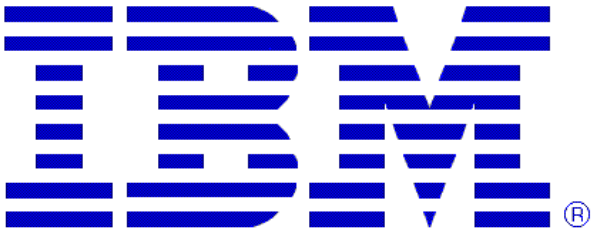

IBM® System Storage™ DS8800™ and DS8700™ Performance with Easy Tier® 3rd Generation

November 2011

Joshua Martin
Nick Clayton
Lee La Frese
Kaisar Hossain
Bruce McNutt
Yan Xu

Document WP102024

Systems and Technology Group

Notices, Disclaimer and Trademarks

Acknowledgements

The authors would like to thank the following colleagues for their comments and insight:

Chip Jarvis – IBM Systems & Technology Group, Tucson, AZ.

Peter Kimmel – IBM Systems and Technology Group, Mainz, Germany.

Allen Marin – IBM Systems & Technology Group, Boulder, CO.

Yang Liu – IBM Systems and Technology Group, Shanghai, China.

Gero Schmidt – IBM Systems and Technology Group, Mainz, Germany.

Brian Sherman – IBM Systems and Technology Group, Ontario, Canada.

A Note to the Reader

This White Paper assumes a familiarity with the general concepts of Enterprise Disk Storage Systems and the DS8000 product line. Readers unfamiliar with these topics should consult the References section at the end of this paper.

Table of Contents

1 Introduction

Over the last eighteen months, Easy Tier has helped numerous DS8700 and DS8800 clients achieve dramatic gains in performance and storage efficiency for their production environments. This paper examines the third generation of Easy Tier which is available with Licensed Internal Code (LIC) Release 6.2. This update further expands Easy Tier to manage hybrid storage pools with up to three tiers including Solid State Drives (SSDs), Enterprise Disks, and Nearline Disks. The Easy Tier management algorithms have been enhanced to optimize capacity utilization of both SSD and Nearline tiers. Auto Rebalance has also been extended to include management of homogenous pools containing only a single tier of storage. This paper will illustrate examples of the performance implications of each of these new features.

Other recently added Easy Tier functions including manual rebalance and automatic management of thinly provisioned storage pools may also enhance performance, but are not covered in this paper.

1.1 Audience

This technical paper was developed to assist IBM Business Partners, field sales representatives, technical specialists, and IBM's clients in understanding the performance characteristics of the IBM 2107 Model 941 and the IBM 2107 951 with the updated Easy Tier feature. The IBM 2107 Model 941 is the DS8700, POWER6 model and shall be referred to as the DS8700 throughout this paper. The IBM 2107 Model 951 is the DS8800, Power6+ model and shall be referred to as the DS8800 throughout this paper. The current update to the Easy Tier function may be referred to as either Easy Tier 3rd Generation or as Easy Tier 3. The previous generation may be referred to as Easy Tier 2, while the initial release of Easy Tier may be referred to as Easy Tier 1.

2 Overview

2.1 New Features

DS8000 LIC Release 6.2 enables Easy Tier management of pools with up to three tiers of storage. The tier definitions remain the same as before. Hard disk drives are separated into two tiers: "Enterprise Tier" which covers 15K and 10K RPM disk drive varieties and "Nearline Tier" (also referred to as SATA Tier) which covers 7.2K RPM disks. SSDs constitute the third tier. At this time, it is not possible to further differentiate the Enterprise Tier into tiers based on disk rotational speed (10K vs. 15K RPM) or RAID configuration (RAID-10, RAID-5 or RAID-6) although the architecture of Easy Tier may permit this distinction in the future. However some level of micro-tiering within the Enterprise Tier is enabled that takes into consideration these types of performance differences and equitably balances the workload across different drives within the tier.

Expanded cold demote is a new Easy Tier feature that will better utilize the bandwidth of drives in the Nearline Tier. This function will look for data that has high bandwidth requirements but very few small I/Os that are residing on the Enterprise Tier. Instead of only using the Enterprise tier, it is better to place some of this high bandwidth data on the Nearline Tier to exploit its bandwidth capabilities and to spread the demand across additional disk spindles.

Easy Tier has always included algorithms to avoid overloading the SSD Tier with too much bandwidth. Although SSDs have far superior performance for small I/Os, the bandwidth capabilities are not substantially better than traditional hard disk drives. If the bandwidth on the SSD Tier is overloaded, performance for small I/Os may be compromised. The prior algorithm for limiting bandwidth was sometimes too conservative and could prevent full exploitation of the SSD tier capacity. The improved algorithm in Easy Tier 3 is more adaptive which allows for improved SSD capacity utilization over time. Another algorithmic enhancement included is a minimum threshold to assure that some amount of data migration for demotion occurs over each hour.

The Auto Rebalance function automatically balances the load on the RAID ranks over time to continuously maintain excellent performance despite variations in workload. Previously Auto Rebalance only applied to storage pools with multiple tiers. With the new release, Auto Rebalance also applies to homogenous pools. Now clients with a single storage tier in their DS8000 may enjoy the automatic management benefits of Easy Tier to maintain balance across their storage resources.

2.2 Disk Magic Support for Easy Tier

One tool available for Easy Tier analysis is Disk Magic™[1], a performance modeling tool that can help customers predict the expected performance of storage subsystems. It has been updated to include support for the new features of Easy Tier 3rd Generation. Using data from existing systems, models can be created to accurately predict the performance of Easy Tier 3rd Generation in DS8800 and DS8700 configurations.

[1] Disk Magic is a registered trademark of IntelliMagic, Inc.

3 Easy Tier Performance Results

3.1 Optimizing SSD Utilization

A DB2 Brokerage Transactional workload was used to evaluate the algorithm improvements for SSD utilization available in Easy Tier 3rd Generation. This lab experiment was designed to simulate a class of applications that facilitate and manage transaction-oriented business processes and are commonly used in a broad range of industry segments including finance, retail, and manufacturing. These applications are characterized by having over 90% read hits in server memory (overall buffer pool hit ratio can be near 97%) and highly random disk I/O. Since it is highly random, this is an ideal workload for evaluating SSD performance with Easy Tier.

The results from these lab experiments demonstrated how customers may benefit from Easy Tier with an SSD tier with similar types of transactional applications. The experiments also showed that the Easy Tier feature greatly simplifies storage management and offers dramatic application performance improvements in real world customer environments without disruption to the application workloads.

