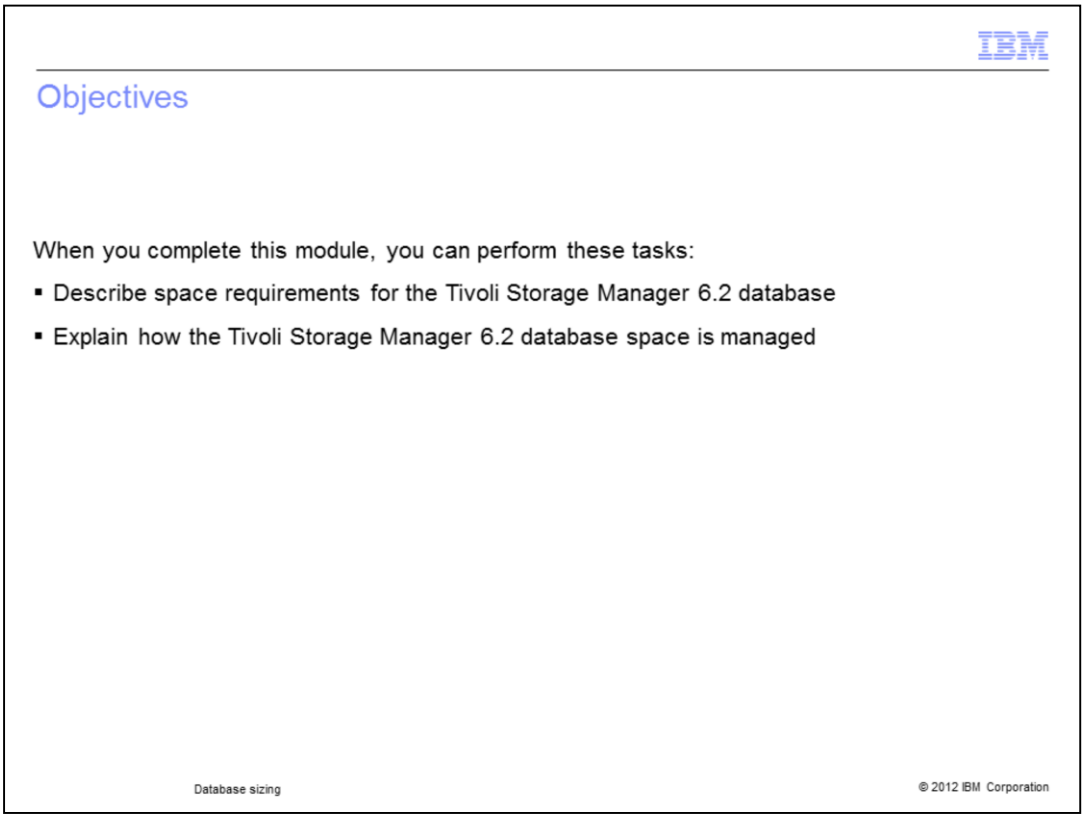
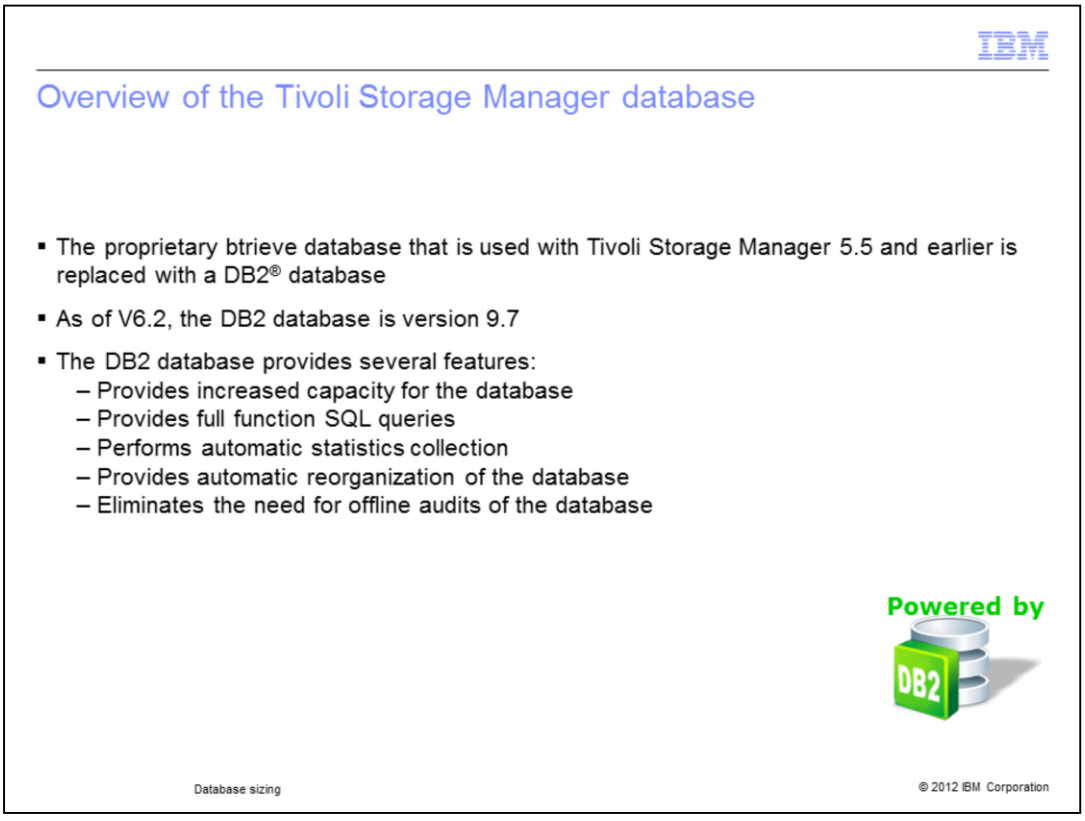
Welcome to the IBM Education Assistant module for Tivoli® Storage Manager version 6.2 database sizing.




When you complete this module, you can describe space requirements for the Tivoli Storage Manager 6.2 database. You can explain how the Tivoli Storage Manager 6.2 database space is managed.




During this presentation, all references to database are the Tivoli Storage Manager version 6.2 database.

- The Tivoli Storage Manager database contains information that is needed for server operations and information about client data that has been backed up, archived, and space-managed. The database does not store client data. Instead, the database points to the locations of the client files in the storage pools.

