Data21

Risks and Hidden Costs of ZIP Compression on z/OS

[2025 Update]

This paper exposes inherent liabilities of legacy z/OS ZIP compression solutions which negatively impact batch window utilization, MLC, SLA's, data security and regulatory compliance.

Data 21 Product Team

Introduction

Despite the growth of computer storage capacity and network speeds, data compression remains an essential tool for storing and transmitting ever-larger collections of data. The .zip format is designed for cross-platform exchange and efficient data storage for a set of related files; combining 85% to 95% average data compression, file management, and data encryption within a portable archive format.

The ready presence of .zip compatible utilities and processes across virtually all computing platforms, and ease of use for end-users of all technical competency levels, makes .zip files a simple solution for quickly facilitating efficient and secure exchange within the enterprise and beyond. Alleviating the need for and complexity of specialized network infrastructures, file transfer or security software at both ends of an exchange. ZIP Utilities also serve as a very efficient and secure file archiving methodology. Producing easily manageable and highly transportable compressed and encrypted archives of any size and number of files.

For the above reasons most z/OS operations have been employing zip file compression in production for many years. However, it is important to recognize that the environment has changed dramatically since such solutions were originally implemented. Most legacy products in operation today lack the 21st century technologies necessary to support strategic business and regulatory mandates, such as mainframe cost reduction and data privacy assurance and compliance, imposed upon today's Mainframe management and the Enterprise in general. The technologies at issue are:

- ☑ Strong AES Encryption
- ☑ zEnterprise® Data Compression (zEDC)
- ☑ **IBM**® z Integrated Information Processor (**zIIP**)

If your current z/OS zip solution does not check ALL THREE boxes, read on. If it does, and it's not Connect/ZIP, you are almost certainly over-paying for it and could benefit by contacting Data 21 about lowering this cost.

Strong AES Encryption

The Zip format provides an easy to implement persistent data encryption feature known as *Password Protection*. The addition of a few parameters is all it takes to virtually eliminate the threat of unauthorized access of the data, in-transit and at rest. Assuming strong encryption is applied that is. And this is the rub. Virtually every Zip product on every other platform supports the Zip AES encryption specification; and for good reason. Most z/OS Zip solutions provide only an old weak form of password encryption known as Zip 2.0 Encryption. This encryption cannot be expected to secure confidential information, is not security regulation compliant, and does not take advantage of the Mainframes cryptographic hardware, making it very slow and costly. Completely secure zip files require two things, well constructed hard to guess pass-phrases, and strong (preferably 256bit) encryption. If confidential data is being shared via weak encrypted or unencrypted zip files, the organization could be subject to the following risks.

"Zip 2.0 encryption format is known to be relatively weak, and cannot be expected to provide protection from individuals with access to specialized password recovery tools."

- WinZip Computing

Data Breach

For the purposes of this paper a data breach is defined as an event in which an individual's name and a medical record and/or a financial record or debit card is potentially put at risk — either in electronic or paper format. The three main causes: malicious or criminal attack (electronic or paper format), system glitch, or human error.

Increasing volume of consumer records lost or stolen annually, though it's not the sole driver—rising per-record costs, regulatory fines, lost business opportunities, and longer breach lifecycles also contribute significantly.

The costs of data breach vary according to the cause and the safeguards in place at the time of the data breach. The study found that the average total cost of a data breach, the average cost for each lost or stolen record (per capita cost), and the average size of data breaches have all increased since 2018.

