IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 8, August 2024

DOI: 10.17148/IJARCCE.2024.13818

EXPOSED VULNERABILITIES OF DATA BACKUP

SARTHAK SANGARE

Independent Researcher

Abstract: Data backups have immeasurable importance in data management. They help protect important data against human errors, viruses, and other such threats. Therefore, unauthorised access to the contents of a backup can prove extremely harmful. Finding vulnerabilities and points of exposure of the backups is crucial to ensure better security. Aim of this paper is to discuss the vulnerabilities regarding the accessibility of data backups, be it cloud or a physical drive. Through surveying the internet forums, we can find out various methods of entry via CVEs. The paper will list the techniques to access data backups as an adversary and then provide potential solutions to each technique.

Keywords: Common Vulnerabilities and Exposures (CVE), Backup, Encryption, Compression

I. INTRODUCTION

A data backup is an archived and compressed copy of an IT system, including protected hidden sensitive files such as system registry files and application configuration files, in a separate location. Backups are useful in case, if any original data is compromised for a company or a user. This is the reason many adversaries target backups frequently. Thus, it is imperative that data backups are protected as they are the last defence against system vulnerabilities, such as human error, malware, or other conditions.

METHODS FOR SERVER BACKUP

Based on the usage of the computers in an organisation, a backup can be configured through the methods listed below:

- **Full Backup:** A full backup is a backup that consists of all the files and folders which were selected for backup. These backups are usually performed first, followed by differential or incremental backups. Full backups occupy a maximum amount of space compared to other backups.

- **Incremental Backup:** Incremental backups save storage space as they allocate backup space for files which were not present in the last backup. This method is also conducted alongside full backup. However, incremental backups demand high network consumption.

- **Differential Backup:** Differential backups are an in-between of incremental backup and a full backup. Differential backup include all the aggregate increments since the most recent full backup. Thus, a differential backup takes less space than a full backup but takes more space than an incremental backup.

- **Image Backups:** This method creates a full disk backup of the entire system, which comprises one single file called Image. This can be helpful when one wants to restore the old system to a new system. They also are the fastest recovery option when restoring an entire system.

- **Copy Jobs:** Copy jobs copies every file and folder in their native uncompressed file format. This method is simple as it literally involves copying and pasting files and folders as backup.

COMMON THREATS AND ATTACKS ON DATA BACKUPS

Due to an increased dependency on technology, technical errors are becoming more commonplace. Despite the comfort that a server backup provides, the entire system can still be compromised. The hackers have caught on and are now frequently targeting backups [5].

Some of the most commonly used attacks are listed as follows:

DOI: 10.17148/IJARCCE.2024.13818

- **Privilege Abuse:** People with higher access to backups can abuse the system. Privilege abuse in databases refers to the improper use of rights to access, manipulate or to perform actions on data beyond the designated capacity.

- **SQL Injection:** SQL injection targets data backup through vulnerabilities. Such an attack occurs when a malicious SQL inquiry is executed in the database. SQL injection can be fatal for the data, such as; unauthorised access, loss of data integrity, corruption of data and loss of control of the database.

- **Unprotected Networks:** Such networks in a database refer to lack of network security measures. These measures are supposed to safeguard data from client servers to database servers. These databases can be modified, accessed, or intercepted by the adversaries.

- **Backup Data Exposure:** Backup data exposure is about the risk of unauthorised access to database backups. The backups are critical for recovery and continuity, thus they become a prime target for adversaries.

- **Defects:** Defects in backups allude to the flaws in the backup systems, data or processes that can hamper the effectiveness and question the reliability of the saved data. The defects hamper data recovery; cause data loss, data leaks and data manipulation.

COMPRESSION

Data Compression is the process of modifying data to reduce its overall size. Such a result is achieved by re-coding the data by using fewer bits to represent data. Compression methods consist of two types - Lossy and Lossless.

- **Lossy Compression:** Discards unnecessary or excess data to reduce complexity of the information, thereby reducing the size of the file. Lossy compressions reduce more data size compared to lossless compressions.

- **Lossless Compression:** Reduces data size without losing any data from the original information, usually by eliminating statistical redundancies. While the original data is wholly preserved, the amount of reduction of data size is less than a lossy algorithm.

