NETAPP PERFORMANCE IN SAN ENVIRONMENTS

TECHNICAL NOTE
P/N 300-004-233
REV A01
October 16, 2006

CONTENTS

Tests performed using a NetApp FAS3050c system revealed major performance shortcomings of NetApp filers in SAN/block environments. The test results contradict NetApp's claims about low TCO, superior performance, and overall life-cycle costs. This technical note covers the following topics:

INTRODUCTION

NetApp has made a number of claims about NetApp filer performance in Fibre Channel or iSCSI SAN environments, many comparing NetApp performance to Dell/EMC® CX Series arrays.[1] For perspective on NetApp's claims, extensive testing was performed on a NetApp FAS3050 system in typical SAN/block data environments, such as OLTP, Exchange, and others. While the tests described here were performed on a particular NetApp filer, the results apply broadly across NetApp systems and software.

The findings demonstrate that NetApp systems exhibit significant performance shortcomings in transaction-intensive block environments. The NetApp capacity utilization and performance issues observed in the tests seriously compromise NetApp's claims about low TCO, superior performance, and life cycle costs.[2]

Specifically, the testing indicates that:

- NetApp users need to maintain very low levels of storage capacity utilization to achieve competitive and predictable performance in transaction-intensive block environments. This factor clearly compromises the NetApp promise of storage "efficiency."
- As a consequence, NetApp systems must be significantly "over-provisioned" in terms of total capacity to provide predictable response times and service levels.
- Because performance also degrades over time, NetApp systems will need to be defragmented on a periodic basis to restore predictable performance. This is a disruptive and time-consuming activity.
- NetApp users will experience the same performance issues regardless of connectivity option—IP, iSCSI, or Fibre Channel.

1 VeriTest, "Network Appliance FAS3020 and Dell/EMC CX500: Comparison of Usability and Performance," May 2005, prepared under contract from Network Appliance.

2 Mercer Consulting, "Total Cost Comparison: IT Decision-Maker Perspectives on EMC, HP and Network Appliance Storage Solutions in Enterprise Database Environments," March 2006, prepared under contract from Network Appliance.

KEY OBSERVATIONS

- NetApp performance degrades as disk capacity utilization increases. Generally, NetApp systems exhibit high performance at low disk utilization capacities—10 percent to 15 percent, the utilization level at which NetApp usually reports benchmark results. As capacity increases, under constant load, NetApp response time and throughput diminish rapidly.
- NetApp performance and response time degrade significantly over time, even at constant load and constant capacity utilization.
- Higher drive count does not reduce or eliminate the observed performance degradation. In fact, test results indicate that response times become less predictable with higher drive counts.
- Performance issues are independent of connectivity—FC, IP, or iSCSI. Performance issues and bottlenecks are related to the NetApp file system and OS, not the interconnect. Utilizing high-performance interconnects such at Fibre Channel does not mitigate the basic performance issues.

PERFORMANCE TEST RESULTS

Performance demonstration #1: NetApp performance degrades as capacity utilization increases

This section presents the results of performance tests conducted on a NetApp FAS3050c system and a Dell/EMC™ CX3-40. The tests used a simulated Exchange application workload based on I/O traces of user environments running Microsoft Exchange and Microsoft's LoadSim tool. The workload consisted of 4 KB random reads and 4 KB random writes with a 2-to-1 ratio of reads to writes.

To demonstrate the impact of capacity utilization on performance, an external load generator was configured to submit multiple concurrent streams of 4 KB random reads and writes to 10 percent, 25 percent, 50 percent, and 100 percent of the usable data capacity of each system. For example, in the 10 percent utilization tests an initial stream of sequential writes was used to fill the LUNs to 10 percent of their capacity before the random read/write tests began. Subsequent tests were done by destroying and re-creating the LUNs and then using sequential writes to fill them to the 25 percent, 50 percent, and 100 percent utilization levels before beginning to submit the random reads and writes.

