Sun's SCSI Disk Drives

Performance Brief

August 1991

Summary

Sun Microsystems recently announced the highest performance SCSI drives available in the industry: a 1.3-GB 5.25-inch disk and a 424-MB 3.5-inch disk. These drives feature performance levels that were once strictly the domain of larger 8-inch drives.

Both of these drives incorporate advanced features to attain this performance. Some of these features are described below:

Zone Bit Recording

The 424-MB and 1.3-GB SCSI drives both feature Zone Bit Recording (ZBR). This method of recording breaks the surface of the disk into different zones, each with a different number of bytes per track, to keep the bit density constant across the disk's surface. Drives with this feature will gain both in transfer rate and in capacity. In the case of the 1.3-GB drive, the disk transfer rate varies from 3 MB/sec at the inner zone to a full 4.5 MB/sec at the outer zone. In addition, the data capacity of the disk is increased by 300 MB through the use of ZBR.

High-Performance Actuators

These drives also feature very low average seek times. The 1.3-GB drive has an average seek time of just 11.5 ms. This is almost 40 percent faster than the average seek time of the 669-MB drive. The 424-MB 3.5-inch drive also has a very fast average seek time of 14 ms. This seek time is not only almost 20 percent faster than the 207-MB drive it replaces but is also faster than most 5.25-inch and 8-inch drives on the market today. Lower seek time is one of the primary ways to improve system performance in database and NFS® applications.

Low Rotational Latency

The other primary way to improve random access performance is to reduce rotational latency. Most drives spin the disks at 3600 rpm. This means that it takes 16.6 ms for one full disk rotation, so the average latency is 8.3 ms. The new 424-MB drive spins the disks at 4400 rpm. This cuts the average latency to only 6.8 ms, or by about 20 percent versus standard rpm drives. The new 1.3–Gbyte disk spins even faster at 5,400 RPM. This reduces the average latency to only 5.5 ms, 35 percent faster.

Fast Transfer Rate

Both of these drives support synchronous SCSI burst rates of 5 MB/sec. They also have very high disk data rates, which are important for long sequential transfers. The 424-MB drive has a raw disk data rate of 2.5 MB/sec to 3 MB/sec. This is up to twice as fast as the 207-MB disk, which has a raw disk data rate of 1.6 MB/sec. The 1.3-GB drive has a disk transfer rate of 3 MB/sec to 4.5 MB/sec, which is almost fast enough to sustain the SCSI synchronous burst rate. This disk has more than twice the transfer rate of the 669-MB drive, which has a raw disk data rate of 1.8 MB/sec.

Disk Drive Comparison Chart

The following chart shows the basic specifications of the new disk drives and the previous generation of products.

	207 MB SCSI	424 MB SCSI	669 MB SCSI	1.3 GB SCSI	911 MB IPI	1.3 GB IPI
Form Factor	3.5 inches	**3.5 inches**	5.25 inches	**5.25 inches**	8 inches	**5.25 inches**
Average Seek	16 ms	**14 ms**	16 ms	**11.5 ms**	15 ms	**11.5 ms**
Average Latency	8.3 ms	**6.8 ms**	8.3 ms	**5.5 ms**	8.3 ms	**5.5 ms**
Raw Disk Data Transfer Rate	1.6 MB/sec	**2.5–3 MB/sec**	1.8 MB/sec	**3–4.5 MB/sec**	6 MB/sec	**3–4.5 MB/sec**
Rotational Velocity	3600	**4400**	3600	**5400**	3600	**5400**
Overall Performance*	509 KB/sec	**1012 KB/sec**	779 KB/sec	**1429 KB/sec**	1368 KB/sec	**1408 KB/sec**
Buffer Size	64 KB	**64 KB**	64 KB	**256 KB**	N/A	**N/A**
Drive MTBF	50,000	**100,000**	50,000	**100,000**	50,000	**100,000**
Controller	N/A	**N/A**	N/A	**N/A**	ISP–80	**ISP–80**
Max Disks/Controller or Channel	4	**4**	4	**4**	8	**8**
Where Used**	DT	**DT**	DT/DS	**DT/DS**	DS/DC	**DC**

*Sun performs a combination of random and sequential benchmarks to develop an overall measurement of performance for mass-storage products. These tests are performed on Sun systems and a geometric mean is calculated to generate a composite of the performance that a typical user might expect. The most important aspect of these tests is not necessarily the absolute values but rather a relative comparison between products.

