Technical Report

Increase Application Performance by Using NetApp Storage Controllers and GridIron TurboCharger Data Accelerator Appliance

Saad Jafri, NetApp

December 2011 | TR-4005

ABSTRACT

This report describes the benefits of using the GridIron® TurboCharger™ appliance combined with NetApp® storage controllers to provide a high-performance environment for OLTP and decision support system (DSS) workloads.

TABLE OF CONTENTS

LIST OF FIGURES

1 INTRODUCTION

This report describes the benefits of using a GridIron TurboCharger (GT) with NetApp unified storage systems. NetApp evaluated the performance and stability of the overall configuration by using both OLTP and DSS type workloads on an Oracle® 11gR2 database.

GT is deployed between the server and NetApp storage controllers, where it continuously profiles the data access behavior of applications and creates a heuristics-driven map of the data space. Input/output (I/O) requests from applications are fulfilled almost instantly based on the real-time data map, which dramatically improves application performance.

2 EXECUTIVE SUMMARY

As a result of this evaluation, NetApp found that configuring GT to work with NetApp unified storage systems was simple and straightforward. Additionally, we found that using GT improved performance and better used database server resources in both OLTP and DSS environments. Specifically, we found the following items of interest:

- We doubled the number of users that could access the OLTP database when using GT while maintaining the same database response times.
- When using GT with a DSS database, query completion times were reduced by approximately 50% compared to when not using GT.

3 CONFIGURING GT WITH NETAPP UNIFIED STORAGE SYSTEMS

NetApp tested two configurations. Figure 1 shows the baseline configuration, and Figure 2 shows the configuration with GT deployed in-line. NetApp ran the workloads on an Oracle 11gR2 single-instance database. NetApp FAS 3170A, a midrange unified storage system, provided the storage.

Figure 1) Initial network configuration without GT.

