loss

First, the loss of data will bring huge economic losses to enterprises, and enterprises rely on data to make correct decisions, whilst the loss of these data will lead to a lack of basis for making decisions, which will make companies unable to make accurate decisions, and then suffer economic losses [15]. In addition, the cost of recovering lost data can be enormous, because companies may need to invest in data recovery tools and services, and even rebuild their infrastructure.

Second, data loss will damage the reputation of the company. Organizations that experience data loss events may lose the trust of their customers, who may consider
them unreliable or untrustworthy. This loss of trust will contribute to the loss of customer loyalty, a decrease in sales volume, and a decline in market profitability, at the expense of ultimately damaging the company’s reputation.

Third, data loss will have significant legal consequences. Organizations experiencing data loss incidents face lawsuits from customers, employees, or regulators if they fail to adequately protect sensitive information. In addition, companies that encounter data leakage may face huge fines and penalties for violating data privacy regulations.

Fourth, data loss will endanger the safety and security of the personal information of users and employees. When sensitive data is lost, it can be used for identity theft, financial fraud, and other malicious activities. Especially in recent years, telecom fraud makes extensive use of personal information to gain trust and commit fraud.

4. CONCLUSION

Data security in aggregate logs is an incredibly important topic. Data leakage, data tampering, and data loss will be inevitably subject to serious consequences. First of all, data leakage will lead to the abuse of information, including theft, malicious attacks, and false advertisements, which has a great impact on the reputation and property of users and enterprises. Secondly, data tampering may lead to inaccurate and unreliable information, which may affect the accuracy and reliability of business decisions and processes. Finally, data loss will lead to business interruption, service interruption, and information loss, resulting in economic losses and customer churn.

On all accounts, the data security of collected logs is a complex problem, which needs to be solved by various measures. Only by taking comprehensive and enhanced data security measures can we ensure the data security of collected logs and reduce the risks of data leakage, tampering, and loss, thus protecting the legitimate interests of individuals and organizations.

REFERENCES

1. Fang Z. L. In 2022, the market share of cloud computing of the three major operators increased rapidly [N]. People's Post and Telecommunications, 2023-02-09 (007). DOI: 10.28659/N.cnki.nrmyd.2023.000365.