

September 10, 2007

RE: Backup performance optimization.

By Thomas Hoops



Summary: If your backups are consuming too much time or your compromising the scope of your backups to cut time, investigating your backup bottlenecks could well be worth your effort. Often the bottleneck may not be where you think.

We use two custom software tools (which we offer for free on our website) to help us make an estimation of system performance. After all, it's hard to figure out the best race strategy if you don't know how fast your car can go. The two free programs we've created, TRMark and FakeBack, help determine how the storage systems of your computer best perform in order to help you devise the best backup strategy or determine where bottlenecks lie.

There are three primary areas for a bottleneck to occur: the source, the destination, the backup program or method. We have found more often that the backup program or the source drive is the bottleneck. To help determine this, we suggest you first run the free program FakeBack (available on our website under downloads) on your backup source. FakeBack will deliver a rate that it was able to read your source tree. For example, copy FakeBack.exe to your C: root (topmost directory). If you're backing up the My Documents folder, "CD" to the My Documents folder then type [c:\fakeback](#) <enter> Fakeback will explore whatever folder you're currently on. Since FakeBack performs no other function, the performance number it gives is in theory the fastest your backup could ever hope to achieve. But, it is unrealistic because in a backup, for every read, there must be a write (which FakeBack does not do). So, a more realistic "ideal" number will be the FakeBack number divided by 2.

Once you know what your theoretical maximum is, you need to know what you're achieving in your actual backup. If your backup software does not already provide this metric for you, then take the total amount of data you backup divided by the total number of minutes it takes to backup. Then, divide that number by 60 (giving you Megabytes per second). If this number is close to (within 15%) of your theoretical maximum divided by 2 from FakeBack, your system is probably as optimum as it can be without the addition of new server hardware (a RAID array for example).

If it's over 15% then it's possible that your backup software or your destination drive is a concern. For this, we also provide for free, TRMark (again available on our website under downloads). TRMark will render the best possible read/write speed achieved on whatever device you request it to test. We typically use 20 GB to surpass all caching mechanisms and give us as true of streaming speed as possible. For example, suppose your source drive is E: and your target drive is Z:.. "TRMark 20000 Z:" would benchmark the speed of your Z: device with 20 GB of data. Should the write speed of this device exceed the theoretical maximum you determined from using FakeBack, then it is likely your server load or backup software is where the bottleneck is. In this case, an increase in server memory, a faster processor or an alternative backup program or configuration (turn off compression) should help.

If it meets your theoretical maximum (within 15%) then the source and destination are fairly matched and any improvement of one will require an improvement of the other to realize any

benefit.

Case Study

A 3.4 Ghz Pentium D machine with 1GB of RAM and a 160GB SATA system drive running Windows 2003 server was only accomplishing 11MB/s (about 40GB/Hr) using NTBackup to backup it's system drive.

The source drive was a RAID 0 array (Windows) of three 120GB WD drives holding about 45GB worth of data. 30GB of which was MP3 files, the remaining 15GB consisted of 3 copies of the Windows folder (5GB) and about 10 GB more of misc .doc files, .ISO, and program files.

Running FakeBack on the system drive showed and read rate of 52MB/s (187GB/Hr). ½ of this is 26MB/s (94GB/Hr). Yet, NTBackup was only accomplishing 11MB/s (40GB/Hr). Using TRMark on the Target drive showed a write speed capable of 20MB/s (72GB/Hr). Server load during backup was at 38% (via task manager) so there was still some headroom. Changing the backup program to RoboCopy rendered a 20MB/s (72GB/Hr) backup which was now bottlenecked at the target drive. Changing from a USB connection to an eSATA connection to the target drive rendered 32MB/s (115GB/Hr) on RoboCopy – a very respectable speed.

RoboCopy is a program supplied by Microsoft available on the tools directory of your distribution disks for your Operating system. The alternative built-in program is Xcopy.

Further testing using the backup program Backup Exec 11d by Symantec (formerly Veritas) on the same system and hardware showed a backup speed of 32MB/s (115GB/Hr) as well. Thus, if Robocopy does not fit the backup requirements, other software does offer performance as well. One trick to remember when configuring Backup Exec: set the “Maximum size for backup to disk files” to the size of your target drive. 11D defaults to 1GB which with 45GB of data produced, before it was

changed, produced 45 media files which really slowed it down.

This test case demonstrates how using these tools along with some simple logic allowed us to increase backup speeds by almost 3 times from 40GB/Hr to 115GB/Hr..