- The DB2 database is embedded in Tivoli Storage Manager V6 and does not require a separate installation or DB2 administration skills. It is packaged, installed, and serviced with Tivoli Storage Manager.

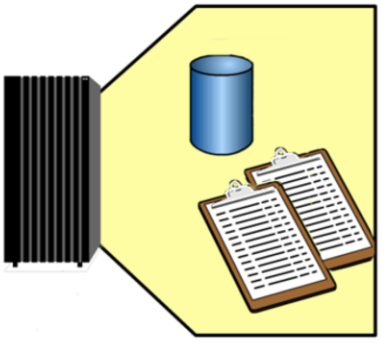
- Using DB2 as the Tivoli Storage Manager database enables next-generation management, scalability, availability, and performance. The maximum supported database size is up to 2 TB. A single Tivoli Storage Manager server can manage as many as 1,000,000,000 objects.

## Managing the database and recovery logs

The use of a DB2 database changes some tasks that are part of the daily management of the Tivoli Storage Manager database and recovery logs.

Examples

- Monitoring the database and recovery logs
- Extending the database space
- Increasing the size of the active log
- Using the database manager





© 2012 IBM Corporation

The server uses the Tivoli Storage Manager database to manage information about client files.

The recovery log ensures consistency and availability of the database.

## Recovery log usage

- Changes to the database are recorded in the recovery log for maintaining a consistent database image
- The recovery log includes updates to information that can include activities as follows:
  - Defining a management class
  - Backing up a client file
  - Registering a client node
- For the best availability and protection, you can specify that active logs be mirrored. Locate those mirror logs on a separate physical device

Database sizing

© 2012 IBM Corporation

The server uses the recovery log as a scratch pad for the database. The log records information about client and server actions while the actions are being performed. The log keeps all transactions since the last database backup.

- The two recovery logs are the active log and the archive log.

- The recovery log helps ensure that a failure does not leave the database in an inconsistent state. The recovery log is also necessary for restoring the database.

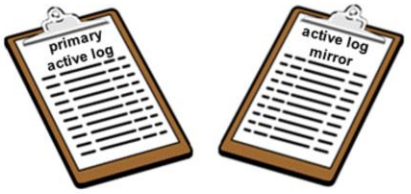
- All changes made to the database since the last database backup are saved in the recovery log. With an intact recovery log, you can recover the database up to its most current state (the point when the database was lost).

- Frequent database backups reduce recovery log storage requirements. After a backup is completed, recovery log records that precede the backup are deleted.

IBM

## Active logs

- Active log: The active log is used for storing in-flight transactions for the server. The transactions are transactions that are not committed yet. The active log always contains the most recent log records
- Active log mirror: The log mirror is a duplicate copy of the active log. All changes that are made to the active log are also written to the log mirror. Having the log mirror for the active log can protect the database. The protection occurs when a hardware failure occurs on the device where the active log is stored. Setting up the log mirror is optional




Place the active log directory and the active log mirror directory on different physical devices

Database sizing

© 2012 IBM Corporation

The active log must be large enough to hold the largest concurrent workload that the server might encounter. It must be able to store the data that represents in-flight transactions for the largest concurrent workload that the server supports.

- The active log is a fixed-size, non-circular log, composed of 512 MB files. A log file cannot be deleted until all transactions in it end.

- DB2 creates a new log file, then deletes the old one. This action requires 512 MB of free space in active log directory. Active log files are copied to archive log directory after the log is full. Transactions might still be active when the log is archived.

## Archive logs

- **Archive log:** The archive log is a closed and stored copy of the active log at an earlier time. Archived log files are saved until they go into a database backup. This log is not needed for normal processing, but is typically needed for recovery of the database.
  Transfer of log files from the active log to the archive log is automatic
- **Archive failover log:** The archive failover log is the directory that the server uses. This log is for storing archive log files that cannot be stored in the archive log directory

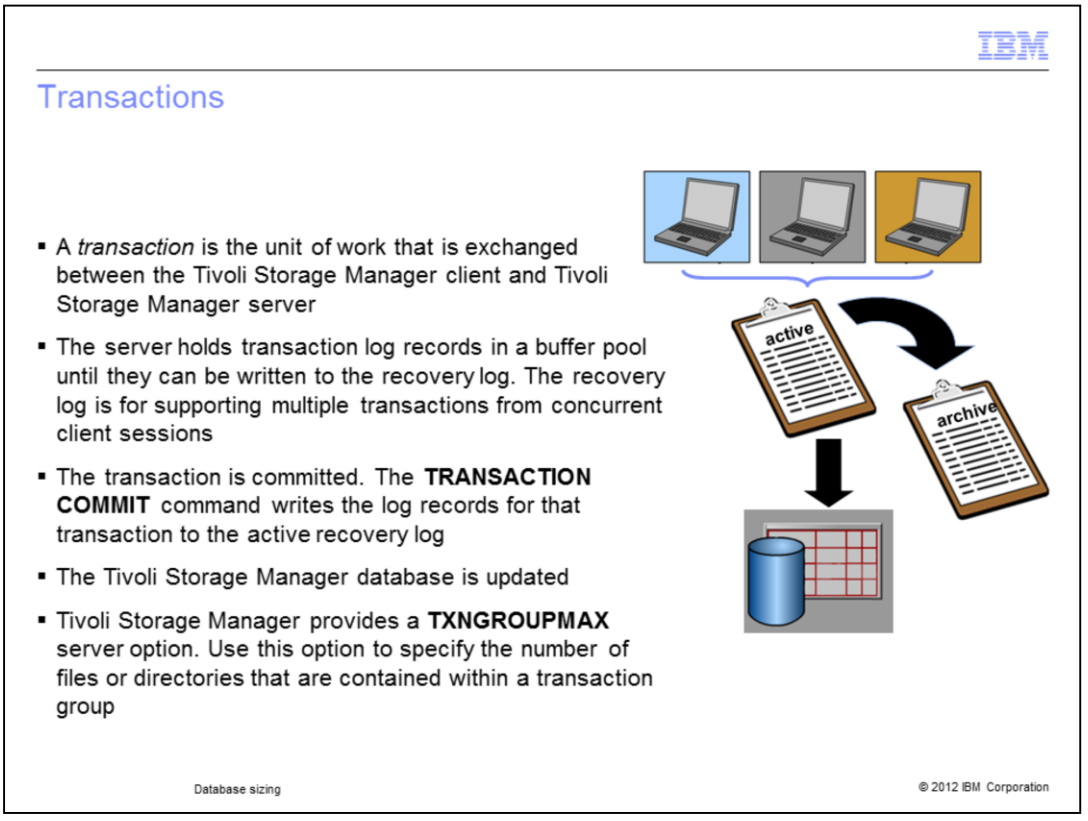
Database sizing

© 2012 IBM Corporation

The archive log is used as a storage repository for log files that were previously active and no longer contain references for in-flight transactions. After a log file in the active log directory is no longer referenced by an in-flight transaction, that file moves to the archive log directory. The movement of log files from the active log directory to the archive log directory occurs automatically.