- The FAS3050c was configured with eight 2 Gb/s Fibre Channel host interconnects and the CX3-40 was configured with four 4 Gb/s Fibre Channel host interconnects.[3]
- The FAS3050c tests were done with data stored on 96 15k rpm 144 GB FC disk drives configured into 12 RAID 4 (7+1) groups. An additional six FC disks were reserved for use by NetApp's Data ONTAP root volume. Similar results were also obtained when the NetApp disks were configured into six RAID double parity (14+2) groups called RAID-DP.
- After subtracting space for the parity disk and ONTAP required space, one LUN with a usable capacity of 792 GB was configured for each of the 12 RAID 4 (7+1) groups.

3 The 4 Gb/s interconnects used by the Dell/EMC CX3-40 provided no performance advantage on this test because the simulated Exchange workload is not bandwidth-intensive.

- For comparison, tests were done on a Dell/EMC CX3-40 with data on 100 15k rpm 146 GB FC disk drives configured as 20 RAID 5 (4+1) LUNs, each with a usable capacity of 534 GB.[4]
- The "Appendix" contains additional information on the tested configurations. Each system was configured and tuned for maximum performance following the vendor's guidelines.

Figure 1 shows the results of the Exchange workload simulation when both storage systems are configured with FC host interconnects. The maximum attainable throughput in I/Os per second (IOPS) was measured at each capacity utilization level. The throughput was then converted into the number of Exchange users the configuration could support by assuming each system would be configured to run at 80 percent of its Exchange IOPS rating and the average Exchange user would perform one I/O per second.[5] Therefore, the formula used was: number of Exchange users = 80% * Exchange IOPS.

- While the NetApp FAS3050c performed better than the Dell/EMC CX3-40 at the very low capacity utilization level of 10 percent, **NetApp throughput capacity quickly degraded as higher percentages of the available disk capacity were utilized**.[6]
- As shown by the next test, "Performance demonstration #2," the initial level of performance for the FAS3050c shown in Figure 1 would rapidly degrade over time.

4 In most cases, the best practice from Microsoft is to configure RAID 1/0 for Exchange applications. However, NetApp does not support RAID 1/0, so RAID 5 was used to get as close to a like-for-like comparison as possible.

5 One IOPS per Exchange user is the assumption used by Microsoft's Exchange Solution Reviewed Program.

6 Similar behavior was observed on other workloads that add or update data over time.

Due to the inherent behavior of physical disks[7], Dell/EMC CX Series throughput also decreased somewhat, but **Dell/EMC CX Series' performance was significantly better than NetApp's at all but the 10 percent capacity utilization level**.

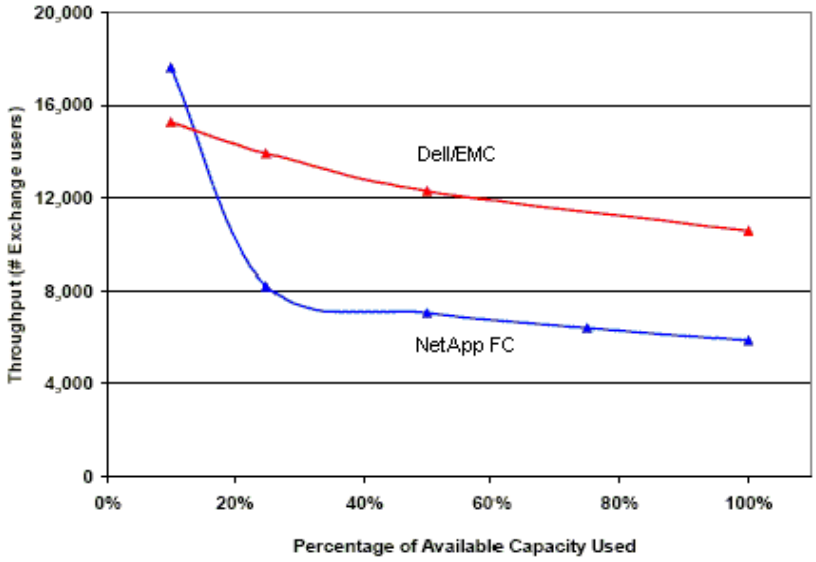

Figure 1. Exchange simulation — NetApp FAS3050c 96-drive RAID 4 (7+1) versus Dell/EMC CX3-40 100-drive RAID 5 (4+1)

7 The three primary components of disk I/O time are seek time, rotational latency, and data transfer time. Seek time is an increasing function of seek distance, which increases as more of a disk's capacity is utilized. As disk I/O time increases, maximum I/O throughput decreases.

