High-Bandwidth Backup and Restore White Paper

by Jay Desai
Storage Management Product Marketing

Jerry Lohr
DSS Technical Marketing

Bill Tuthill
Senior Technical Writer

Summary

In less than three hours, using an Origin2000™ server and 8 mm tape drives, employees of Silicon Graphics, Inc. and Spectra Logic Corp. performed a hot backup and complete restore of a 1TB Oracle8™ database. Using only four tape libraries, they performed a hot backup and restore of the same database in about eight and a half hours. Backup and restore performance of a more typical 320GB database was also measured.

Storage Trends

Market predictions forecast that the total consumption of disk drives will grow from 76,000TB in 1995 to 3,000,000TB before 2000. Customer requirements show that data repositories will double or triple each year. Proper management of rapidly expanding data repositories is critical. As the value of data increases, so does its cost of maintenance. More than ever, databases need backup protection. In competitive global enterprises, technology managers must provide continuous access to large online databases and data warehouse servers under centralized corporate control. At the same time, users of these resources want to shrink the backup window and reduce their cost of data ownership.

The server of choice for this type of customer must be powerful and scalable. It must have configuration headroom to accommodate growing data capacities and shrinking backup windows. It must also be highly reliable and affordable in order to justify replacement of the traditional mainframe.

Performance Goals

During recent years, Silicon Graphics has seen a market trend toward the implementation of VLDB (very large database backup) on its systems. Origin2000 servers, configured to run multi-terabyte database applications, are often required to perform fast backup of a live database, because it is critical for enterprises to maintain 24-hour access to these data repositories while protecting them with backup. The loss of database access for even one hour can cost a company thousands of dollars in lost productivity and revenue. Enterprises are looking for better VLDB solutions and cost-saving alternatives to the traditional mainframe. This project demonstrates that the I/O subsystem of an Origin2000 server scales well and can protect live data using only a small backup window.

Insurance companies, hospitals, large banks, credit card companies, and other financial institutions that

depend on extremely large databases can be confident about the latest tape backup and restore technologies. These tests demonstrate many advantages of the latest Fibre Channel and Advanced Intelligent Tape (AIT) technology, under control of an eight-processor Origin2000 server, for backing up and restoring large data sets.

Methodology

Four tests were performed with various combinations of database sizes and media library configurations.

Two tests were conducted with a 1TB (1024GB) Oracle ® 8.0.4 database, designed using a TPC-C schema. The database was 80% populated and contained 64 data files. Two more tests were conducted with a 320GB Oracle 8.0.4 database, also using a TPC-C schema, and also 80% populated. Backups were live, with 100-user connections sustained. Restores were accomplished with the database offline. The first test employed 16 Spectra Logic 10000 tape libraries containing four Sony AIT drives each. The second and third tests employed four similar SL 10000 libraries. The fourth test employed a single SL 10000 library. Backups and restores were done with Spectra Logic Alexandria software. All benchmarks were performed on an Origin2000™ server, which offers exceedingly high bandwidth. The first two tests required a pair of 4-CPU Origin rack modules, while the third and fourth tests required only a single 4-CPU Origin deskside module.

Test Results

Performance results of the four tests are summarized in Table 1.

Table 1: Test Performance Results

Number of AIT drives	Total Data Moved	Duration of Backup	CPU Backup Overhead	Throughput during Backup	Duration of Restore	CPU Restore Overhead
64	1TB	75 minutes	22.9% average	818GB/hour	95 minutes	16.5% average
16	1TB	4 hours 25 minutes	5.2% average	232GB/hour	4 hours 12 minutes	3.7% average
16	320GB	1 hour 45 minutes	7.9% average	183GB/hour	2 hours 30 minutes	5.5% average
4	320GB	4 hours 23 minutes	3.5% average	73GB/hour	4 hours 28 minutes	3.0% average

Notes:

1. Duration measures the total elapsed time to move 1TB from disk to tape and complete backup index operations.
2. Backup time included time to copy archive log files to tape. Restore time was measured as time to copy data from tape to disk.

3. This document uses the following definitions: 1024 bytes = 1KB, 1024KB = 1MB, 1024MB = 1GB, and 1024GB = 1TB.

These benchmark results demonstrate that the Origin2000 I/O subsystem offers more than enough performance to back up a live 1TB Oracle database in about an hour, with spare headroom for additional devices. Even with a single tape library, a 320GB database could be backed up or restored within a 4½ hour nightly backup window. Using off-the-shelf software applications from Oracle and Spectra Logic, Origin2000 performs extremely well as a VLDB backup server.