- The archive log is included in database backups, and is used for roll-forward recovery of the database. At the end of a full database backup, space is recovered by the automatic pruning of older archive log files that are no longer needed. The archive log files that are included in a backup are automatically pruned after two more full database backups have completed.

- If the archive log directory is full, active log files cannot be copied. If active log files cannot be copied, they cannot be deleted. This leads to out-of-log space in the active log. It is very important that this directory does not fill up. The failover archive log directory is used, if defined.



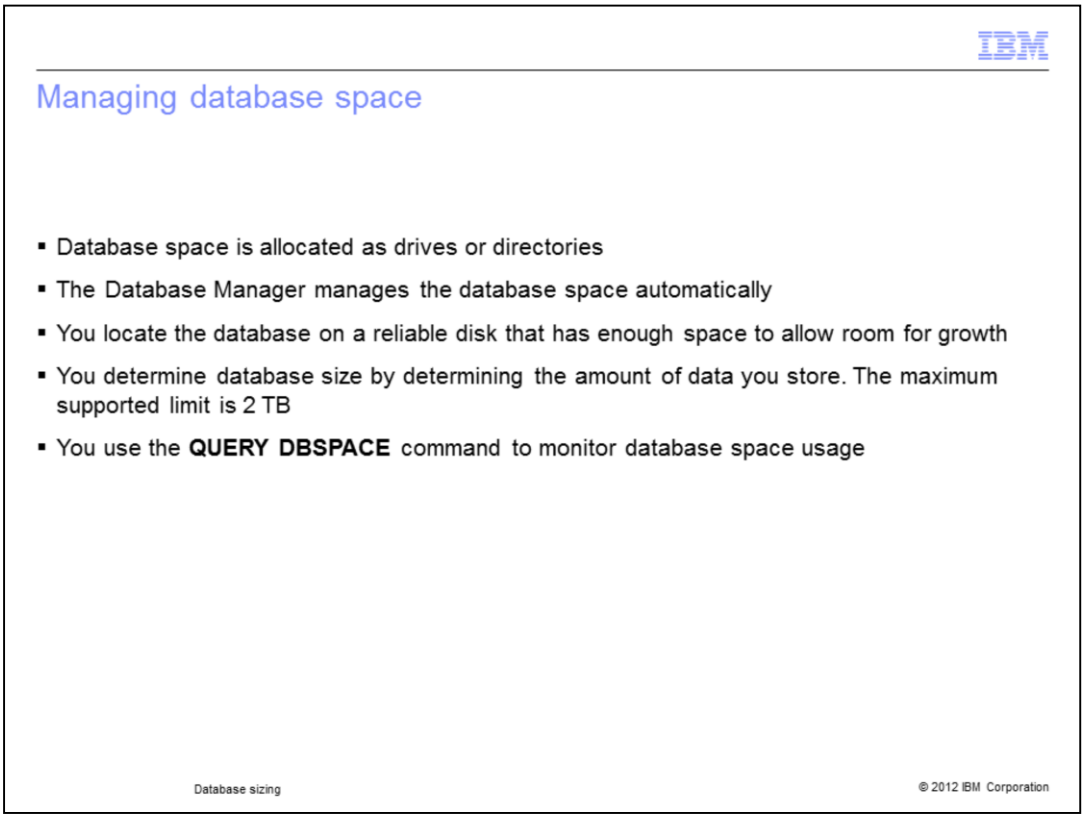

When a transaction occurs, the Tivoli Storage Manager server performs these actions:

1. It reads a database page into the database buffer and updates it. A page is a 4096 byte block that transfers as a unit between memory and disk storage.

2. A transaction commit occurs. It writes a transaction log record to the recovery log. It describes the action that is occurring and associates it with the database page in case the database page requires rollback during recovery.

3. The Tivoli Storage Manager database updates. The database page is written to the database, releasing it from the buffer pool. The page remains in the buffer pool until buffer space is needed for another page.

**Note:** A transaction can contain more than one file or directory. This is called a transaction group. Tivoli Storage Manager provides a TXNGROUPMAX server option so that you can specify the number of files or directories contained within a transaction group. You can affect the performance of client backup, archive, restore, and retrieve operations by using a larger value for the TXNGROUPMAX option. You can use the TXNGROUPMAX option to increase performance when Tivoli Storage Manager writes to tape. This performance can be considerable when a user transfers multiple small files. Be sure to monitor the effects on the recovery log if you increase the value of TXNGROUPMAX by a large amount. The larger value can increase utilization of the recovery log and the length of time for a transaction to commit.




The size of your Tivoli Storage Manager database depends on these factors:

- The number of client files to be stored.

- The number of versions of backed up files you keep.