Data backups generally consist of critical information. It is of utmost importance that the data is preserved. Therefore it is recommended that a Lossless compression method is used as follows [3]:

- **LZ4:** It is a Lossless compression known for its compression and decompression speed. It uses a dictionary-matching stage only. LZ4 reduces compression time greatly, which makes it ideal for incremental backups.

- **ZStandard:** ZStandard, or zstd, is a Lossless compression algorithm that was designed to provide a compression ratio like that of the DEFLATE algorithm. Zstd is tuneable with compression levels ranging 22 (the slowest but provides the best compression ratio) and negative 7 (fastest). Zstd also has an ‘adapt’ feature which can regulate the compression speed based on the speed at which it can write the output.

- **XZ:** XZ is Lossless data compression command-line software. It can compress and decompress xz and lzma files. It has a better compression rate compared to gzip and bzip2.

- **GZip:** Gzip is software used for compression and decompression. Gzip is based off DEFLATE algorithm. It uses gzip file format, and it compresses data by representing repeated sequences with shorter sequences.

- **Bzip2:** Bzip2 is an open-software compression program that employs Burrows-Wheeler. While it only compresses a single file, the problem can be fixed by third party extension. Bzip2 is particularly good for text files.

ENCRYPTION

Encryption is the process of manipulating a message such that only authorised persons can read the message. Encryption is a subset of cryptography. Following are the common methods of encryption used in backups [2]:

- **AES:** AES, or Advanced Encryption Standard, is a variation of the Rijndael block cipher. It is a NIST approved encryption method. AES can adapt to a range of key sizes while being quicker than legacy algorithms like DES. AES is one of the most trusted algorithms.

DOI: 10.17148/IJARCCE.2024.13818

- **DES:** DES, or Data Encryption Standard, uses a symmetric key algorithm to encrypt data. While it is not the most secure method of encryption as its key length is only that of 56 bits, it is still a viable method for physical appliances. DES was made for hardware; thus, would perform well on a physical copy of backup.

- **RSA:** RSA, or Rivest–Shamir–Adleman, is a public key cryptosystem. It is one of the most frequently used encryption techniques. RSA is a slow algorithm and is primarily used for bulk encryption. There are no proposed methods to defeat the system if a large enough encryption key is used, which makes RSA one of the best encryption methods.

II. METHODOLOGY

Ransomware frequently targets data backups as security best practices are generally neglected [1]. The methodology used in this paper consists of a qualitative approach [7], where all the samples were collected based on data backup attacks that occurred from the year 2020-2024.

The paper uses these findings as the basis for analysing common behaviour of the adversaries. The process began with a survey categorising backup CVEs namely; software vulnerability, hardware vulnerability and service vulnerability. This observation is presented in tabular format (Table 3) and graphs (Fig 1 and Fig 2). Mitigation policies are further proposed followed by conclusion with an outcome and future scope.

III. OBSERVATION

Based on the survey [6], the CVEs have been categorised year wise.

Year wise details (till June 2024)	CVEs in total	Top 5 CVEs
2024 (till 20th June)	70	CVE-2024-5947, CVE-2024-5599, CVE-2024-5551, CVE-2024-5264, CVE-2024-4469
2023	122	CVE-2023-7236, CVE-2023-7232, CVE-2023-7204, CVE-2023-7201, CVE-2023-7165
2022	89	CVE-2022-4932, CVE-2022-4931, CVE-2022-48482, CVE-2022-47911, CVE-2022-47732
2021	84	CVE-2021-47164, CVE-2021-46960, CVE-2021-46957, CVE-2021-45732, CVE-2021-44255
2020	102	CVE-2020-9474, CVE-2020-9289, CVE-2020-8427, CVE-2020-7912, CVE-2020-7241
Previous Years (1999-2019)	694	-

