# Evaluation of STEC s1120 PCIe Accelerator and EnhanceIO™ SSD Cache Software

***Evaluation report prepared under contract with STEC***

## Introduction

Solid state storage continues to gain acceptance in enterprise data centers, providing increased performance and lower latency for mission-critical applications. Enterprises have a variety of choices to make when it comes to deploying SSD technology, including SSD form factors and data placement strategies.

STEC commissioned Demartek to evaluate its s1120 PCIe Accelerators in both primary storage and caching configurations with its EnhanceIO SSD Cache Software, and to compare these solutions to a more traditional Fibre Channel SAN hard disk drive storage solution. This evaluation was conducted in the Demartek lab in Colorado.

## Executive Summary and Key Findings

Demartek evaluated the STEC s1120 PCIe Accelerators in both primary storage and caching configurations, and compared these results to a well-known Fibre Channel hard disk drive (HDD) solution. The STEC PCIe Accelerators provided outstanding performance in terms of database transactions per second, bandwidth and latency.

For the primary storage configurations, using the s1120 980GB capacity, we found that:

- Database transactions per second were up to 16x higher than the HDD solution.
- Bandwidth was up to 7.5x higher than the HDD solution.
- Latency was more than 100x better (lower) than the HDD solution.

For the caching configurations, using the s1120 480GB capacity, we found that:

- Database transactions per second were up to 8x higher than the HDD solution.

We recommend that enterprises consider the STEC s1120 PCIe Accelerators and EnhanceIO Caching software for any application that needs significantly improved performance.

## Solid State Storage – PCI Express Adapter Form Factor

Solid State Storage can be implemented in several form factors. STEC provides both the "disk drive" form factor and PCI Express (PCIe) form factor storage units. This report focuses specifically on the PCIe form factor STEC s1120 PCIe Accelerator.

PCI-Express (PCIe) stands for Peripheral Component Interconnect Express and is the computer industry standard for the I/O bus for computers introduced in the last few years. The first version of the PCIe specification, 1.0a, was introduced in 2003, and the first products supporting PCIe 1.0 began appearing in the marketplace in 2004. Since then, servers designed for the enterprise datacenter have included PCIe slots designed for higher-speed devices.

The STEC s1120 PCIe Accelerator provides outstanding performance due to its use of high performance SSD technology and ASIC-based SSD controller. It has low latency, as this device connects directly to the PCIe bus, with no external network or storage protocols between the CPU and the storage.

Figure 1 – STEC s1120 PCIe Accelerator

The s1120 series PCIe Accelerator is available in 240GB, 480GB, 980GB and 2TB capacities. The 240GB model is available with SLC NAND flash, and the other models are available with MLC flash. The s1120 PCIe Accelerator comes with a five-year warranty.

The s1120 PCIe Accelerator requires a PCIe 2.0 x4 slot and will also work in PCIe 3.0 systems.

## Data Placement with Solid State Storage

Once a decision has been made to incorporate SSD technology into an enterprise, the next major decision focuses on how to use this SSD technology, or more specifically, how to place data on these devices. There are two general ways to use SSDs: for primary storage or for caching. Both methods are effective at improving I/O performance. Both of these two methods can work with any SSD form factor.

The overall goal of SSD solutions, both primary storage and caching, is to improve the performance of applications that depend on storage systems, because the storage systems are generally the slowest parts of a computing environment.

Each method of data placement has advantages, which are outlined below.

### SSDs as Primary Storage

When SSDs are used as primary storage, the following observations can be made:

- The administrator (or user) decides which data to place on the SSDs.
- The administrator decides when to place the data on the SSDs.
- The administrator is responsible for adjusting applications and backup configurations in order to find the data at its new location on the SSDs.
- Only the applications whose data resides on the SSDs gain a performance improvement.
- The performance gains are immediate and significant.

### SSDs as a Cache

When SSDs are used as a cache, the following observations can be made:

- The caching solution places a ***copy*** of "hot" data into the cache.
- The caching solution decides when to place and replace the copy of the hot data into the cache.
- Multiple applications can gain a performance benefit.
- The aggregate performance gains occur over time as the cache "warms-up."
- Applications do not need to be modified to take advantage of the SSD cache.
- Some caching solutions only cache reads, others cache both reads and writes.
- Management of a caching solution is relatively simple.

## Database Performance and Workloads

Considerable effort is devoted to optimizing database performance to overcome the limitations of electro-mechanical hard disk drives. While this effort is appropriate, another approach to improving database performance is to upgrade the data storage devices to an entirely new type of storage capable of considerably improved performance. SSD technology is the technology that can accomplish this purpose.

