**ISSN:** 2583-9888
**DOI:** doi.org/10.51219/JAIMLD/prashanth-kodurupati/56

URF PUBLISHERS
connect with research world

Journal of Artificial Intelligence, Machine Learning and Data Science

https://urfpublishers.com/journal/artificial-intelligence

Vol: 1 & Iss: 1

*Literature Review*

# Zero Byte Data Files: A By-Product of Interrupted File Transfer and Mitigation Strategies

**Prashanth Kodurupati***

Information Technology, Managed File Transfer Engineer, PragmaEdge LLC, Alpharetta, USA

**Citation:** Prashanth Kodurupati. Zero Byte Data Files: A By-Product of Interrupted File Transfer and Mitigation Strategies. *J Artif Intell Mach Learn & Data Sci 2022*, 1(1), 120-122. DOI: doi.org/10.51219/JAIMLD/prashanth-kodurupati/56

**Received:** December 10, 2022; **Accepted:** September 23, 2022; **Published:** September 25, 2022

***Corresponding author:** Prashanth Kodurupati, Information Technology, Managed File Transfer Engineer, PragmaEdge LLC, Alpharetta, USA, E-mail: prashanth.bachi21@gmail.com



## ABSTRACT

IZero-byte data files are created when a file transfer is interrupted due to a network error. Timeout is one of the most common reasons behind this error, but there are others as well, including physical disruptions in the network and packet loss during transition. The solutions include increasing the timeout duration, network optimization, setting up retries and resumable transfers, and fixing any incompatibility issues that may prevent successful file transfers.

**Keywords:** Zero-byte data, Zero-byte data files

## 1. Introduction

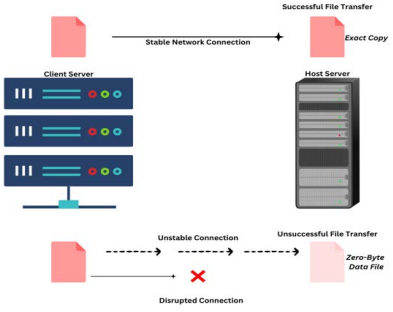
Trillions of gigabytes of data are transferred every day - between people, businesses, and a wide range of other entities online. However, not all transfers are successful. There are numerous reasons a data transfer can fail, including disconnection. The chances of interrupted file transfers rise/grow with larger files and volumes of data. An interrupted file transfer can have several unintended consequences, one of which is zero-byte data, also called a zero-byte file or zero-length file. It's essentially a damaged or corrupted file that may have the name of the original file but no data, hence the name zero-byte.

## 2. Literature Review

The literature references of zero-byte data/zero-byte files are sporadic. It's never the focal point of academic literature and is often mentioned in the context of other areas of discussion, primarily file systems, data transfers, and other related topics[1]. But there are instances when it's discussed comprehensively, and multiple reasons behind the creation of a zero-byte data file have been discussed, including signifying the beginning of a file transfer, when it isn't necessarily a sign of a problem or a result of an interruption event[2]. There is more data pertaining to zero-byte data files being created during unsuccessful or interrupted file transfers on various support pages, blogs, and other online resources[3].

However, if we zoom out a bit and focus on interrupted file transfers, their reasons, and remediation measures, the volume of available literature is significantly higher.

## 3. Problem Statement: Zero Byte Files and Interrupted File Transfers



A zero-byte file can be created on a drive for a number of reasons. In the context of file transfer, it may signify the initiation of a file transfer or accidental termination of the transfer before the file was fully moved between the host device/server and yours. An intentionally executed termination of a file transfer may not result in the creation of a zero-byte data file. While these files are not perceived as a threat by cybersecurity systems, they do represent certain underlying problems. The most common of which is an internet connection disruption/interruption:

### 3.1 Disrupted connection/timeout

If the internet connection is disrupted during a file transfer, it may lead to the creation of a zero-byte data file. This disruption may happen for a number of reasons, including a physical disconnection, but a common reason is that the internet connection times out. If your internal server starts receiving a file from a client-server (after a secure connection is established), beginning with the metadata (file name and size), but the transfer isn't completed, or the confirmation of the packet transfer isn't achieved within the pre-defined time limit, a connection timeout will occur, severing the file transfer. The initial information sent ahead of the actual file will appear as a zero-byte data file on your servers/drives. Timeout is also connected to a critical factor influencing file transfer success and failure, i.e., throughput[4].

### 3.2 Incompatibility issues

The occurrences are not as common as timeouts, but zero-byte data files can also be created when there are incompatibility issues between host and client servers, including file system and data transfer protocol incompatibilities. In rare cases, these may lead to the successful transfer of the metadata of the file but not the data itself, resulting in the creation of zero-byte data.

### 3.3 Packet loss

Network congestion or weak internet connection may also lead to packet losses during data transfer, even when secure file transfer protocols like SFTP (SSH File Transfer Protocol) and FTPS (FTP over SSL/TLS) are in place. The reasons primarily concentrate on network-related weaknesses like congestion and latency. The reason is that even though these secure file transfer protocols focus on the safety of the packets being transferred, they may have limitations when it comes to the transfer itself[5].

