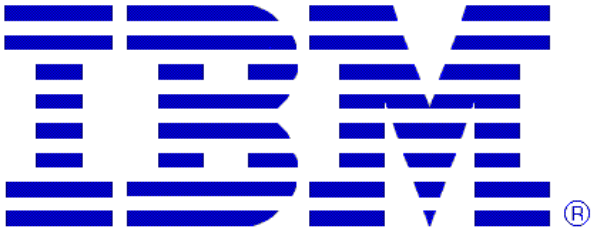

Planning for Easy Tier® with IBM® System Storage™ Storwize V7000 and SAN Volume Controller

May 2013

Nick Clayton
Carlos Fuente

Document WP102295

Systems and Technology Group

Notices, Disclaimer and Trademarks

Acknowledgements

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

- Phil Bryan
- Lawrence Chiu
- Chip Jarvis
- Bruce McNuttt
- Brian Sherman
- Yan Xu

We would also like to thank Fiona Pudge for her work providing the V7000 configuration and pricing data.

A Note to the Reader

This White Paper assumes a familiarity with the general concepts of Enterprise Disk Storage Systems ,the Storwize V7000 product line and the SAN Volume Controller. Readers unfamiliar with these topics should consult the References section at the end of this paper.

Table of Contents

Table of Figures

1 Introduction

Over the last two years, Easy Tier has helped numerous of IBM's clients using SAN Volume Controller and IBM® Storwize® V7000 to achieve dramatic gains in performance and storage efficiency for their production environments. This paper aims to provide an overview of the Easy Tier functionality and show how typical configurations can help improve performance and Total Cost of Ownership for real world production workloads.

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 Storwize V7000 and SVC when used in an Easy Tier environment.

2 Overview

IBM® Storwize® V7000 and Storwize V7000 Unified are virtualized storage systems designed to consolidate workloads into a single storage system for simplicity of management, reduced cost, highly scalable capacity, performance and high availability.

They offer improved efficiency and flexibility through built-in solid state drive (SSD) optimization, thin provisioning, Real-Time compression and non-disruptive migration of data from existing storage. They can also virtualize and reuse existing disk systems offering a greater potential return on investment.

2.1 Solid State Storage in Storwize V7000 and SVC

SAN Volume Controller and Storwize V7000 come with a broad range of options for integrating solid state storage into a storage environment.

Storwize V7000 can be ordered with industry standard SSDs with configurations starting as small as two 200GB drives, growing to many TB's of SSD RAID storage using 200GB or 400GB drive options. Data can be protected with any of the supported RAID levels, but it is likely that RAID-5 will be the best choice for most environments as is generally the case with traditional HDDs.

IBM has submitted SPC-1 benchmarks for the Storwize V7000 both with and without SSDs and a link to these results can be found in the references section. The response time and throughput information from these benchmarks is shown in Figure 1 below. This shows both the significant response time improvements possible when using SSDs and the ability of the Storwize V7000 to drive significantly more IO than is possible purely using traditional HDDs.

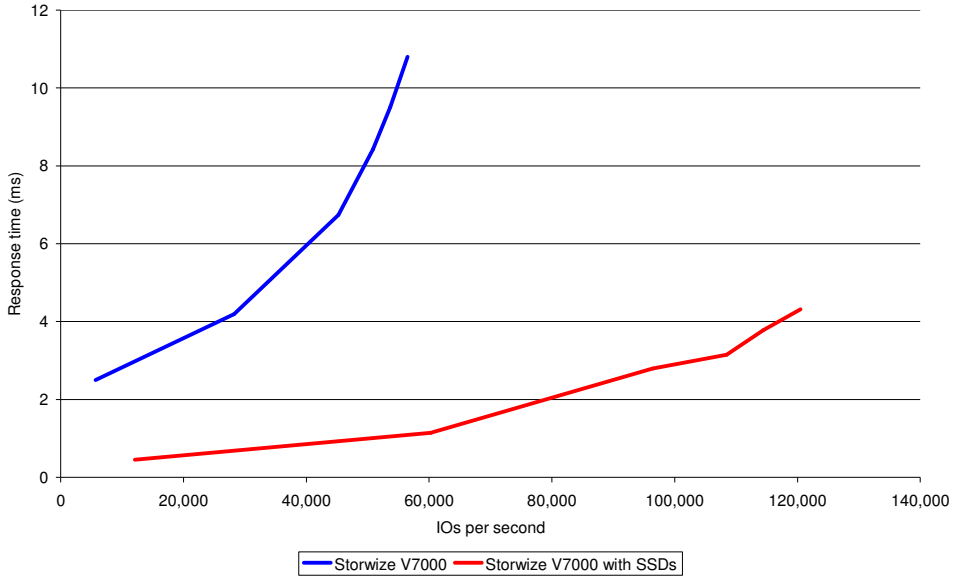

Figure 1. Storwize V7000 SPC-1 benchmark response time curves

SVC supports up to four 200GB or 400GB SSDs in the latest CG8 node, which can be protected using RAID-10 and provides a very convenient way of adding small amounts of solid state storage to a large environment. Up to 1.6TB of usable RAID-protected capacity can be added to an I/O group or up to 6.4TB for a clustered system.

