AhsayCBS v8.3 – Enhanced Backup Performance, Robustness and Data Integrity with Redesigned Backup Process

Ahsay Systems Corporation Limited

25 September 2020

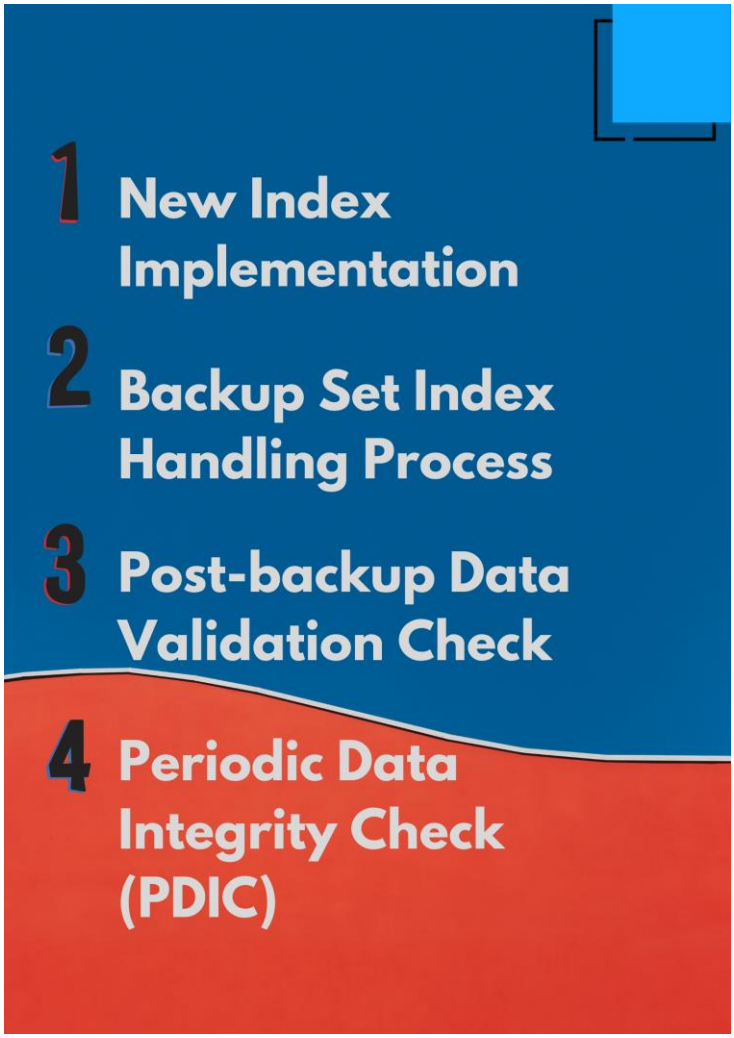
Revision History

Date	Descriptions	Type of modification
11 February 2020	Initial draft	New
25 September 2020	Modified Overview on the Backup Process for Run on Client and Run on Server backup job; Modified the detailed process of Periodic Data Integrity Check (PDIC)	Modifications

Table of Contents

New Features

Four new features were introduced in the **AhsayCBS v8.3.0.0 (or above)** to provide **better backup performance**, **robustness** and **data integrity of backup data**.



New Index Implementation

v8.3 supports a **new index design** to provide **better backup performance** and **robustness**. However, for customers who are going to upgrade from older versions (v6, v7, and pre-v8.3), they still need to undergo a mandatory index conversion process before the new index can be used and normal backup jobs can continue.

After the v6, v7 and pre-v8.3 AhsayOBM and AhsayACB clients are upgraded to v8.3, the backup set index conversion will occur **immediately** on the first run of backup job.

Index Conversion

For AhsayOBM/AhsayACB client

v6 backup sets:

➔ Index.bdb and r-index.bdb will be converted to the new index file structure:

index.db* and **backupInfo.db***

1 Old index file (index.bdb) will be converted to the new index file structure: **(index.db)**

2 Data Integrity Check will be performed once the index conversion is finished

v7 and pre-v8.3 backup sets:

➔ index.b2b*, index.xml* and index-s0 will be converted to the new index file structure:

index.db* and **backupInfo.db***

1 Old index file (index.b2b) will be converted to the new index file structure: **(index.db)**

2 Data Integrity Check will be performed once the index conversion is finished

For AhsayCBS Run on Server (Office 365 and Cloud File) backup

v7 and pre-v8.3 backup sets:

- index.b2b*, index.xml* and index-s0 will be converted to the new index file structure:

  index.db* and **backupInfo.db***

1 Old index file (index.b2b) will be converted to the new index file structure: **(index.db)**

2 Data Integrity Check will be performed once the index conversion is finished

Index Conversion requirements

For AhsayOBM/AhsayACB client

Client Version	Index Conversion Needed	Data Migration Needed
v6	✓	✓
v7	✓	✗
Earlier than v8.3.0.0	✓	✗
v8.3.0.0 (or above)	✗	✗

- Index conversion cannot be disabled.
- For Run on Client backup sets, index conversion process will automatically start during the first backup job after upgrading to **AhsayOBM/AhsayACB v8.3.0.0** (or above).
- After index conversion, the size of the index will not be the same. For large data index, the new index will become smaller since duplicate information will be grouped. While for small data index, the new index may be bigger since additional information may be included in the new index.

For AhsayCBS Run on Server (Office 365 and Cloud File) backup

Client Version	Index Conversion Needed	Data Migration Needed
v7	✓	✗
Earlier than v8.3.0.0	✓	✗
v8.3.0.0 or above	✗	✗

For Run on Server (Office 365 and Cloud File) backup sets, index conversion process will automatically start during the first backup job after upgrading to **AhsayOBM/AhsayACB v8.3.0.0** (or above).

For backup sets which contain large number of files and/or folders, the v8.3 index conversion process could take several hours to complete. For example: File, Cloud File, MS Exchange mail level, and Office 365 backup sets. In some cases, backup sets containing several million files and/or folders could take days to complete the v8.3 index conversion process. Kindly take this into consideration when planning your AhsayOBM/AhsayACB client upgrade to v8.3.0.0 (or above).

Backup Set Index Handling Process

The main purpose of the **new backup set index handling process** is to provide an **improved backup user experience** by automating the handling of corrupted index related issues. Therefore, it can be guaranteed that the backup job process will NOT be interrupted or prevented from running when index related issues are encountered.

This whole process is done without the need for intervention by the end customer or backup service provider, which **reduces the overall cost of support**.

The new backup set index handling process will perform the following:

1. All available index files (index.db) from the current directory (e.g. cloud destination, AhsayCBS, local drive, and FTP or SFTP) will be sorted according to the modified date and time.
2. If the local temporary directory and backup job folder also contain index files, then these index files will also be added to the list.
3. Latest index file in the list will be downloaded according to the most current modified date and time.
4. The latest modified index file will be verified if it is **valid** and can be opened.
5. If it CAN be opened, then it will be used to compile file list for backup.
6. If NOT, then the index handling process will proceed to the next index file from the list (sorted by latest modified time to oldest) until a valid index is found that can be compiled for backup.