** Legend:
DT = Desktop systems
DS = Deskside systems
DC = Data Center systems

Multiple Performance Levels of Disks

At any given capacity, a number of sizes and types of disks are available. This is particularly true in the 1+ GB area. For example, today the market is flooded with SCSI disk drives called 1.2-GB drives. These drives are actually 1.0 GB formatted capacity and have performance characteristics very similar to Sun's 669-MB drive. Rotational speeds are still 3600 rpm with seek times typically in the range of 15 to 16 ms. Next are a few drives in the 1.3-GB formatted (1.6 unformatted) capacity range. Although these drives do have the same capacity as Sun's 1.3-GB drive, the performance is again lower. For example, the HP–97560 is generally considered to be one of the hottest 1.3-GB SCSI drives on the market today. This drive features ZBR with a raw disk data rate of 2.2 to 2.9 MB/sec, rpm of 4002, and seek time of 13 ms. As you can see from the basic specification comparison on the following page, this drive is still lower performance than Sun's new 1.3-GB SCSI drive.

	Wren 7	HP–97560	Sun 1.3 GB	Sun 424 MB
Form Factor	5.25 inches	5.25 inches	**5.25 inches**	**3.5 inches**
Raw Capacity	1.2 GB	1.6 GB	**1.6 GB**	**494 MB**
Formatted Capacity	1 GB	1.3 GB	**1.3 GB**	**424 MB**
Average Seek	15 ms	13 ms	**11.5 ms**	**14 ms**
Average Latency	8.3 ms	7.5 ms	**5.5 ms**	**6.8 ms**
Raw Disk Data Transfer Rate	1.8–2.9 MB/sec	2.2–2.9 MB/sec	**3–4.5 MB/sec**	**2.5–3 MB/sec**
Rotational Velocity(RPM)	3600	4002	**5400**	**4400**

As shown above, Sun's new 1.3-GB drive is a higher performance drive in all categories. In addition, the 424-MB drive compares favorably in performance with the larger competitive drives. It is important that your customers understand the performance and real capacity of the drives they are considering.

Performance

This brief compares the performance of Sun's new 1.3-GB disk drive and 424-MB disk drive with other disk drives that Sun currently offers. Single disk and multiple disk configurations were tested. The following tests were performed:

- Filesystem I/O throughput
- Musbus Benchmark and SunMark Benchmark — Single disk configurations
- Musbus Benchmark and SunMark Benchmark — Two disk configurations
- TPC-B Database Benchmark

The other disk drives tested were:

- Sun's 669-MB SCSI disk drive
- Sun's 207-MB SCSI disk drive
- A 5.25-inch drive of similar capacity — the 1.0 GB Wren 7

Note: All capacities referenced in this report are formatted capacities.

As the following summaries indicate, the 1.3-GB disk drive performed better than the other drives.

Filesystem I/O Throughput

With the highest I/O throughput compared to all other disks tested, the 1.3-GB disk is:

- **1.83** times better than the 669 MB drive
- **1.50** times better than the 1.0 GB Wren 7 drive
- **1.41** times better than the 424 MB drive
- **2.81** times better than the 207MB drive

Musbus and SunMark Benchmarks — Single Disk Configurations

The 1.3-GB drive performed the best on both system throughput benchmarks. The other disks' Musbus and SunMark ratios are compared to the 1.3-GB disk.