Workload Configuration

The experiment was composed of two tests: a baseline test to establish the performance of the system without Easy Tier active and a test with Easy Tier active to show the effects of the feature. Both tests contained a 30 minute ramp up time (RUT) period followed by 24 hours of steady state run time. In both tests the DB2 Brokerage workload was set to run at a peak I/O intensity.

DS8800/DS8700 Configuration

The baseline test was run with 14.4 TB of storage capacity attached to two Device Adapter (DA) pairs. The storage capacity consisted of 128 x 146 GB 15K HDDs configured as 16 RAID-5 ranks. The Easy Tier 3 tests used an additional 16 SSDs on a separate DA pair. Each test was run on both a DS8800 and a DS8700 with workload generated by the attached POWER7 host server (detailed configuration available in Appendix 8.B.1).

To shorten the experiment duration, the Easy Tier short-term decision window was decreased from the default 24 hours to 2 hours and the migration rate was set to the fastest allowed. The workload characteristics were stable over time so the performance outcome would have been the same if the settings had been kept at the default values.

DB2 Brokerage Workload with Peak I/O Intensity

As shown in Figure 1, Easy Tier 3 improved the Overall Transaction Rate (OTR) by almost 3 times when compared to the baseline on the DS8800. This clearly demonstrates the efficiency of the Easy Tier algorithms in moving the most active (hot) extents of this transactional workload to the SSD ranks. The algorithm improvements included in Easy Tier 3 are also displayed in this chart allowing for the modest improvement over the previous Easy Tier generation.

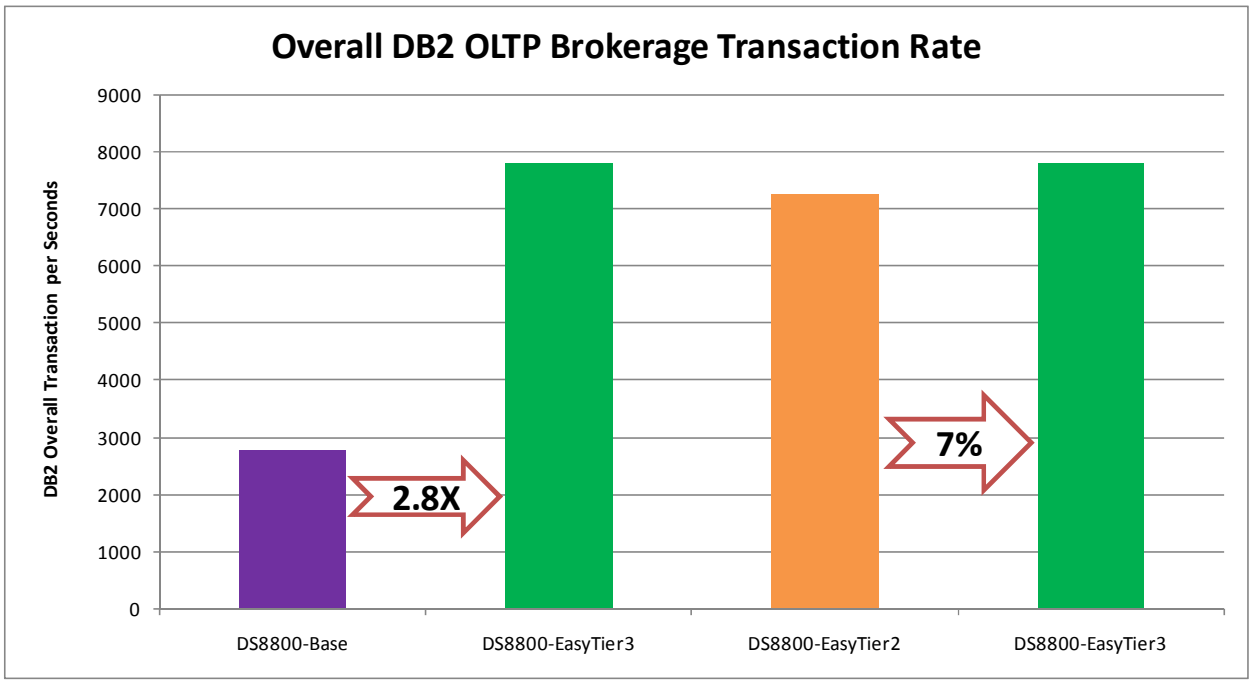

Figure 1. Measurement of Baseline, EasyTier 2, and Easy Tier 3 runs on DS8800

The Overall Transaction Response Time (OTRT) displayed in Table 1 shows an improvement of 52% in the Easy Tier 3 test compared to the Baseline despite the significantly increased throughput shown in Figure 1.

Weighted Average RT (ms)				
Trade Activity	**Lookup**	**Order**	**Update**	**Overall Transaction**
DS8800 Base	9671	318	13289	11205
Easy Tier 3	4349	177	6650	5361
Benefit (%)	55.03%	44.43%	49.96%	52.16%

Table 1. Weighted Average Overall Transaction Response for the DS8800

Figure 2 shows a comparison of the lab results of the DS8700 Baseline run to the runs of each generation of Easy Tier. Overall, the DS8700 saw an improvement of 2.5 times the OTR from the Baseline run to the Easy Tier 3 run. A 75% improvement was observed between Easy Tier 1 and Easy Tier 3 and a 24% improvement from Easy Tier 2 to Easy Tier 3. The DS8700 saw a larger improvement in Easy Tier 3 from the previous generation (24% improvement) compared to the DS8800 (7% improvement) due to the algorithm updates for managing storage bandwidth. The SSDs used in the DS8700 have lower bandwidth capability compared to those on the DS8800, thus there was more room for improvement using the new Easy Tier algorithms.