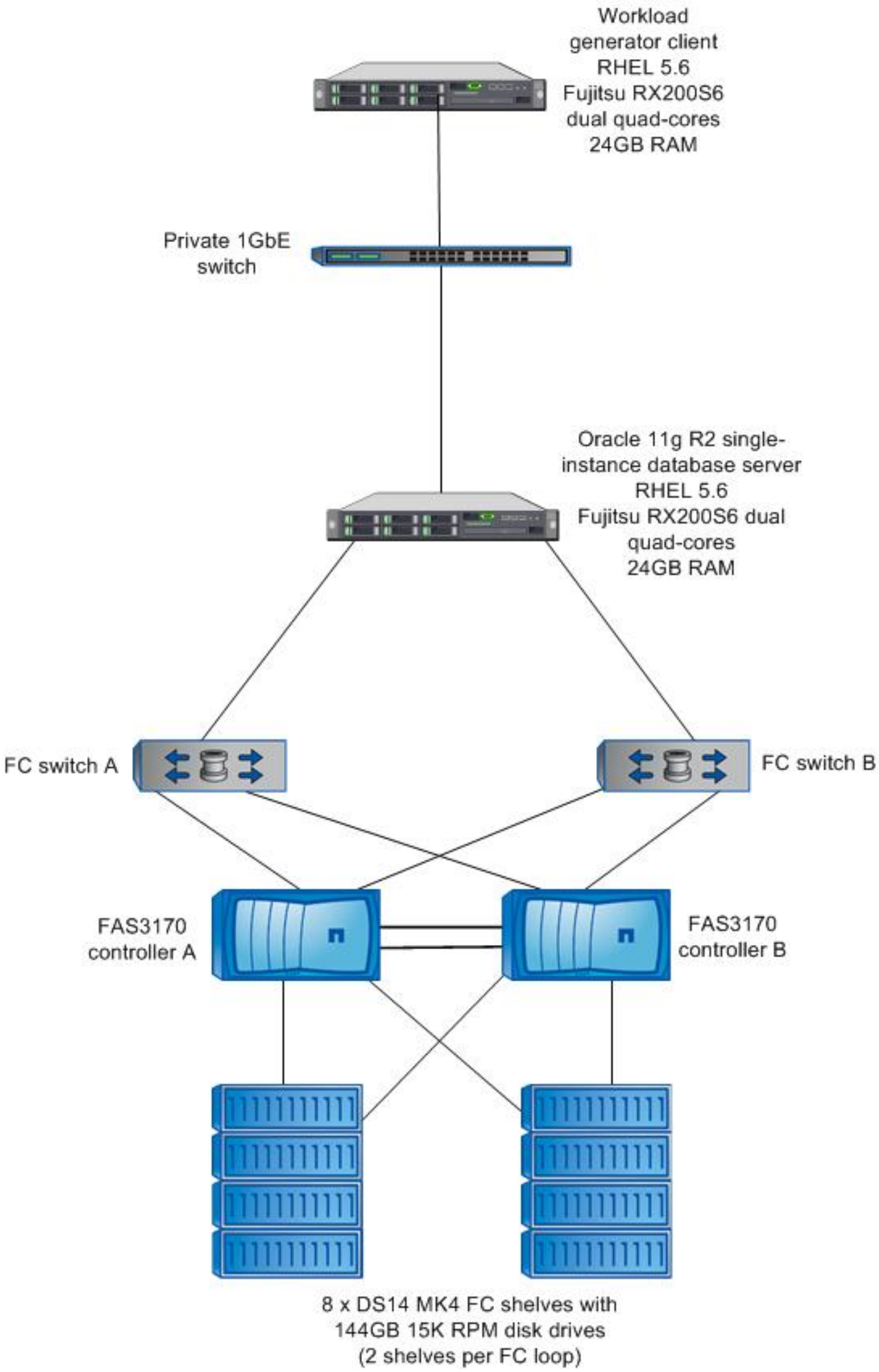
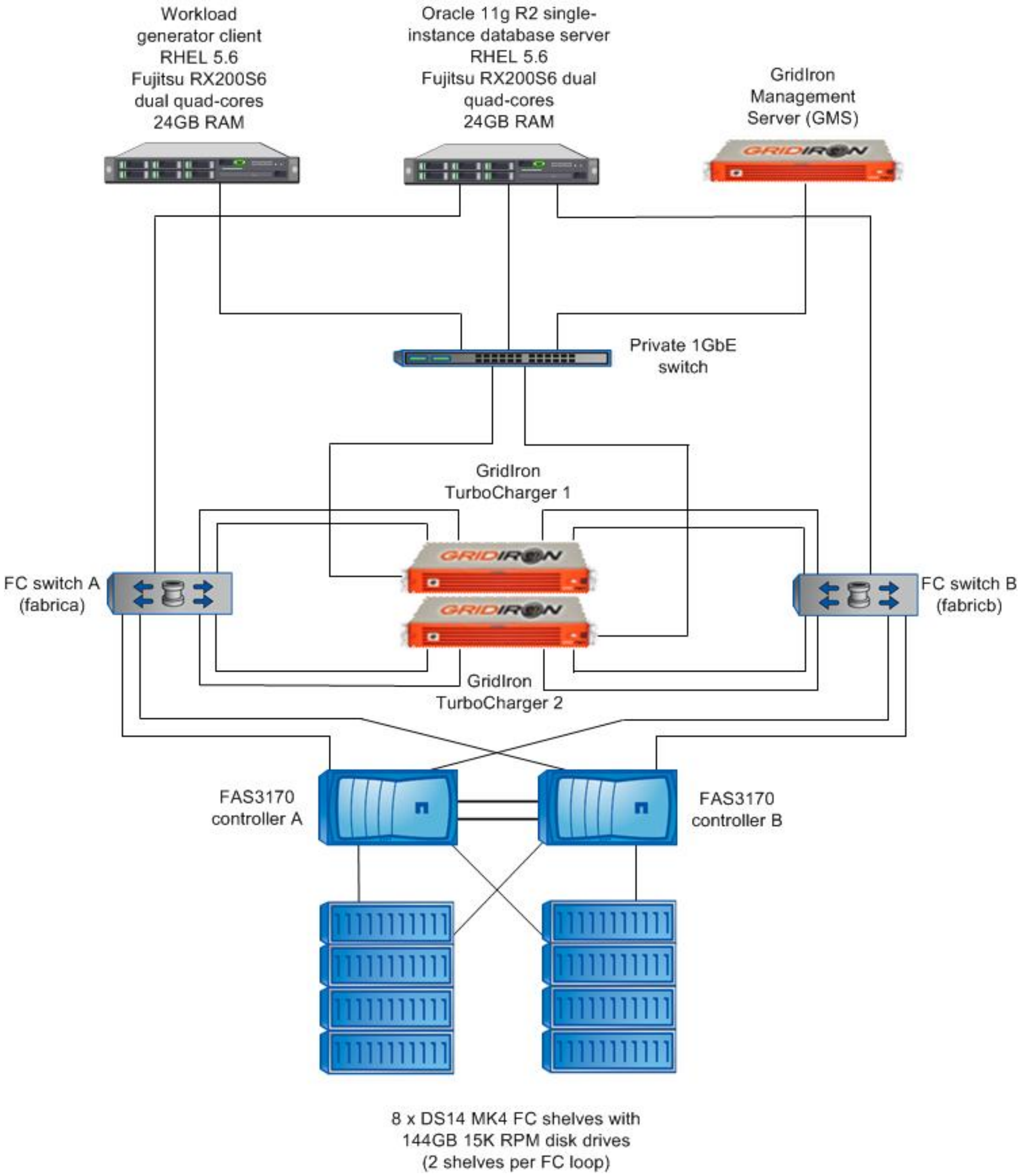