Disk Type	Musbus Ratio	SunMark Ratio
1.3 GB	**1.00**	**1.00**
424 MB	**0.77**	**0.83**
669 MB	**0.64**	**0.70**
1.0 GB	**0.70**	**0.75**
2 x 207 MB	**0.70***	**0.86***

Musbus and SunMark Benchmarks — Two Disk Configurations

The combination of a 1.3-GB disk and a 424-MB disk was tested using the Musbus and SunMark benchmarks. The results were compared to a 207-MB and 669-MB disk combination as well as the single disk configurations. The ratios were measured relative to the single 1.3-GB disk:

Disk Type	Musbus Ratio	SunMark Ratio
1.3 GB	**1.00**	**1.00**
424 MB	**0.77**	**0.83**
669 MB	**0.64**	**0.70**
1.0 GB	**0.70**	**0.75**
2 x 207 MB	**0.70***	**0.86***
207 MB + 669 MB	**0.99**	**0.94**
424 MB+ 1.3 GB	**NA**	**> 1.07**

TPC-B Database Benchmark

The combination of 5 x 1.3-GB + 2 x 424-MB disks performed 6 percent better than the combination of 10 x 669-MB + 2 x 207-MB disks on the TPC-B database benchmark.

Note – due to the lower capacity of the 207–Mbyte disk, Musbus and Sunmark testing could only be done with two disks. 207–Mbyte ratios are artificially high since the test results represent the performance of two disks and not one.

Test Methodology

For the I/O throughput, MusBus and SunMark tests, the disks were attached to a SPARCstation™ 2, with 64 MB memory, running SunOS™ 4.1.1. The 207-MB and 424-MB disks were attached as the second internal disk, while the other three disks (669 MB, 1.0 GB, 1.3 GB) were attached as external disks using the internal SCSI adapter.

The **I/O throughput** was measured using the I/O bench benchmark to measure five access patterns:

- Filesystem Sequential Read (FSR)
- Filesystem Sequential Update (FSU)
- Filesystem Sequential Write (FSW)
- Filesystem Random Read (FRR)
- Filesystem Random Update (FRU)

All the tests used a 10-MB file with a transfer size of 8 KB. In these I/O bench tests, the operating system ran off the SPARCstation 2's standard 207-MB internal SCSI disk, while the tested filesystem and the file were on the disk under test.

To further characterize the performance of each disk under regular system operation, they were tested with the **Musbus** and **SunMark** benchmarks. Both SunOS and the benchmark programs were loaded on each of the disks under test (in different partitions), and the tests were run with 1 to 45 users in increments of 5. Because the multi–user RTE test environment (MURTE) was used, rather than the original Musbus and SunMark benchmarks, the results may not be directly comparable to results reported elsewhere. However, they are probably more accurate, because the RTE test environment imposes essentially no overhead to the system being tested.

In both the Musbus and SunMark workloads, each simulated user logs in to the system undergoing a test with the *rlogin* command, performs a set of predetermined tasks (job script) in a random order three times, and then exits. To make sure that the system is driven to its maximum capacity, all the simulated users had no typing delays and only 2 seconds think time delays on average. As the results will show in later sections, this kind of workload indeed stressed the system to its maximum efficiency and differentiated how the disks performed under heavy workloads.

For the **TPC-B benchmark** tests, a client-server configuration was used. (The server was a SPARCstation 2; the client was either a SPARCstation 1 or Sun-4/280. This difference does not affect the results.) The database was scaled for 55 TPS, and the benchmark ran for 16 minutes, with one checkpoint. The operating system was SunOS 4.1.1, running DBE 1.1.

Test Results

I/O Throughput

Figure 1 shows the I/O throughput of all the disks tested under five different categories. The 1.3-GB disk has the highest throughput in all five categories, the 424-MB disk is the second in most categories, and the 1.0-GB disk is the third. Table 1 shows their relative performance.