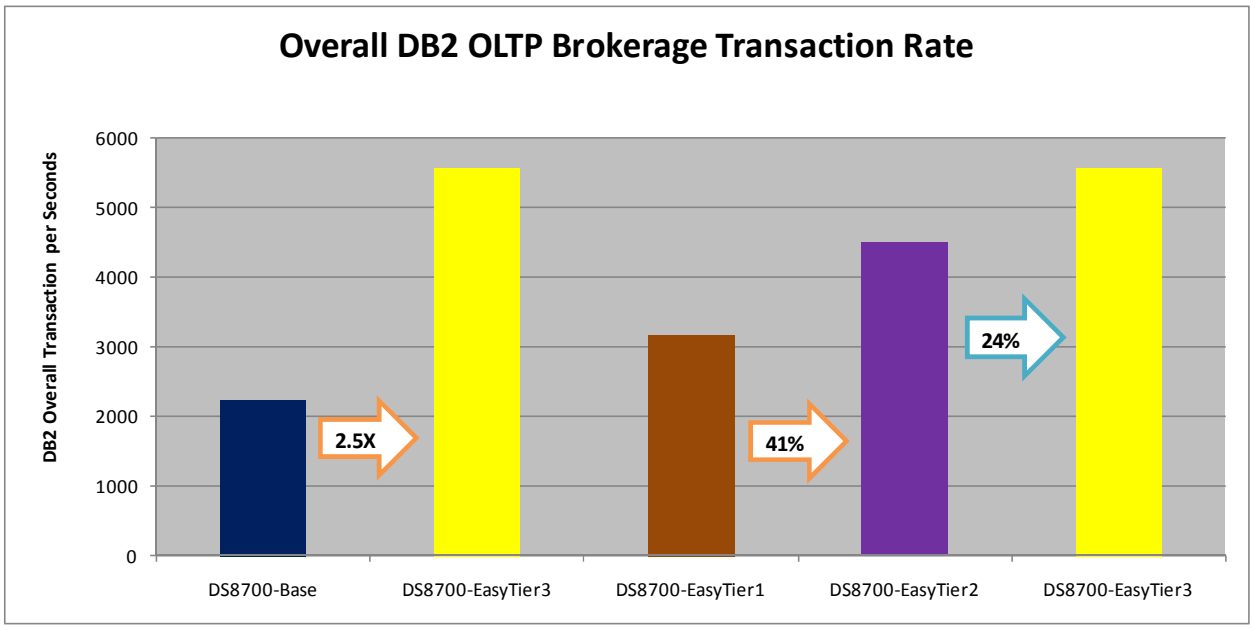

Figure 2. Measurement of Baseline, EasyTier 1, EasyTier 2, and Easy Tier 3 runs on DS8700

Similar to the DS8800 results, an improvement of 49% OTRT with Easy Tier 3 was also seen on the DS8700 as shown in Table 2.

Weighted Average RT (ms)				
Trade Activity	**Lookup**	**Order**	**Update**	**Overall Transaction**
DS8700 Base	12208	359	16746	14212
Easy Tier 3	5850	205	8853	7202
Benefit (%)	52.08%	42.84%	47.13%	49.33%

Table 2. Weighted Average Overall Transaction Response for the DS8700

3.2 Moving to a Multi-tiered Storage Environment

In the past, the introduction of a new, high-capacity or high-performance disk technology typically involved a full replacement of the preceding disk technology. By contrast, Easy Tier 3 provides automated support for a mix of disk technologies, for example a mix of 15K RPM, Nearline, and SSD drives.

Historically, in some installations, a considerable effort has been made to manually manage data placement within a multi-tier storage environment. Such endeavors to match each application to its appropriate drive technology did allow for some storage cost savings and notable performance improvements. However, cost savings were often undercut by the significant effort required to continually manage these environments.

Easy Tier enables a client to adopt a more effective approach to intelligently managing application performance, capacity growth, and storage costs. In a multi-tier storage environment, Easy Tier automatically moves the most active data to the high-performance drives while less active data is migrated to high-capacity drives. Additionally, new high-performance or high-

capacity disk drives can be introduced in an incremental manner and Easy Tier will continue to appropriately manage data placement. This approach yields not merely lower storage costs, but also simpler planning and administration even during growth in the environment.

This section examines performance results obtained using a 50, 30, 20% mixture of OLTP, sequentially accessed, and unused data respectively. Sequential read/write operations simulating a batch workload were applied to the sequentially accessed data giving a total aggregate data rate of 100–500 MB/s. A high level of I/O activity was maintained in the OLTP data, producing a total aggregate I/O throughput level as reported in the following paragraphs. Also, the level of I/O activity varied widely within the OLTP data producing "hot spots". Due to the variations in the test data, the results are representative of a typical production environment which would also be expected to exhibit similar "hot spots".

Within this framework, we started with a configuration of 96 high performance drives (300 GB, 15K RPM). To that configuration, 16 high capacity drives (3 TB, 7200 RPM) were added. The system was then permitted to run under Easy Tier 3 management for several days. The expected result of this experiment is that Easy Tier 3 should free up space on the high-performance drives by migrating data with low activity to the lower storage tier. We should therefore expect the performance of the system to be similar before and after this test, even though we have now introduced a high capacity storage tier of lower RPM drives.

Figure 3 presents the results obtained from the test just described. The dashed green line shows the batch I/O activity added to the system corresponding to the OLTP workload throughput. The dark blue and pink curves respectively show the performance of the system before and after the introduction of the Nearline drives. This confirms our expectation that the Nearline drives can be introduced while also maintaining the previous level of performance.

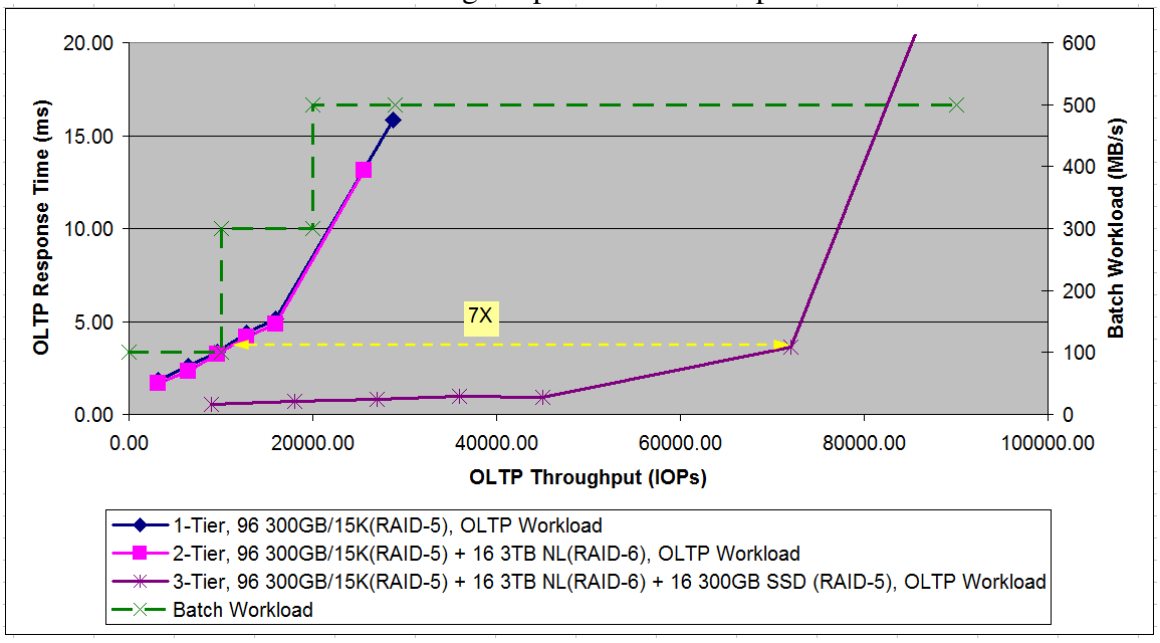

Figure 3. Easy Tier makes it easy to combine capacity growth with higher performance.