Both SVC and Storwize V7000 can integrate storage from external storage systems into storage pools. This includes the family of IBM FlashSystem products developed by Texas Memory Systems.

Solid State Storage can also be integrated into the external storage systems, such as the IBM® System Storage DS8870™. This storage can be presented as Tier 0 storage and integrated into the storage pools.

Any of these forms of storage, from IBM or other vendors, can benefit from the advanced virtualization capabilities of SVC and Storwize V7000 including features like Easy Tier. All this is accessible from the easy to use management UI, or from the CLI..

2.2 Easy Tier

Storwize V7000 and SVC both support Easy Tier, IBM's industry leading technology for optimizing use of solid state storage. This technology was first released for the DS8000, and has since been integrated into Storwize V7000 following its release in 2010.

Easy Tier provides a very easy to configure means of integrating solid state storage into existing data center environments and optimizing the use of that storage, providing maximum return of investment.

Many applications exhibit a very significant skew in the distribution of I/O workload: a small fraction of the storage is responsible for a disproportionately large fraction of the total I/O workload of an environment. This paper will present data from real environments that illustrate this distribution.

Easy Tier acts to identify this skew and to automatically place data to take advantage of it. By moving the 'hottest' data onto solid state storage, the workload on the remainder of the storage is significantly reduced.

By servicing the majority of the application workload from the fastest storage, Easy Tier acts to accelerate application performance, and increase overall server utilization. This can reduce costs in servers and application licenses.

Easy Tier also presents a wide range of opportunities for reducing the cost of the remainder of the storage, without losing the performance advantage. This reduced cost can come from both lower acquisition cost, by using devices that cost less $/GB, and also from savings in terms of power and rack space that contribute significantly to the hidden costs in a data center. This paper will present sample calculations to demonstrate this.

2.3 Disk Magic

One tool available for Easy Tier analysis is Disk Magic™[1], a performance modeling tool used by IBM that can help predict the expected performance of storage subsystems. It has been enhanced to include support for the new features of Easy Tier. Using data from existing systems, models can be created to predict the performance of Easy Tier in Storwize V7000 and SVC configurations.

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

3 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 using a single tier of enterprise drives and striping the data over the available drives? This provides a good view as to how active or inactive the overall workload is.
2) What is the skew of my data and how much SSD capacity would be needed to handle say 50%, (or even 80% or 99%) of the random workload?
3) Given particular capacity of SSD storage, 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.
4) Is there an opportunity 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 case the workloads are on other disk subsystems we would need to use a combination of overall workload performance data and to compare the environment against the range of workloads for which we have collected real world performance data to provide an estimate of how the workload skew might be distributed.

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

3.1 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.

The key metric that Disk Magic uses to determine the optimal drive size is the ‘access density’ of an application. This compares the peak IO workload that the application generations to the total amount of data the application needs to access.

We need to remember that the SVC and Storwize V7000 are able to perform wide striping of data using striped mode VDisks, which can significantly improve balance of workload and allow for larger drives than might have been used in a traditional disk subsystem with volumes assigned to a subset of the available backend resources.

For example, analysis in customer environments have shown that implementing wide striping of data using SVC would have allowed for equal or better performance with twice the drive size as was being used in a traditionally configured environment.

The table below shows some example access densities (IOPS/GB) that can be supported for a particular workload on different drive types based on Disk Magic modelling. We have chosen a 16kB average IO size with 80% read, 15% sequential IO and a 50% cache hit ratio as a reasonably representative workload.

Drive Type	Access density (IOPS/GB) – 16kB workload
200GB SSD	60
400GB SSD	30
146GB 15K	2.00
300GB 15K	1.00
450GB 10K	0.48
600GB 10K	0.36
900GB 10K	0.24
1TB 7.2K	0.14
2TB 7.2K	0.07
3TB 7.2K	0.03

Figure 2. Example maximum access densities by drive type

Note : RAID6 is assumed for the Nearline drives and RAID5 for Enterprise SAS drives and SSD.

For example, if we had a workload doing 50,000 IOPS on 60TB of data then this would be doing 50,000/60,000 = 0.83 IOPS/GB and so we might choose a single tier configuration using 300GB 15K drives. This would be a relatively active workload in typical mixed workload commercial environment.

The main point here is the comparison of the capability of the different drive types, which span a wide range and show a dramatically higher access density for SSD compared to other tiers. Real application workloads are a complex mix of IO sizes but the 16KB IO size is a representative average and much more realistic than some other benchmarks such as 512byte 100% read workloads.