For these tests, Demartek deployed a real database, Microsoft SQL Server, and database workload onto a server in the Demartek lab using a well-known Fibre Channel hard disk drive storage array. We then compared that performance to deploying the same database using the STEC s1120 PCIe Accelerator in both primary storage and caching configurations.

The workload chosen is a variation of an On-Line Transaction Processing (OLTP) workload (TPC-E) that models a brokerage firm with customers who generate transactions related to trades, account inquiries, and market research. The brokerage firm in turn interacts with financial markets to execute orders on behalf of the customers and updates relevant account information.

The benchmark is “scalable,” meaning that the number of customers defined for the brokerage firm can be varied to represent the workloads of different-size businesses. The benchmark defines the required mix of transactions the benchmark must maintain. The TPC-E metric is given in transactions per second (tps). It specifically refers to the number of Trade-Result transactions the server can sustain over a period of time.

This TPC-E workload was not used to generate an official TPC benchmark score, but was used to provide an actual database workload for the purposes of comparing the performance of the storage systems. The performance data shown in this report was taken from PerfMon, the standard performance monitoring application that is provided with the Windows operating system.

For these tests, we limited Microsoft SQL Server to 32GB RAM even though there was more RAM in the server under test. This is because we wanted to limit the effects of memory caching functions that most database applications employ, including SQL Server, so that we could put more stress on the storage devices.

## Primary Storage Using SSDs Directly

One of the ways to deploy SSD technology is to use the STEC s1120 PCIe Accelerators as the primary storage for database applications. These PCIe 2.0 x4 SSDs provide outstanding performance and low latency for mission-critical database applications, and one or more of these devices can be deployed, depending on the capacity and performance requirements of the application.

In these tests, Demartek deployed four of the 980GB STEC PCIe Accelerators in two different configurations, RAID0 and RAID5, and compared the performance to a hard-disk drive storage array (the "baseline"). The standard performance tool, PerfMon, was used to capture SQL Server statistics during these tests.

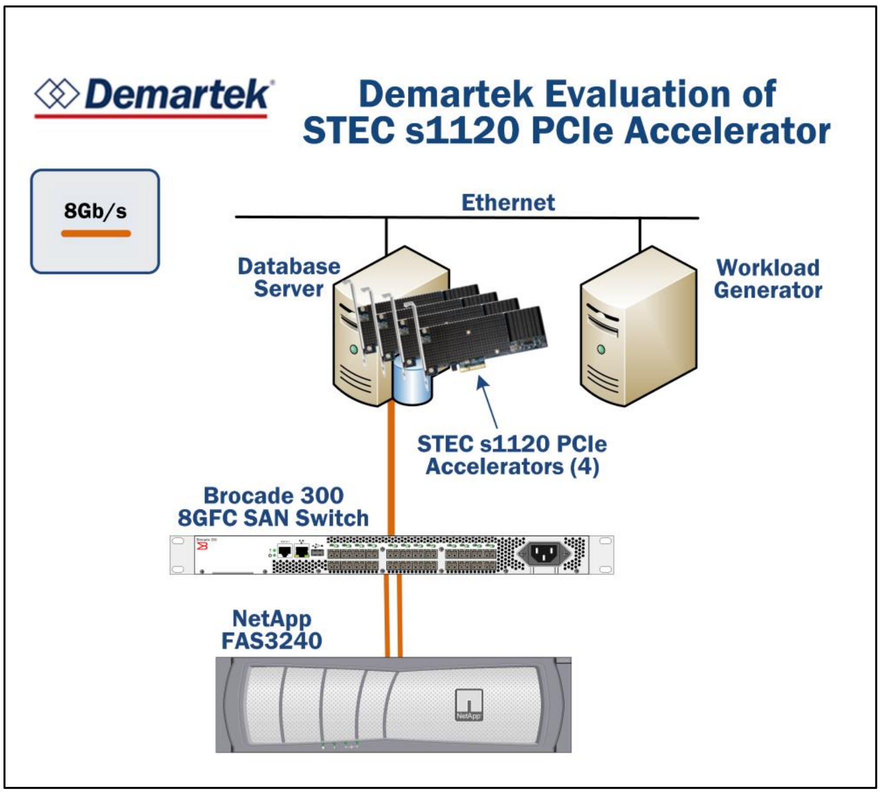


**Figure 2 – Demartek Test Configuration: s1120 PCIe SSD**

When multiple s1120 PCIe accelerators are installed into the same server, various types of RAID or striping technologies can be used to increase performance. RAID0 is often used to achieve the maximum performance of a set of devices, but RAID0 provides no additional parity or other data protection mechanisms in the event of device failure.