Figure 3 also illustrates the incremental adoption of SSDs into the same environment. The purpose of SSDs is to increase the overall throughput of the system. As shown by the purple curve, the addition of 16 SSDs results in a throughput improvement of 7 times and a significant reduction in average I/O response times.

Figure 4 and Figure 5 provide additional details about data distribution in the multi-tier environment. Figure 4 was obtained using the Storage Tier Advisor Tool (STAT). It presents the initial conditions with all data located on 15K drives. The Heat Distribution column presents the distribution of activity within a given volume. The red bars represent capacity of hot (highly active) data, the orange bars represent the capacity of warm (moderately active) data, and the blue bars represent the capacity of cold (slightly or not active) data. Figure 5 shows data allocation across tiers when the configuration had changed from a single-tier to a 2- and then 3-tier configuration. All of the "cold" data with no I/O operations moved to the lower tier; part of the data with sequential I/O operations also moved to the lower tier as part of expanded cold demote, while "hot" data (high I/O activity) moved to the upper tier.

Volume Heat Distribution

Volume ID *5	Configured Size *6	Tier	Capacity on Tier *7	Heat Distribution *8
0x4300	900G	Enterprise Tier	900G	72G 774G 54G
		NL Tier	0G	
0x4301	900G	Enterprise Tier	900G	72G 774G 54G
		NL Tier	0G	
0x4302	900G	Enterprise Tier	900G	80G 766G 54G
		NL Tier	0G	
0x4303	900G	Enterprise Tier	900G	78G 768G 54G
		NL Tier	0G	
0x4304	900G	Enterprise Tier	900G	79G 767G 54G
		NL Tier	0G	
0x4305	900G	Enterprise Tier	900G	888G 12G
		NL Tier	0G	
0x4306	900G	Enterprise Tier	900G	888G 12G
		NL Tier	0G	
0x4307	900G	Enterprise Tier	900G	888G 12G
		NL Tier	0G	
0x4308	900G	Enterprise Tier	900G	900G
		NL Tier	0G	
0x4309	900G	Enterprise Tier	900G	900G
		NL Tier	0G	

Figure 4. Initial data layout and heat distribution on 15K RPM Tier

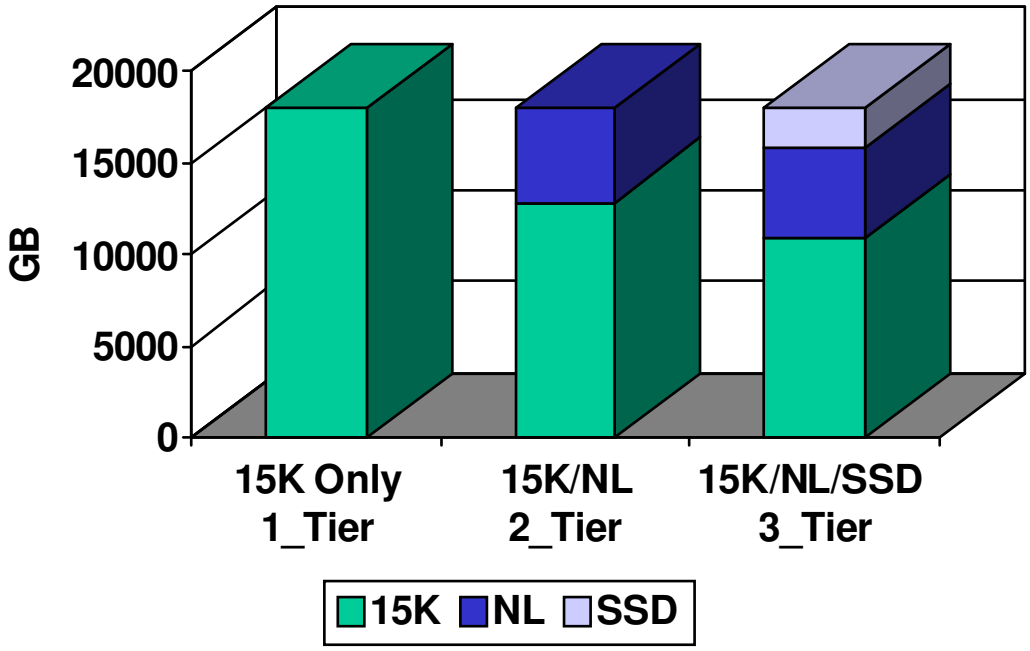

Figure 5. Capacity Allocation Across Tiers

For this particular experiment, some internal Easy Tier settings were changed in order to shorten the test cycle. For example, the short-term decision window was reduced from the default 24 hours to 2 hours, the data migration rate was changed from the default 12 extents every 5 minutes to the maximum allowed by the system, and the guaranteed cold demote rate was changed from the default of 1 extent to 6 extents every 10 minutes. The workload characteristics were stable over time so the performance outcome would have been the same if the settings had been kept at the default values.

The above results suggest that high capacity drives can be used to hold "cold" data with no or low I/O throughputs without sacrificing the application performance. Easy Tier 3 can automatically migrate data from enterprise drives (15K/10K RPM) to Nearline drives and free up space on the enterprise drives. The amount of cold data that can be migrated to Nearline drives depends on the application. With some customers, Easy Tier was able to move over 70% of the "cold' data from the enterprise drives to the Nearline drives. With more available space in enterprise drives, a user can expand application data or add new applications to the system. By default, Easy Tier 3 will initially place added data onto the enterprise drives. In an environment with multiple tiers of drive technology, this assures that the new data will be written to the enterprise tier first instead of a slower drive tier which may have more available capacity. With 3-tier support in Easy Tier 3, customers can address capacity growth and achieve a performance boost by simply adding Nearline and SSD drives to the system. Easy Tier can optimize the use of all three drive tiers in the system to ensure the right balance of performance and cost objectives.

3.3 Managing Data Placement for Performance Within a Single Tier

Included in the release of Easy Tier 3 is an update to the Auto Rebalance function to allow it to function within homogenous (single-tier) pools. With this added functionality in addition to the

three-tier support in Easy Tier, the user is now provided with a complete option to fully automate data placement balancing within a DS8000 storage system.