Sizing a Backup Solution

When sizing a VLDB backup solution, start by determining the length of your required backup window. Divide the total database size (in Megabytes) by 80% of the tape drive speed to account for backup overhead. Tape speed is usually rated in megabytes per second; multiply by 60 to get megabytes per minute. Divide again by the length of your backup window (in minutes) to arrive at the number of tape drives required.

DataSize ÷ (80% tape speed × 60) ÷ backup window = tape drives required

Ideally, each tape drive should be cabled to its own SCSI channel. This is especially true of tape drives whose speed approaches theoretical SCSI limits.

Disk drive performance is seldom a limitation on backup performance. Fibre Channel loops offer less bulky connections than SCSI cable, with more flexibility of device placement distance.

One-Hour Terabyte Backup

Many customers have benchmarked the performance of the Origin2000 server against their requirements, testing its scalable I/O subsystems, high-performance IRIX device drivers, and high-availability features. Now the Origin2000 server has been measured for its scalable I/O capabilities, using VLDB hot backup as a testbed.

The goal of a previous project was to demonstrate that an active 1TB database could be backed up in less than an hour. Breaking the one-hour barrier was a joint undertaking by Silicon Graphics and its storage management partners: IBM Corporation, Oracle Corporation, Legato Systems Corporation, and Spectra Logic Corporation. Thanks to these alliances, team members were able to build the world's largest and fastest VLDB system and set the world record for high-performance hot database backups. This project is described in the One-Terabyte per Hour Backup white paper. (http://www.sgi.com/Technology/teraback/teraback.html)

One-Terabyte Recovery

The tests described in this white paper complete the one-hour terabyte backup study, showing that a database restore can be done about as quickly as the backup, and it can be done with more affordable technology. Figures 1 to 4 show performance numbers generated by these tests.

