

# LOKKBOX

## DOCUMENTATION IN-FILE DELTA TECHNOLOGY

### Copyright Notice

The use and copying of this product is subject to a license agreement. Any other use is prohibited. No part of this publication may be reproduced, transmitted, transcribed, stored in a retrieval system or translated into any language in any form by any means without the prior written consent of us. Information in this manual is subject to change without notice and does not represent a commitment on the part of the vendor. We do not warrant that this document is error free. If you find any errors in this document, please report to us in writing.

This product includes software developed by the Apache Software Foundation (<http://www.apache.org/>).

### Trademarks

Microsoft, Windows, Microsoft Exchange Server and Microsoft SQL Server are registered trademarks of Microsoft Corporation.  
Sun, Solaris, SPARC, Java and Java Runtime Environment are registered trademarks of Sun Microsystems Inc.  
Oracle, Oracle 8i, Oracle 9i are registered trademarks of Oracle Corporation.  
Lotus, Domino, Notes are registered trademark of IBM Corporation.  
Red Hat is registered trademark of Red Hat, Inc.  
Linux is registered trademark of Linus Torvalds.  
Apple and Mac OS X are registered trademarks of Apple Computer, Inc.

All other product names are registered trademarks of their respective owners.

### Disclaimer

Lokkbox LLC will not have or accept any liability, obligation or responsibility whatsoever for any loss, destruction or damage (including without limitation consequential loss, destruction or damage) however arising from or in respect of any use or misuse of or reliance on this document. By reading and following the instructions in this document, you agree to accept unconditionally the terms of this Disclaimer and as they may be revised and/or amended from time to time by Lokkbox LLC without prior notice to you.

# 1 In-File Delta Technology

This chapter describes what in-file delta technology is and how in-file delta can be used to backup large database files (e.g. a 10GB Outlook.pst file) without uploading the whole database file everyday.

## 1.1 Overview

In-file delta technology is an advanced data block matching algorithm, which has the intelligence to pick up changes (delta) of file content between two files. Using this algorithm, daily backing up of large files (e.g. a 10GB Outlook.pst file) over low-speed internet connections is made possible because it requires only the changes of information (typically marginal) since the last backup to be sent over a low-speed internet connection to complete the backup of a large file (here we assume that the full backup of the file has been saved on the backup server already).

This is what will happen during the backup of a 10GB Outlook.pst file when Lokkbox OBM backs it up with in-file delta technology.

- i. The whole file (10GB), along with its checksum file, is backed up to the backup server. This can be done directly through the Internet or indirectly using the seed loading utility on a removable hard disk.
- ii. When the backup runs again later, Lokkbox OBM will download a checksum listing of all data blocks of the full backup file (or last incremental backup file) from the backup server and use it to pick up all changes that have been made to the current Outlook.pst file.
- iii. Changes detected are then saved in a delta file which is uploaded to the backup server. A new checksum file is generated and is also uploaded. (This delta file is assumed to be small because the content of PST files doesn't change a lot during daily usage)
- iv. Subsequent backups of this 10GB Outlook.pst file will go through step ii and step iii again. As explained, only a small delta file and checksum file will be uploaded to the backup server.
- v. With in-file delta technology, daily backing up of large file over low-speed internet connection is now possible

### Incremental in-file delta type

Incremental Delta will facilitate ease of backup. The delta is generated by comparing with the latest uploaded full or delta file. The delta file generated is the smallest possible and uses the least bandwidth during backup. For restoration, the full file and all delta files up to the required point-in-time is required to restore the file to a specific point-in-time.

Example: You are adding around 100MB changes to a 10G Outlook.pst everyday.

After the first full backup job, Lokkbox OBM will continue to generate and upload delta files instead of uploading the full file until either one of the following two rules are true:

1. Number of deltas generated since the last full backup is greater than the [No. of Delta] setting.
2. The delta ratio (the ratio of the delta file size against the full file size) is greater than the [Delta Ratio] setting

Using incremental delta, the delta generated by comparing the current file with the latest uploaded full or delta file is around 100MB. Since the delta file size is around 100MB daily, the delta ratio (around 1%) will not trigger a full file upload. Instead Lokkbox OBM will continue to generate and upload delta files until day 102 (assuming the full file was uploaded on day 1) when the number of deltas generated exceeds the [No. of Delta] setting and triggers the upload of the full Outlook.pst file. You can disable the [No. of Delta] and [Delta Ratio] if you don't want to upload the full file.

All delta files are generated with respect to changes made since the last incremental or full backup. This means that the last full backup file and **ALL** incremental delta backup files are required to restore the latest snapshot of a backup file.

### Differential in-file delta type

Differential Delta will facilitate ease of restore. The delta is generated by comparing with the latest uploaded full file. The delta file generated grows daily and uses more bandwidth during backup. For restoration, the full file and a single delta file is required to restore the file to a specific point-in-time.

Example: If you are adding around 100MB changes to a 10G Outlook.pst everyday.