Performance experiments were run on a single homogenous pool consisting of 80 x 146 GB 15K RPM disk drives configured as 10 RAID-5 ranks. The OLTP workload was arranged to operate with a stable skewed distribution over the ranks. The run started in the skewed distribution state and then the Auto Rebalance function was enabled and allowed to reach a new stable state. Since the experiment was performed in a lab environment some Easy Tier default settings were changed to reduce the duration of the experiment: the short term Easy Tier decision window was reduced from the default of 24 hours to 1 hour, the Auto Rebalance decision window was reduced from the default of 6 hours to 0.5 hours, and the migration rate was set to run at the fastest rate possible. The workload characteristics were stable over time so the performance outcome would have been the same if the settings had been kept at the default values.

Figure 6 displays the utilization of each individual rank observed in the configuration over the duration of the test. The workload was stable at the start of the test, but clearly skewed with the majority of the workload on three of the ten ranks. After letting the workload run at that stable, skewed rate, the Auto Rebalance function was enabled and data redistribution began over the available ranks. The utilization of the ranks that started with the bulk of the workload decreased as data was migrated to the less utilized ranks, which then increased in utilization. Throughout the test, the host I/O rate remained constant.

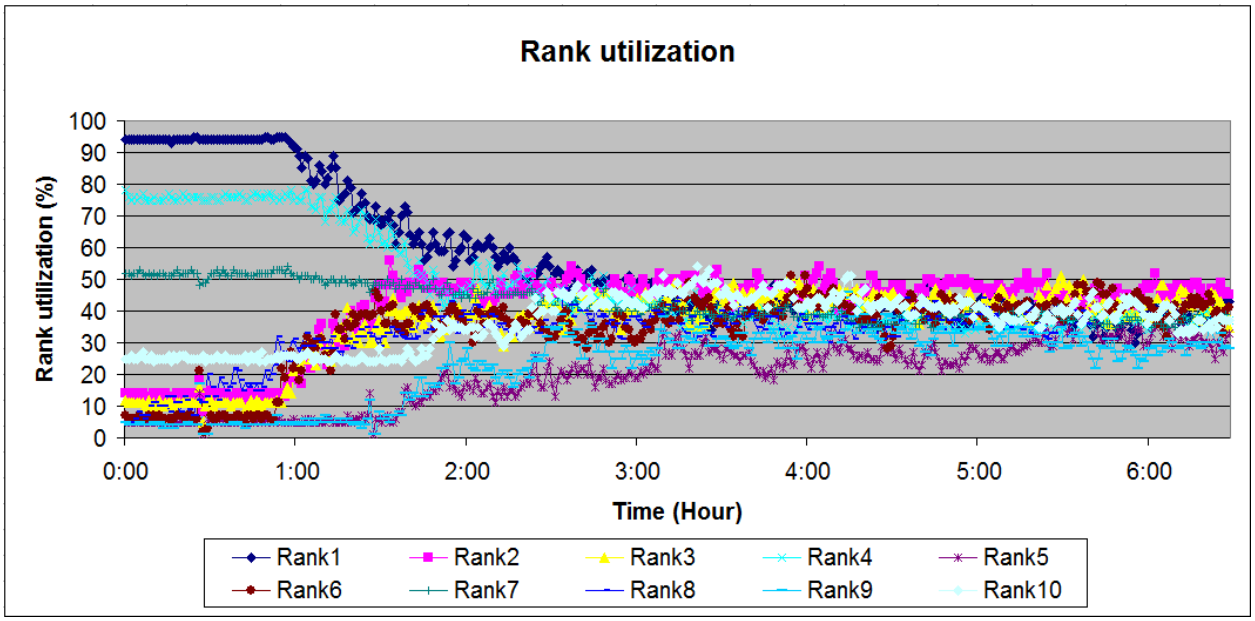

Figure 6. Effect of Auto Rebalance on individual rank utilizations in the system

The capability of the system was also measured both before Auto Rebalance was enabled and then after when the system had reached a stable state. The results of those measurements are shown in Figure 7. As expected, the balanced system was capable of a significantly greater throughput and equal or better response times.

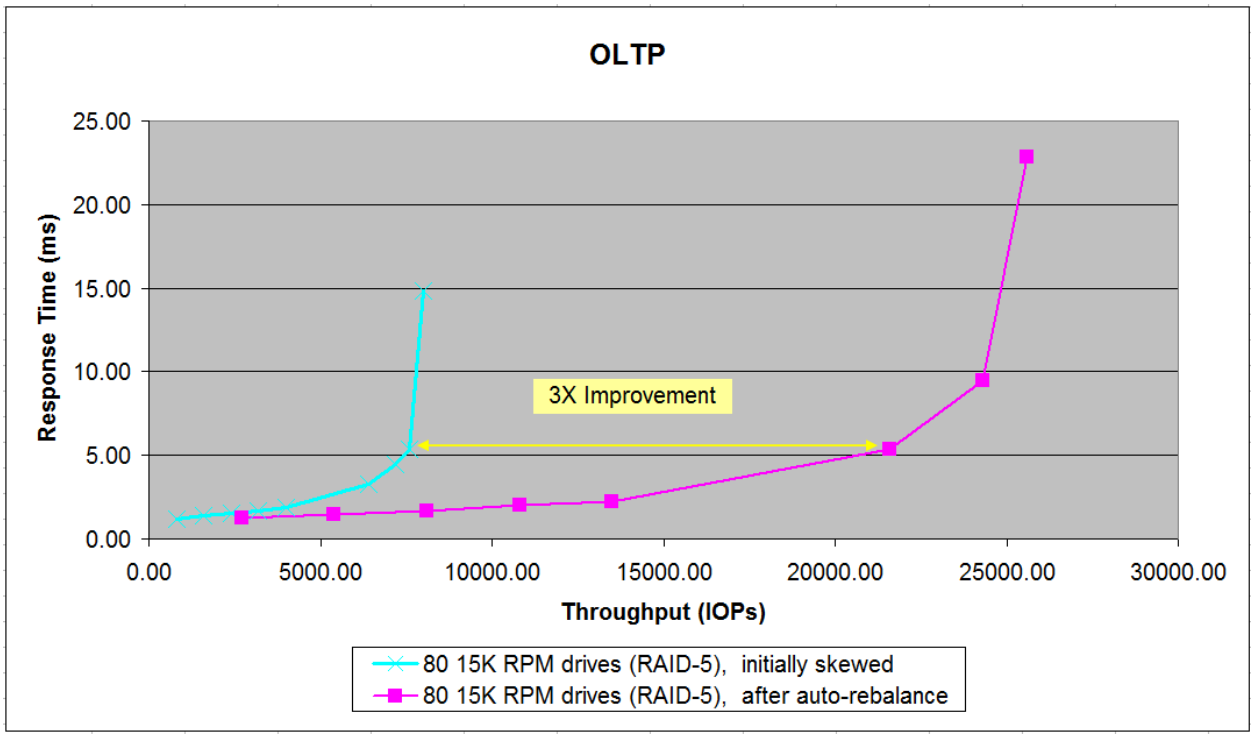

Figure 7. Throughput capability of the system before and after Auto Rebalance

This experiment demonstrated the capability of Auto Rebalance within an extent pool composed of only 15K RPM disk drives. However, the Easy Tier algorithms (including Auto Rebalance) are aware of the performance capabilities of various drive types. If the extent pool had included both 10K and 15K RPM drives, the algorithms would be able to balance the I/O operations across the ranks according to the capability of each rank.

