**Reduce Data Backup Loads**
By: Christian Perry
May 6, 2011
*Processor*

Attempting to lasso the relentless stampede of data in the modern enterprise is a venture fraught with peril for IT managers unprepared for the challenges of storage. Throw backups into the mix, and those same managers might find themselves buried under huge backups that drain resources and budgets. But by reducing the data backup load, it's possible to relieve the pressure and create a sensible backup process that
doesn't create big problems for IT and the organization's other employees.

"The reduction of data backup loads benefits several areas of the business, the most significant of which is the reduction in IT administrator support time as data volume is decreased," says Jeff Pederson, operations manager for Kroll Ontrack's (www.krollontrack.com) Data Recovery division. "The available backup window is even smaller in the globally connected world of business today, so a reduced backup load
not only makes it possible to work within these windows, but provides flexibility for patches, upgrades, and other maintenance."

The benefits don't end there. In fact, the initial effects of reducing a data backup load can snowball into larger benefits, Pederson adds, including servers and systems that perform at higher levels, which in turn reduce power and cooling costs. These smaller loads also reduce the costs of backup tapes, transportation, offsite storage, and retrieval; result in faster data restore times; and narrow the set of content potentially subject to regulatory investigations and legal discovery, Pederson says.

**Investigate First**
When considering the size of backup data loads, remember that storage has an error rate, explains Vincent Berk, CEO of ProQueSys (www.proquesys.com). Backups aren't immune to errors that can plague storage—in fact, backup media fails at the same rate as the media that holds the original copy of the data, he says, and when using many devices for backup storage, you can increase the opportunity for data loss. Further, big backup loads mean more complex management requirements, but Berk says that a compact set of carefully chosen backups that can easily be restored is more valuable than keeping many copies of the same data.

Instead of taking a slash-and-burn approach to cutting data backup loads, experts recommend a careful approach that keeps several
factors in mind. For example, organizations first need to understand what they're trying to accomplish before attempting to implement
the technologies and procedures that can help achieve this task, notes Tom McCaffrey, director of archiving at Kroll Ontrack.

"A cross-functional team should be responsible for setting the retention policies for proper data governance to ensure regulatory obligations, legal discovery requirements, and . . . the needs of the business are met. From here, the implementation should first focus on what needs to be kept and for how long. Next, the implementation will address the optimal systems for maintaining data—for example, where should the data
reside and when should it move from one system to another," McCaffrey says.

**Tools Of The Trade**

Understanding your environment is a key step in the process of reducing data backup loads (see sidebar for more information), but once you have a handle on that, it's time to investigate technologies that can help further solve the problem of cumbersome loads. According to Mario Blandini, senior director of products and marketing at Drobo (www.drobo.com), technologies such as disk-based backup and deduplication
aren't new but can make a difference when used in new backup solutions.

For example, when used as the target for backup data, disk technologies can improve data streaming to the storage location and can improve the speed of backups. And when it comes to truly reducing the loads themselves, deduplication can be a major player in the process when added to disk-based backup. While tape offers data compression, deduplication can improve the overall backup efficiency more than 10 times that of tape, Blandini says.

There are multiple flavors of deduplication that can vary in their ability to reduce the amount of data stored to the equivalent of one copy or less, says Brian Greenberg, president and CEO of General System Dynamics (www.gsysd.com). For instance, client-based deduplication sends less data over the network; server-based puts the processing load completely on the deduplication engine; inline processing requires higher performance engines and may not be as fast or efficient; and post-processing requires a larger landing site (disk storage) as the data is deduplicated.

Other practices can also help to reduce data loads, Greenberg adds, including de-junking the data. "If you know what your data is and have it properly classified, you can de-junk your data by only backing up the right data and not junk. This can reduce the amount of data in many organizations by 50% or more," he says.

Greenberg also recommends putting a good reporting system in place that provides a clear window into the storage environment to provide
a good perspective on the state of storage efficiency and data management capabilities. His other suggestions include keeping data
and OSes separate and preserving only OS images or virtual machines, and don't back up the OS of every machine in the company.
Finally, he recommends leveraging snapshots and replication through a multitiered data protection strategy that can reduce the load,
capacity, and/or frequency of backups.

**Reduction Costs**
Gauging the cost of reducing data backup loads can be a difficult endeavor without knowing the amount of data involved, but McCaffrey from Kroll Ontrack says that with moderate storage requirements, organizations can expect hosted archiving services to range from $1,000 to $5,000 per month after setup and legacy ingestion costs (which themselves can range from $2,000 to $20,000).

Drobo's Blandini adds that a disk-based backup for a small to midsized enterprise can start around $2,000 for a DAS solution, while an iSCSI SAN system that supports multiple backup servers can start at $5,000. Meanwhile, he says that organizations can expect to spend more than $30,000 on entry-level storage and a deduplication application along with backup software—and even up to six figures for a midrange solution. Smaller shops, he says, can deliver deduplication on a backup media server for less than $3,000 per server.