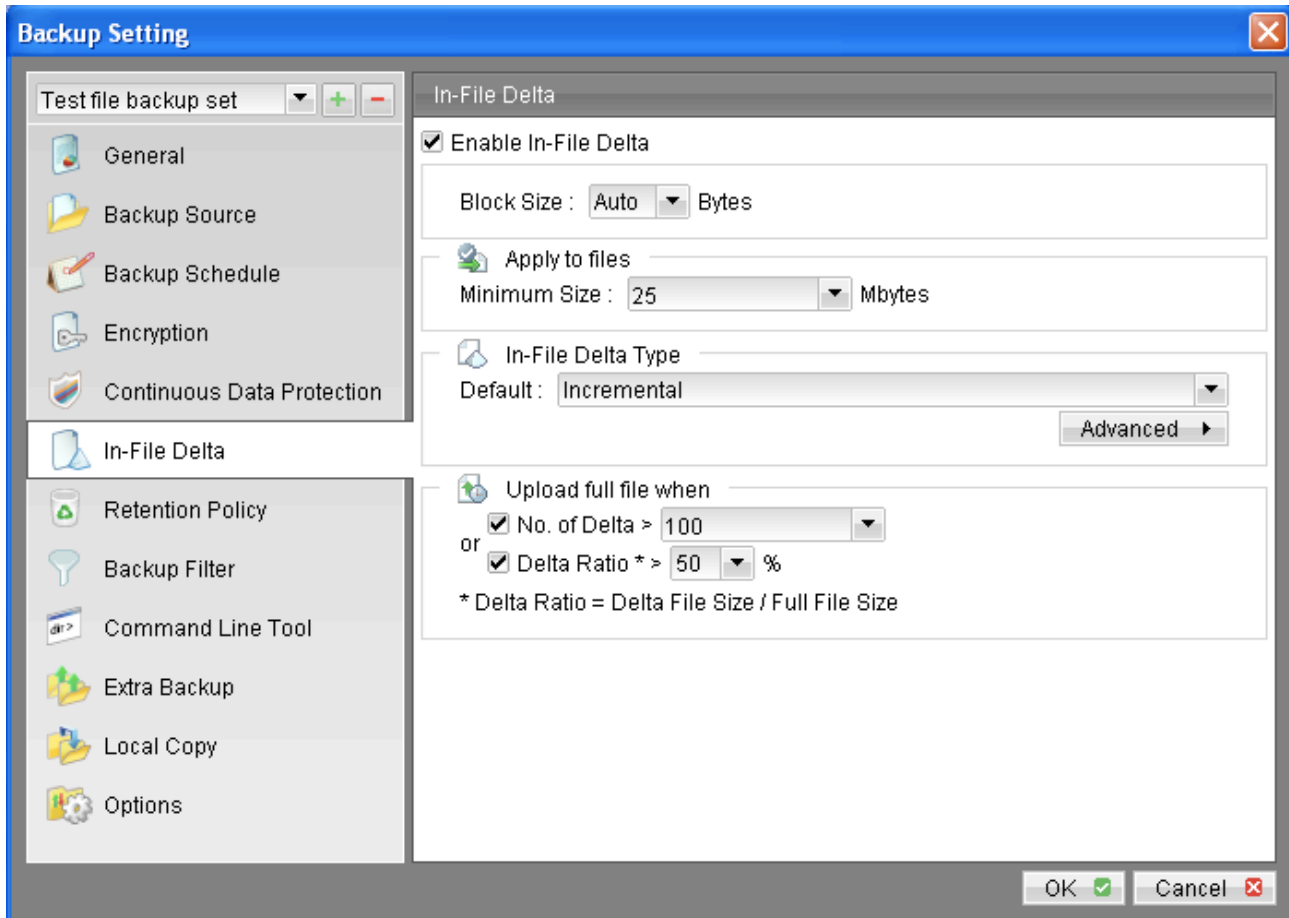
After the first full backup job, Lokkbox OBM will continue to generate and upload delta files instead of uploading the full file until either one of the following two rules are true:

3. Number of deltas generated since the last full backup is greater than the [No. of Delta] setting.
4. The delta ratio (the ratio of the delta file size against the full file size) is greater than the [Delta Ratio] setting

Using differential delta, the delta generated by comparing the current file with the latest uploaded full file is around 100MB for the 1<sup>st</sup> delta, 200MB for the 2<sup>nd</sup> delta, 300MB for the 3<sup>rd</sup> delta and so on. Since the delta file grows by 100MB daily, the delta ratio for day 52 (delta file size is approximately  $51 \times 100 = 5.1\text{GB}$ , assuming the full file was uploaded on day 1) being over 50% exceeds the [Delta Ratio] setting and triggers the upload of the full Outlook.pst file. You can disable the [No. of Delta] and [Delta Ratio] if you don't want to upload the full file.

All delta files are generated with respect to changes made since the last full backup file (i.e. differential backup). This means that only last full backup file and the last delta file are required to restore the latest snapshot of a backup file. This means that other intermediate delta files are only required if you want to restore other snapshots of a backup file.

Differential in-file delta backup has the benefits that a corrupted delta file would only make one particular version of a backup file non-recoverable and all other backups created by other delta files of the same file would still be intact.