Figure 1. Disk Throughput (KB/sec)

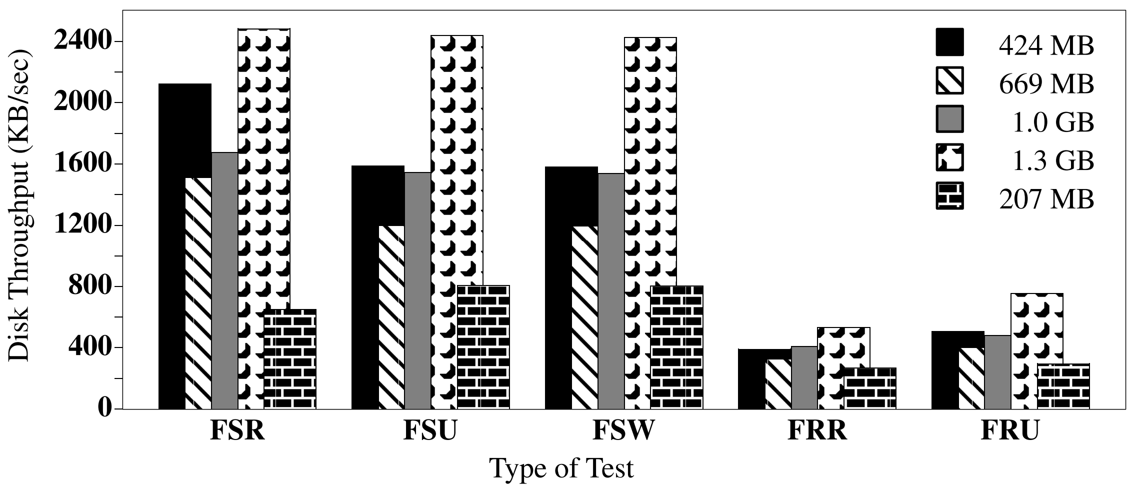


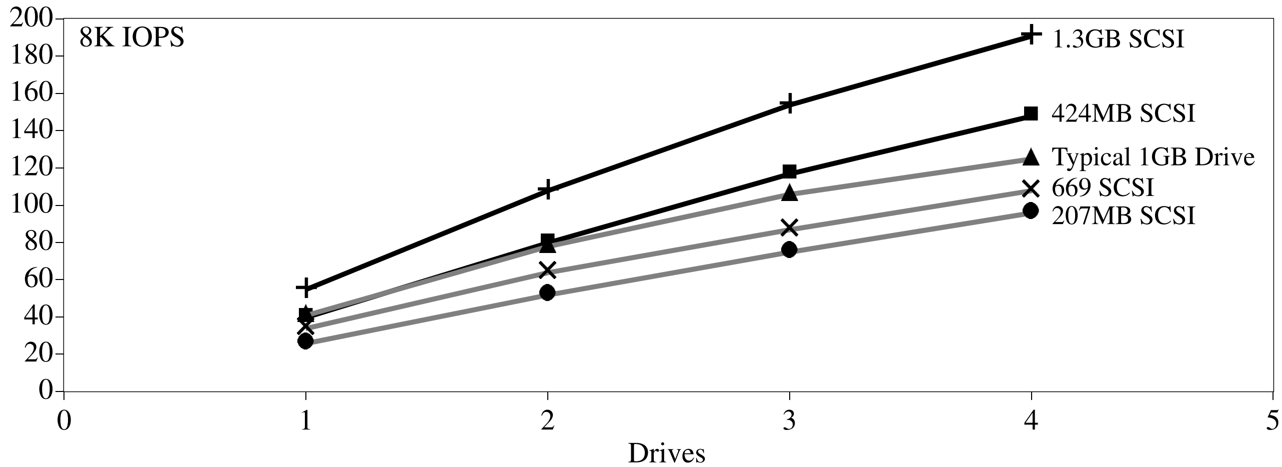
Table 1. Disk Throughput (KB/sec)

Disk	FSR	FSU	FSW	FRR	FRU	Geometric Mean	Ratio
424 MB	2124	1590	1583	392	507	1012	.71
669 MB	1515	1200	1198	327	404	779	.55
1.0 GB	1677	1546	1538	411	482	954	.67
1.3 GB	2485	2441	2428	535	757	1429	1.00
207 MB	651	809	806	269	298	509	.36

Random String Performance

The following chart shows random string performance of the different drives, which were all tested on a SPARCstation 2. As shown in the chart below, the 1.3-GB drive delivers twice the random string performance of the 669-MB drive, and is the highest performer of all the disk drives tested. This performance is primarily because of the fast seek time and low rotational latency of the drive and, to a lesser extent, the higher transfer rate.