Because RAID0 does not calculate and require any added parity, it has lower latency than RAID5 configurations using the same number of devices.

## Transactions per Second – SQL Server

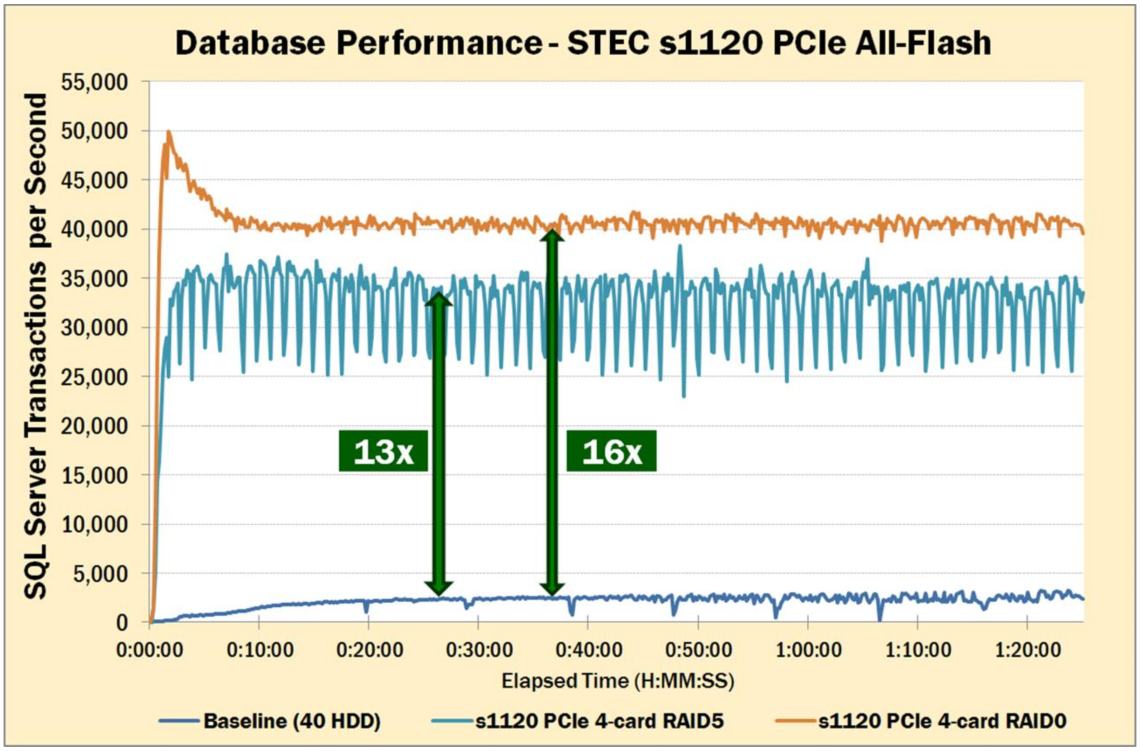


Figure 3 – SQL Server Transactions per Second

For database transactions per second, the RAID5 configuration of the PCIe Accelerators achieved **13x** the average performance of the baseline and the RAID0 configuration achieved an average of **16x** the performance of the baseline.