Figure 2 shows the same configuration as Figure 1, but with GT installed to act as a cache for the Oracle database server.

Figure 2) Network configuration with GT.



The GT deployment consisted of two GT appliances in a high-availability configuration and a management server (GMS). Each GT contained SSD-based flash drives totaling 6.5TB.

The GMS was connected to a 1 Gigabit Ethernet switch used primarily for system management, performance statistics data collection, and performance monitoring through a Web-based GUI.

4 PERFORMANCE TESTING AND VALIDATION METHODOLOGY

NetApp validation tests measured and compared performance, with and without GT, on Oracle 11gR2 single-instance databases by using both OLTP and DSS workloads.

The following sections describe the OLTP and DSS databases used for the testing, as well as the methodology used to capture the performance data.

4.1 OLTP WORKLOAD

The database created for the OLTP workload used a data model designed for order entry transaction (OET) processing. The OLTP database was approximately 1TB and contained approximately 6,000 warehouses. A separate application server ran the client processes for the OLTP application.

To determine the workload used for testing, we ran a series of tests using different numbers of users and, through an iterative process, increased the load generated against the database until we observed the read latencies reported by the Oracle database to be approximately 12ms. Response times in this range indicate an upper limit on acceptable performance in this configuration. NetApp required a load of 250 users to reach this threshold, and we used that threshold as our baseline performance without GT. NetApp FAS3170As were fully used in this configuration and were the primary factor limiting performance.

When we tested with GT deployed, we started the load with the same 250 users. Using an iterative process, we increased the load on the database up to 500 users. With this process, we determined that deploying GT between the database server and storage improved performance and lowered the response time.

The I/O mix of the OLTP workload was approximately 60% reads and 40% writes.

4.2 DSS WORKLOAD

The database created for the DSS workload used a data model designed for business-oriented ad hoc queries. The DSS database was approximately 1TB. To generate a DSS workload, we executed a query with some degree of complexity that read large volumes of data sequentially. The query performed a full table scan of the largest table in the database, resulting in a workload that was primarily sequential-read operations using large request sizes.

To achieve a steady state, we ran the query 10 times consecutively without deploying GT and used the results as a baseline. We then executed the same query with GT and compared the results to the baseline. The query was executed on a separate system from the DSS database server running in client-server mode. We used the query completion time as the primary metric to measure the application performance.

5 RESULTS AND ANALYSIS

This section discusses the results of performance and validation testing, with and without GT, on Oracle 11gR2 single-instance databases using both OLTP and DSS workloads on FA3170A unified storage controllers.