4 Drive Selection with Easy Tier

While the operation of Easy Tier is extremely simple, the number of drive combinations possible in an Easy Tier environment provides a much greater degree of choice than when considering a single tier of storage.

The key questions we might ask for an environment are

1) What size drives would I need to satisfy the workload if striping the date over a single tier of enterprise drives? This provides a good view as to how active or inactive the overall workload is.
2) How much of my data is doing very little random I/O (but perhaps some sequential I/O)? This indicates what data would be suitable for high capacity Nearline drives.
3) What is the skew of my data and how much SSD capacity would be needed to handle 50% (or even 80% or 99%) of the random workload?
4) Given a use of Nearline or SSDs what size of Enterprise drives should be used to handle the workload remaining on the Enterprise tier? In some cases it may even be that enterprise drives are not required and a combination of SSD and Nearline drives is appropriate but given current SSD prices and typical workload skew this is probably not typical.
5) Is there a need to provide significantly better backend performance in the target environment to accelerate application workloads and is the workload type (cache hit ratios / percentage sequential workload etc) such that SSD storage will provide significant application acceleration?

For environments already running on an Easy Tier-capable disk subsystem, we can answer these questions with a good degree of accuracy as we can use the monitoring capability of Easy Tier to understand the environment in detail.

In the case of workloads on other disk subsystems, we would need to compare the overall workload performance data of the environment to previously collected real world performance data. Through that comparison, an estimate could be generated of how the workload might then be distributed in an Easy Tier configuration.

In the sections below we will discuss each of these five points in turn.

4.1 Drive selection using a single tier

Sizing disk subsystems in a single-tier environment is a well understood process; where, given a particular capacity and workload requirement, tools such as Disk Magic are used to determine the optimal drive size and other configuration factors to meet both of these requirements. Generally, more disk drives with smaller capacities are needed to contain high workload requirements, while fewer high capacity disk drives may be sufficient for lower workload requirements. However, we need to remember that the DS8000 is able to perform wide striping of data using Storage Pool Striping, which can significantly reduce disk hotspots and allow for larger drives than might have been used in a traditional disk subsystem with volumes assigned to limited backend resources. For example, analysis in one customer environment showed that

restriping data with the Storage Pool Striping feature would have allowed for equal or better performance with twice the drive size (half the number of drives) as was being used in a traditionally configured environment.

We might for example have a workload where 300 GB drives would be required to handle the workload. This would be a relatively active workload in a typical commercial environment but there are, of course, workloads that require smaller drives and other workloads where larger drives are entirely appropriate.

4.2 Percentage of cold data

Figure 8 shows skew for a range of production environments showing that the random workload is concentrated on the top 25% or less of the extents and there is a significant percentage of the capacity that also does very little sequential activity.

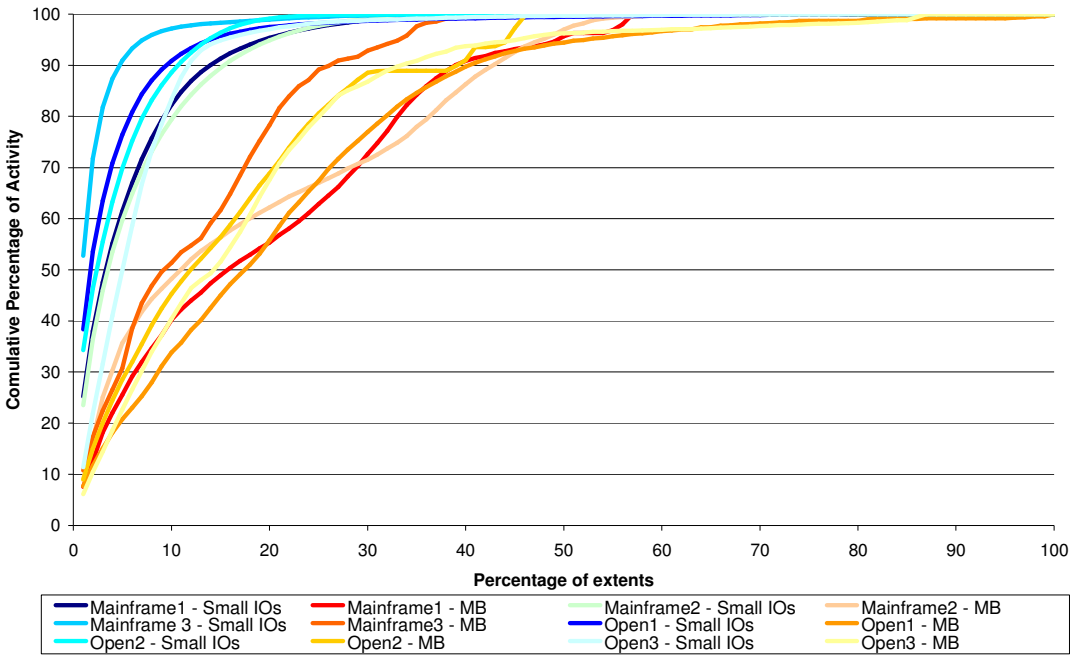

Figure 8. Skew for several workloads

Depending on the workload mix, it is possible in many environments to use between 50% and 75% Nearline capacity while the vast majority of the small, random workload is concentrated on the enterprise drives. When configuring this type of environment the drive size for the enterprise drives should be reduced commensurate to the concentration of workload on these drives. For example we might reduce the enterprise drive capacities by half to give smaller capacity drives when using 50% Nearline drives. Table 3 shows how the percentage of Nearline drives can change the required drive size when moving to a 2-tier configuration from a proven starting point single-tier configuration.