## 4. Suggested and Implemented Solutions

Since the problem of zero-byte data file creation mostly stems from network timeouts, disruptions, and limitations, the bulk of the solutions are concentrated around server settings and management.

### 4.1 Modifying server timeout settings

The most commonly implemented and straightforward solution to zero-byte data files resulting from timed-out connections is to increase the timeout duration. This gives ample time for data to be safely transferred to and from the host server, but it may come at the cost of performance. If the timeout duration is increased to accommodate the highest data transfer volume without interruption, it may lead to significant performance degradation[6]. Increasing the timeout window may also make the detection of underlying network quality issues a bit difficult.

However, with adequate server monitoring, finely tuned file-transfer protocols and systems in place, and data transfer taking place among trusted and well-understood entities (B2B), increasing timeout may be the most optimal solution to prevent the creation of zero-byte data files.

### 4.2 Retries and resumable transfers

A server can be set to allow multiple retries to reinitiate the file transfer, with or without appropriate time delays. The idea is that if a file transfer is interrupted because of network congestion or a low timeout window, it may resume once the connection is re-established, and a complete file transfer can be achieved instead of a zero-byte data file being created.

A more difficult-to-implement solution, especially for high-volume transfer instances, is file transfers that aren't suspended when the network connection is broken or timed out; they are simply paused and can resume once the connection is re-established. It may require dividing the file into smaller segments, each constituting an independent transfer[7].

### 4.3 Optimizing file transfer protocol and resource allocation

Incompatibility and insufficient file space can also lead to the creation of zero-byte data files, so individuals responsible for managing file transfers (typically MFT engineers) should ensure that the servers, virtual machines (VMs), file-transfer protocols, etc., are all in sync between host and client servers so that all file transfers are successful, preventing the creation of zero-byte data files.

### 4.4 Network overhaul

A network overhaul or at least optimization is necessary to ensure that everything, from timeout settings to adequate bandwidth, is available for a high-volume data transfer. The network should be optimized for peak data transfer volumes; otherwise, you may encounter higher instances of zero-byte data file creation during some transfers and almost no instances in others, which may be a problem from a predictability perspective.

## 5. Summary of Problems and Proposed Solutions.

| Problems | Possible Solutions |
|---|---|
| Disrupted Connection/ Timeout | 1. Modifying Server Timeout Settings<br>2. Retries and Resumable Transfers<br>3. Optimizing File Transfer Protocol |
| Incompatibility Issues | 1. Optimizing File Transfer Protocol<br>2. Data Validation<br>3. Standardization |
| Packet Loss | 1. Network Overhaul<br>2. Error Correction Mechanisms<br>3. Packet Retransmission |

## 6. Best Practices to Avoid Zero Byte Data

The best practices to avoid zero-byte data include implementing the solutions and preventive measures preemptively. However, it's important to ensure that a possible solution does not contribute to network degradation or mask network problems. This is one of the reasons why increasing timeout can be a complex solution to implement in certain cases.

Another good practice is to have adequate network resources at your disposal and optimize the network to mitigate any

inefficiencies that may lead to congestion, delays, timeouts, and, consequently, the creation of zero-byte data files[8]. This offers a much wider range of benefits and prevents several other unwanted occurrences as well.

Even with all these best practices to ensure that no new zero-byte data files are created, it's important to have a monitoring system in place for data integrity and identification of zero-byte data file creations so that proper measures can be taken to restart the transfers.

## 7. Potential Use Cases

The solutions designed to prevent the creation of zero-byte data files, including increasing the timeout limit, can be implemented in a wide array of file transfer scenarios. This includes (but is not limited to):

- Transferring large files or larger volumes of data with a weak connection or limited bandwidth.
- Transfers take place on high-latency connections.
- Sensitive data transfer between servers of financial institutes, research facilities, etc., with stringent timeout thresholds for maintaining throughputs.
- Data transfer takes place between incompatible (different operating systems, file transfer protocols, etc.) servers and entities.

## 8. Conclusion

Zero-byte data files are not a vulnerability per se but a reflection of improper configuration or network limitations pertaining to the handling of large data transfer volumes. Identifying these files is critical to unearthing these issues and ensuring that the desired file transfer takes place as intended because, if undiscovered, zero-byte data files may lead to a false sense of operational success.

## 9. References

1. Mizusawa N, Kon J, Seki Y, Tao J, Yamaguchi S. Performance improvement of file operations on overlayfs for containers**.** 2018 IEEE International Conference on Smart Computing, 2018.
2. Montfort N. No code: Null programs. A technical report from The Trope Tank, Massachusetts Institute of Technology, 2013.
3. Mathai A. What are zero-byte files and how do I deal with them? goldfynch Blog 2018.
4. Li J, Malathi Veeraraghavan, Steve Emmerson, et al. File multicast transport protocol (FMTP). 2015 15th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing, 2015.
5. https://ieeexplore.ieee.org/abstract/document/7732070
6. https://www.ojs.jctecs.com/index.php/com/article/view/11
7. https://ieeexplore.ieee.org/abstract/document/7207376
8. https://ieeexplore.ieee.org/abstract/document/7568330