## Bandwidth (MB/sec) – Logical Drive

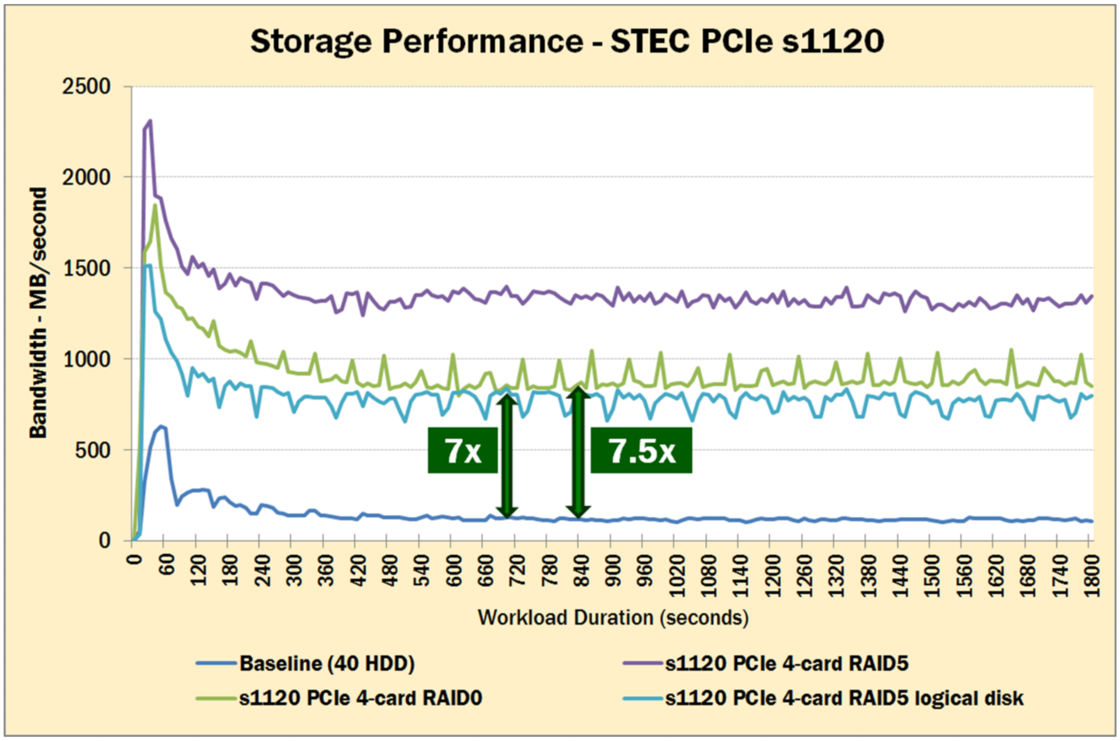


**Figure 4 – Bandwidth (MB/sec)**

### RAID5 Parity Calculations on Physical Disk

Both the physical and logical disk performance data are shown in the chart above, because RAID5 performs additional I/O operations for its parity calculations. The most appropriate comparison is to use the logical disk data for RAID5.

### Bandwidth Comparisons

For bandwidth, the RAID0 configuration achieved more than **7.5x** the average bandwidth of the baseline and the RAID5 configuration (logical disk) achieved more than **7x** the average bandwidth of the baseline.

## Latency – Physical Drive

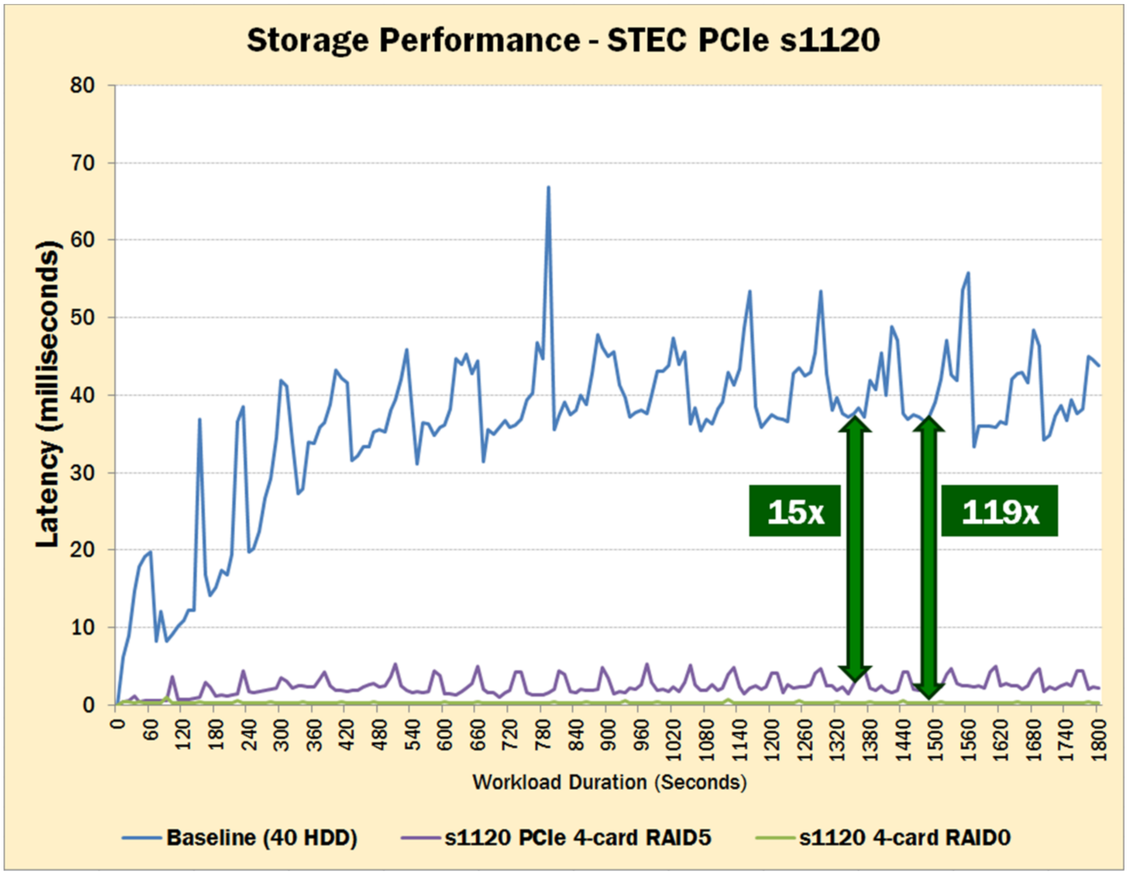


**Figure 5 – Latency**

The overall average latencies (read and write combined) were:

- Baseline: **40.77 ms**
- RAID5: **2.71 ms**, a **15x improvement** over the baseline
- RAID0: **0.34 ms**, a **119x improvement** over the baseline

The RAID0 configuration provided a sub-millisecond latency (34 microseconds) that was two orders of magnitude better than the baseline configuration. With latencies this low, it is not surprising that the amount of work processed was significantly higher than the baseline in the same amount of time.

## Using SSDs as a Cache

Another way to improve database performance is with a server-side SSD cache. This solution can be much less costly because the entire database does not have to fit onto the SSD cache. The tradeoff is that the performance will be somewhat less than the performance of an SSD used in a primary storage configuration, because the cache has some processing overhead. In addition, because the entire database is not stored on the SSD, a smaller quantity of SSD can be deployed, resulting in lower overall costs.

EnhanceIO works by using the server-side SSD as a cache in order to accelerate read and write operations that are intended for the storage under management, without having to change the back-end storage configuration. When the host processor performs a read operation to the back-end storage, EnhanceIO places a copy of the requested data block into the cache. If there is no room for this new block in the cache, an older data block is removed from the cache

For this set of tests, two different capacity SSDs were used – a 240GB SSD and a 480GB SSD. The tests were run longer than in the primary storage configuration, in order to allow the cache to "warm up" to its full performance capability.

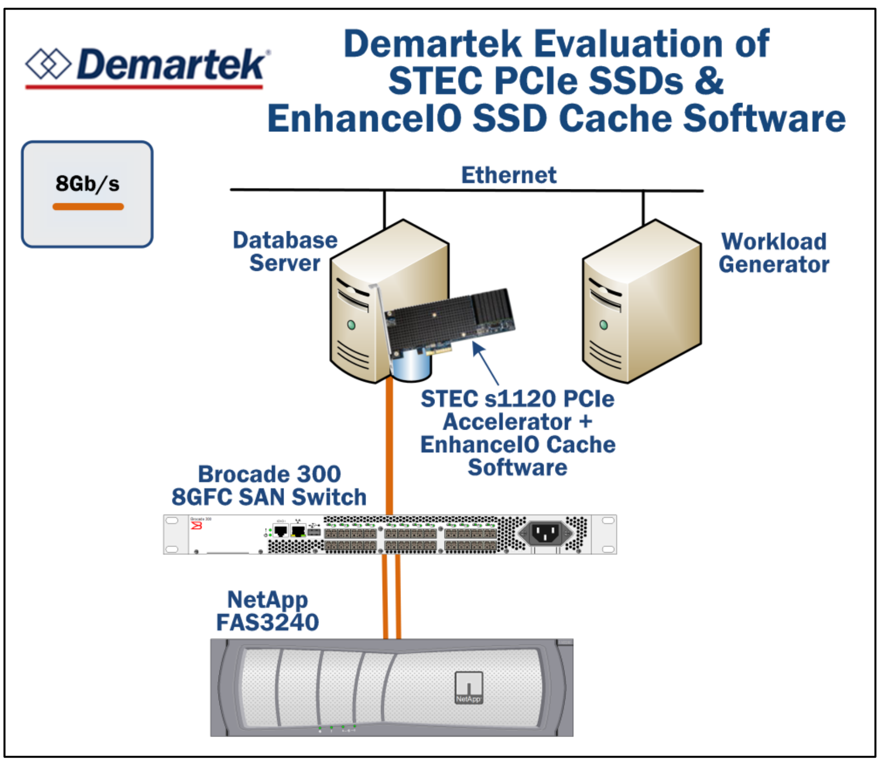


**Figure 6 – Demartek Test Configuration: EnhanceIO SSD Cache**

With the smaller cache (240GB) we observed a 4x improvement in performance in transactions per second, as observed by the SQL Server application, once it reached steady state. A second run was made with the larger cache after it had achieved steady state ("ss"), and it achieved approximately 8x performance improvement in database transactions per second.

As is typical of SSD caches, there is a gradual increase in performance over time until the cache reaches a steady state. The smaller cache reached a steady state during the test time period. The larger cache required a longer time period to reach steady state.

