TOLIS Group

Backup You Can Trust℠

bru backup & restore utility

Tape: Not Your Father's Oldsmobile

General Motors' Oldsmobile division needed to pump up their image in the 1980's. Their cars had become passé in the minds of young buyers and something had to be done. Enter the "This is not your father's Oldsmobile" campaign, designed to entice a new generation into looking at Oldsmobile.

This scenario of Oldsmobile closely mirrors that of tape technology today. In the minds of many CTOs and system/network administrators: tape is something to be avoided. Yes, tape of old had its issues but those days are long gone. No longer is tape a performance, usability, or cost gating factor. LTO-4 technology is actually faster than the filesystems that feed it, the reliability is excellent, and the life expectancy of written tapes is measured in decades. The same can be said for other fine tape formats, though throughput performance varies.

While this paper will contrast the considerations of implementing hard disk drives in lieu of tape technology in larger, networked system environments, tape is also appropriate for standalone systems being used in the small office.

In any environment, tape can effectively serve both near-line backup and long-term archival storage duty. Depending on the tape format used, performance can come close to that of hard disk drives so the backup completion time delta is now significantly minimized. For example, we have users running LTO-4 media at consistently around 100MB/sec on Intel Mac's with Fibre-Channel or SAS connectivity. This would allow for the backup of 16 TB in 46 hours + tape changes. Or, if one can split the data into smaller chunks and run multiple drives (say 4TB to 4drives) with SAS or F-C you'll see an even more useful backup window reduction. Tape cartridges can also easily be taken off-site for storage - a distinct advantage compared to removing hard drives that will be detailed below.

The backup tool being used has a pronounced effect on throughput performance when writing to either disk, or tape. We know that our BRU technology is not a bottleneck to performance - it is so fast in fact that we have actually not been able to determine its throughput threshold.

In a lab environment (read: not real world), using a Dual G5 Xserve with 2GB RAM and 2Gb FC direct connect, we have created an artificial environment involving a 1GB cached filesystem from which BRU was able to consistently read and write data at the rate of 1.7TB/Hour (~540MB/sec). BRU's throughput performance is still gated by the throughput of the filesystem. In a real-world test, when reading and writing directly from an Xserve RAID, BRU is able to sustain over 150MB/sec (527GB/Hour).

Some might argue that restoring from tape is a slow process. Again, depending on the backup software tool being used, this can be true. When using BRU Server, the specific location of every file is cataloged and when the tape containing the target file is in the tape drive, the tape is streamed to the specific location of the file and then restored to the original system or re-directed to another system in a single step. Restoring from a BRU Server tape is almost as fast as restoring from hard disk.

There has been a lot of hype in recent years extolling the virtues of implementing D2D (disk-to-disk) as the sole backup strategy in large organizations. This flawed approach is based on the perceived performance, pricing, and capacity points of hard drives, all of which look good at first glance. But what danger lies beneath the surface?

Let's take a look at an example networked system environment that encompasses 16TB of data - a huge number by yesterday's standard but not an uncommon one today.

16TB of RAID-5 backup drives would require 21, 1-terabyte drives (1,099,511,627,776 Bytes, not 1,000,000,000,000 Bytes per drive) plus 2 RAID parity drives and a chassis capable of supporting 21+ drives. Most currently available configurations will only support 16 drives and cost more than $20,000, and the cost escalates significantly should one specify server-class hard drives instead of the end-user oriented drives one see advertised in the $100 price range. While disk may seem the overall cost-effective alternative to tape of late, disk comes with the expense of keeping it powered and what you're plugging them into. RAID architectures help assure subsystem availability, yet the user accessible capacity in one of these 16 drive chassis is reduced to ~13TB, not 16TB.

In contrast cost wise, a 2x24 LTO-4 tape library (2 drives - 24 slots) will cost approximately $10,000. Add in the cost of 20 - 800GB native tape cartridges (~$1,400) and the total cost for a true user accessible 16TB storage subsystem tops out at $13,000: max.

There are a couple of important hard disk drive (HDD) reliability characteristics that come into play. First, a simple fact of life is: HDDs break. Now picture your net 13TB disk subsystem comprised of the 21+ drives, most if not all of the drives coming from the same manufacturer, and most probably, the same or close manufacturing date. Hard drives of any particular manufacturing run tend to fail within a surprisingly tight time band. If you lose a disk, make sure the hot spare is replaced ASAP, otherwise if a second drive fails and a new hot spare is not available, RAID will be of little help. You then have lost all of the data.

The risk of RAID-5 can be reduced somewhat by the use of RAID-6 (stripped disks with dual parity) so two disk failures can be handled. The dual parity drives will reduce the amount of total accessible capacity by 2-drives, versus 1-drive for RAID-5, however additional considerations for HDD-only backup still exist.

The second hard drive consideration is a condition known as "sticktion." Backup strategies that involve filling up hard drives and placing them on shelves for a protracted period of time place the recovery of that data at risk.

Today's hard drives employ very close tolerances and special aqueous coatings on the platters to help achieve ultra high-recorded bit density (tighter the bits - higher the capacity per given area). Over time while drives are powered-on, centrifugal force migrates a portion of the aqueous coating outward, to the rim of each platter where it forms a physical ridge. When the drive is powered-down, its heads return to a home position that is outside of each platter's edge. When the drive is again powered-on, the heads must pass beyond that ridge to be positioned over the platters. For every 6 months a drive is left powered-down, there's an increasing probability that the read/write heads cannot be moved inward, over the ridge - thereby rendering the drive unable to access its data.

When this happens, the drive must be taken to a data reclamation firm that can charge thousands of dollars (capacity of drive dependent) to retrieve the data - if it is possible to do so at all. If drives must be sequestered on shelves, it is recommended they be spun-up for a couple of hours on a weekly basis. Given the nature of humans, adhering to such a schedule to assure successful data availability requires a significant amount of discipline. Should one's organization fall under the mandate of Regulatory Governance, be sure your data retention strategy can deliver the necessary compliance. One must take into account the added cost of alternative drive power chassis as well as the actual power consumed. This adds to the overall disk cost.

In contrast, should a tape device fail, one's data is still available on the tape cartridges. Hard disk drive and tape hardware failures pose very much different risks to data. When using tape, simply repair/replace the device, and recover the data safely stored on the tapes. Additionally, once a tape is written, it requires nothing more than a secure and controlled environment (such as a bank safe deposit box, or bonded off-site storage facility) and no additional power to retain their usability.

Tape cartridge technology has evolved as well. Unlike the issue of hard disk drive sticktion arising, an LTO-4 tape can be placed on a shelf for 5 years or 80 years and the data will still be recoverable.

Time is most usually a precious commodity, and there is one final consideration regarding the example of the 16TB gross, 13TB net RAID-5 array - namely generating the array before one can write the first bit to it. On our bruAPP appliance, it takes our manufacturing group 42 hours per system to generate a 10TB RAID-5 assembly. And that's with a hardware array controller, not software RAID. Doing the math, to generate the 16TB array will take a minimum of 67 hours. Should something go awry, the process must be started again. When using tape, just plug it in and begin to use it.

TOLIS Group utilizes DAT and LTO tape technology both solely for its on-site and off-site backups, but also takes advantage of hard disk drives which are an appropriate fit in our and others backup strategies to provide near-line access to data in concert with the tape (D2D2T). Don't be fooled by misleading marketing that HDD backup is necessarily faster, cheaper, and just as reliable as tape - it is not. The use of HDDs solely is akin to playing Russian roulette with one's data.