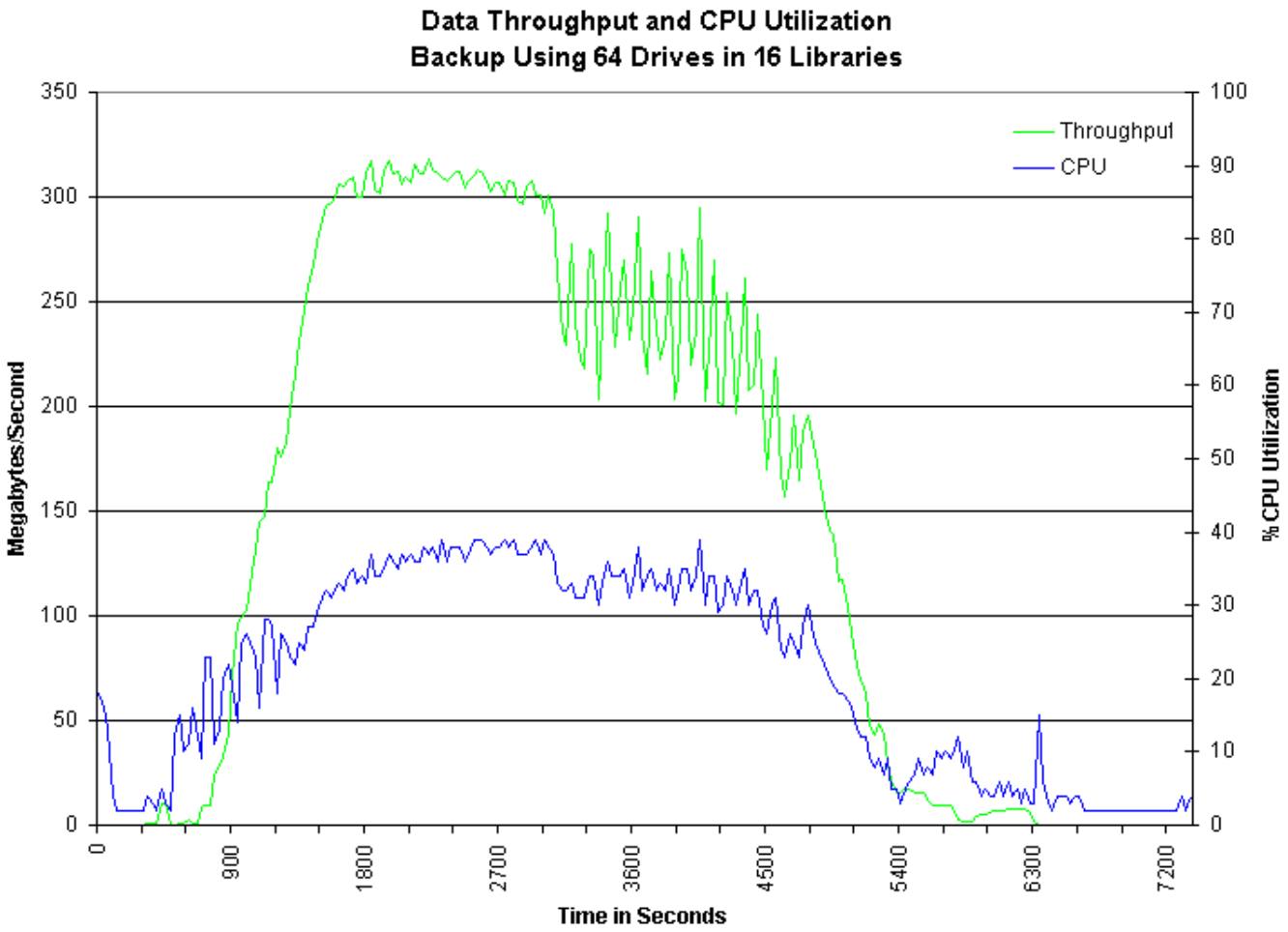

Figure 1: Terabyte Backup Test (64 Drives)

