The Availability Corner

The Great Tape Backup Paradigm Shift

May/June, 2005

Dr. Bill Highleyman
Dr. Bruce Holenstein
Paul J. Holenstein

The capacity of disk drives has grown at a phenomenal rate. 144 gigabyte disks are now becoming available, and the prediction is that terabyte disks will be common in just three years. A terabyte is a tremendous amount of data. An interesting vignette demonstrates this problem. Researchers worldwide are now conducting studies that involve sharing terabytes of data. Jim Gray, who specializes in massive databases at Microsoft, is one such researcher. He and his peers have realized that it is faster to transfer a terabyte of data by mailing the disks rather than by transmitting the data, and that is exactly what they are doing! (Think about it. Using a one megabyte/sec. communication channel, it will take one million seconds, or over 11 days, to transmit one terabyte.)

What does this have to do with availability? Availability is the measure of the proportion of time that a system is up. High availability requires not only a long time between failures, but just as importantly, a very fast *recovery time.* If an ultra-reliable system fails just once every ten years and is to have six 9s availability (up 99.9999% of the time), then it must be able to recover in less than five minutes.

How long will it take to load a terabyte database? How long will it take to do the audit roll-forward? Hours, if you are lucky. This is absolutely unacceptable if we are striving for highly reliable, mission critical systems.

Physical tape will just not do it as we move to very large databases in critical systems. Tapes may have to be moved to the backup site. Then massive backups take massive amounts of time to restore. This not only significantly reduces availability but also violates many RTO (Recovery Time Objective) goals. The hard fact behind the on-going paradigm shift is that we must re-consider the way in which we back up our systems – after all, it is the recovery time that matters when making an application available after an outage.

Virtual tape helps to ease this transition. With virtual tape, backups are made directly to disk rather than to tape. Virtual tape is a faster, easier to manage approach. However, even virtual tape suffers from the serialized nature of each virtual tape device during recovery – each can only restore so fast, and recovery can take an unacceptable amount of time when massive databases need to be restored and possibly “rolled forward” through audit trails.

In order to achieve fast recovery times, it is common practice today to configure systems for disaster recovery. This approach entails real-time replication of primary database changes to a remote standby system. Thus, if the primary system fails, the standby system can be ready to quickly take over the application functions – just as soon as applications can be started, transactions recovered, and users switched over. But even this could take the better part of an hour.

Here is where the new backup paradigm comes in – active/active systems. In an active/active configuration, multiple independent nodes are participating as peers in the application; and each is replicating its changes to the other nodes in the application network. If one node should fail, all that is necessary is to switch users from the failed node to another surviving node in the network to restore services to those users. Properly implemented, this can be done in seconds. Furthermore, the other (unaffected) users are substantially unaware that a failure has occurred except that they may find their system is now more heavily loaded.

True, a disaster recovery configuration could run as a "hot" standby with all applications up and running, ready to take over on a moment's notice. But active/active systems bring many additional benefits to a system:

- All of the systems in the network are productively contributing to the application processing. There are no passive standby nodes.
- There are no software license fees for idled system capacity.
- Disaster recovery comes as a free by-product of an active/active configuration.
- Active/active systems are scalable by adding additional nodes.
- On-line zero downtime migrations can be accomplished one node at a time to upgrade hardware, operating systems, applications, and databases, essentially eliminating "planned downtime" from the availability profile.

Active/active systems are being deployed today for applications which just cannot go down. Watch for this new backup paradigm to become ever more pervasive.