Corrupted index files will no longer be recoverable. Corrupted index files will remain in the backup destination(s) and can only be removed by manually running a **Data Integrity Check (DIC)** on AhsayOBM/AhsayACB client or AhsayCBS Web Console for Run on Server (Office 365 and Cloud File) backups.

It is recommended to perform a Data Integrity Check **regularly** to ensure that NO corrupted data will remain in the backup destination(s).

Post-Backup Data Validation Check

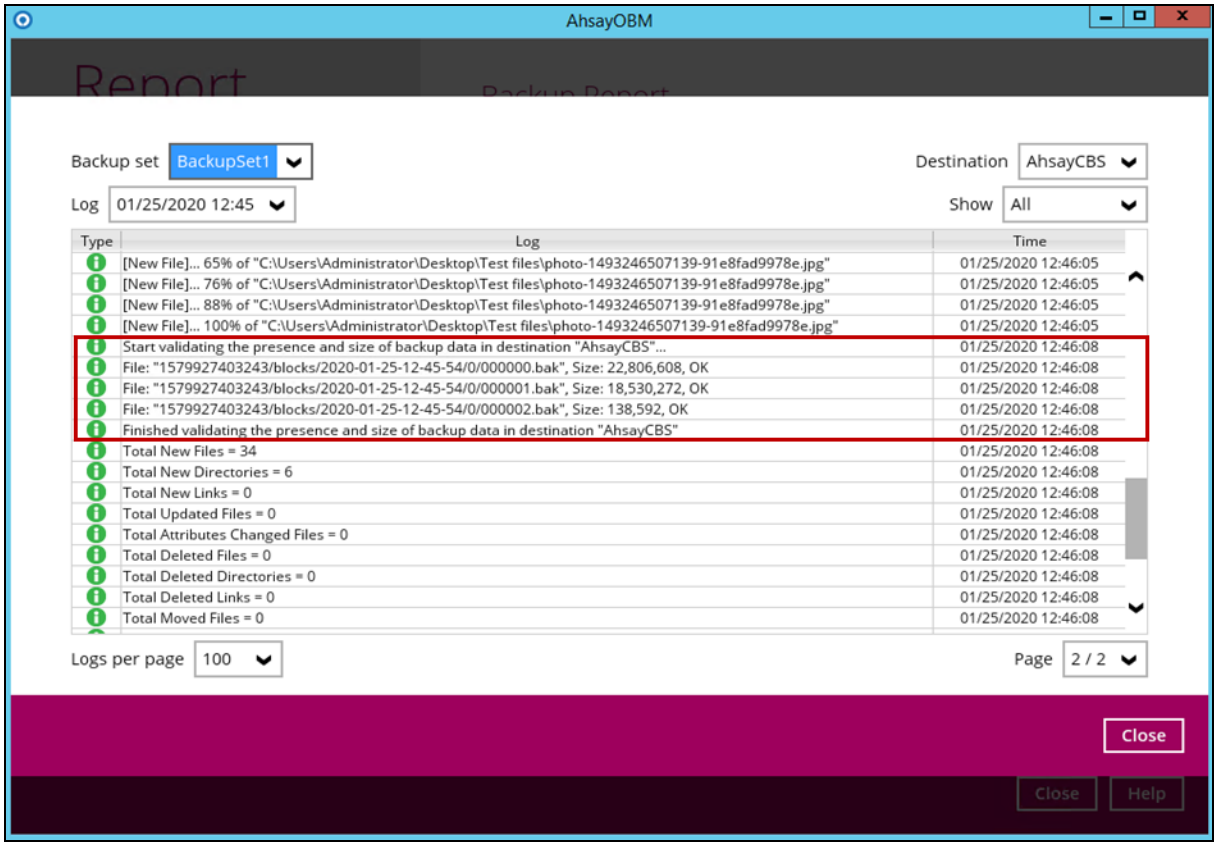
Although as part of the current backup process, a **checksum validation** is already performed on the backup files to ensure data integrity and recoverability, in addition, to provide peace of mind, an additional **post-backup data validation check** is now performed in v8.3.

The post-backup data validation check will perform the following:

1. The number of 16 or 32 MB data blocks in the backup destination(s) is identical to the number of blocks transferred
2. The individual sizes of each data block in the backup destination(s) is identical to the sizes of each block transferred

After performing a backup job, click the button to display the detailed backup log. This will show if the post-backup data validation check has completed successfully.

The screenshot below shows an example of the post-backup data validation check process in a backup job.



Periodic Data Integrity Check (PDIC)

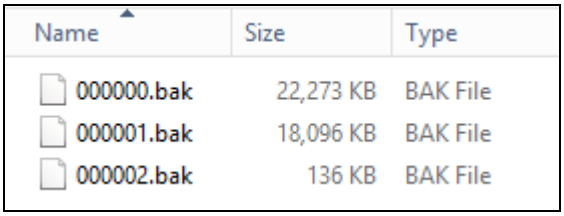
Data integrity is a core focus of the Ahsay backup software. It is important to maintain the overall accuracy, completeness, and consistency of the backup data throughout its lifecycle. To ensure that these objectives are met, Periodic Data Integrity Check (PDIC) is implemented which performs automated error-checking and validation procedures throughout a backup job process.

Also, as part of the PDIC, the storage statistics for the backup set are recalculated to maintain up-to-date and accurate data usage which guarantees consistent customer billing experience.

During a backup job, the PDIC is performed to ensure the validity and recoverability of the backup data.

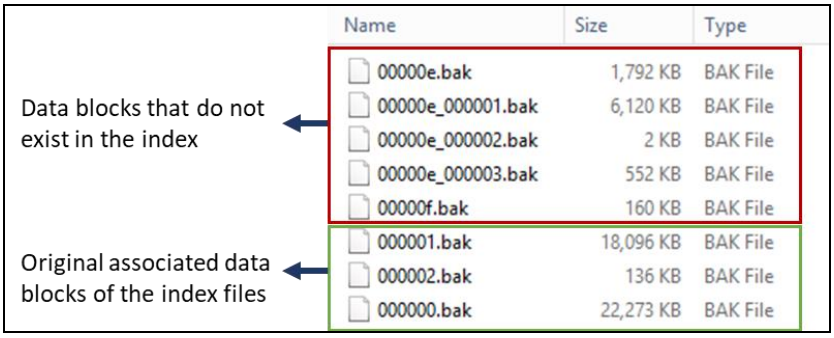
The following process is performed during a Periodic Data Integrity Check (PDIC):

1. At the start of the backup job, the PDIC will check for the physical data blocks (.bak files) in the backup destination(s) if they exist in the index.



