[V7][V8] FAQ: Why is there a discrepancy between AhsayCBS storage statistics vs physical data size on a storage device?

Article ID: 8067
Reviewed: 2020-02-07

Product Version:
Ahsay Software: 7.3 to 7.x
Ahsay Software: 8.1.0.24 to 8.1.1.174, 8.3.0.30 to 8.x

Audience

This article is for AhsayCBS Administrator's (partners) who have inquiries about the discrepancy between AhsayCBS storage statistics vs physical data size on a storage device.

Description

Starting from version 7 and continuing on version 8, Ahsay implemented a new storage format for backup data on AhsayCBS and cloud destinations to enhance the backup performance and improve the overall performance of the AhsayCBS backup server. All files backed up are stored in either 16 MB or 32 MB blocks depending on the file size on the backup destination(s), which can be on AhsayCBS server, local, and or cloud storage destinations.

  • The backup and storage of multiple small files (< 16 MB) are combined and stored in 16 MB data blocks or "shared blocks". Depending on the file type and size of the source file, a 16 MB data block can contain a few small files or even hundreds of small files.

  • The backup and storage of large files (>32MB) are broken down and stored in 32 MB data blocks. For example: a file with the size of 100 MB would be split up and stored as 4 'block' files: 3 x 32 MB and 1 x 4 MB.

Due to the new data storage format, there can be discrepancies between storage statistics reported by the Operating System which is considerably larger than the storage statistics reported by AhsayCBS as:

  • AhsayCBS storage statistics (on an application Level) represents the files/folders in a backup set(s), which are restorable in both the data area and retention area.

  • The storage statistics reported by the Operating System which is the size of the backup set data on the disk or cloud storage represents the physical 16 MB and 32 MB block files.

The discrepancy stems from the handling of small files stored in the 16 MB blocks or "shared blocks" in version 7 / 8.

   • When a file has been backed up is subsequently deleted from a customer machine, on the next backup job AhsayOBM/AhsayACB will detect the change and update the related index file to mark the file as deleted and change the location of the files from the data area to retention area. To minimize disk I/O load on the AhsayCBS server, the deleted remains in the 16 MB block.

   • Once the deleted file in the retention area exceeds the retention policy settings, the reference to the deleted file is removed from the index file only, to minimize disk I/O load on the AhsayCBS server it is not physically removed from the 16 MB block file, therefore the AhsayCBS storage statistics will report small size.

A 16 MB block or "share block" file can only be physically deleted from the storage device, once all the files stored in the block are changed from the customer machine, moved to the retention area, and exceed the retention policy.

However, in the backup and storage of large files the physical 32 MB block files are removed from the storage device, once the deleted file in the retention area exceeds the retention policy settings.

Affected Backup Sets

As the issue stems from the handling of small files stored in the 16 MB blocks or "shared blocks" in version 7 / 8, this issue visible on the following backup set types; *File, Cloud File, MS Exchange Mail level, and Office 365* where the possibility of backup of small files is common. This can be further compounded if there are frequent changes in the content backup source which will result in files frequently being moved to the retention area.

While other backup types are not affected; *MS Exchange database, MS SQL Server, Oracle, VMware, Hyper-V, MS Windows System backup, MS System State backup, and Lotus Domino* usually involve backup of large files which are stored in 32 MB blocks.

Problem

The discrepancy between the storage statistics presented challenges to some partners who invoiced their customers based on the amount of physical data storage on their backup servers and not the amount of data, which is recoverable. The affected customers may receive invoices for data usage, which is considerably larger than the data shown in their customer's backup user accounts on the AhsayCBS web console. Which would include; index, data blocks (.bak), and checksum (.bak) files.

Possible Solution

Due to feedback from our partners relating to the storage statistics discrepancy, Ahsay attempted to provide a solution to this issue using a hotfix, where the calculation of storage statistics was re-

calibrated to accurately reflect the physical size of the 16 MB and 32 MB data blocks stored on AhsayCBS server, local, and, cloud destinations.

This version was made able to partners who applied the following hotfix version 8.3.0.50 to 8.3.0.131 from the Ahsay Partners Portal. After upgrading, the storage statistics for each backup set were re-calculated based the physical size of the files located in the **%CBS_user_home%\%username%\%backupset_id%\blocks** folder.

There was no option is choose another statistics calculation method.

Results

After upgrading to AhsayCBS **version 8.3.0.50 to 8.3.0.122** the following issues were encountered.

- The storage statistics for each backup user account with the affected backup sets increase significantly.

- If the newly calculated storage statistics exceeds the storage quota of the user account, this will cause the backup jobs on the affected user accounts to end prematurely end with a "quota exceed" backup status.

- If customer billing is based on AhsayCBS storage statistics, then customers receive large increases in their invoices after upgrading to the latest version.

After feedback from our partners, it was decided to temporarily suspend the new storage statistic calculation method in the next release 8.3.2.0 or above, to continue to use the original storage statistic calculation method until further notice. Therefore, partners who have been using **version 8.3.0.50 to 8.3.0.131** hotfix, after they apply next official release **8.3.2.0 or above** will find the storage statistics for the affected backup sets and the user will be reduced and calculated based on restorable data and not by physical storage size.

Addressing Physical Storage Provisioning

The new storage format for backup data written to AhsayCBS, local, and cloud destinations containing "shared blocks", could affect your storage provisioning, you should closely monitor your storage to avoid reaching capacity. Unfortunately, there is no workaround to reduce the shared blocks usage from the affected backup set, or method to list the block's content or report the gap between restorable data vs physical storage.

Future Direction

Ahsay is currently looking into the possibility of implementing a solution that allows partners the option of selecting the storage statistics calculation method.

Keywords

statistics, storage, retention, disk size

From:
https://wiki.ahsay.com/ - **Ahsay Wiki**

Permanent link:
https://wiki.ahsay.com/doku.php?id=public:8067_faq:discrepancy_between_ahsaycbs_storage_statistics_vs_physical_data_size

Last update: **2021/12/22 16:59**