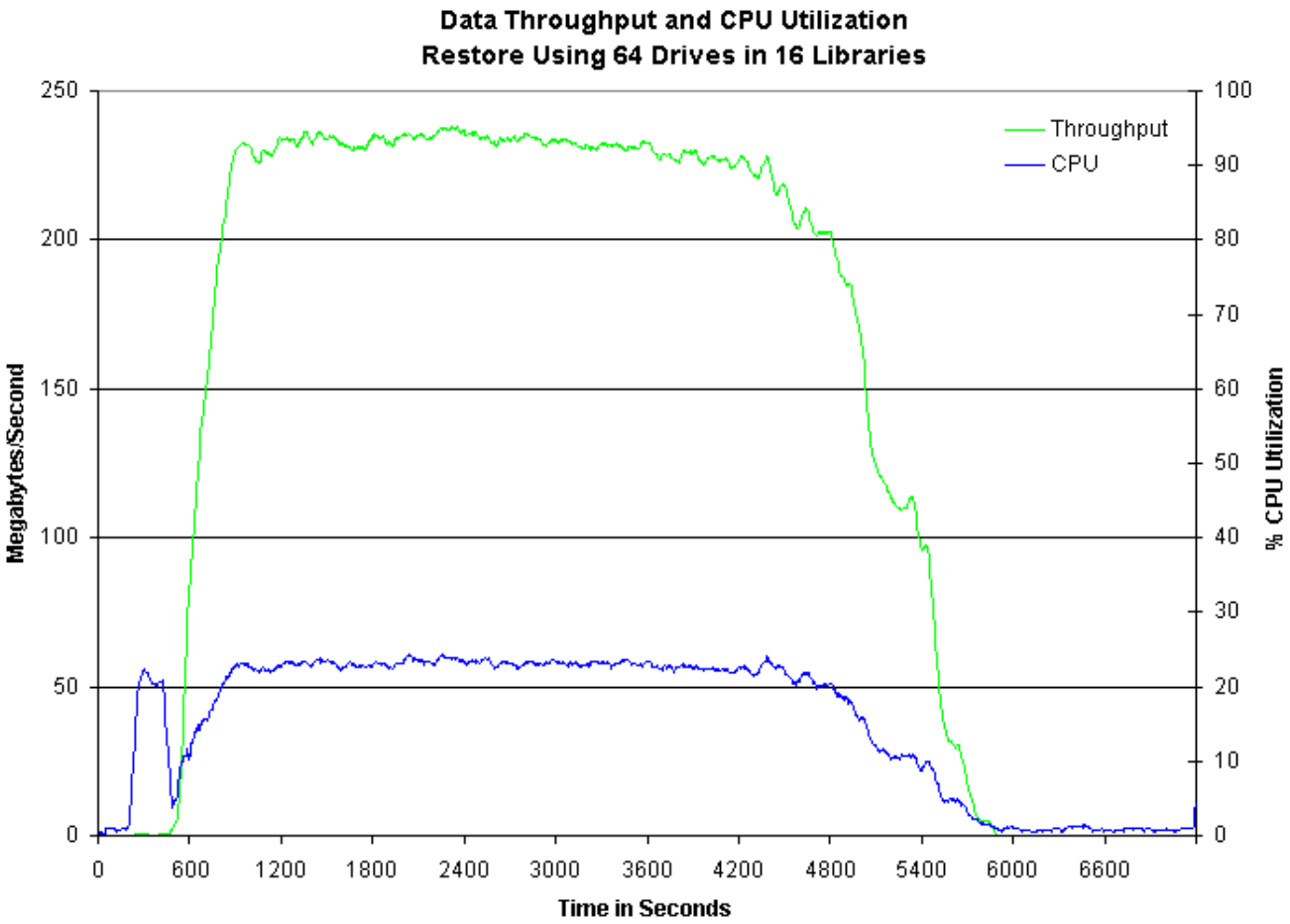

Figure 2: Terabyte Restore Test (64 Drives)