**Assuming these are the original associated data blocks of the index files*

2. If there are physical data blocks (.bak files) found which do not exist in the index (i.e. data blocks with NO related index), then these physical data blocks will be **automatically removed** from the backup destination(s).



3. The statistics on the Data area and Retention area are recalculated.
4. After the PDIC is completed, the backup job process will continue.

NOTE: This behavior only applies to v8.3.0.30 – v8.3.2.xx

The PDIC only checks for the physical data blocks (.bak files) directly instead of checking each file in the index. So, if there are index files found to be corrupted, then these index files and its associated data blocks can only be removed from the backup destination(s) by manually running a **Data Integrity Check (DIC)** on AhsayOBM/AhsayACB client or AhsayCBS Web Console for Run on Server (Office 365 and Cloud File) backups. **This behavior only applies to v8.3.0.30 – v8.3.2.xx**

It is recommended to perform a Data Integrity Check **regularly** to ensure that NO corrupted data will remain in the backup destination(s).

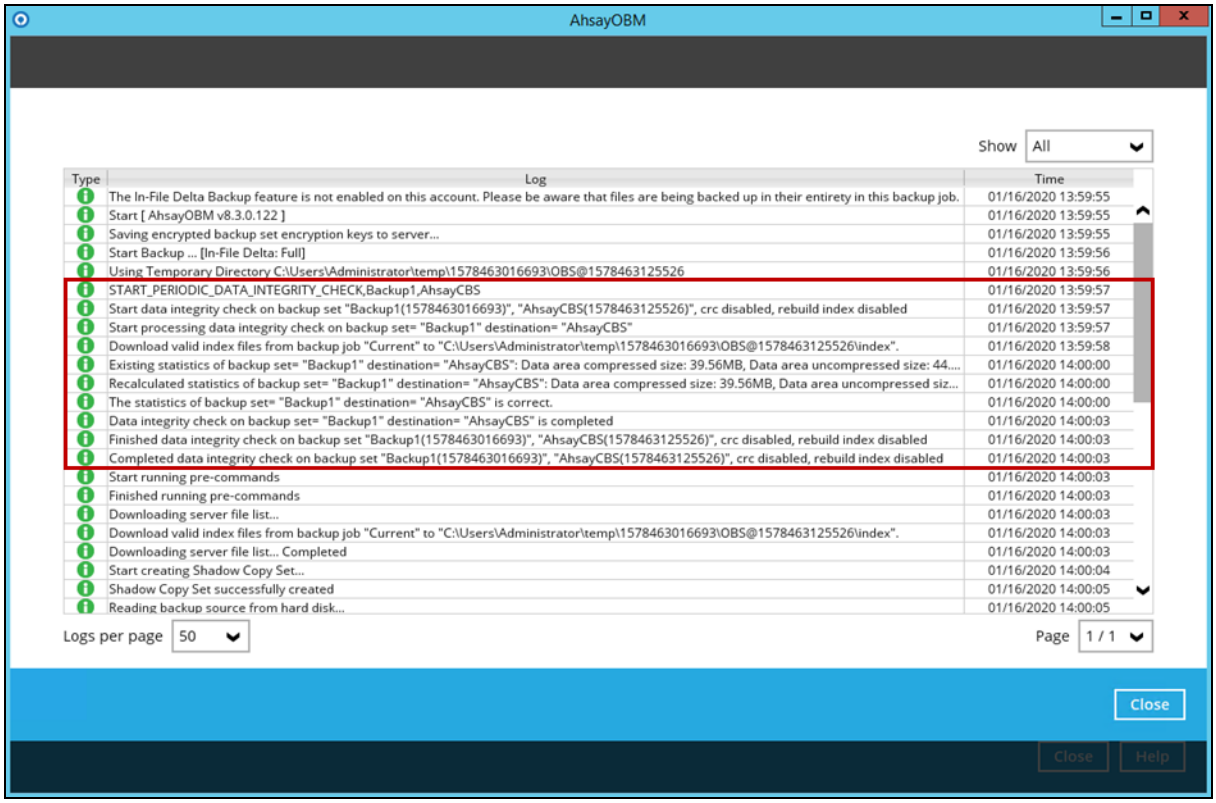
After the backup job, click the button to display the detailed backup log. This will show if the PDIC had run and completed successfully.

There are two possible outcomes in performing a PDIC:

1. The PDIC job is completed without data blocks related issues encountered
2. The PDIC job has detected data blocks with NO related index

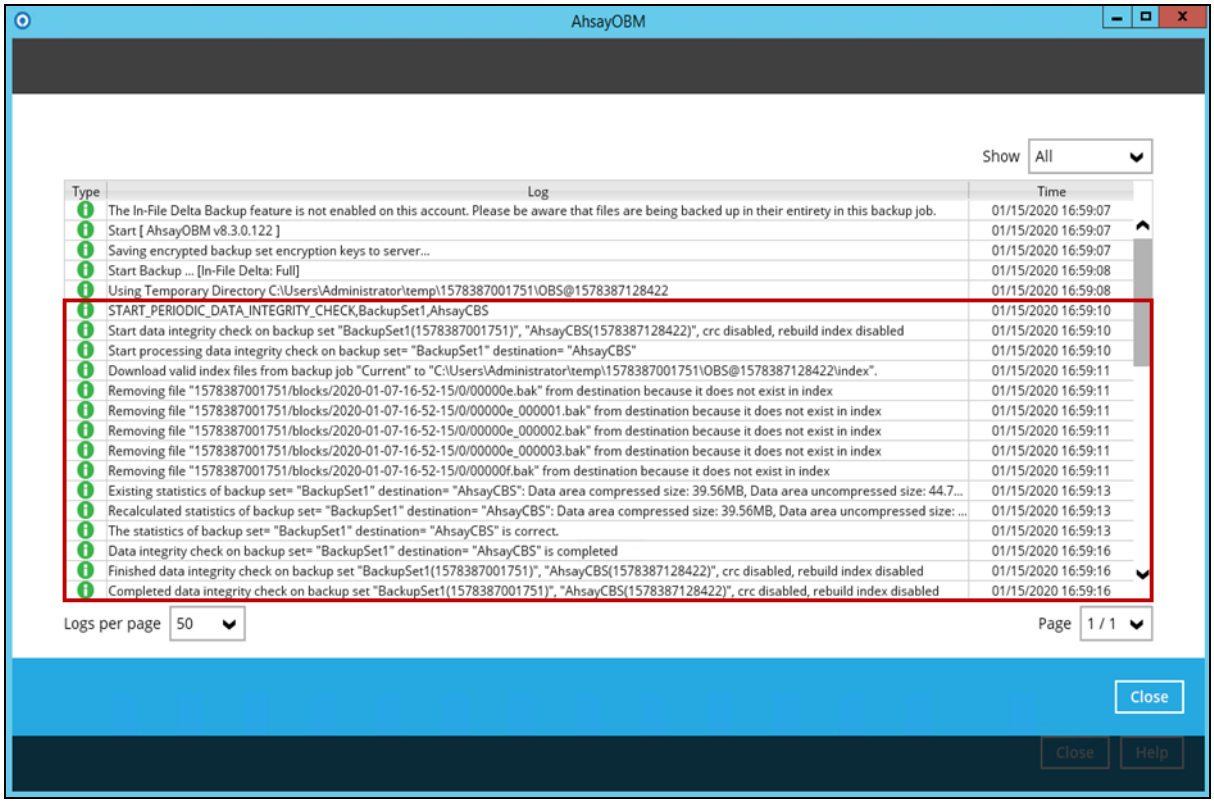
Result 1

The screenshot below shows an example of a backup job where the PDIC has run and has NOT detected any issues on the physical data blocks (.bak files).