Key Trends in Costs and Records Exposed

To illustrate, here's a summary of global average breach costs (from IBM reports) alongside U.S.-specific data on total records exposed and average per breach (pieced from Statista, Varonis, and other reports; note that U.S. data often proxies global trends due to high breach volume there):

Year	Global Avg. Breach Cost	U.S. # of Breaches	U.S. Total Records Exposed	Avg. Records per Breach (U.S.)
2018	$3.86M	~1,244	~2B (global estimate; U.S. subset ~500M)	~400K
2019	$3.92M	1,279	~1.4B	~1.1M
2020	$4.24M	~1,100	~800M	~727K
2021	$4.45M	1,862	~293M	~157K
2022	$4.35M	1,802	422M	~234K
2023	$4.45M	3,202	~418M	~130K
2024	$4.88M	3,158	1.35B	~427K
2025	$4.44M (prelim.)	N/A	N/A	N/A

Increasing records lost: Total U.S. records exposed have fluctuated but trended upward overall, surging to 1.35 billion in 2024—a 211% jump in victim notices from 2023 alone. Globally, breaches exposed over 7 billion records in the first half of 2024. Larger-scale incidents (e.g., those exceeding 200 days to resolve or involving multi-environment data storage) often expose more records and cost up to $5.01 million on average.

Not just volume: The average cost per lost/stolen record has also risen modestly, from $148 in 2018 to ~$160–$169 for customer PII by 2024–2025. However, average breach size (records per incident) has varied without a clear upward trajectory, suggesting the cost escalation ties more to the sheer number of incidents (up ~150% since 2018) and ancillary expenses like notification ($1.20M average in 2025, down 6% YoY) and lost business ($1.47M, down 6%).

2025 dip: The recent cost decline (9% from 2024) stems from faster detection/containment (down to 241 days total lifecycle, a 9-year low), aided by AI tools saving ~$1.9M per breach—not from fewer records. Still, high-risk factors like supply chain attacks or shadow AI use continue amplifying costs for larger record exposures.

In short, while escalating record volumes play a role in driving up total costs (especially via more frequent mega-breaches), multifaceted factors like regulatory pressures and remediation delays are equally—if not more—responsible for the "costlier" trend through 2024.

Non-Compliance

While compliance requires ongoing investment, it's a fraction (often 1/3) of noncompliance's total toll, making it a net positive for risk-averse organizations. Tailor strategies to your scale and regulations for optimal value.

In the context of data privacy regulations like GDPR and CCPA, as well as broader cybersecurity practices, the costs of noncompliance—encompassing fines, breach remediation, lost revenue, and reputational damage—typically far exceed the upfront investments required for compliance.

Recent 2024-2025 data shows noncompliance can be 2.7 times more expensive on average, with preventive measures often yielding a strong ROI by avoiding breaches altogether. For small and midsize businesses (SMEs), the disparity is even starker: annual compliance might cost $5,000–$50,000, while a single breach could exceed $120,000–$3.3 million. Larger organizations face higher absolute figures but similar ratios.

Key drivers of noncompliance costs include regulatory fines (e.g., GDPR's up to 4% of global revenue) and amplified breach expenses when non-compliance is a factor. Compliance, meanwhile, involves audits, training, tools like encryption and access controls, and incident response planning—costs that have remained relatively stable while breach totals have risen.

Most organizations today recognize that protecting only the data that is required to achieve compliance is a minimum threshold and that a move from selective encryption (protecting only specific types of data) to pervasive encryption (encrypting all data) is needed. A recent Data 21 poll of its z/OS customers revealed that the majority polled routinely zip confidential data for transfer to servers located within and/or outside the company. Given the nature and percentage of enterprise data resident and processed on the Mainframe, this was not surprising. What was surprising, in light of the all the above, is in most cases these files are not being password encrypted. Given how easy it is to implement and the benefits gained, the question is why?

One reason is many organizations today are focused on *Transmission Protection* rather than *Data Protection*. The fact is that data is more vulnerable to unauthorized access while stored on a server than during transmission. Secure FTP and Virtual private networks (VPN's) provide only transmission protection and therefore fall short of delivering complete privacy assurance. Zip AES Encryption secures confidential data during transmission and storage.

The fact is that data is more vulnerable to unauthorized access while stored on a server than during transmission.