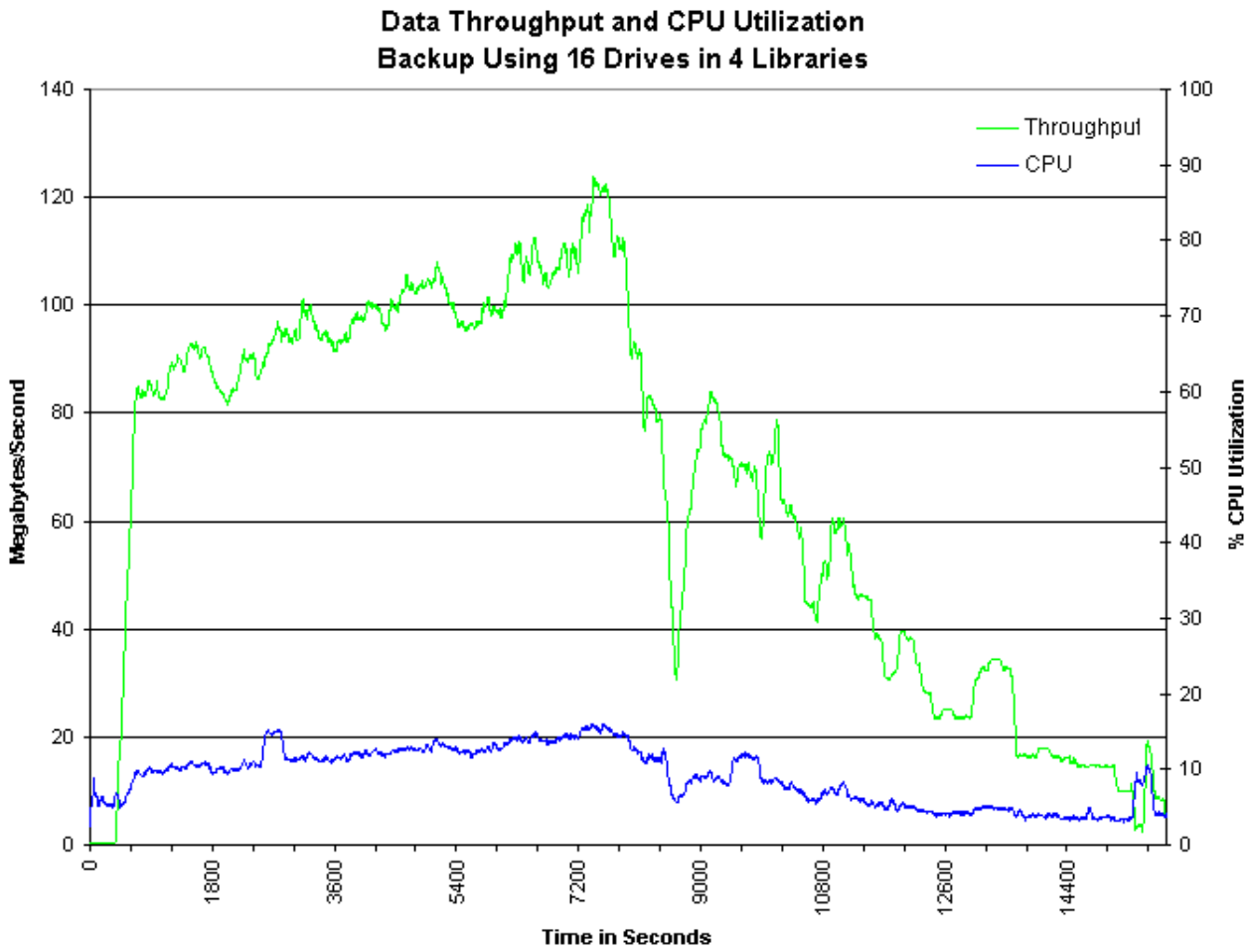

Figure 3: Terabyte Backup Test (16 Drives)

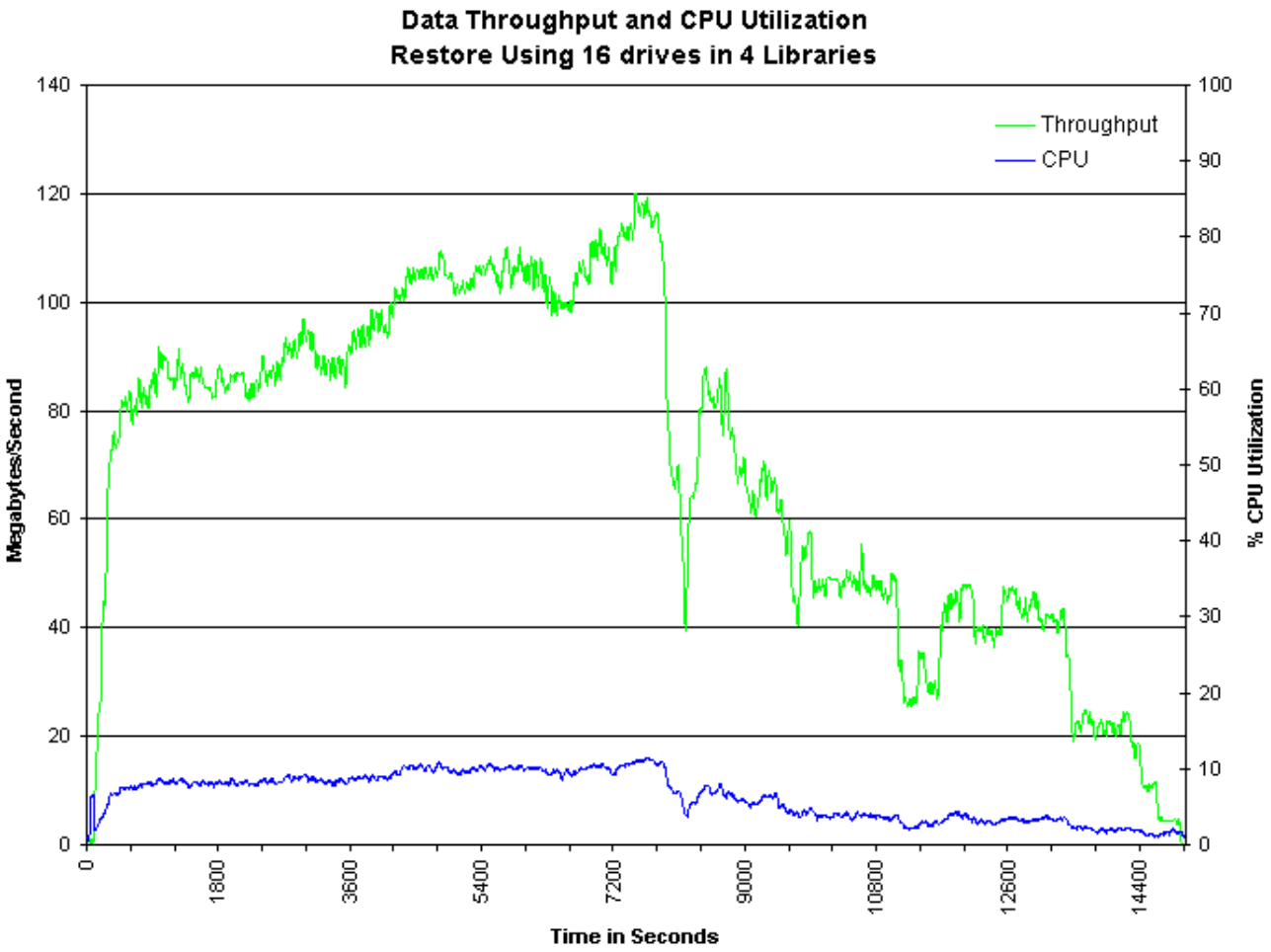

Figure 4: Terabyte Restore Test (16 Drives)