Figure 2. SCSI Disks – File System Random Reads



Multi-User Benchmarks — Single Disk Configurations

Figures 3 and 4 (on the following page) show the behavior of different disks under the Musbus and SunMark workloads, respectively. The key numbers here are the throughput numbers as measured by how many scripts the system can finish in an hour. The curves increase smoothly until the capacity of the system is reached, at which point a steady decrease occurs. Such behavior is normal. The CPU utilization peaked at around 70 percent. Clearly, the disk is the bottleneck of the system in these cases.

Figure 3. **Throughput vs. Number of Users — Musbus Workload**

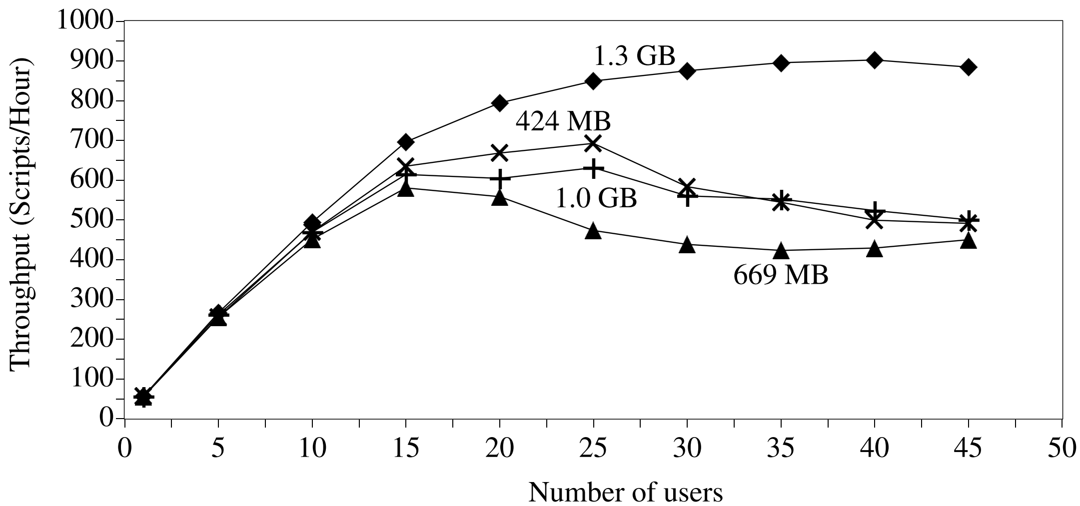


Figure 4. **Throughput vs. Number of Users — SunMark Workload**

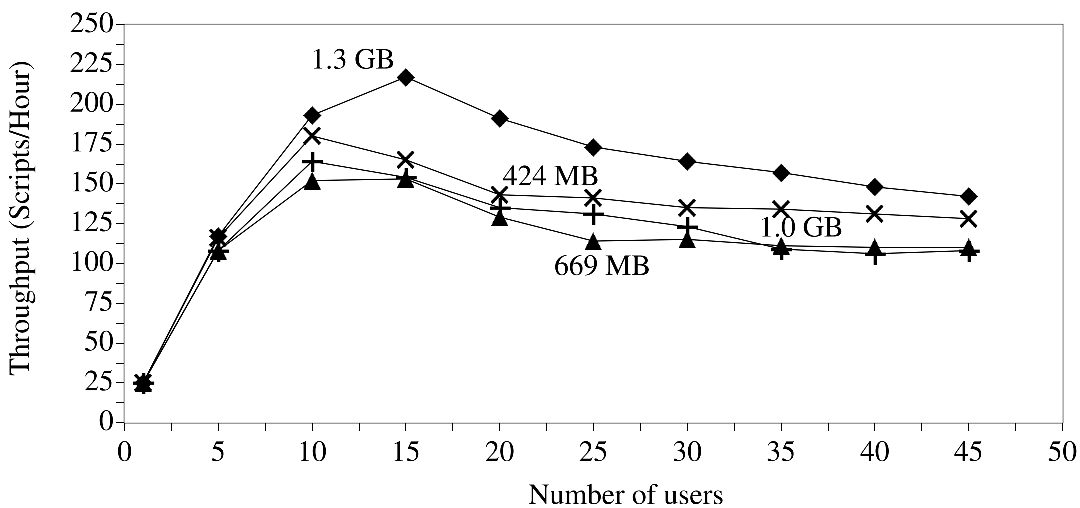


Table 2 compares the peak points of the throughput curves of the Musbus and SunMark workloads for the 1.3-GB disk relative to the other disks.

Table 2. Musbus and SunMark Peak Throughput Ratios

Test \ Disk	424 MB	669 MB	1.0 GB	1.3 GB
Musbus Ratio	0.77	0.64	0.70	1.00
SunMark Ratio	0.83	0.70	0.75	1.00

Multi-User Benchmarks — Two Disk Configurations

Multi-user tests were done on the combination of 207-MB plus 669-MB disks and 424-MB plus 1.3-GB disks. In this case, the smaller disk was configured as the system disk, with the SunOS operating system and swapping area on it. The larger disk acted as the user disk, with all the user data on it. Figures 5, 6 and 7 show the load throughput, average response time, and CPU utilization for the Musbus and Sunmark tests.

Figure 5. Musbus and SunMark Load Throughput (Scripts/Hour)