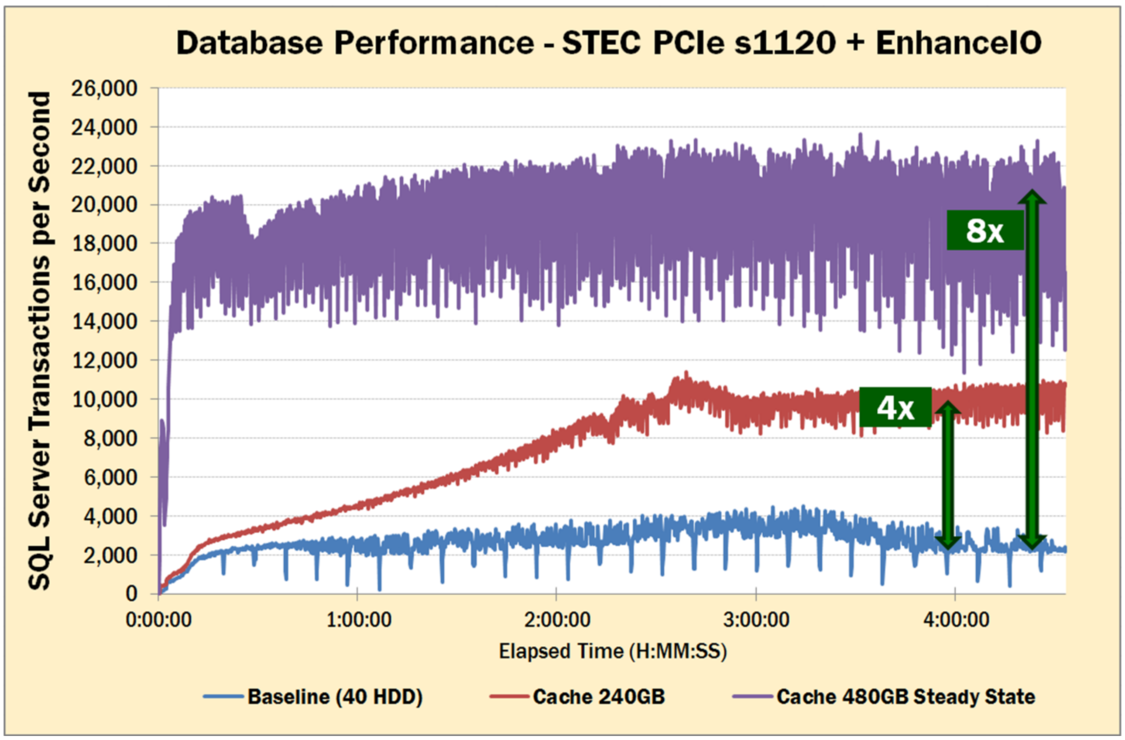


**Figure 7 – STEC s1120 PCIe + EnhanceIO Performance over Time**

## Average vs. Peak Cache Performance

The chart in Figure 5 above shows the cache performance over time. The comparison figures displayed on the chart (4x and 8x) indicate the improvement once the cache has reached steady state.

The average number of transactions per second during the warm-up period is less than that achieved at the end of the test period or the cache has reached steady state. This is because the average factors in the initial period when the cache was "cold" and began its steadily increasing performance.

The steady state run using the 480GB cache was manually terminated early, relative to the other runs, because it had reached steady state performance.

The SSD cache warm-up can be observed in the following chart. The smaller, 240GB cache, achieved steady state at slightly more than 2.5 hours into the test period. The larger, 480GB cache performance climbed steadily throughout the test period and was still increasing after 7.5 hours. The larger cache produced an approximately 7x performance improvement during the last 20% of the 7.5 hour test period. As described above, we ran a second test with the 480GB cache fully warmed to show the maximum performance for this test.