320GB Recovery

As discussed above, this set of tests extended the original study into smaller database sizes. Figures 5 and 6 show performance numbers generated during recovery of the 320GB database.

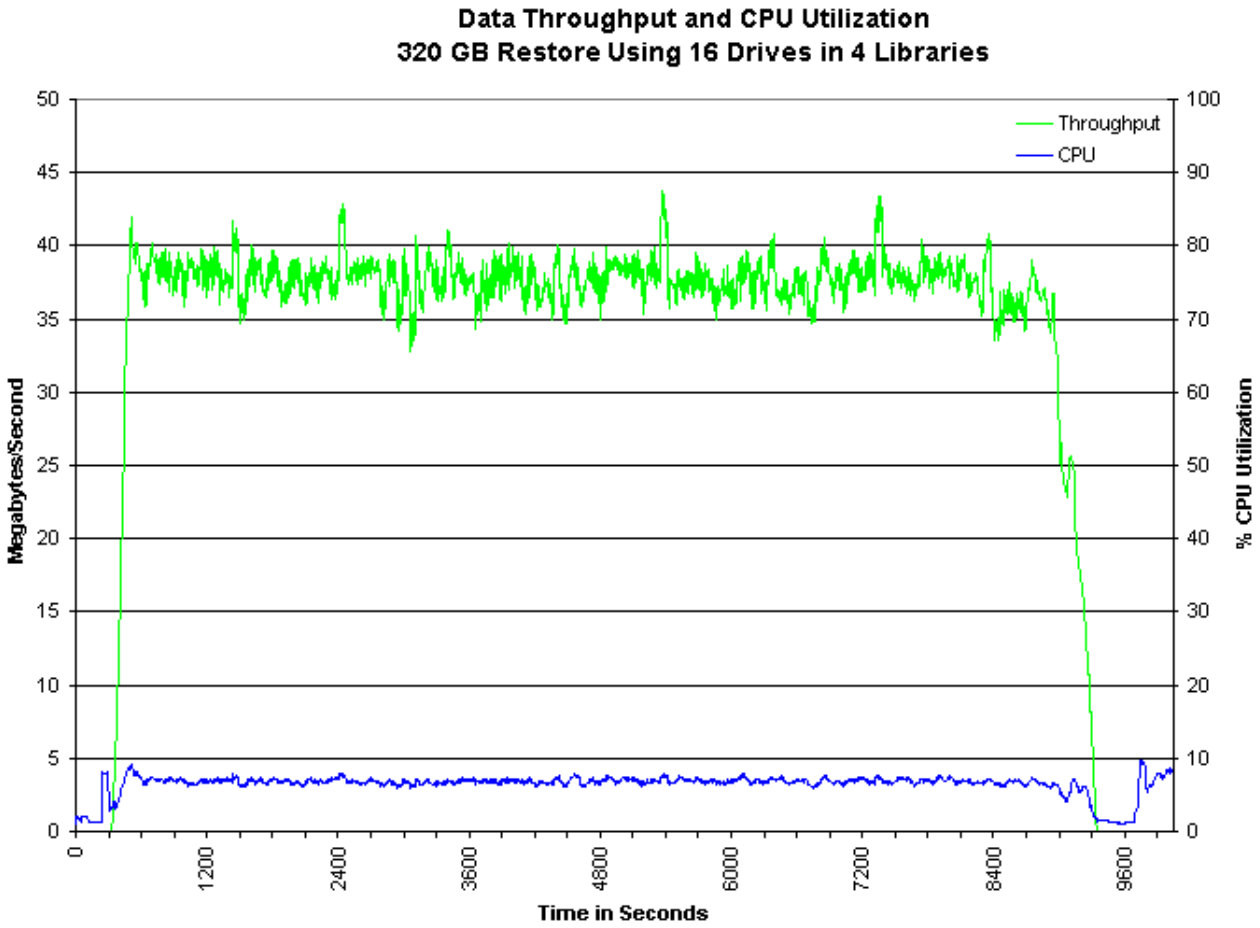

Figure 5: 320GB Restore Test (16 Drives)