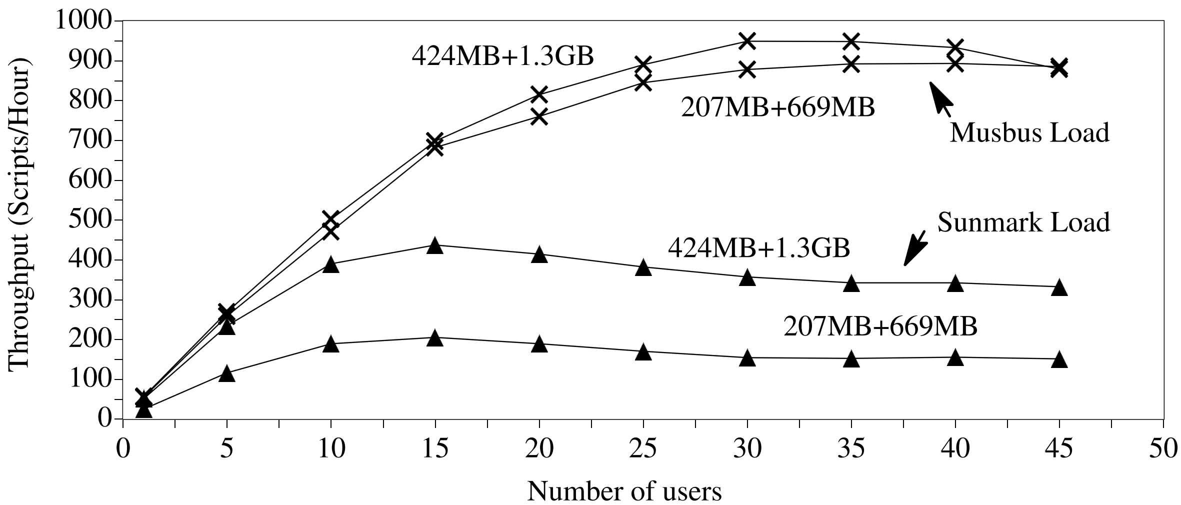


Figure 6. **Musbus and SunMark Average Response Time (Seconds