

Optimizing backup with HP TurboSTORE/iX

By Paul Raj D. and Bijo S. Kappen

Objective

This paper aims at identifying the software options for optimum backup performance using HP TurboSTORE/iX with different system configurations and data patterns.

Using TurboSTORE options

The following are the TurboSTORE options that can be considered for performance optimization:

- | Using high software compression (COMPRESS=HIGH option):

Typically, this option takes about 80% of CPU time on a single CPU machine for a high-end target device (e.g. DLT8000) and is CPU intensive. This option gives the best performance when the system load is very low (e.g. nightly backups when all users/sessions are logged out). It is recommended not to use this option, if the CPU load is expected to peak during the backup window. Also, using HIGH compression along with hardware compression may degrade the backup performance.

- | Using low software compression (COMPRESS=LOW option):

Typically, this option takes about 30% of CPU time on a single CPU machine for a high-end backup device (e.g. DLT8000). To achieve a better performance, this option can be used along with hardware compression. On a heavily loaded system, it is recommended to use only hardware compression.

- | Interleaving of files on backup (INTER option):

Interleave option gives best performance:

- | when files are evenly distributed across disks.
- | when backup file-set contains many large files.

This option may degrade the backup performance when used with many small files. Presently TurboSTORE is capable of interleaving up to four files.

- | Using different buffer sizes (MAXTAPEBUF option):

TurboSTORE supports 16KB and 32KB buffer sizes. Default buffer size is 16KB. For backups to remote devices, only 16KB is supported. Typically, performance doubles with 32KB buffer (MAXTAPEBUF) when compared to 16KB.

- | Doing parallel backup (STORESET option):

Doing parallel backups helps to reduce the backup window by several folds, especially on multi-processor systems. When using high-end devices like DLT8000, a general rule for doing parallel backup is to have one less target (tape) drive in the parallel set than the number of processors. For example, on a 4-way system, perform a 3-device parallel backup. This is because STORE uses two additional processes to control the overall backup. If you are doing a three device parallel backup, it is possible that you could get three times the performance of a single backup on a four-way multiprocessor system.

On a single processor system, one may not see the same performance improvement with parallel backups. However, depending on the system configuration, one should still see an increased backup performance.

Optimal h/w configuration

I H/W compression:

Enabling hardware compression alone gives a small percentage of performance improvement. For best backup performance, it is recommended to use HIGH software compression, or LOW software compression along with hardware compression.

I System configuration:

The backup performance not only depends on the backup software and target drives used, but also on the host machine configuration and the system load. TurboSTORE creates one process per target drive in parallel storeset. Hence it is recommended to have one-more processor than the target drives in the parallel storeset. I.e., on a n-CPU machine its recommended to have n-1 tape drives. The type of compression should be selected depending on the CPU load.

I Load balancing:

It is very important to balance the load of each tape drive during parallel backup. If some tape drives finish their jobs much earlier than the others do, it does not improve the overall backup performance. TurboSTORE uses a simple algorithm to divide the files among parallel tape drives. It stores these files in the sequence they are passed to store command and never moves them around to balance the loads. It is recommended to load balance by providing proper mix of files in the fileset if the backup contains some very large files.

Conclusions

Best performance is achieved with COMPRESS=HIGH and MAXTAPEBUF options. The HIGH compression option is CPU intensive.

When the CPU load is moderate, it is recommended to use COMPRESS=LOW and MAXTAPEBUF along with hardware compression.

Proper load balancing with fileset can give best performance in parallel backup.