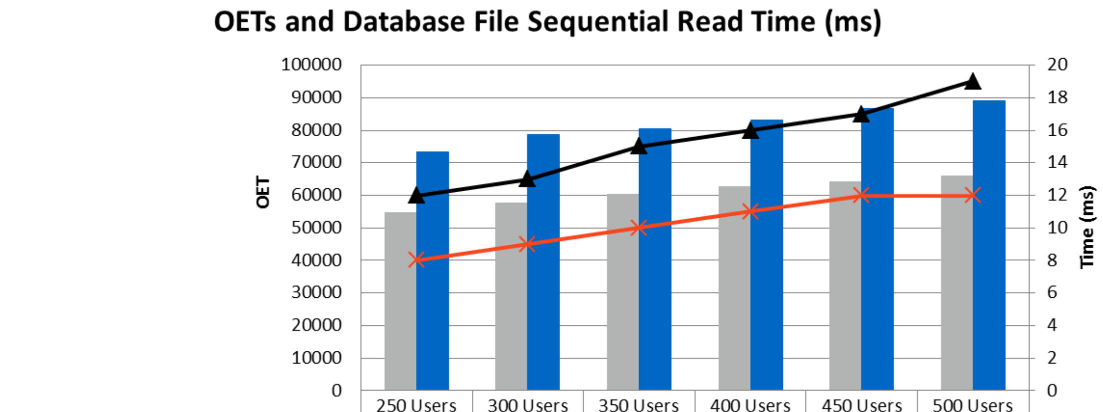
5.1 DETAILED OLTP WORKLOAD TEST RESULTS

Figure 3 shows the total number of OETs and the average read response time as reported by the Oracle database for OLTP tests. We observed the following results:

- With GT, the number of users who could access the OLTP database doubled from 250 to 500, while maintaining the database read latency of 12ms.

- Based on the number of users, OETs increased from 33% to 37% with GT.
- Based on the number of users, read latency decreased from 31% to 37% with GT.

Figure 3) OETs and database file sequential read time.

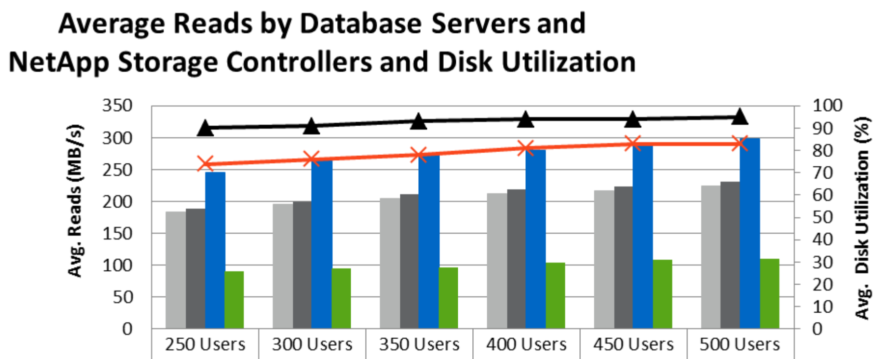


	250 Users	300 Users	350 Users	400 Users	450 Users	500 Users
OETs -Without GT	54591	57578	60216	62629	64286	66037
OETs - With GT	73388	78646	80407	83265	86729	89063
db File Sequential Read - Without GT	12	13	15	16	17	19
db File Sequential Read - With GT	8	9	10	11	12	12

Figure 4 shows the amount of data read by the database server from the NetApp storage controllers, with and without GT, using different numbers of users. Figure 4 also shows the disk utilization of the corresponding storage controllers. We observed the following results:

- Without GT, database reads were within 2% to 3% of data read from the NetApp storage controllers, which indicates that limited host-based caching occurred in the configuration.
- With GT, database reads decreased 62% to 64% in data read from the NetApp storage controllers, which indicates that GT cached and served a significant amount of the data read by the database server.
- With GT, storage controller disk utilization decreased from 11% to 16%, depending on the load, which indicates that GT reduced the load being placed on the storage system.

Figure 4) Average reads and storage disk utilization.