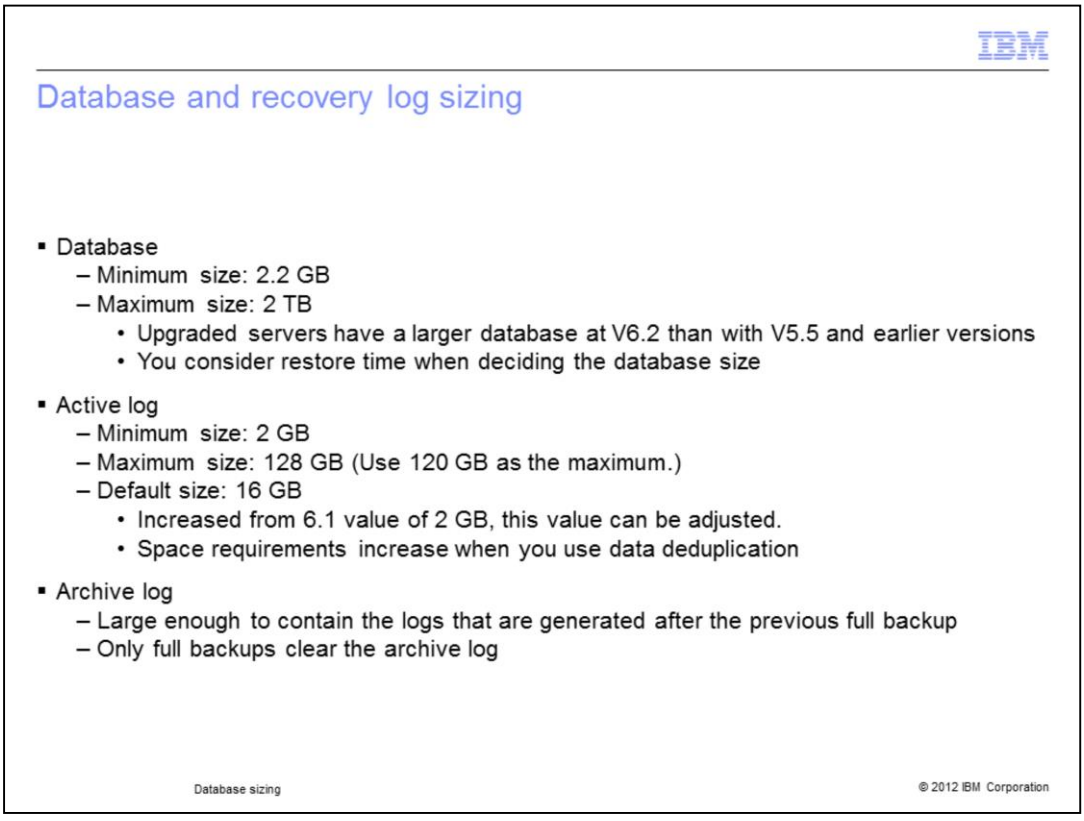
- Whether you use caching.

**Note:** You create and designate directories or drives that the server can use for the database. Afterwards, the database manager, which is part of the server, automatically manages the space available to the directories as database space.

- Locate the database on fast reliable storage, such as Redundant Array of Independent Disks (RAID) hardware.

- Do not place the database in directories that might run out of space.

- Locate the directories across different physical devices (separate LUNs).



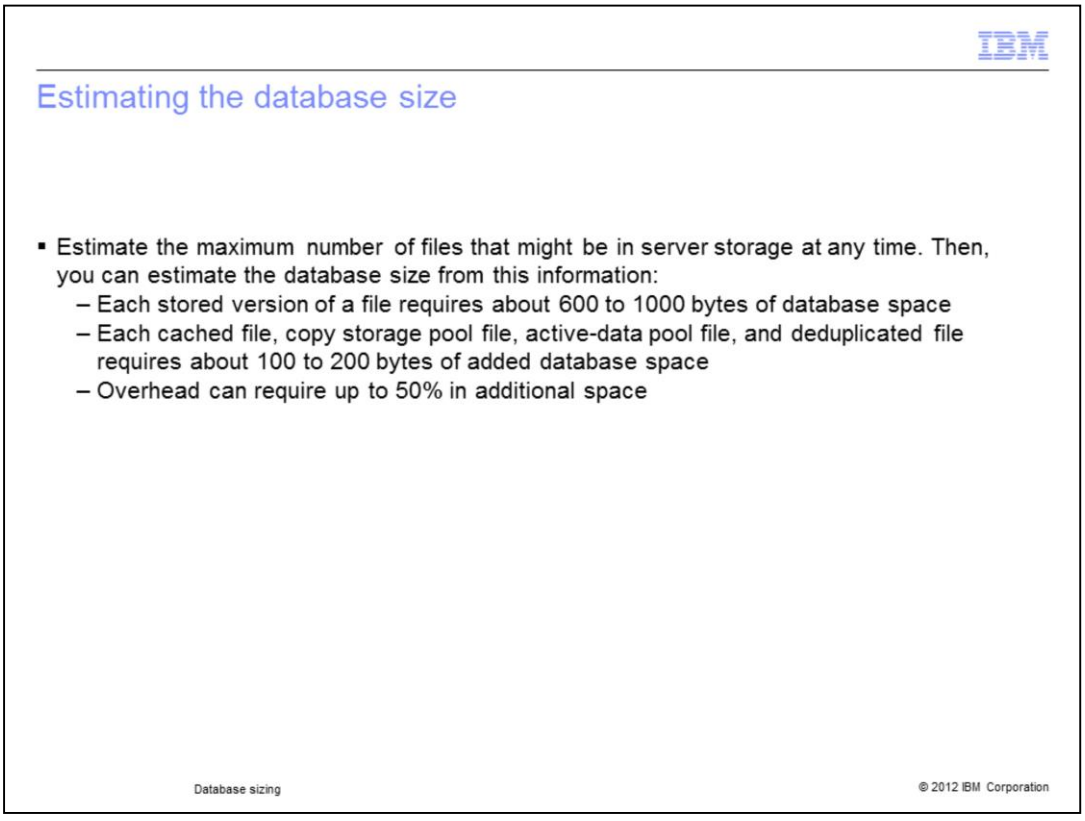

If space for the recovery log is limited, you might need to perform more frequent full backups of the database so that the automatic pruning operation can recover the space.

- The size of the archive log depends on the number of objects stored by client nodes over the period of time between full backups of the database. A full backup of the database causes obsolete archive log files to be pruned, to recover space.