## 1.2 Block Size

The block size defines the size of data block being used to detect the changes between the last full or delta backup file and the file sitting on the local computer right now. In general, the smaller the block size, the more likely a matched data block can be found between the last backup file and the file on local computer. It, therefore, produces a smaller delta file but it would require more processing power to detect these changes. On the other hand, in-file delta backup running with larger block size will run faster but this will generally produce a larger delta file.

In most case, the default setting [Auto] will choose the optimal block size for each file (depending on the size of the file) for you.

## 1.3 Minimum File Size

The [Minimum Size] setting defines the smallest file size a file must have before the use and application of in-file delta backup technology.

If the size of a file that is being backed up is smaller than the [Minimum Size] setting, in-file delta backup technology won't be applied to this file and the whole file, instead of just the delta file, will be uploaded to the backup server. It is not necessary to perform in-file delta backup on small files because backing up the whole file doesn't take too long anyway. Backing up the whole file instead reduces the time required to restore a backup file.

## 1.4 Uploading full file again

### No. of Delta

The [No. of delta] setting defines the maximum number of delta files from the same full backup file to be generated and backed up to the backup server before a full backup (the whole file) of this file is uploaded to the backup server instead.

For example, if you have created 100 delta files from the full backup file already and the [No. of delta] setting is set to 100, the next backup will upload a full backup file (the whole file) instead of just the delta file. However, if the [No. of delta] rule is disabled, it will keep generating delta files and uploading these delta files to the backup server until the other delta rule forces a full backup (i.e. delta ratio exceeded). This setting is here to make sure that there will always be a full backup file after a certain number of delta files have been generated.

### Delta Ratio

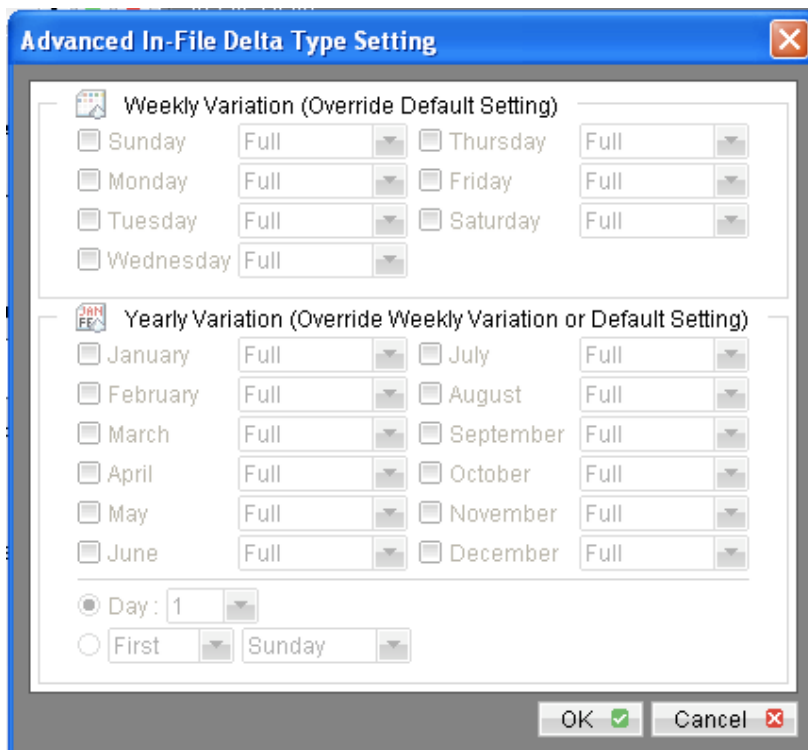
The [Delta Ratio] setting is defined to be the size of a delta file divided by the size of its full backup files (i.e. the percentage of changes detected between the last full backup file and the current file). If the delta ratio calculated from the size of the generated delta file and the size of the full backup file is greater than the [Delta Ratio] setting, the whole file, instead of just the delta file, will be backed up to the backup server. It is done this way because the difference between backing up the whole file

and the delta file is not significant and it is better to upload the whole file instead to reduce the time required to restore the file.

The default setting of [Delta Ratio] is 50%. This means that if a file changes by more than 50%, the whole file, instead of just the delta file, will be backed up and uploaded to the backup server.

## 1.5 Advanced In-file delta type

The [In-file delta] -> [Advanced] setting allows users to override default in-file delta type when on a certain number of days (e.g. all Sundays or the 1<sup>st</sup> day of each month). This is useful if you want all in-file delta backups to be incremental but you want to do a differential in-file delta backup on Saturdays as well as the full file backups on the 1<sup>st</sup> day of every month. If you want to do this, simply press the [Advanced] button and the [Advanced In-File Delta Type Setting] dialog will be popped up as shown below.



With this setting, all backup Jobs on Saturdays will be differential in-file delta backups, backups for the 1<sup>st</sup> day of each month will run a full backup. This ensures that all backup files will be backed up in full at a regular interval. One benefit of this is that restore time will run faster because of less delta merging. Another benefit is that the risk of a corrupted incremental delta file resulting in data loss is much lower because a full backup is always available periodically.