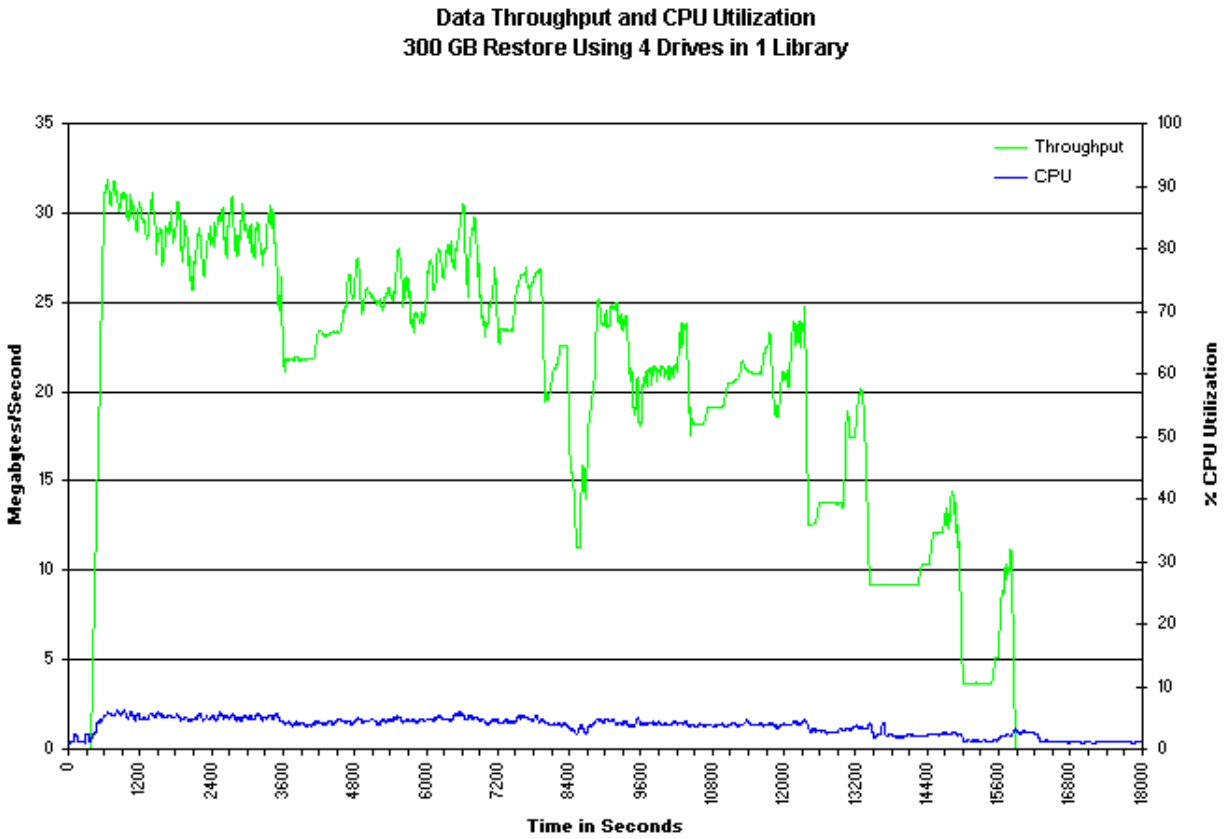

Figure 6: 320GB Restore Test (4 Drives)

Configuration Information

Database Description

The database was built using the standard TPC-C schema. For a detailed description of this schema, see the TPC Web site at www.tpc.org.

The TPC-C schema uses the following business model: a wholesale parts supplier runs a business through warehouses and associated sales districts. Each warehouse supplies 10 sales districts, and each sales district supplies 3,000 customers. For the purposes of our tests, we chose 1,000 warehouses to build the database, or the equivalent of 30 million customers. Additional stock tables were also added to round out data volume. Because this was not a TPC-C benchmark, we did not conform to some of the TPC-C timing constraints. However, we did simulate an order-entry system.