Performance demonstration #2: NetApp performance degrades over time at constant capacity utilization

For this simulated Exchange test, the NetApp FAS3050c used 96 15k rpm 144 GB FC disk drives configured into six RAID-DP (14+2) groups. (Similar results were observed when the disks were configured as RAID 4 (7+1) groups.)

- Two LUNs, each with a usable capacity of 792 GB, were configured on each RAID-DP group. Snapshots were not enabled.
- **An initial stream of sequential writes was used to fill the LUNs to 25 percent of their capacity.** Then 312 concurrent streams of 4 KB random reads and writes—26 per LUN with a 2-to-1 ratio of reads to writes—were submitted to the previously written data space to simulate a use case in which an Exchange database is held constant at 25 percent of the usable disk space.
- IOPS throughput was collected at 10-second intervals and logged for more than 84 hours.

The simulated Exchange throughput data in *Figure 2* demonstrates how the performance of the NetApp FAS3050c system degrades over time.

- After starting out at a maximum throughput capacity of 8,150 users, **throughput degrades 21 percent** to 6,450 users after 24 hours, **30 percent** to 5,690 users after 48 hours, and then **38 percent** to 5,070 users after 72 hours
- NetApp's Write Anywhere File Layout (WAFL) file system always writes new or updated data blocks to a new block location on disk. To reduce physical I/O, WAFL batches individual writes together and tries to use free blocks in the same RAID stripe to reduce updates to the parity disk.
- As NetApp's free blocks become scarce and scattered over different RAID stripes, more physical I/O is required to write a block and user throughput decreases.
- **In comparison, Dell/EMC CX Series' performance at constant utilization will not degrade over time because data blocks are overwritten in place.** For example, tests confirmed that the Dell/EMC CX3-40 throughput capacity of approximately 13,900 users at the 25 percent utilization level, as shown in Figure 1, did not go down over time as data blocks were updated.

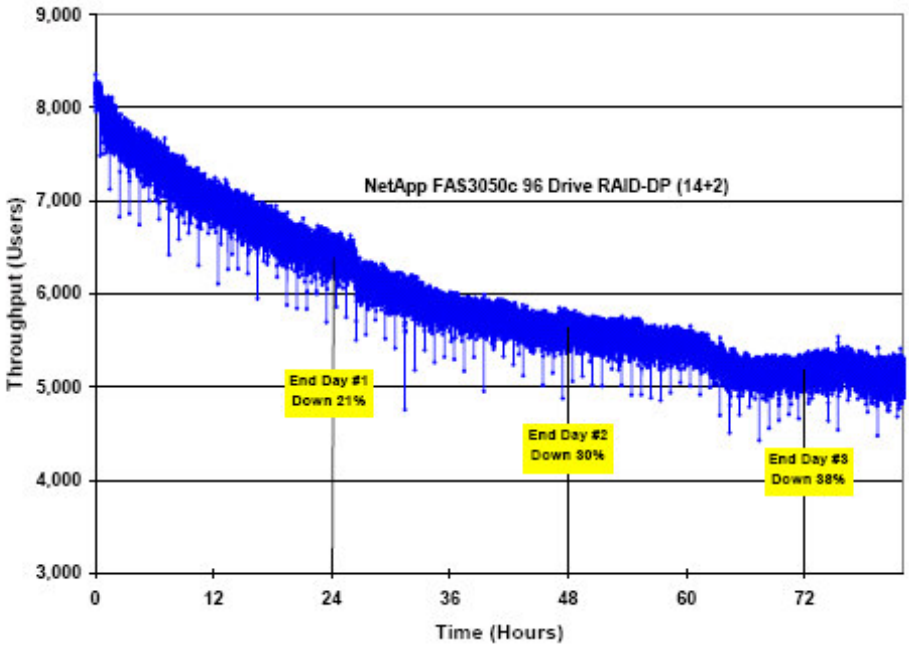

Figure 2. Exchange simulation over time—NetApp FAS3050c 96-drive RAID-DP (14+2)

Performance demonstration #3: Higher drive count does not reduce or eliminate performance degradation

FAS3050c random write tests were done with 32 and 96 data drives configured as RAID 4 (7+1) groups, one LUN per group. The host driver was configured with 26 concurrent threads per LUN, where each thread submitted a synchronous stream of 8 KB random writes to half of the total usable address space of each newly initialized (empty) LUN.