per Script)**

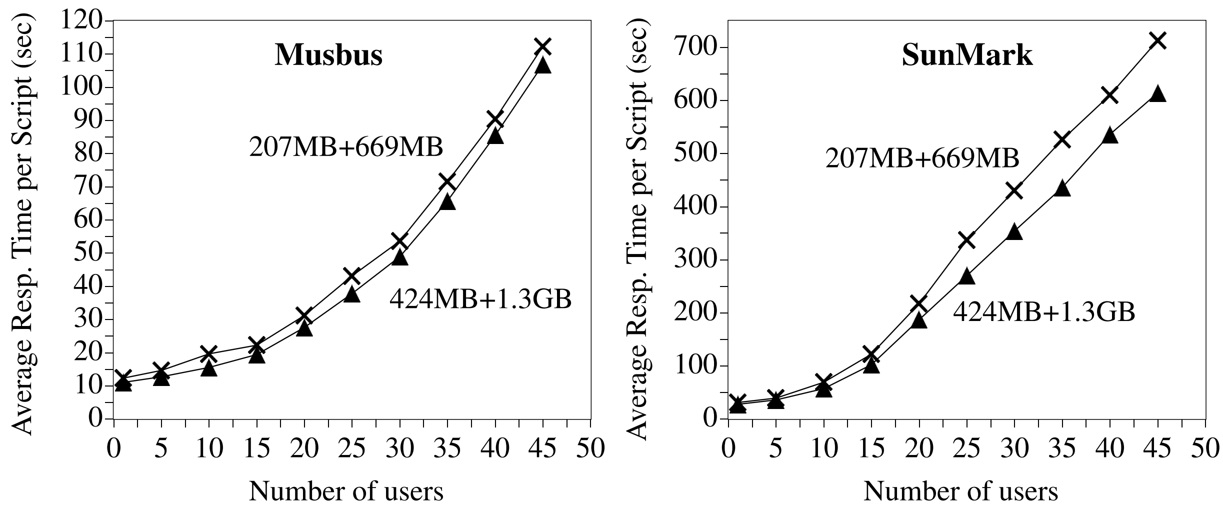
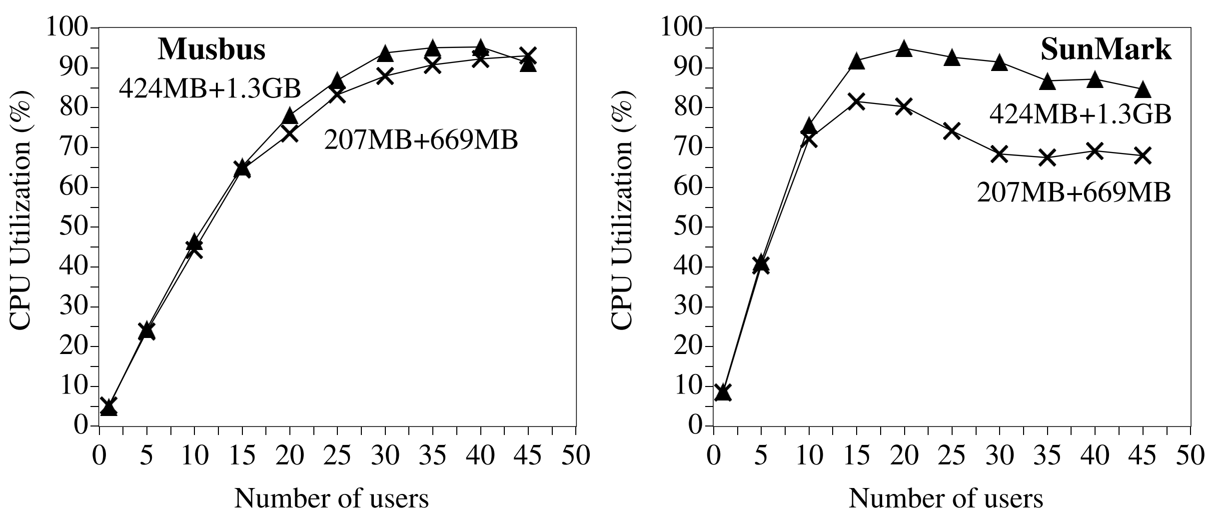