Result 2

The screenshot below shows an example of a backup job where the PDIC has run and has identified data blocks (.bak files) which do not exist in the index file. These physical data blocks are automatically deleted from the backup destination(s).



Corrupted data blocks (.bak files) can only be deleted by manually running a full **Data Integrity Check (DIC)** on AhsayOBM/AhsayACB client or AhsayCBS Web Console for Run on Server (Office 365 and Cloud File) backup sets.

It is recommended to perform a Data Integrity Check regularly to ensure that NO corrupted data will remain in the backup destination(s).

What conditions will trigger a Periodic Data Integrity Check

During a backup job, a Periodic Data Integrity Check (PDIC) will be triggered under the following conditions:

- Will be triggered on a weekly basis, usually on the first run of backup job that falls on any one of these days: **Friday, Saturday, or Sunday**

  OR

- If there is no active backup job(s) running on Friday, Saturday, or Sunday, then the PDIC will be triggered on the next available backup job

  NOTE: This behavior only applies to v8.3.0.30 – v8.3.2.xx

Which types of backup jobs will initiate a Periodic Data Integrity Check?

Periodic Data Integrity Check (PDIC) can be initiated by the following types of backup job:

Manual backup job on:

- AhsayOBM/AhsayACB Client
- AhsayCBS Web Console for Run on Server (Office 365 and Cloud File) backups

Scheduled backup job on:

- AhsayOBM/AhsayACB Client
- AhsayCBS Web Console for Run on Server (Office 365 and Cloud File) backups

Continuous backup job on:

- AhsayOBM/AhsayACB Client
- AhsayCBS Web Console for Run on Server (Office 365 and Cloud File) backups

For AhsayOBM/AhsayACB running on Windows platform:

- Windows System Tray Icon backup job
- Windows Logout Reminder-initiated backup job
- *RunBackupSet.bat* script

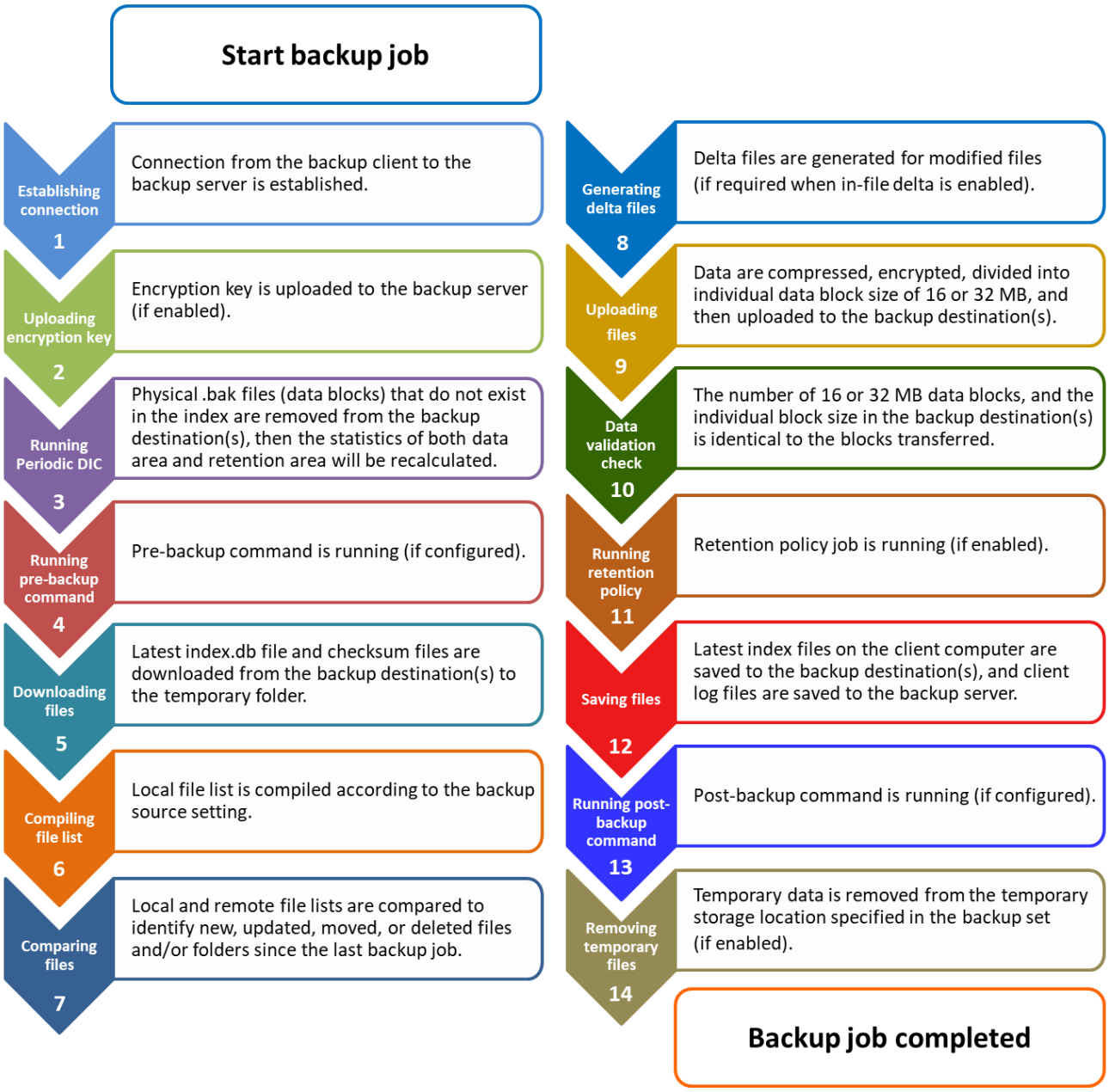
For AhsayOBM running on Linux/FreeBSD platform:

- *RunBackupSet.sh* script

AhsayCBS Web Console server-initiated backup job

Overview on the Backup Process (for Run on Client backup)