Figure 3 shows random write throughput over time of the 32 and 96 data drive configurations.

- Short-term performance was higher with 96 drives but performance did not scale well with increased drive count.
- **Over time, performance of the 32- and 96-drive configurations degraded to the same low level.** For example, after 12 hours the 96-drive system did about 9,500 random writes/s while the 32-drive system did 6,300 random writes/s, or only 50 percent more performance with three times the drives.
- After 48 hours the performance of both configurations decreased to about 5,200 writes/s and performance was still decreasing with time.

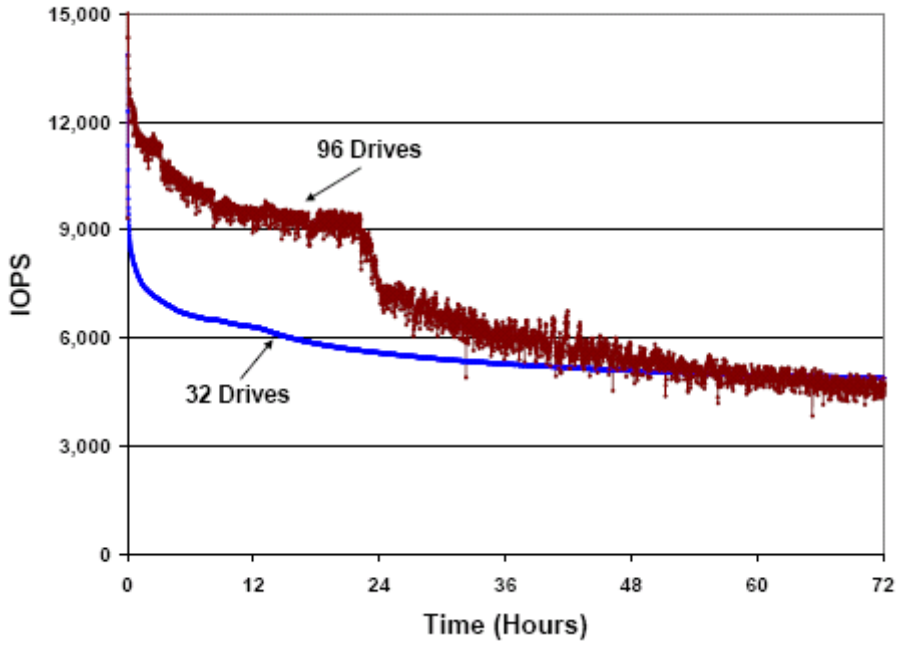

Figure 3. 8 KB random writes over time—NetApp FAS3050c 96-drive versus 32-drive RAID 4 (7+1)

Performance demonstration #4: Performance issues are independent of connectivity—FC, IP, iSCSI

This section compares the results of simulated Exchange application workload tests conducted on the FAS3050c using FC, iSCSI, and FC+iSCSI combination host connections versus the Dell/EMC CX3-40 with FC and FC+iSCSI combination connections.

- The FAS3050c tests were done with data stored on 96 15k rpm 144 GB FC disk drives configured into 12 RAID 4 (7+1) groups.
- The Dell/EMC CX3-40 tests were done with data on 100 15k rpm 146 GB FC disk drives configured as 20 RAID 5 (4+1) LUNs.
- In the combination tests the I/O was divided evenly between the FC and iSCSI ports.

The test results are plotted in *Figure 4*.

- **The FAS3050c showed virtually the same performance on the Exchange workload with the different interconnects and the same rapid performance degradation as the percentage of utilized capacity was increased from 10 percent to 100 percent.**
- The Dell/EMC CX3-40 showed similar performance on the FC and FC+iSCSI combination tests and a gradual performance decrease as more of the disk capacity was utilized.