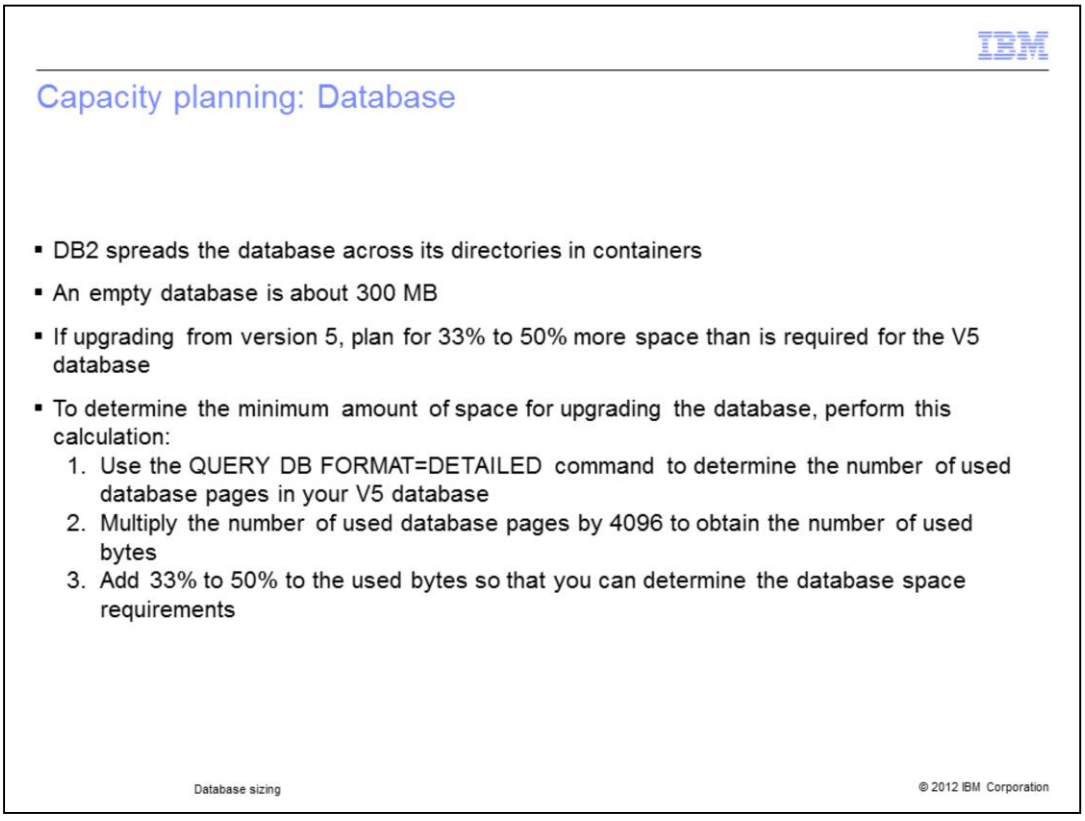
- The archive log files that are included in a backup are automatically pruned after one full database backup has completed. It is important to maintain adequate space for the archive log directory. If the drive or file system where the archive log directory is located becomes full and there is no archive failover log directory, the data remains in the active log directory. This condition can cause the active log to fill up, which causes the server to stop.

- For simple backup and archive activity with no data deduplication, 20 GB for the active log is adequate. If you use data deduplication, and if you deduplicate large objects (for example, image backups), use an active log size that is 20% of the database size.

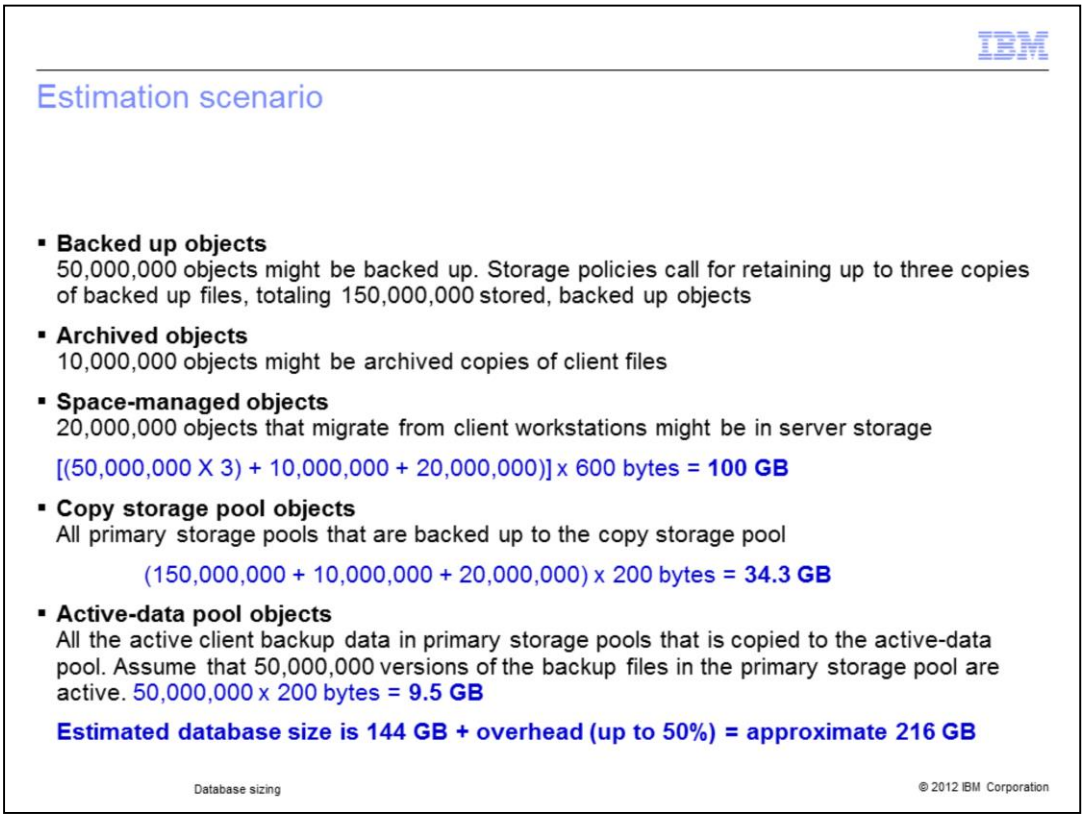
- Raw logical volumes are not supported. If you are upgrading from a system that used raw logical volumes for the database or logs, you can reuse those volumes for storage pool space.