Table 1: Year-wise CVEs

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 8, August 2024

DOI: 10.17148/IJARCCE.2024.13818

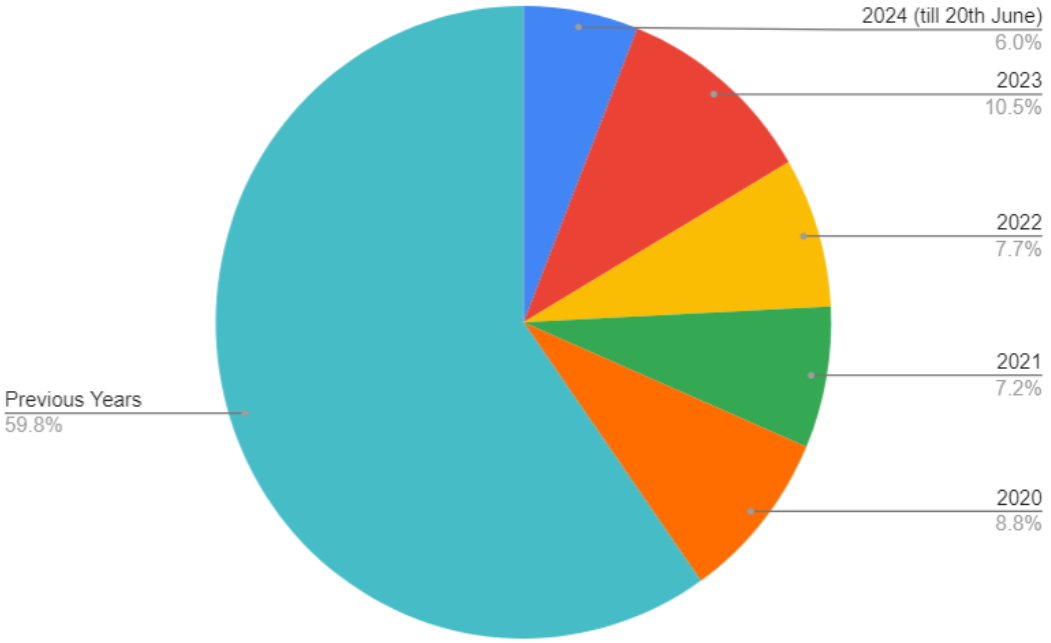

Fig 1: Total CVEs year-wise

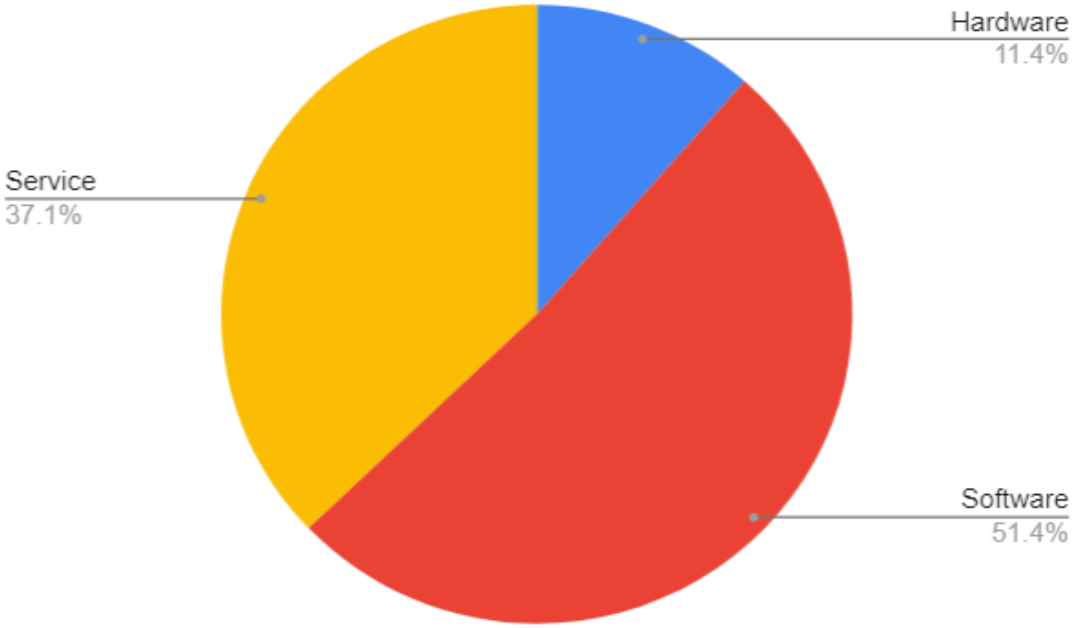

Fig 2: Categorization of 2024 CVEs (till 20th June 2024)

Maximum vulnerabilities, about 51.4% of the cases occur due to a software bug or an error. Errors have been found in plug-ins, OS and external backup software. However, WordPress plug-ins errors happen to dominate the number of cases.