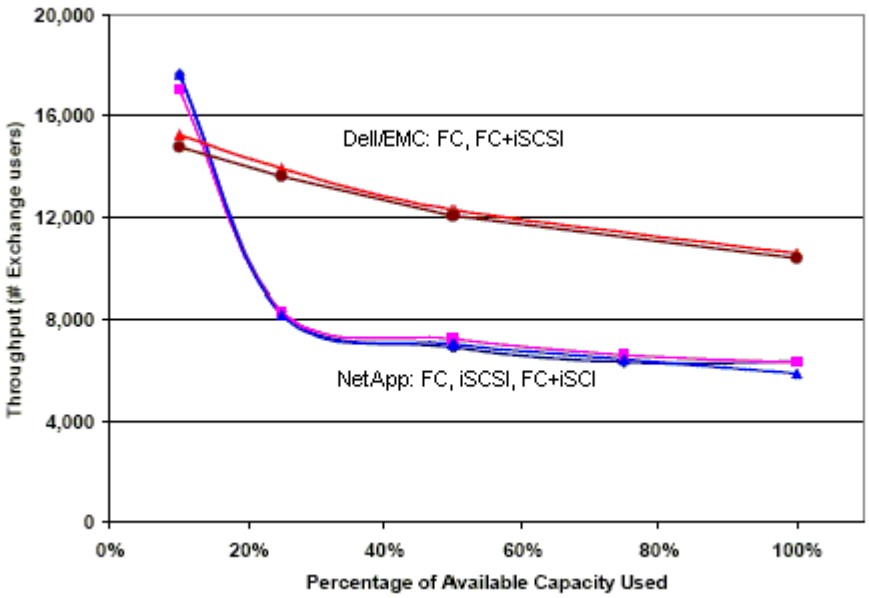

Figure 4. Exchange simulation with FC, iSCSI, and FC+iSCSI connectivity

SUMMARY

Fibre Channel and iSCSI SAN tests of a NetApp FAS3050c filer demonstrated that NetApp systems exhibit significant shortcomings in transaction-intensive block environments. In particular, tests showed that performance degraded rapidly as:

- Storage capacity utilization increases
- File system ages, even when storage capacity utilization is constant

In both cases, adding more disk drives did not reduce or eliminate performance degradation.

APPENDIX

NetApp FAS3050c configuration tested	
Storage processor	FAS3050c active/active dual-node cluster in single-image cluster mode
System ECC Memory	8 GB (4 GB/node)
Nonvolatile Memory (NVRAM)	1 GB (512 MB/node)
Host-attach	a) 8 Fibre Channel @ 2 Gb/s (4 FC per node) , or b) 8 iSCSI @ 1 Gb/s (4 iSCSI per node), or c) 8 Fibre Channel @ 2 Gb/s plus 8 iSCSI @ 1 Gb/s (4 FC + 4 iSCSI per node)
Disk drives	102 15k rpm 144 GB (6 system, 96 user)
Host LUNs	a) 12 @792 GB each, one per RAID 4 (7+1) disk group, or b) 12 @792 GB, 2 per RAID-DP (14+2) disk group
Back-end FC disk loops	4 @ 2 Gb/s (1 active and 1 passive per node)
Operating system	Network Appliance Data ONTAP version 7.2

Dell/EMC CX3-40 configuration tested	
Storage processor	Dell/EMC UltraScale CX3-40 dual storage processor (SP)
System ECC Memory	8 GB (4 GB/SP)
Host-attach	a) 4 Fibre Channel @ 4 Gb/s (2 per SP), or b) 4 Fibre Channel @ 4Gb/s plus 8 iSCSI @ 1 Gb/s (2 FC + 4 iSCSI per SP)
Disk Drives	100 dual ported 15k rpm 146 GB
Host LUNs	20 RAID 5 (4+1) @ 534 GB each
Back-end FC disk loops	4 @ 4 Gb/s (2 active per SP)
Operating system	Dell/EMC FLARE® release 22

The information contained in this document, including all instructions, cautions, and regulatory approvals and certifications, is provided by EMC Corporation and has not been independently verified or tested by Dell. Dell cannot be responsible for damage caused as a result of either following or failing to follow these instructions. All statements or claims regarding the properties, capabilities, speeds or qualifications of the part referenced in this document are made by EMC and not by Dell. Dell specifically disclaims knowledge of the accuracy, completeness or substantiation for any such statements. All questions or comments relating to such statements or claims should be directed to EMC Corporation.