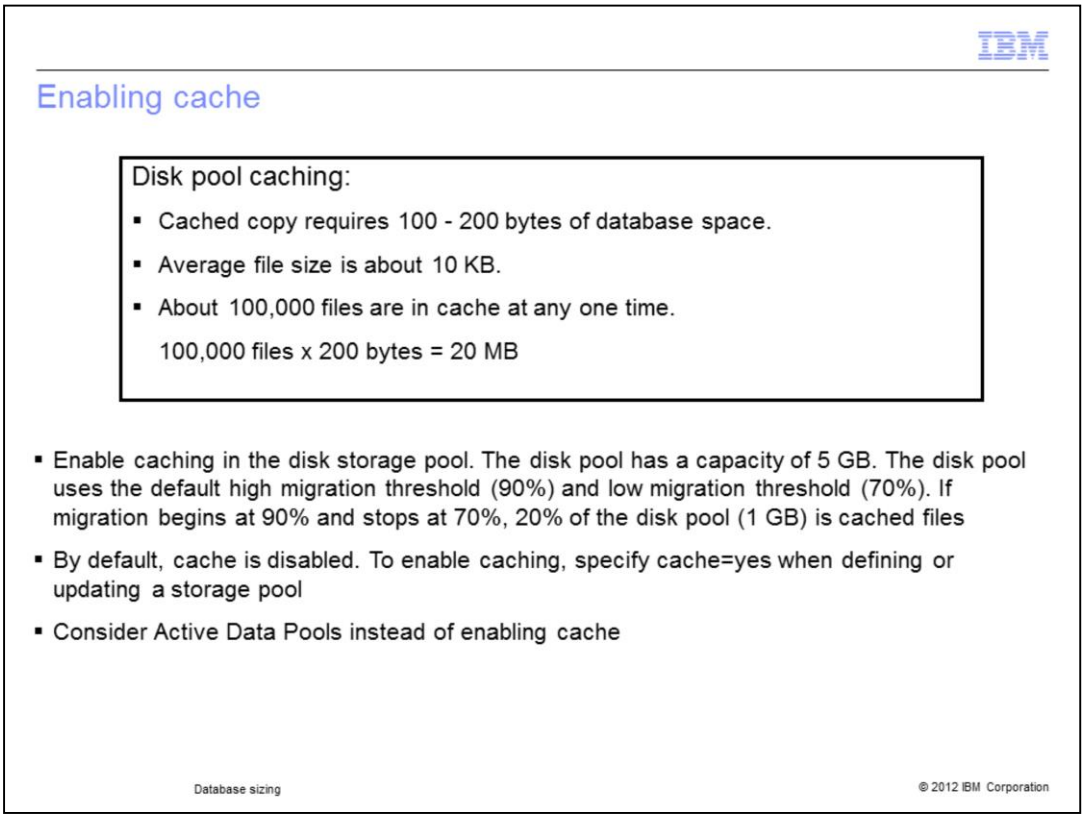

If you cannot estimate the numbers of files, you can roughly estimate the database size as from 1% to 5% of the required server storage space. For example, if you need 1,000 GB of server storage, your database should be between 10 GB and 50 GB.




Adding a new database directory after the initial load causes a database reorganization to occur, which is time-consuming and can result in server down time.




In this example, the computations are probable maximums. The numbers are not based on the use of file aggregation. The more that small files are aggregated, the less database space that is required.




When using cache, consider these advantages and disadvantages:

- When cache is enabled, the migration process leaves duplicate copies of files behind after the server migrates these files to the next storage pool in the hierarchy.

- Using cache can improve the speed that the server retrieves some files with. Consider enabling cache for space-managed files that clients frequently access. If you need space for storing new data in the disk storage pool, cached files are erased, and the space is used for those new files.

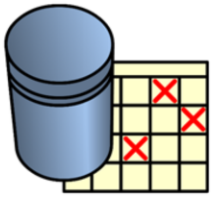
- Using cache can require more space for the server database because the server maintains two copies of the file. If you want to use caching, you cannot also enable shredding for that disk storage pool.

## Increasing the size of the database

- You can increase the size of the database by creating new directories and adding them to the database
- You issue the **EXTEND DBSPACE** command to add one or more directories to the database. Examples are as follows:

```
extend dbspace e:\tsm_db\dbs4
extend dbspace /tsmdb/dbs8
```

- The directories must be empty and accessible to the user ID of the database manager
- The *IBM Tivoli Storage Manager Performance Manual* has more information



Database sizing

© 2012 IBM Corporation

The size of the database depends on the number of client files to be stored and the method that the server manages them with. If you can estimate the maximum number of files that might be in server storage at any time, you can estimate the database size from this information:

- Each stored version of a file requires about 400 to 600 bytes of database space.

- Each cached or copy storage pool file requires about 100 to 200 bytes of database space.

- Overhead can require up to 25% in additional space.

**Note:** The Administrator's Guide for your platform has more information about how to size your database.

## Increasing the size of the active log

To increase the size of the active log, complete these steps:

1. Issue the DSMSERV DISPLAY LOG offline utility to verify the size of the active log
2. In the **dsmserv.opt** file, update the ACTIVELOGSIZE server option to the new maximum size of the active log. For example, to change the active log to 120 GB, enter the server option: activelogsize 122880
3. If you want to use a new active log directory, update the directory name that is specified in the ACTIVELOGDIR server option

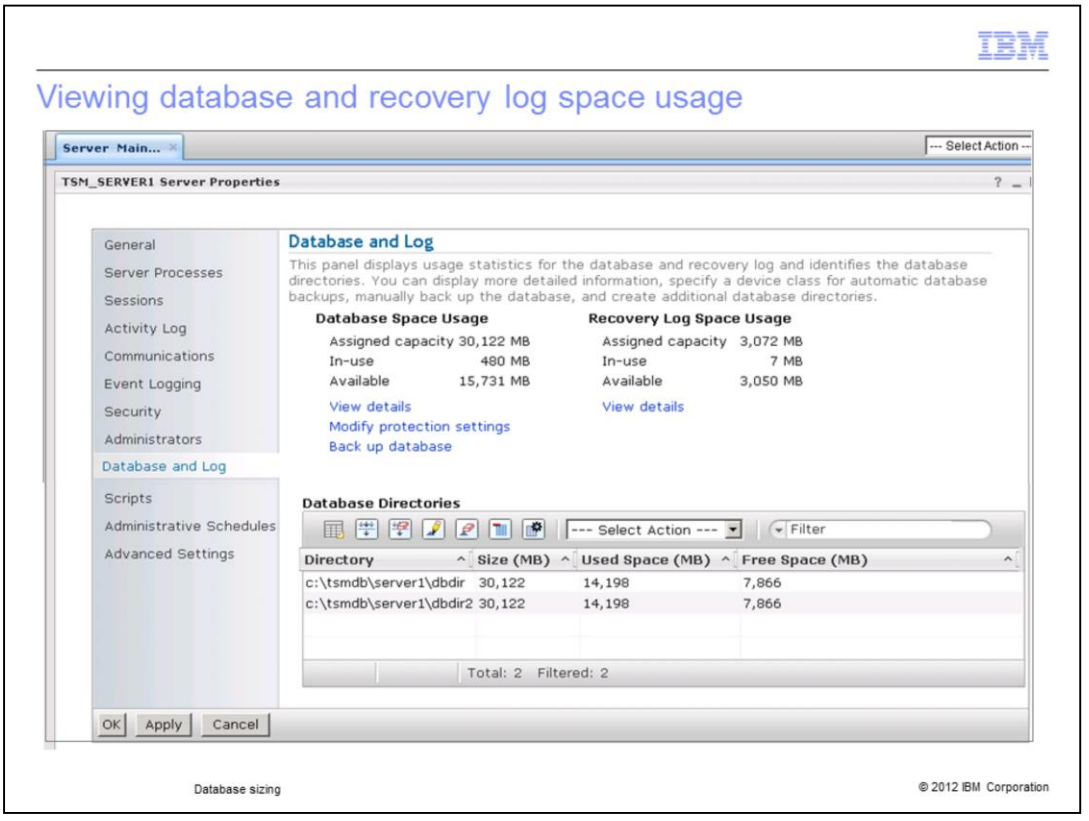
   The changes take effect when you restart the server