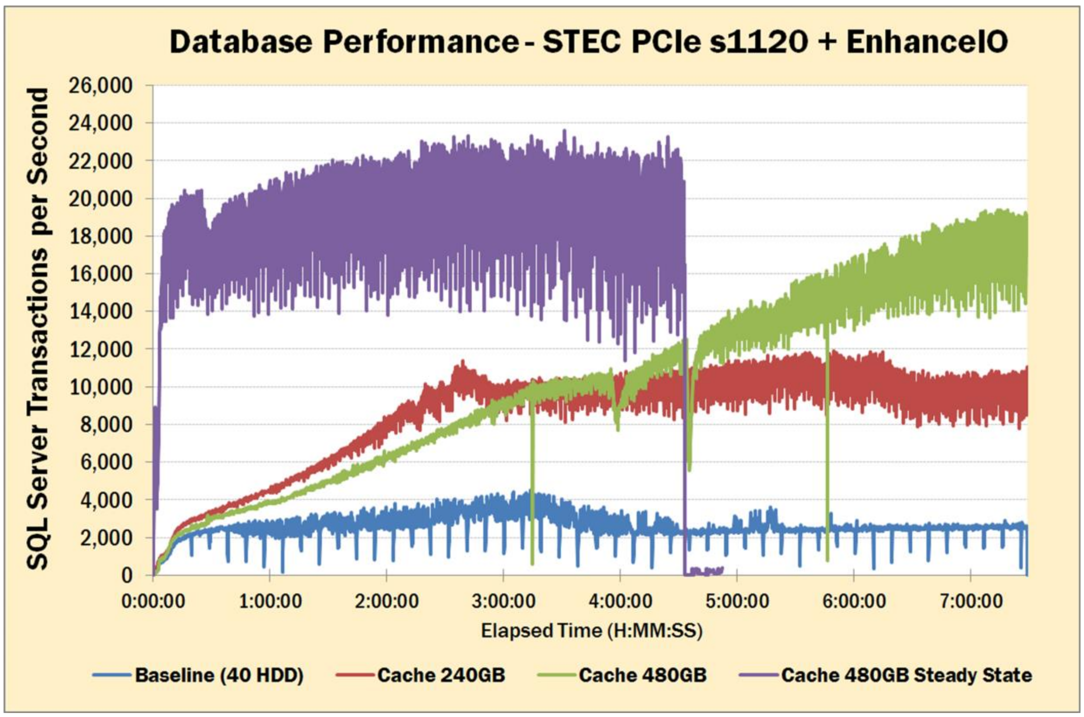


**Figure 8 – EnhanceIO Cache Warm-up Performance**

## Summary and Conclusion

The STEC s1120 PCIe Accelerator provides significant performance advantages over traditional hard disk drive storage solutions when it comes to database transactions per second, bandwidth and latency results.