An increased focus on pervasive data encryption, including robust methods like AES-256 encryption for ZIP files (commonly referred to as "ZIP AES"), significantly helps mitigate the impact and costs of data breaches. Encryption renders stolen data unreadable without the key, limiting damage from exfiltration, reducing regulatory fines (e.g., under GDPR/CCPA for "unsecured" breaches), and lowering notification expenses—often by 20-50% for affected records. In 2025, organizations using encryption as a core security practice saw average breach costs drop by ~$360,000 compared to those without, per recent analyses.

How Encryption Mitigates Breaches

Data Protection at Rest/Transit: Tools like Connect/ZIP, 7-Zip or WinZip with AES-256 secure archived files (e.g., ZIPs) against brute-force attacks—cracking a strong-password-protected symmetric AES-256 ZIP could take years with current tech. This prevents "useful" data leaks in 70%+ of theft scenarios.

Cost Savings Breakdown (2024-2025 Trends): While global breach costs fell 9% to $4.44M in 2025 (driven by faster detection), encryption specifically cuts ancillary expenses like lost business ($1.47M average) and post-breach recovery.

Factor	Impact of Encryption	Avg. Savings (2025)
Notification/Legal Fines	Avoids "reportable" status for encrypted data	$200K–$500K
Remediation & Downtime	Limits data usability for attackers	$300K+
Overall Breach Cost	Reduces total by 8-10%	~$360K
Healthcare-Specific	Most effective mitigation; encrypts PHI to prevent breaches	Up to 50% cost cut

In high-risk sectors like healthcare, encryption is the top mitigator, preventing 80% of common exposure risks. For ZIP files, always use AES-256 over legacy Zip 2.0 Crypto to ensure compatibility and strength.

Another reason may be the time and cost of compressing and encrypting. This is the subject of the following section.

Additional Hidden Costs

Reducing the usage of General Processors (GP's) delivers immediate direct savings. In most organizations mainframe capacity is continuing to grow. MIPS are trending upward, putting more pressure on the budgets of organizations that rely on the mainframe. Based on recent statistics more than 61 percent say mainframe cost reduction is a top priority. Therefore it is important that the solutions deployed incorporate zSeries technologies that directly support this priority.

zEnterprise® Data Compression (zEDC)

zEDC data compression provides up to 118X reduction in CPU and up to 24X throughput improvement. The new "on chip" zEDC feature of the IBM z15 platform replaces the zEDC cards of the past, making zEDC available to all and un-enabled mainframe zip compression solutions obsolete. If your current mainframe zip solution does not support zEDC, replacement of the un-enabled solution with an zEDC enabled one should be seriously considered.

Up to 80% reduced CPU cost compared to non-enabled compression solutions.

The significantly lower compression cost and incredible compression speed of zEDC not only serves to improve all the metrics of mainframe zip processing, but enables a more pervasive use of native mainframe compression.

IBM® z Integrated Information Processor (zIIP)

If zEDC is not elected for ZIP compression processing, zIIP is the next best option for reducing compression cost. According to recent information zIIP is 7-10x cheaper per MIPS than CP's, primarily because IBM (and most ISV's) assess no software license charges on ZIIP capacity. As processes go, data compression is one of the most cpu intensive. Offloading compression cycles to lower cost zIIP's is the next best thing.

zIIP eliminates associated MLC charges and frees up more expensive GP white-space.

Furthermore, in cases where GP's are knee-capped, compression time may be reduced as well. Since most installations have zIIP's available, running a solution that does not support them represents a considerable extra cost over one that does.

Our Solution

Connect/ZIP is a relatively low cost fully supported Zip compression solution for z/OS and USS that delivers all three key technologies presented in this White Paper, and more. Superior features and attractive competitive product replacement pricing makes Connect/ZIP the ideal replacement for more expensive and/or less technologically advanced solutions.

Visit https://www.data21.com/products/options/connectmp/connectzip.html for more information.

About Data 21

Data 21, Inc. has specialized in the development and support of IBM Mainframe software solutions continuously, and under the same management, since 1980. Data 21's commitment to timely implementation of relevant IBM zSeries hardware and software advances within its products, combined with highly competitive pricing, results in superior products which consistently deliver the best cost performance in their respective niches.