If the log is running out of space, the current transaction is rolled back. The server issues an error message and halts. You cannot restart the server until you increase the active log size.

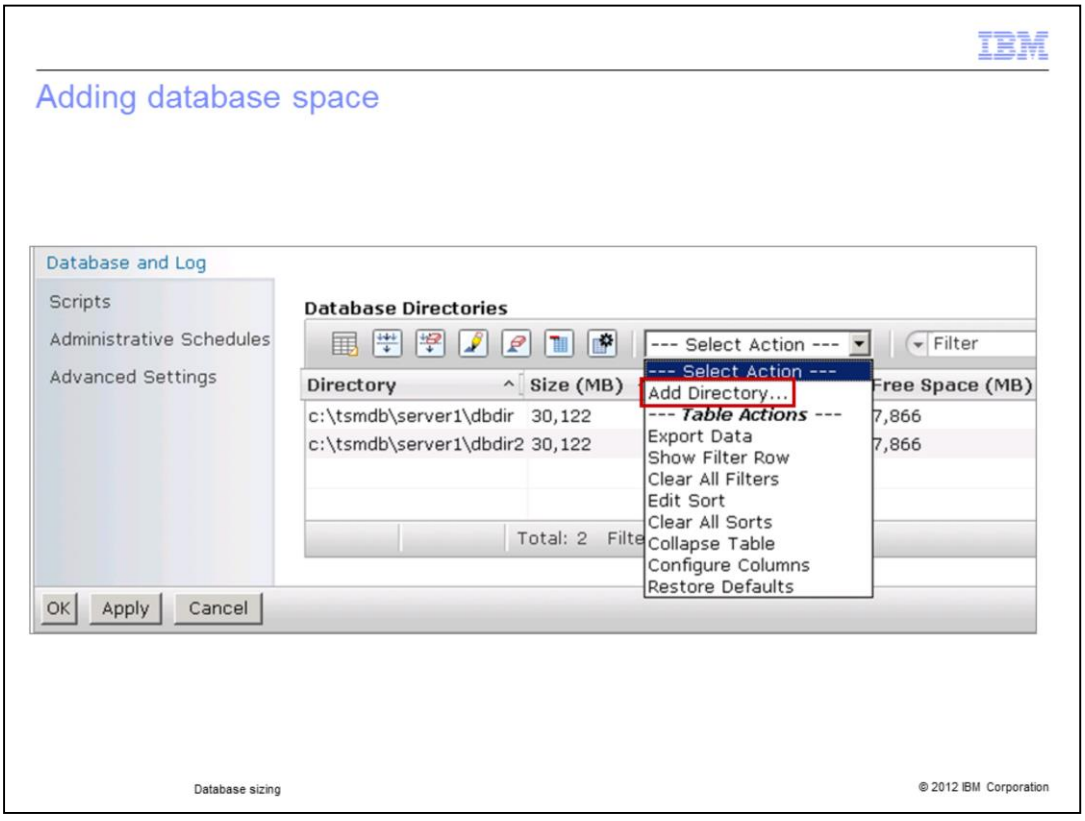
Database sizing

© 2012 IBM Corporation

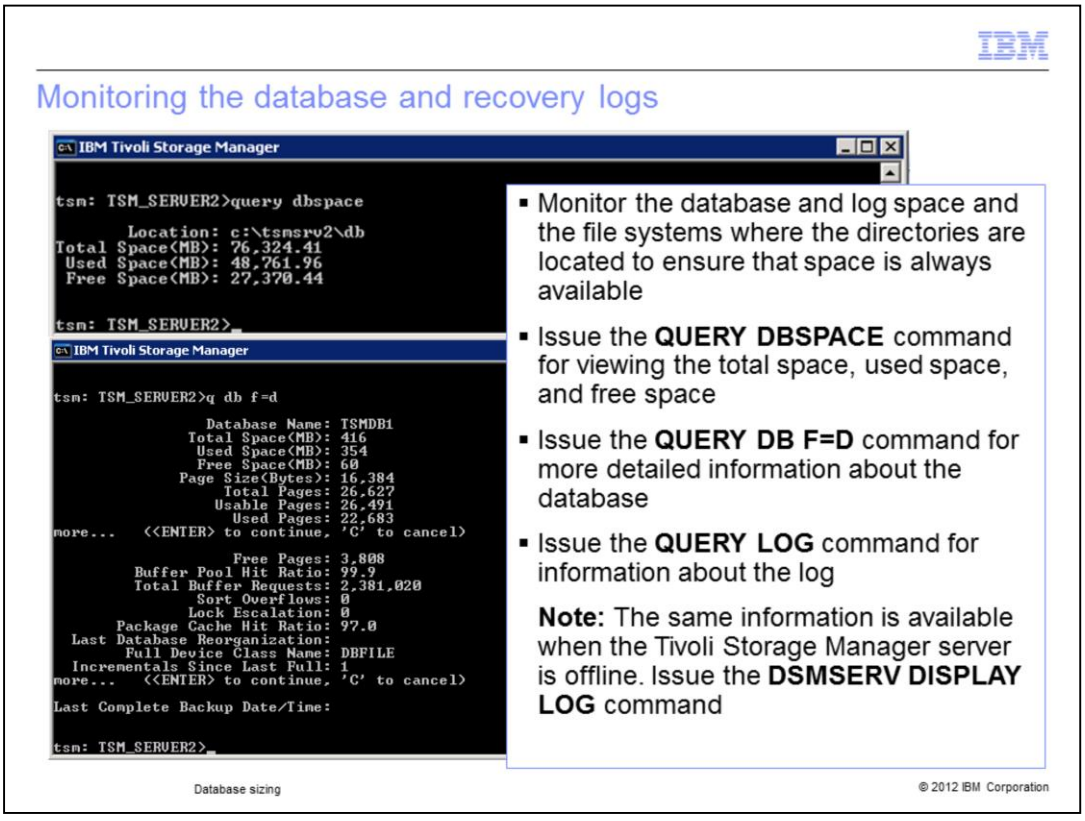
If you use the active log mirror, it increases to the new size of the active log.