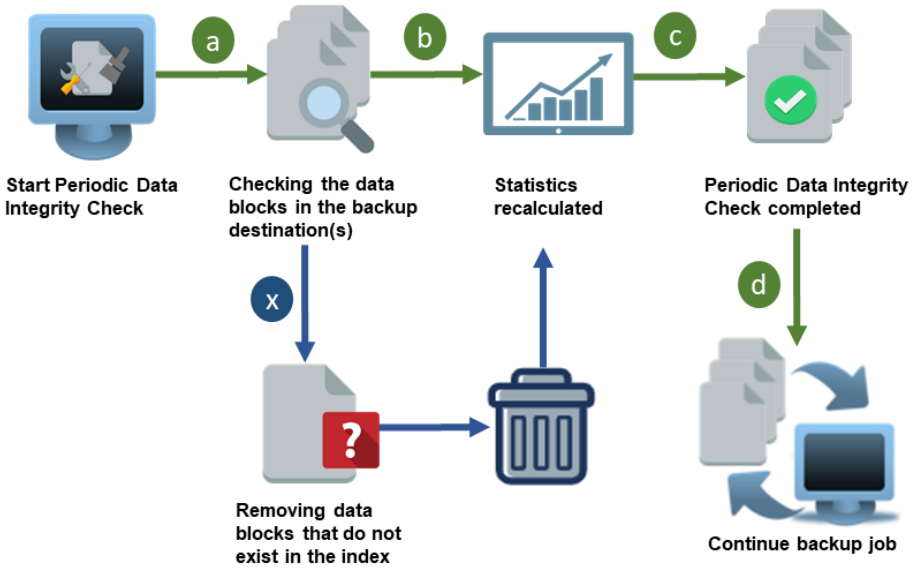
The following steps are performed during a Run on Client backup job. The three new features will take place on steps **3 (PDIC)**, **5 (Backup Set Index Handling Process)** and **10 (Post-Backup Data Validation Check)**.



How it works

Periodic Data Integrity Check (PDIC)

The diagram below shows the detailed process of periodic data integrity check.



a Check the data blocks (.bak files) located in the backup destination(s) if they exist in the index.

➔ If **YES**, proceed to b

➔ If **NO**, proceed to x

b Data area and Retention area statistics are recalculated.

c Periodic Data Integrity Check is completed.

d The backup job process will continue.

x Data blocks (.bak files) that do not exist in the index will be removed from the backup destination(s).

Proceed to b

Periodic Data Integrity Check (PDIC) Will run under the following conditions:

- Will be triggered on a weekly basis, usually on the first run of backup job that falls on any one of these days: **Friday**, **Saturday**, or **Sunday**.

  OR

- If there is no active backup job(s) running on Friday, Saturday or Sunday, then the PDIC will run on the next available backup job.

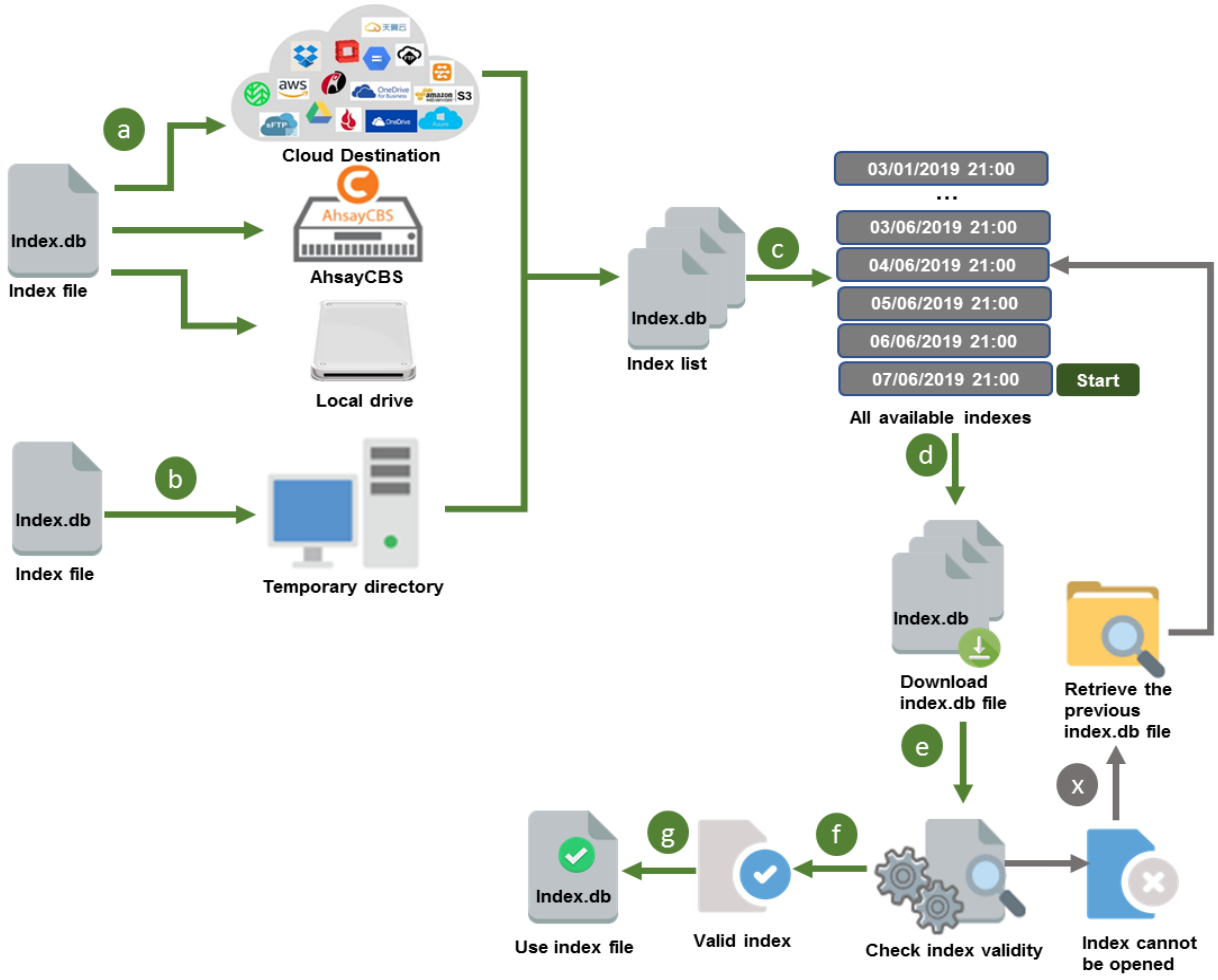
Corrupted data blocks (.bak files) can only be deleted by manually running a full Data Integrity Check (DIC) on AhsayOBM/AhsayACB client or AhsayCBS for Run on Server (Office 365 or Cloud File) backup sets.