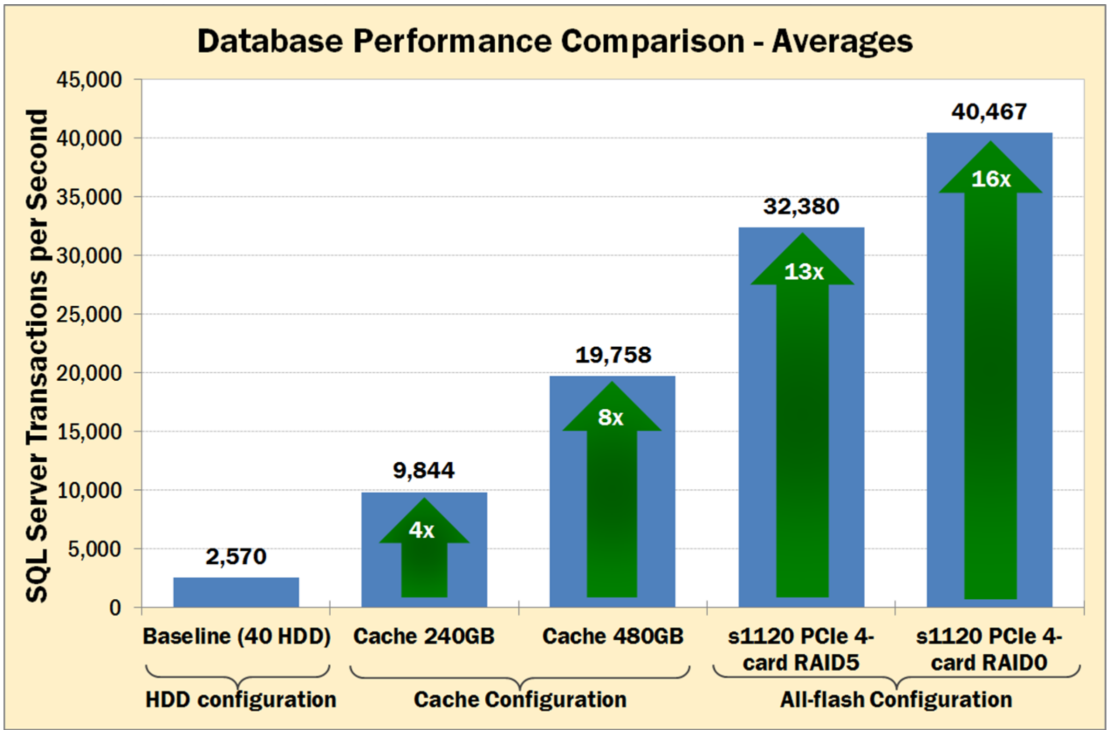


Figure 9 – Database Performance Comparison Averages

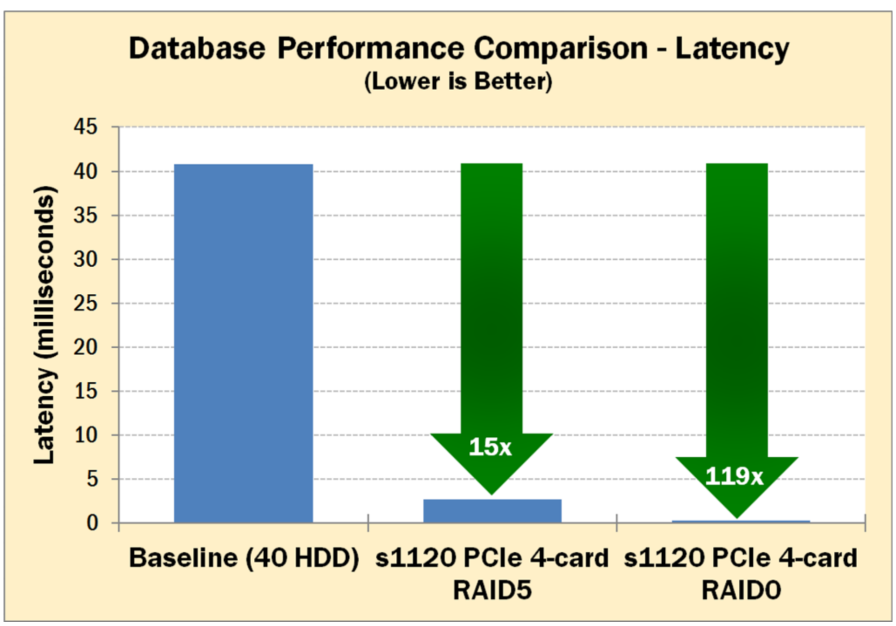


Figure 10 – Database Performance Comparison - Latency

In the primary storage configuration, where the database was moved over to the STEC s1120 PCIe Accelerators and run directly there, there were several key results:

- Database transactions per second, as measured by SQL Server, were **16x higher** using the RAID0 configuration and **13x higher** using the RAID5 configuration that they were for the 40-drive HDD database configuration.
- Bandwidth was **7.5x higher** using the RAID0 configuration and **7x higher** using the RAID5 configuration.
- Latency was **15x better** in the RAID5 configuration and a whopping **119x better** in the RAID0 configuration.

The STEC s1120 PCIe Accelerator with EnhanceIO Caching solution provided the following key results:

- A **4x higher** database transactions per second (as measured by SQL Server) with the smaller cache (240GB) once the cache reached steady state. The average number of transactions per second during the test period was approximately **3x higher** than the baseline.
- A **7x higher** database transactions per second with the larger cache (480GB) during the last 20% of the test period. The average number of transactions per second during the test period (while the cache was warming) was approximately **4x higher** than the baseline.
- An average of **8x higher** database transactions per second with the larger cache after the cache achieved steady state.

## Appendix – Evaluation Environment

The tests were conducted in the Demartek lab in Colorado.

**Server – Database Server** (DMRTK-SRVR-L)

- Supermicro X9DR3-LN4F+ (PCIe 3.0)
- 2x Intel Xeon E5-2690, 2.9 GHz, 16 total cores, 32 total logical processors
- 192 GB RAM
- Internal SSD boot drive
- Windows Server 2008 R2
- SQL Server 2012 – limited to 32 GB RAM usage

**Server – Workload Generator** (DMRTK-SRVR-K)

- Supermicro X9SCM-F (PCIe 2.0)
- 1x Intel Xeon E3-1280, 3.5 GHz, 4 total cores, 8 total logical processors
- 32 GB RAM
- Internal SSD boot drive
- Windows Server 2008 R2

**Storage – Internal (Server)**

- 4x STEC s1120 PCIe Accelerators 980GB each

**Storage – External (via Fibre Channel)**

- NetApp FAS3240, 8Gb FC host interfaces, 48x 10K RPM, 450GB

**Database Configuration**

- HDD configuration: database spread across 40 HDD (10K, 450GB) RAID-DP (NetApp's RAID6)
- PCIe Accelerators primary storage: 4x 980GB configured as RAID5 and separately as RAID0
- EnhanceIO for Windows Server 2008 R2
  - Caching configuration: single PCIe Accelerator configured as 240GB and 480GB for the SSD cache

---

The most current version of this report is available at http://www.demartek.com/Demartek_STEC_S1120_PCIe_Evaluation_2013-02.html on the Demartek website.