IV. MITIGATION POLICIES

As a result of the analysed data, following mitigation strategies are proposed in accordance with the 2024 survey of the CVEs [6] and ISO 27001 standard [4]:

DOI: 10.17148/IJARCCE.2024.13818

- Backup data should be encrypted, and key management services (KMS) are recommended to manage the keys.
- Multiple authentication processes such as Two Factor Authorization should be used.
- Frequent backups should be taken with an immediate verification via checksum validation.
- Software must be updated regularly, while hardware should have planned checkups.
- Backup data ought to be compressed first and then encrypted using the RSA encryption method.
- Cloud backups should take a snapshot every hour and the snapshots should be retained for at least two days.
- Logs should be audited, and access should be provided based on the principle of least privilege.
- Each backup update should undergo a restoration test.
- It is recommended to attempt penetration testing on backups, especially plug-ins, to ensure security.
- Frequency as well as the type of backup should be based on the organisation's requirement.
- It is recommended to follow a 3-2-1 plan. The plan recommends having 3 copies of all the data, 2 of which should be on different storage media, and 1 of which should be offsite – physically distant at least a few kilometres away.
- Physical copies of the backup should be kept under constant surveillance in a safe location.
- Multiple Backups should be made to offset any redundancies in a copy.
- Storage media should be protected from electromagnetic radiation via tinfoil casing.

V. CONCLUSION

Backups are vulnerable as evident by the sheer amount of CVEs present. Hackers are identifying these CVEs and making their way into backups. Timely intervention as well as regular auditing is recommended to minimise risks from various contingencies. Based on the CVEs, the results have been analysed till 20th June 2024 and solutions have been proposed accordingly.

VI. FUTURE SCOPE

The survey can be extended to specific factors as discussed in the paper; such as software vulnerabilities of backups. The research can be expanded to more avenues like Blockchain to improve backup integrity and AI monitoring of backups.

ACKNOWLEDGEMENTS

I would like to thank **Dr. Anup Girdhar**, CEO - Seduility Solutions & Technologies and Editor-in-Chief at Cyber Times Newspaper, Delhi, for his insights into the subject matter. His pointers were instrumental in shaping my paper.

My gratitude also extends to **Dr. Manisha Nene**, Director - School of Computer Engineering and Mathematical Sciences at Defense Institute of Advanced Technology (DIAT), DU, DRDO, Pune, for her guidance and constructive feedback. Her comments were critical in enhancing the quality of my research. I am grateful to **Mr. Sunil Hinge**, Delivery Head - ICT Sector at British Standards Institution (BSI), Mumbai for his expert guidance on ISO 27001 ISMS Standard.

REFERENCES

[1] Computer Weekly, 2023, *"Almost All Ransomware Attacks Target Backups, Says Veeam: Computer Weekly."*,www.computerweekly.com/news/366538492/Almost-all-ransomware-attacks-target-backups-says-Veeam, [Accessed 26 March 2024]

[2] Cloudian, 2024, *"Data Encryption: The Ultimate Guide"*, https://cloudian.com/guides/data-protection/data-encryption-the-ultimate-guide/, [Accessed 31 July 2024]

[3] Datamation, 2023,*"What is Data Compression & How Does it work?"*, https://www.datamation.com/big-data/data-compression/, [Accessed 31 July 2024]

[4] Info-Savvy, *"ISO 27001 Annex: A. 12.3 Backup"*, https://info-savvy.com/iso-27001-annex-a-12-3-backup/, [Accessed 2 August 2024]

[5] LinkedIn, 2023, *"Why your backup server is a Hacker's first target"*, https://www.linkedin.com/pulse/why-your-backup-server-hackers-first-target-toss-corporation#:~:text=Why%20Are%20Backup%20Servers%20So,essential%20to%20an%20organization's%20operations [Accessed 24 March 2024]

[6] Mitre, 2024, *"Search Results - Backup"*,https://cve.mitre.org/cgi-bin/cvekey.cgi?keyword=backup, [Accessed 15 June 2024]

[7] Scribbr, 2023, *"What is qualitative research?"*, https://www.scribbr.com/methodology/qualitative-research/, [Accessed 26 June 2024]