This behavior only applies to v8.3.0.30 – v8.3.2.xx

Backup Set Index Handling Process

Run on Client

The diagram below shows the backup set index handling process.



a Index file(s) will be listed in the current directory (e.g. cloud destination, AhsayCBS, local drive, FTP or SFTP).

b If the local temporary directory contains index.db file which has a later modified date and time, then the current directory (e.g. cloud destination, AhsayCBS, local drive and FTP or SFTP) will also be added to the list.

c All available indexes will be sorted according to the modified date and time.

d Latest index file in the list will be downloaded according to the most current modified date and time.

e Index file will be checked to see if it can be opened.
- → If YES, proceed to f
- → If NO, proceed to x

f If it can be opened, the index will be used for the current backup job.

g Use the index.db file to compile file list for backup.

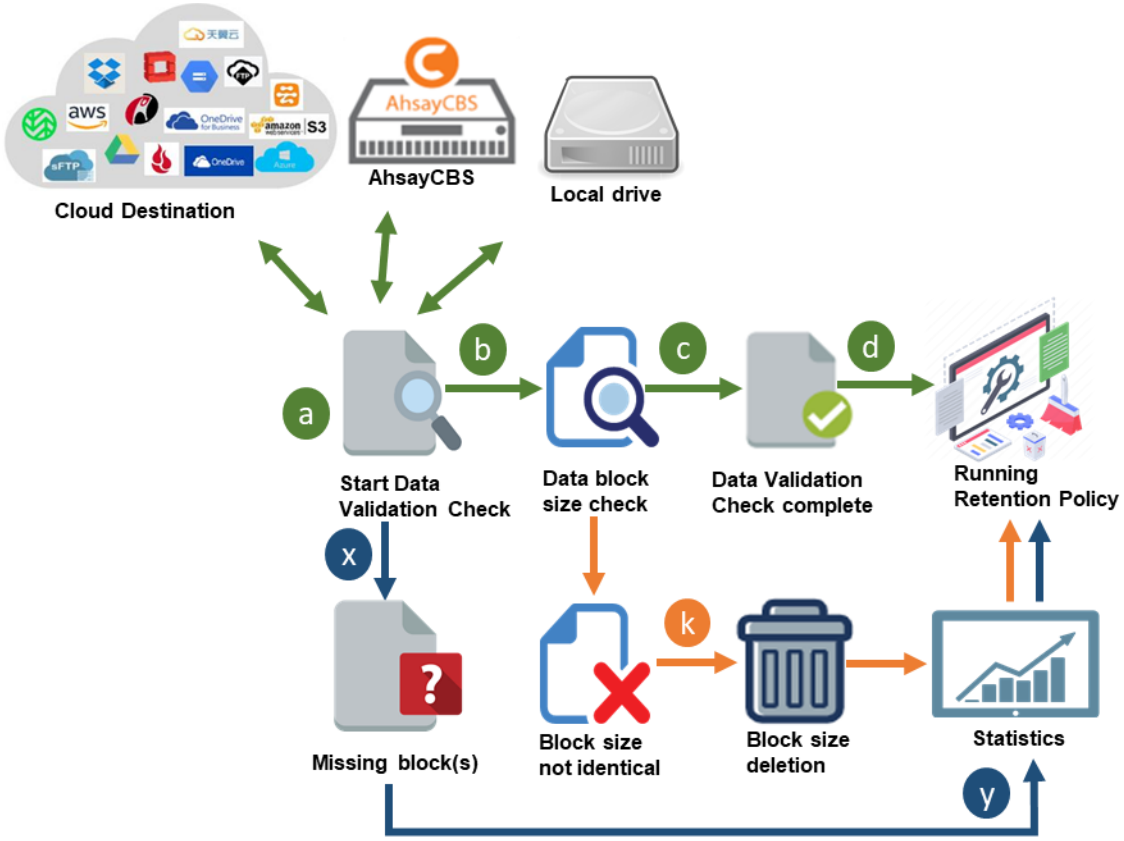
x If it cannot be opened, it will proceed to the previous index.db file in the list until it finds a valid index that can be opened.

For index.db files found to be corrupted and its associated data blocks will no longer be restorable. They will not be removed automatically from the backup destination(s) and can only be removed by manually running the Data Integrity Check on the AhsayOBM/AhsayACB client.

Post-Backup Data Validation Check

Run on Client

The diagram below shows the detailed process of post-backup data validation check.



a Check the number of 16 or 32 MB data blocks in the backup destination(s) is identical to the number of blocks transferred.
- ➔ If YES, proceed to b
- ➔ If NO, proceed to x

b Check the individual sizes of each data block in the backup destination(s) is identical to the sizes of each block transferred.
- ➔ If YES, proceed to c
- ➔ If NO, proceed to k

c Data validation check is complete.

d Retention policy will run (if enabled).

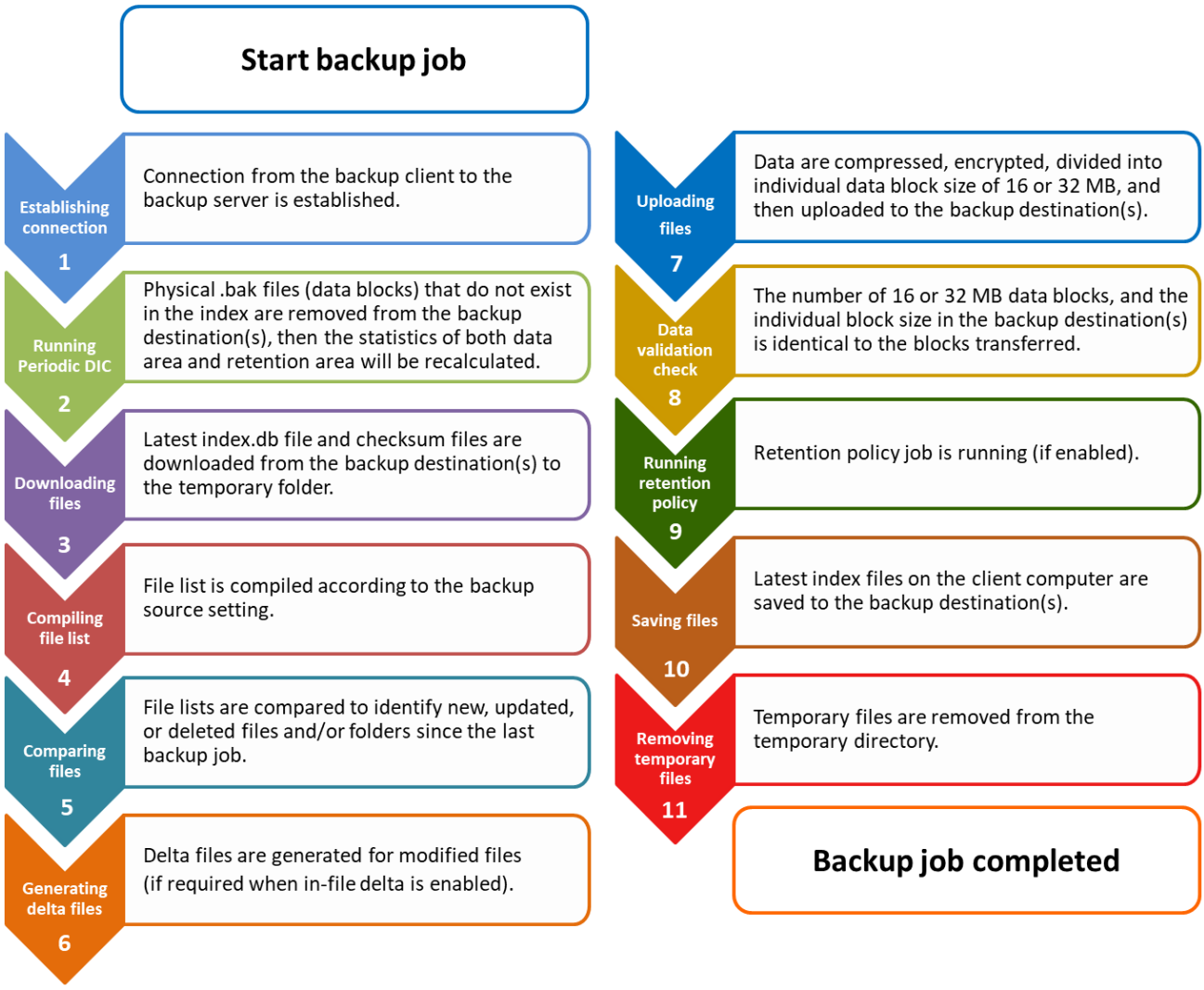
x Files in the missing block(s) will be removed from the index.db file.