Figure 7. **CPU Utilization During Musbus and SunMark Tests**



- ▶ The combination of 207-MB and 669-MB disks has slightly less peak performance than the single 1.3-GB disk, but can sustain heavy workload better. Its peak throughput ratio to 1.3-GB disk is .99 for the Musbus benchmark and .94 for the SunMark benchmark, respectively.
- ▶ The combination of 424-MB and 1.3-GB disks creates a much better performance than all other disks. In fact, with this combination, the CPU became the bottleneck during the test. Thus, the peak throughput observed in this test may not be a true performance indicator of this disk

combination. To estimate the relative performance of this disk combination, examine the number at 45 users, where its throughput ratio for the sunmark benchmark is 1.07. (See Figure 4 for peak throughput of the 1.3-GB disk and Figure 5 for peak throughput of the 424-MB+1.3-GB disks.) The peak performance ratio should exceed this number.

Table 3. Musbus and SunMark Peak Throughput Ratio

Test \ Disk	424MB	669MB	1.0GB	1.3GB	207MB + 669MB	424MB + 1.3GB
Musbus Ratio	0.77	0.64	0.70	1.00	0.99	NA
SunMark Ratio	0.83	0.70	0.75	1.00	0.94	> 1.07

Database Benchmark Results

Database benchmark tests done on a combination of 1.3-GB and 424-MB disks compared results with earlier test results on a configuration of 669-MB and 207-MB disks. The tests performed on the configuration of the 1.3-GB and 424-MB drives were not audited and the TPC numbers are not official, whereas the tests on the 669-MB/207-MB combination were audited. This is why the relative performance is discussed in this report, rather than actual numbers being quoted.

The two configurations being compared were:

Disk	Configuration # 1	Configuration # 2
207 MB	2	0
424 MB	0	2
669 MB	10	0
1.3 GB	0	5
Total # of disks	**12**	**7**

The workload and testing methodology were identical for the two configurations.

Configuration 2 had 6 percent *higher* transactions per second (TPS) than Configuration 1, even though Configuration 2 had 5 *fewer* disks than Configuration 1.

The faster seek and latency times of the 1.3-GB and 424-MB disks mean that fewer data disks are required to maintain the same TPS rate. The lower rotational latency results in reducing queuing at the log disk and thus permits higher throughput results.

Conclusion

Sun's new SCSI disk products offer much better performance than any drives previously offered by Sun. These drives have characteristics that move them into an entirely new performance class.

These benchmarks provide a sense of the type of performance you can expect from these products. Although different applications and different benchmarks will yield somewhat different results, clearly the new 1.3-GB and 424-MB SCSI drives will provide the best performance for their form factor, capacity, and price classifications.

Sun Microsystems, Inc.
2550 Garcia Avenue, Mountain View, CA 94043
415 960–1300   Fax 415 969–9131

For U.S. sales office
locations, call:
800 821–4643
In California:
800 821–4642

European Headquarters
Sun Microsystems Europe, Inc.
Bagshot Manor, Green Lane
Bagshot, Surrey GU19 5NL
England
0276 51440
Fax 0276 51287

Australia: (02) 413 2666
Belgium: +32 2 730 38 11
Canada: 416 477–6745
Finland: +358–0–502 27 00
France: (1) 30 67 50 00
Germany: (0) 89–46 00 8–0
Hong Kong: 852 5–8651688
Italy: 039 60551
Japan: (03) 3221–7021
Korea: 2–563–8700

Latin America: 415 688–9464
Netherlands: 033 501234
New Zealand: (04) 499 2344
Nordic Countries: +46 (0)8 623 90 00
PRC: 1–8315568
Singapore: 224 3388
Spain: (1) 5551648
Switzerland: (1) 825 71 11
Taiwan: 2–514–0567

UK: 0276 20444
Europe, Middle East, and Africa,
call European Headquarters:
0276 51440
Elsewhere in the world,
call Corporate Headquarters:
415 960–1300
Intercontinental Sales:
415 688–9000

Specifications are subject to change without notice.

For: sucadmin
Printed on: Thu, Aug 22, 1991 11:08:52

Document: 8571
Last saved on: Thu, Aug 22, 1991 11:08:12