You can view database and recovery log space usage in the Administration Center. This page displays the same information that you can obtain by using the query dbspace and query log commands.




You can use the Administration Center to add database space while the server runs.

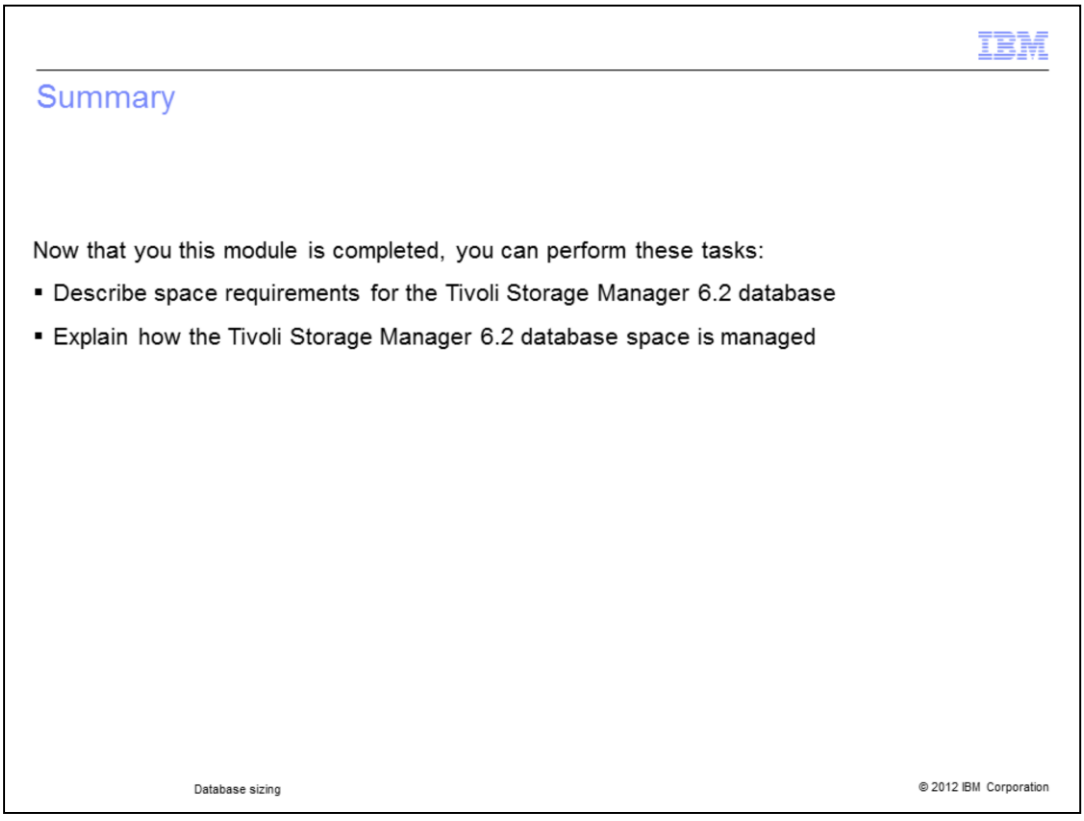



Information about the database can be displayed on the service console and in the activity log. You can set the level of that information by issuing the set dbreportmode command. Three options for the dbreportmode command are as follows:

- NONE: No diagnostic information is displayed

- FULL: All diagnostic information is displayed

- PARTIAL (the default): Only exceptions that might represent errors are displayed




This concludes the IBM Education Assistant module for Tivoli Storage Manager version 6.2 database sizing.

Now that this module is completed, you can perform these tasks:

- Describe space requirements for the Tivoli Storage Manager 6.2 database

- Explain how the Tivoli Storage Manager 6.2 database space is managed

IBM

# Trademarks, disclaimer, and copyright information

IBM, the IBM logo, ibm.com, DB2, and Tivoli are trademarks or registered trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of other IBM trademarks is available on the web at "Copyright and trademark information" at http://www.ibm.com/legal/copytrade.shtml

THE INFORMATION CONTAINED IN THIS PRESENTATION IS PROVIDED FOR INFORMATIONAL PURPOSES ONLY. THE INFORMATION CONTAINED IN THIS PRESENTATION IS PROVIDED FOR INFORMATIONAL PURPOSES ONLY. WHILE EFFORTS WERE MADE TO VERIFY THE COMPLETENESS AND ACCURACY OF THE INFORMATION CONTAINED IN THIS PRESENTATION, IT IS PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED. IN ADDITION, THIS INFORMATION IS BASED ON IBM'S CURRENT PRODUCT PLANS AND STRATEGY, WHICH ARE SUBJECT TO CHANGE BY IBM WITHOUT NOTICE. IBM SHALL NOT BE RESPONSIBLE FOR ANY DAMAGES ARISING OUT OF THE USE OF, OR OTHERWISE RELATED TO, THIS PRESENTATION OR ANY OTHER DOCUMENTATION. NOTHING CONTAINED IN THIS PRESENTATION IS INTENDED TO, NOR SHALL HAVE THE EFFECT OF, CREATING ANY WARRANTIES OR REPRESENTATIONS FROM IBM (OR ITS SUPPLIERS OR LICENSORS), OR ALTERING THE TERMS AND CONDITIONS OF ANY AGREEMENT OR LICENSE GOVERNING THE USE OF IBM PRODUCTS OR SOFTWARE.

© Copyright International Business Machines Corporation 2012. All rights reserved.

© 2012 IBM Corporation