y Statistics will be updated according to the files removed.
Proceed to d

k Block size that are not identical will be removed from the backup destination(s).
Proceed to y then
Proceed to d

Overview on the Backup Process (for Run on Server backup)

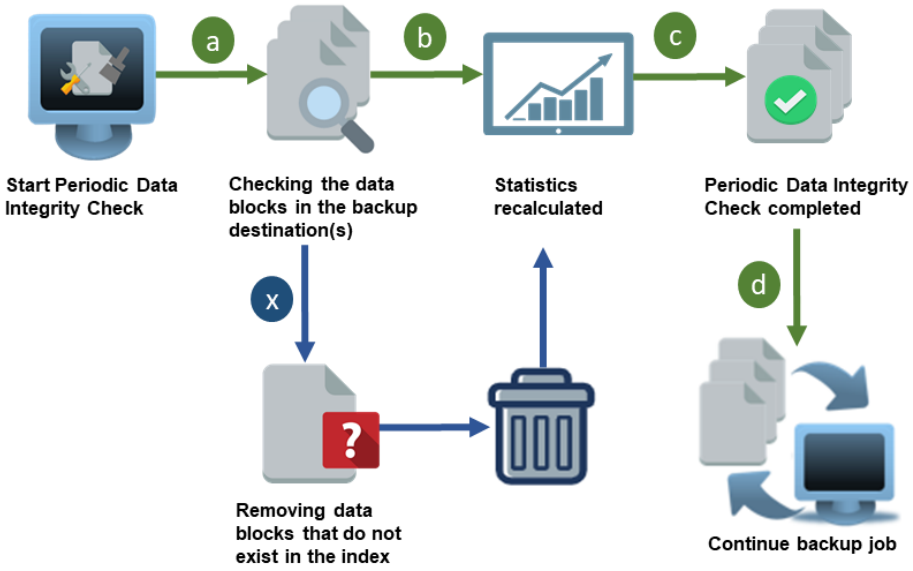
The following steps are performed during a Run on Server backup job. The three new features will take place on steps **2 (PDIC)**, **3 (Backup Set Index Handling Process)** and **8 (Post-Backup Data Validation Check)**.



How it works

Periodic Data Integrity Check (PDIC)

The diagram below shows the detailed process of periodic data integrity check.



a Check the data blocks (.bak files) located in the backup destination(s) if they exist in the index.

➔ If **YES**, proceed to b

➔ If **NO**, proceed to x

b Data area and Retention area statistics are recalculated.

c Periodic Data Integrity Check is completed.

d The backup job process will continue.

x Data blocks (.bak files) that do not exist in the index will be removed from the backup destination(s).

Proceed to b

Periodic Data Integrity Check (PDIC) Will run under the following conditions:

- Will be triggered on a weekly basis, usually on the first run of backup job that falls on any one of these days: **Friday**, **Saturday**, or **Sunday**.

  OR

- If there is no active backup job(s) running on Friday, Saturday or Sunday, then the PDIC will run on the next available backup job.

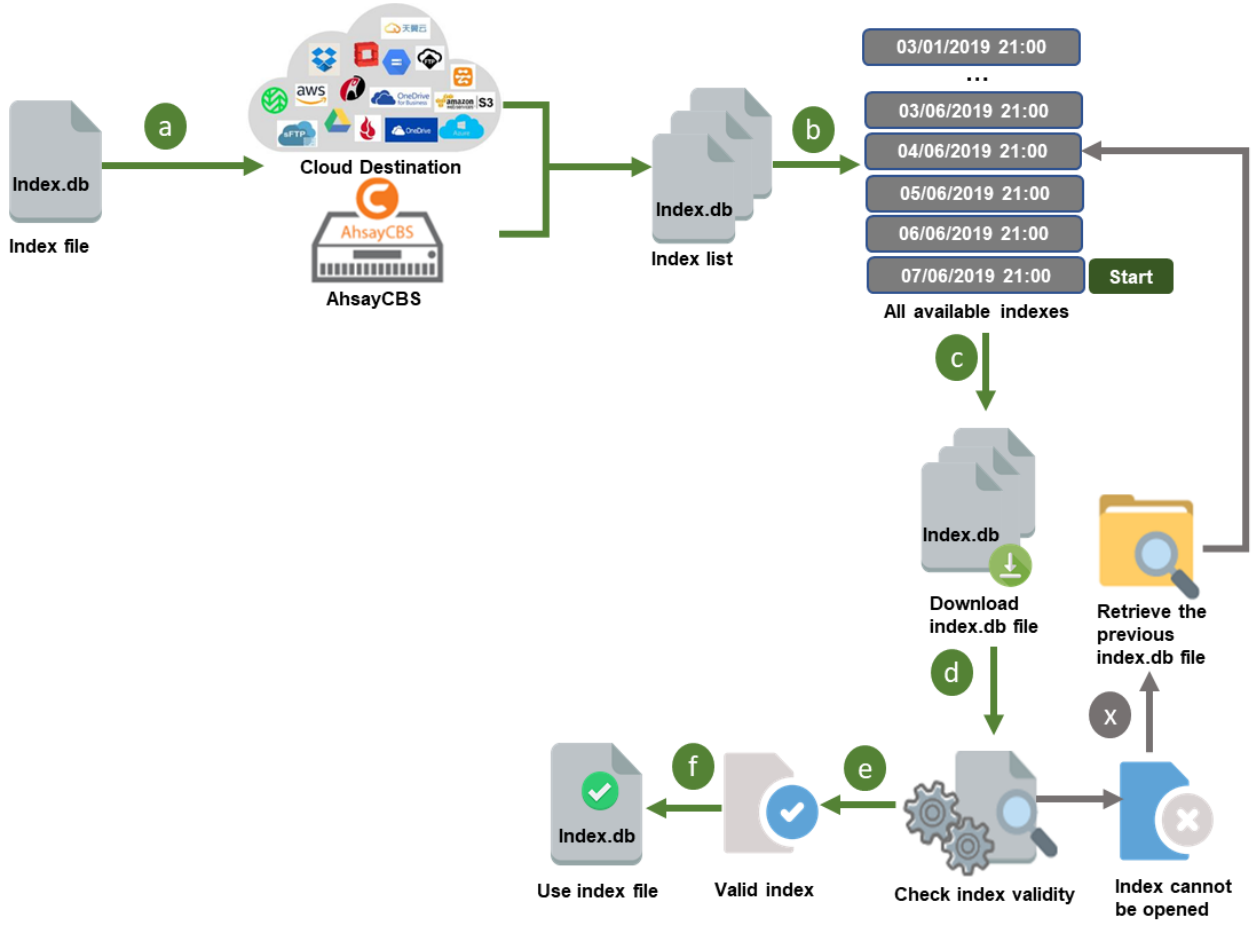
Corrupted data blocks (.bak files) can only be deleted by manually running a full Data Integrity Check (DIC) on AhsayOBM/AhsayACB client or AhsayCBS for Run on Server (Office 365 or Cloud File) backup sets.

