Transbase®

WHITE PAPER

Transbase® on Network Filesystems (NAS)

NAS disks provide non-volatile storage over fast LAN connections. For various reasons NAS disks have become very popular and this raises the question if Transbase® can be run on NAS disks. The network protocols used with NAS disks include NFS and SMB.

Transbase® guarantees transactional consistency which means that each transaction either becomes committed or rolled back completely, even after crashes of the database application, the database system or the database host. Transbase®, however, does not protect against disk crashes. Therefore, administrators have to run database backups regularly. Transbase® provides tools that allow a backup to be run while the database is operational. After a disk crash such backups can be used to redo the database to its latest state (or an earlier state defined by users in case the database had been corrupted by software). The amount of transactions missing after a crash typically depends on the frequency of backup operations.

1. Functional Aspects

To implement transactional consistency, Transbase® must know from time to time that an IO has successfully completed and the data persistently has been written to disk. It is particularly not sufficient that the data be written into caches driven by modern operating systems or by disk controllers. In other words, the data must have been written through all caches onto disk.

Technically Transbase® uses fsync calls that force the host operating system to flush all buffers belonging to a given file. A fsync call is particularly issued on log files for each COMMIT operation of a transaction.

Both NFS and SMB guarantee that a fsync call not only flushes its local buffers but also the NAS server flushes the file to disk. This behaviour means that Transbase® also can run on NAS disks without suffering any data inconsistencies.

2. Performance Aspects

NAS disks tend to be a little slower than directly attached disks. Therefore performance degradation may be observed particularly on transactional query profiles with many short transactions to be committed. Some administrative measures may help to prevent such degradation:

- To increase the Transbase® cache size may help to simply reduce the number of physical IO operations, in particular for all reading transactions.
- Logfiles could be placed on locally attached disks, so that commit processing is still fast. Data files still can reside on NAS disks.
- If logfiles also reside on NAS disks, the internal log buffers controlled by Transbase® should be increased so that data transfers between the host operating system and the NFS server are less frequent and move larger data sizes each time.
- In future versions of Transbase® it will be possible to process log IO in background so that the file latency is reduced. However, the COMMIT still requires operation be performed over the network the result of which must be waited for synchronously.

3. Backup Strategies

As mentioned, Transbase® databases must be backed up in order to protect against disk failures or other data losses. As long as the database is shut down while a backup is running, there are no special considerations to achieve a consistent backup of a database. Be sure that all files belonging to the database are backed up, in particular: all configuration files (`dbconf.ini` and `dblist.ini`) and all diskfiles (`tbdsk<n>`) which may be distributed over more than a single file system or disk node.

When however, the database shall be backed up while it is operational, the database must be run in so-called delta-logging mode and the backup can only be performed by the Transbase® tool tbadmin -drec which is described in the Transbase® system guide in detail. It is important that such backups are consistent only if both diskfiles and logfiles are saved.

In addition, tbadmin -drec may also be used to produce incremental backups simply by adding logfiles to a previous backup. In this case it is important that the sequence of logfiles is continuous since a later restore operation would have to stop upon the first hole in this sequence.

NAS disks provide a facility to produce so-called snapshot backups which guarantee that a consistent state of a disk is taken. Such snap-shot backups can be used to backup Transbase® databases even when in operation, provided that all files belonging to a Transbase® database are part of that snapshot and are taken at the same time. In particular, all such files should reside on the same physical disk.

Contact

Transaction Software GmbH
Willy-Brandt-Allee 2
81829 Muenchen, Germany

Tel.: +49 89 / 627 09 - 0
Fax: +49 89 / 627 09 - 11

info@transaction.de
www.transaction.de
www.transbase.de