Single-tier drive size driven by performance requirements	2-tier with ~25% Nearline	2-tier with ~50% Nearline	2-tier with ~75% Nearline
900 GB	600 GB	450 GB	300 GB
600 GB	450 GB	300 GB	146 GB
450 GB	300 GB	300 or 146 GB	146 GB
300 GB	300 or 146 GB	146 GB	Too much activity on enterprise drives for 2-tier with this percentage of Nearline
146 GB	Too much activity on enterprise drives for 2-tier with Nearline		

Table 3. Effect of different percentage of Nearline drives in a 2-tier configuration

Single-tier environments which meet required performance with larger enterprise drives are likely to realize significant cost benefits from a 2-tier configuration of somewhat smaller enterprise drives and a significant percentage of Nearline drives. However, if the performance requirements of a workload in the single-tier environment would already require small enterprise drives, we need to consider the use of SSDs as there is no benefit to add Nearline drives if we are not able to reduce the capacity of the Enterprise drives.

4.3 Skew and percentage of hot data

In typical environments somewhere between 2–5% of SSD can result in the movement of 50% of the small, random I/O workload from enterprise drives to SSDs as shown in Figure 9. This would potentially allow for a doubling of the enterprise drive size compared to a single-tier configuration to provide the same capacity and equivalent or better performance. In some environments this can be significantly more and we can see up to 80% of the small, random I/O moved to SSDs for a very small percentage of SSD capacity.

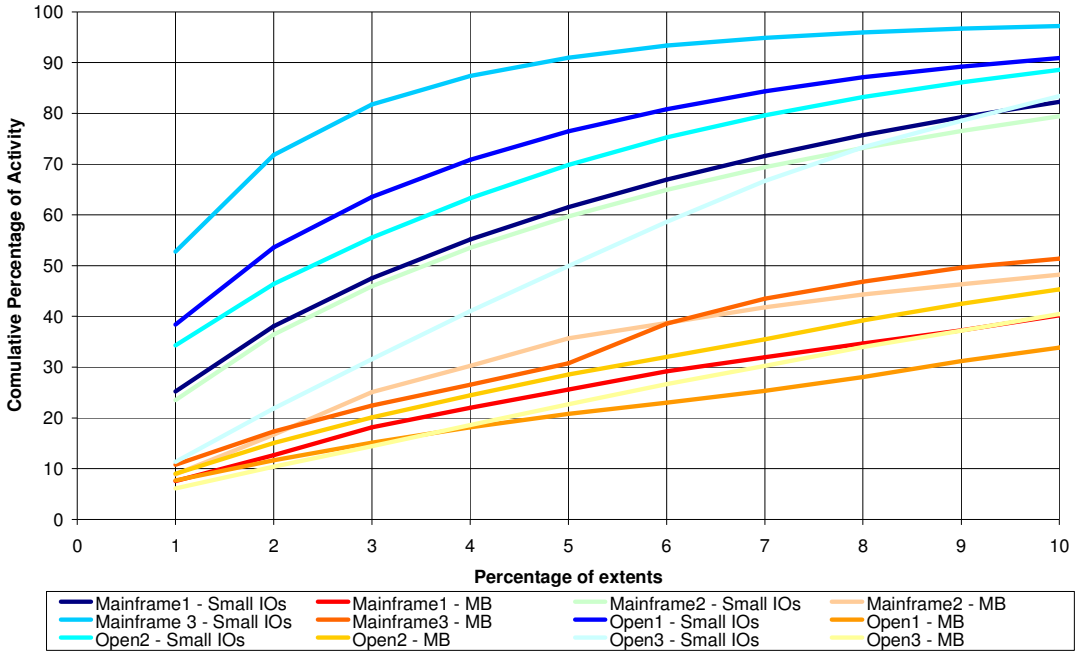

Figure 9. Hot data skew for several workloads

Table 4 gives some examples of the effect of adding SSDs to a configuration. It assumes that 2–5% of SSDs will result in 50% of the small, random workload on SSDs and 5–10% of the SSDs will result in 75–80% of the small, random workload on SSDs which is a reasonable assumption for the skewed workloads shown in Figure 9.

Single-tier drive size driven by performance requirements	2-tier with 2–5% SSDs	2-tier with 5–10% SSDs
900 GB	Perhaps Nearline	Nearline
600 GB	900 GB	Nearline
450 GB	900 GB	Perhaps Nearline
300 GB	600 GB	900 GB
146 GB	300 GB	600 GB

Table 4. Effect of different percentage of SSDs in a 2-tier configuration

If the environment is one where there is a significant amount of very small granularity striping, such as Oracle or DB2 tablespace striping, then the skew of the workload may be significantly reduced. This is also true of at least some System i environments. In these cases there may be less benefit from smaller amounts of SSDs and it may be more appropriate to size SSD capacity to meet application requirements.

Figure 10 shows the skew for a System i environment. It is clear to see that System i is very effective in spreading the workload over the extents although there is a small degree of skew for the small, random write activity.

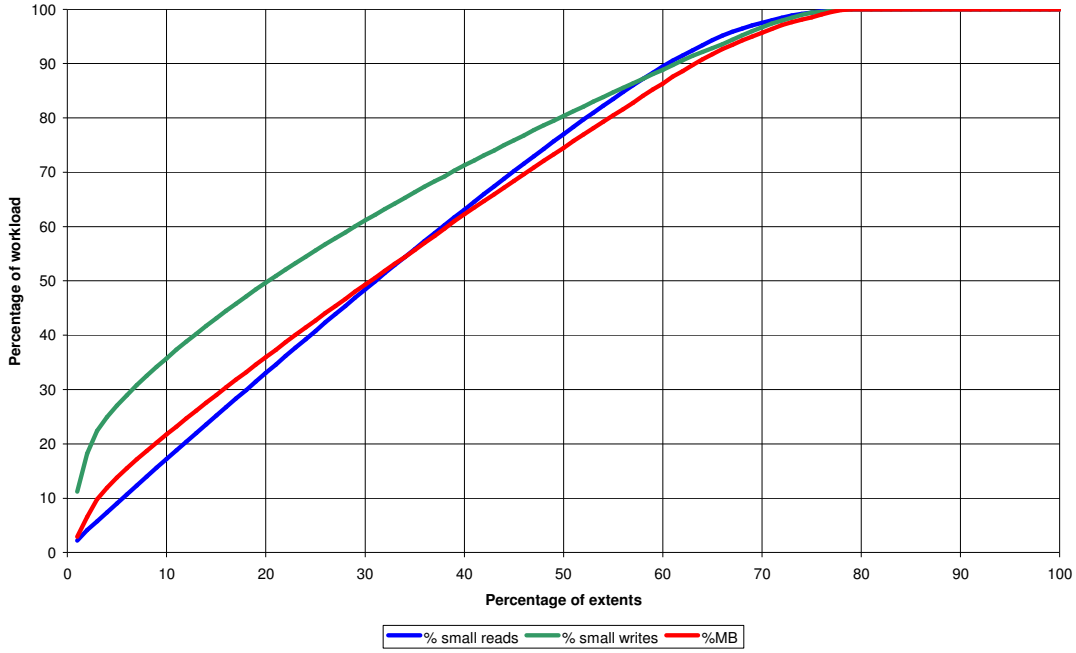

Figure 10. Skew for System i workload

4.4 Combining for three tiers

With Easy Tier 3rd Generation, we can now combine both Nearline and SSDs with Enterprise drives to create a 3-tier configuration. The effects of Nearline drives decreasing the required size of enterprise drives and the effect of SSDs increasing the required size of enterprise drives are combined.