This behavior only applies to v8.3.0.30 – v8.3.2.xx

Backup Set Index Handling Process

Run on Server

The diagram below shows the backup set index handling process.



a Index file(s) will be listed in the current directory (e.g. cloud destination and AhsayCBS).

b All available indexes will be sorted according to the modified date and time.

c Latest index file in the list will be downloaded according to the most current modified date and time.

d Index file will be checked to see if it can be opened.

→ If YES, proceed to e

→ If NO, proceed to x

e If it can be opened, the index will be used for the current backup job.

f Use the index.db file to compile file list for backup.

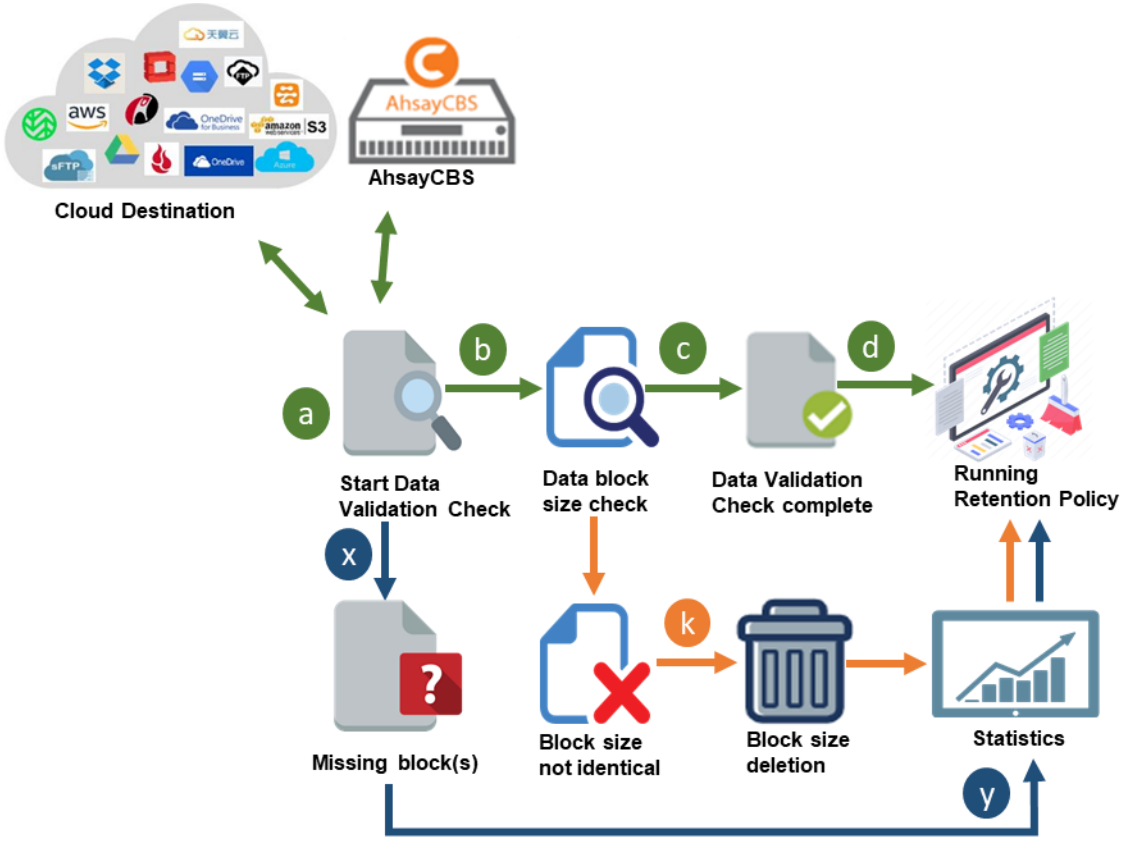
x If it cannot be opened, it will proceed to the previous index.db file from the list until it finds a valid index that can be opened.

For index.db files found to be corrupted and its associated data blocks will no longer be restorable. They will not be removed automatically from the destination(s) and can only be removed by manually running a full Data Integrity Check (DIC) on the AhsayCBS Web Console for Run on Server (Office 365 and Cloud File) backup.

Post-Backup Data Validation Check

Run on Server

The diagram below shows the detailed process of post-backup data validation check.



a Check the number of 16 or 32 MB data blocks in the backup destination(s) is identical to the number of blocks transferred.
- ➔ If YES, proceed to b
- ➔ If NO, proceed to x

b Check the individual sizes of each data block in the backup destination(s) is identical to the sizes of each block transferred.
- ➔ If YES, proceed to c
- ➔ If NO, proceed to k

c Data validation check is complete.

d Retention policy will run (if enabled).

x Files in the missing block(s) will be removed from the index.db file.

y Statistics will be updated according to the files removed.
Proceed to d

k Block sizes that are not identical will be removed from the backup destination(s).
Proceed to y then
Proceed to d