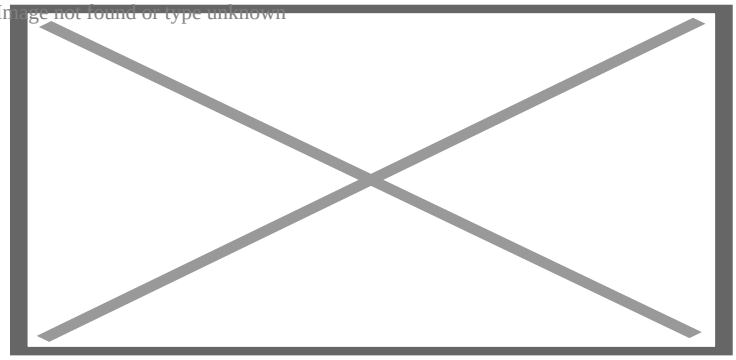
Chain of Custody in Google Docs: Using ETag Verification and Hash Values

Description

Image not found or type unknown



Google Docs is a popular cloud-based office suite that offers users features and capabilities. One of the key benefits of using Google Docs is the ability to collaborate and share files with others in real time. However, when files need to be collected for eDiscovery or forensic purposes, it's essential to have a process in place to ensure that downloaded files are authentic and have not been modified since they were last accessed.

To address this need, Google Docs uses an ETag download verification process for native Google office files such as Docs and Sheets. The ETag value is a unique identifier generated by the server. It is based on factors such as the file's content, conversion parameters, and the API version used. Google Docs use this ETag value to verify the integrity of the downloaded file and ensure that it matches the original file.

The ETag calculation process used by Google Docs is not publicly documented but is known to be consistent and reliable. While there are hashing algorithms such as MD5 and SHA-1 that can be used to create a common hash value at the time of download, these hash values do not consider the specific factors used by Google Docs to generate the ETag value. As a result, a common hash value cannot be used to verify the integrity of the file at a later time, nor can it be used to replicate the ETag process.

In practical terms, this means that if a third-party application needs to verify the integrity of a downloaded file, it cannot rely on the ETag value generated by Google Docs. Instead, the application would need to use a standard hashing algorithm to create a hash value of the downloaded file at the download time. Pinpoint Cloud software immediately hashes files after downloading using the Google API. This hash value can then be compared to the hash value of the original file to ensure that the downloaded file has not been modified. Additionally, Pinpoint Cloud is the only Google Drive collection tool that provides the downloaded file time stamps that match those used in the Google drive and creates a chain of custody log for the entire process.

It's important to note that while a traditional hash value can be used for chain custody purposes, it does not provide the same type of verification as the ETag value. The ETag value considers the specific factors used by Google Docs to generate the value, which means that it provides a higher level of assurance that

the downloaded file is authentic and has not been modified.

In conclusion, the ETag download verification process used by Google Docs for native Google office files is an essential tool for ensuring the integrity of downloaded files. While third-party applications cannot replicate the ETag process, a typical hash value can be used for the chain of custody purposes. However, it's essential to understand that a common hash value does not provide the same verification level as the ETag value and should only be used with other methods of verifying file integrity.

Date
12/22/2024
Date Created
03/14/2023