During the backup, the STOCK table generated a load of 100 TPC-C users sustained. With RAID-3 and disk management utilities, the STOCK table was striped across disks to improve performance. This ensured that the backup and transaction processes would be in a state of minimum contention for disk resources.

In all tests, a total of 64 data files were involved, and over 80% of the total table space was populated

with real data, index, and overhead information, while under 20% was composed of null pages (table spaces with no data).

Processor and Memory

For the first two tests, the server was an eight processor Origin2000 rack system (two Origin2000 modules each with four 195 MHz processors) running IRIX 6.5. Each module had a 4MB cache and 4GB of shared main memory.

For the third and fourth tests, the server was an Origin2000 deskside system with four 195 MHz processors running IRIX 6.5. This system had 4MB cache and 4GB of shared main memory.

I/O Subsystem

Each Origin2000 module provides a total of 12 XIO slots. Two are reserved for base I/O and PCI cards, and 10 adjacent general-purpose slots can be configured with either 4-port UltraSCSI cards or 2-port Fibre Channel cards. Table 2 shows how the available XIO slots were configured for the 1TB database tests.

Table 2: Dual Origin2000 Test XIO Slot Configuration

XIO Slots per Module	Total XIO Slots Available	Total Channels	Explanation of XIO Slots
1 Base	2	4 SCSI	Base I/O card with two UltraSCSI channels, local disk
1 PCI	2	4 PCI	PCI extension with three slots, unused in this test
2 XIO	4 x 2 ports	8 FC	Fibre Channel cards, each with two Fibre Channel loops
4 XIO	8 x 4 ports	32 SCSI	UltraSCSI cards, each with four SCSI channels
4 XIO	8	32 SCSI 16 FC	UltraSCSI cards with four channels each, or Fibre Channel cards with two loops each, unused in this test

There are 80 available UltraSCSI channels (20 XIO slots) in a two-module Origin2000 configuration. The system could have 32 additional SCSI channels available for expansion, or the XIO slots could be used for Gigabyte Ethernet, Fibre Channel, HIPPI, or other high-speed I/O channels. In other words, 40% of the I/O capacity was still available as headroom.

Disk Drives

The 1TB Oracle8 database was constructed on a total of 160 disks, all 9GB Fibre Channel drives (7200 RPM). The RAID-3 disk drive subsystem was never a limiting factor during backup or restore.

The 320GB Oracle8 database was constructed on a total of 50 disks, all 9GB Fibre Channel drives (7200

RPM). Again, RAID-3 disks were never a limiting factor.

Media Libraries

The first test was done with 16 Spectra Logic 10000 libraries, each containing four AIT drives for a total of 64 tape drives. Each library was connected to a SCSI channel, imposing little overhead. Two drives of each library were daisy-chained to the same SCSI channel. The other two drives were daisy-chained to a different SCSI channel. Thus, 32 SCSI channels were required for the first test. Each tape drive has an uncompressed data rate of 3MB per second and a compressed data rate of 6MB per second, assuming a 2:1 compression ratio. Thus, the maximum theoretical throughput for this test was 192MB per second uncompressed or 384MB per second compressed, and the actual backup graph (Figure 1) peaks around 315MB per second.

The second and third tests were done with only four Spectra Logic 10000 libraries, each containing four AIT drives for a total of 16 tape drives. Again, each library and two drives were connected to a SCSI channel, while the remaining two drives were connected to another SCSI channel. Thus, only eight SCSI channels were required for these tests. Because database size did not change for the second test, all Fibre Channel cards remained in use. Maximum theoretical throughput in these tests was 48MB per second uncompressed.

The fourth test was done with a single Spectra Logic 10000 library containing four AIT drives. The library and two drives were on one SCSI channel, while two more drives were on a second SCSI channel. The maximum theoretical throughput in this test was 12MB per second uncompressed.

Figure 7: Spectra Logic 10000

Application Software

Silicon Graphics chose a key backup software partner for this 1TB database restore benchmark: Spectra Logic Corporation (www.spectralogic.com).

Alexandria is a commercially available off-the-shelf product. As a first-tier ISV and developer with the Silicon Graphics ® server division, Spectra Logic has long offered its Alexandria product for IRIX systems. Alexandria is distributed by Spectra Logic resellers.

Part number 007-3975-001