For example if we had a workload that could be adequately supported by a single tier of 300 GB drives, then adding 50% Nearline capacity would mean that 146 GB drives would be the appropriate enterprise drive size in a 2-tier configuration. Further, adding 2–5% SSDs to build a 3-tier configuration would mean we could again use the 300 GB enterprise drives.

4.5 Further Application acceleration

We have seen situations where a client wanted SSD performance for an entire environment rather than simply to optimize the price, power, and footprint of their environment. We have seen that it is possible to move almost the entire active workload for an environment onto SSDs with a relatively small percentage of SSDs. One particular client is running with Easy Tier and with 20% of their database environment on SSDs and has equivalent performance to that of a 100% SSD solution.

Looking at the skew information from a range of clients we can see that adding more SSDs to a configuration can very quickly result in the majority of the small, random workload being moved to these drives. However, there are diminishing returns so unless there is an application and business benefit for significantly improved read miss performance the higher percentages of SSDs may not provide the best return on investment.

5 Conclusions

Easy Tier 3rd Generation offers an intelligent, automated solution for managing data placement to help optimize performance and cost requirements in single-tier or multi-tier configurations. The addition of support for three tiers allows Easy Tier to manage both growth in capacity with the addition of Nearline drives and growth in throughput with SSDs. The addition of homogenous (single-tier) pool support in the Auto Rebalance function makes it possible for any pool to be balanced and remain balanced regardless of whether there are multiple tiers in the system or not.

In this paper we used a DB2 Brokerage Transactional Workload to demonstrate the algorithm improvements available in Easy Tier 3rd Generation over what was available in the 2nd Generation. Then, with an OLTP workload environment, we demonstrated the combined benefit of added capacity with the Nearline tier and increased throughput performance with the SSD tier, which was made possible with three tiers in Easy Tier 3. In our last experiment, we showed the efficiency of the Auto Rebalance function within a single tier and pool. Finally, in Section 4 we gave guidance toward building a storage configuration with Easy Tier including suggestions for how to best plan which tiers to use and how to size those tiers.

6 References

[1] Martin, J., Clayton, N., La Frese, L., Hossain, K., McNutt, B., Xu, Y. "IBM® System Storage™ DS8700™ and DS8800™ Performance with Easy Tier® 2nd Generation." June 2011.

[2] La Frese, L., Hossain, K., Hyde, J., Lin, A. W., McNutt, B., Sansone, C., Xu, Y., Zhang, Y. "IBM® System Storage™ DS8700™ Performance with Easy Tier®." May 2010.

[3] La Frese, L., Sutton, L., and Whitworth, D. "IBM® System Storage™ DS8000® with SSDs: An In-Depth Look at SSD Performance in the DS8000." April 2009.

[4] Lin, A. W., Whitworth, D., Williams, S. E., Xu, Y. "IBM® System Storage™ DS8800® Performance Whitepaper." December 2010.

[5] La Frese, L., Lin, A. W., Martin, J., Williams, S., Xu, Y. "IBM® System Storage™ DS8700® Performance Whitepaper." August 2010.

[6] Ripberger, R. and Xu, Y. "IBM System Storage, DS8000 Storage Virtualization Overview, Including Storage Pool Striping, Thin Provisioning, Easy Tier", WP101550 V2.0, May 2010.

[7] Altman, J., Sutton, L., and Sutton, P. z/OS Hot Topics, article 20-48: z/OS Support for Solid State Drives in the DS8000. February 2009.

7 Appendix

8.A Appendix A: Workload Characteristics

- *DB Open*: 70% reads, 30% writes, 50% read hits. This workload is designed to be comparable to typical applications with transactional workloads. Read/Write Ratio = 2.33, Read Hit Ratio = 0.50, Destage Rate = 17.2%, Transfer size = 4 KB.
- *OLTP workload*: simulates the workload of transaction processing systems that require small, mostly random, read and write operations (for example, database systems, OLTP systems, and mail servers). It resembles the mix of I/O workload components as defined in the SPC-1 specification.

8.B Appendix B: DS8000 Hardware Configurations

8.B.1 Configuration for DB2 Brokerage Measurements

- DS8700 Configuration
  - 941 4-Way 128 GB Cache
  - 128 x 146 GB / 15K RPM drives configured as RAID-5
  - For experiments with SSDs: 16 x 600 GB SSDs configured as RAID-5. (16 x 146 GB SSDs were used in Easy Tier1)
  - 3 DA Pairs with SSDs on a separate DA pair
  - 4 x 4 Gb HA ports
- DS8800 Configuration
  - 951 4-way 128 GB Cache
  - 128 x 146 GB / 15K RPM drives configured as RAID-5
  - For experiments with SSDs: 16 x 300 GB SSDs configured as RAID-5
  - 3 DA Pairs with SSDs on a separate DA pair
  - 4 x 8 Gb HA ports
- DB2 Configuration
  - DB2 9.7 FP1, 4 Instances, 4 x 2 TB DBs, 4 Buffer Pools at 54 GB each
    - 8 x 1.6 TB volumes were allocated for database, Temp files and Data Generation
    - 8 x 50 GB volumes were allocated for log files
- Server Configuration
  - P770 (AIX 6.1.5.0), 8 Eight Core P7 (3GHz)
  - 512 GB Cache
  - 4 x 8 Gb FC Ports

8.B.2 Configuration for OLTP Measurements with SSD, 15K, and 7.2K RPM Drives

- DS8800 Configuration
  - 951 4-Way 256 GB Cache/8 GB NVS (reduced logically to 64 GB Cache/2 GB NVS)
  - 96 x 300 GB / 15K RPM drives configured as RAID-5, 16 3 TB / 7.2K RPM Nearline drives with RAID-6, 16 x 300 GB SSDs with RAID-5
  - 4 DA Pairs
  - 8 x 8 Gb HA ports
- Server Configuration
  - P770 (AIX 7.1.0.15), 8 Eight Core P7 (3GHz)
  - 256 GB Cache
  - 8 x 8 Gb FC Ports