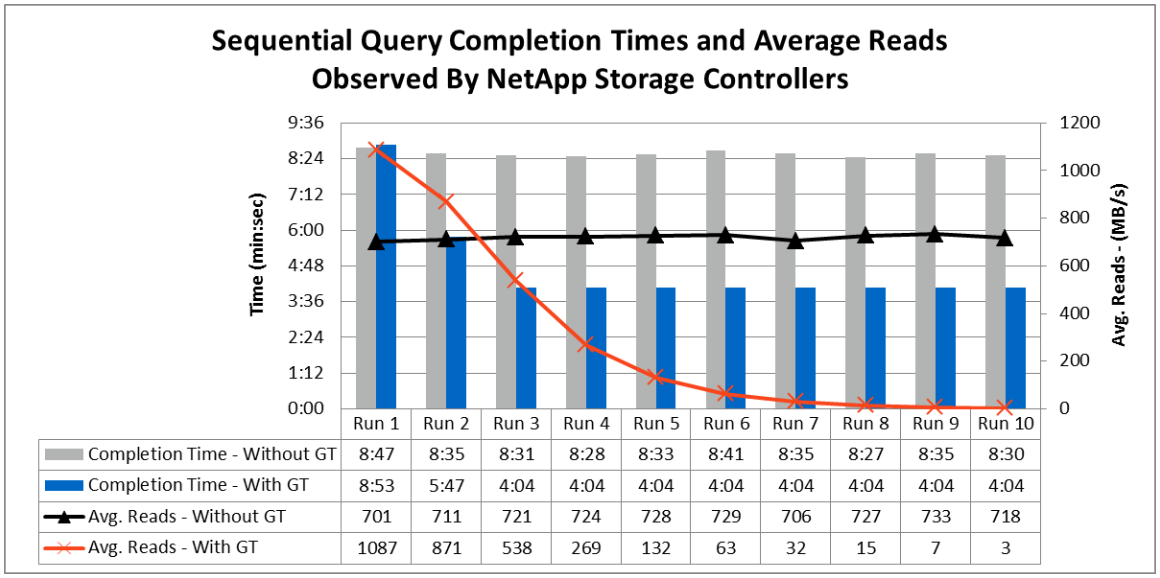
	250 Users	300 Users	350 Users	400 Users	450 Users	500 Users
Reads by DB Srv - Without GT	183	195	205	212	216	225
Reads by Strg Cntrlrs - Without GT	189	200	211	219	223	231
Reads by DB Srvr - With GT	246	263	272	280	291	298
Reads by Strg Cntrlrs - With GT	90	95	97	104	108	110
Strg Cntrls Disk Utilzation - Without GT	90	91	93	94	94	95
Strg Cntrls Disk Utilzation - With GT	74	76	78	81	83	83

5.2 DETAILED DSS WORKLOAD TEST RESULTS

Figure 5 shows query completion times for each run and the corresponding data read throughput in megabytes per second (MB/sec) as observed by NetApp storage controllers. The query performed a full table scan of the largest table in the database 10 consecutive times and reported the elapsed time required to complete each query. We observed the following results:

- With GT, query completion times were reduced approximately 50% after consecutively executing the same query and allowing the GT cache to warm up, which indicates that there was at least 50% I/O overhead in the DSS query.
- The average amount of data read from the NetApp storage controllers was reduced from 1087MB/sec to 3MB/sec after consecutively executing the same query 10 times, which indicates that GT served most of the data read by the database server from its cache.
- GT converged on the important data by the third consecutive run of the query (run 3). After this run, additional caching by GT did not decrease the query time; however, GT continued to cache more data, which resulted in additional reductions in back-end storage read load well into the eighth run. After the first three runs, the additional read off-load meant more spare storage resources for other processes.
- With GT, the storage system read load was reduced by 99%.

Figure 5) Query completion times and average reads by storage controllers.



	Run 1	Run 2	Run 3	Run 4	Run 5	Run 6	Run 7	Run 8	Run 9	Run 10
Completion Time - Without GT	8:47	8:35	8:31	8:28	8:33	8:41	8:35	8:27	8:35	8:30
Completion Time - With GT	8:53	5:47	4:04	4:04	4:04	4:04	4:04	4:04	4:04	4:04
Avg. Reads - Without GT	701	711	721	724	728	729	706	727	733	718
Avg. Reads - With GT	1087	871	538	269	132	63	32	15	7	3

6 CONCLUSION

The results described in this document demonstrate that using GT appliances in the data storage path works as intended by intelligently caching data in its SSD-based flash drives and improving performance of both OLTP and DSS type workloads. The GT data acceleration appliance, when combined with NetApp storage controllers, is easy to install and operate, significantly reduces response times for database queries, and significantly increases the concurrent query capacity of the system.