3.2 Skew and percentage of hot data

When implementing Easy Tier one of the key characteristics of the workload is the skew of the random and small sized IO activity in the environment. By this we mean how evenly or unevenly the IO activity is spread over the capacity. The skew curve of a number of real workloads (Figure 3) shows that in most environments the random IO activity is concentrated on a very small percentage of the capacity and a significant percentage of capacity has very little random IO workload. The skew is in the range of 5% of allocated capacity receives 50% (SVC #4) to 87% (SVC#2, V7000#1) of small IO accesses.

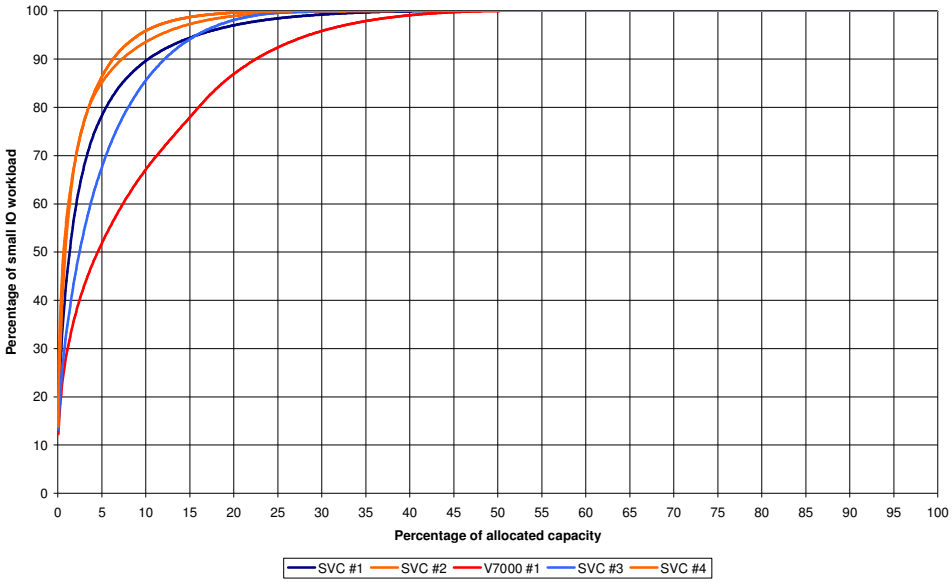

Figure 3. Skew for several workloads

In typical environments somewhere between 2–5% of SSD can result in the movement of 50% of the small/random IO workload from enterprise drives to SSDs as shown in the zoomed graph in Figure 4 which shows only the hottest 20% of the workload. 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 IO moved to SSDs for 5% or even less of SSD capacity.

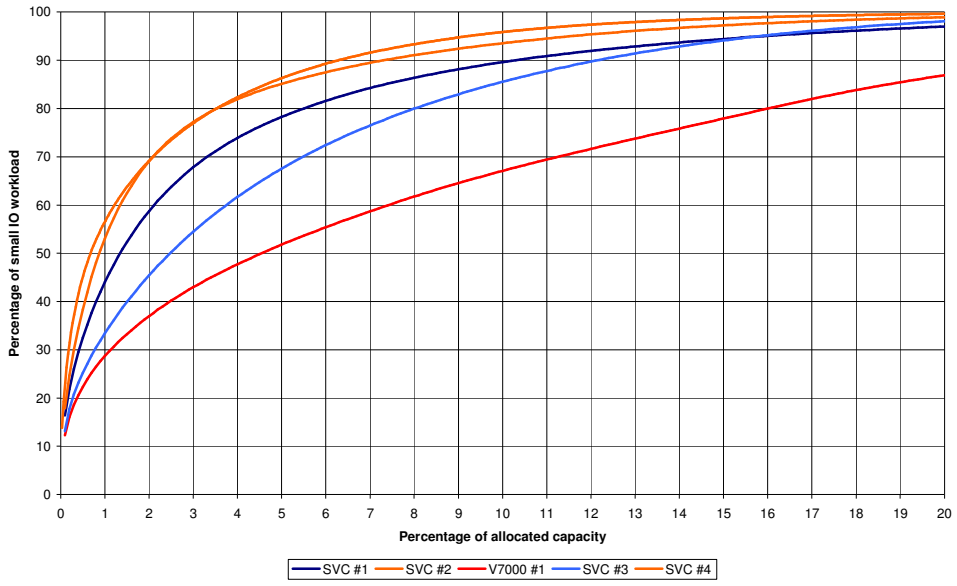

Figure 4. Skew for several workloads showing hottest 20% of data

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 not be economical to implement an Easy Tier solution.

Figure 5 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 some noticeable skew for the small random write activity.

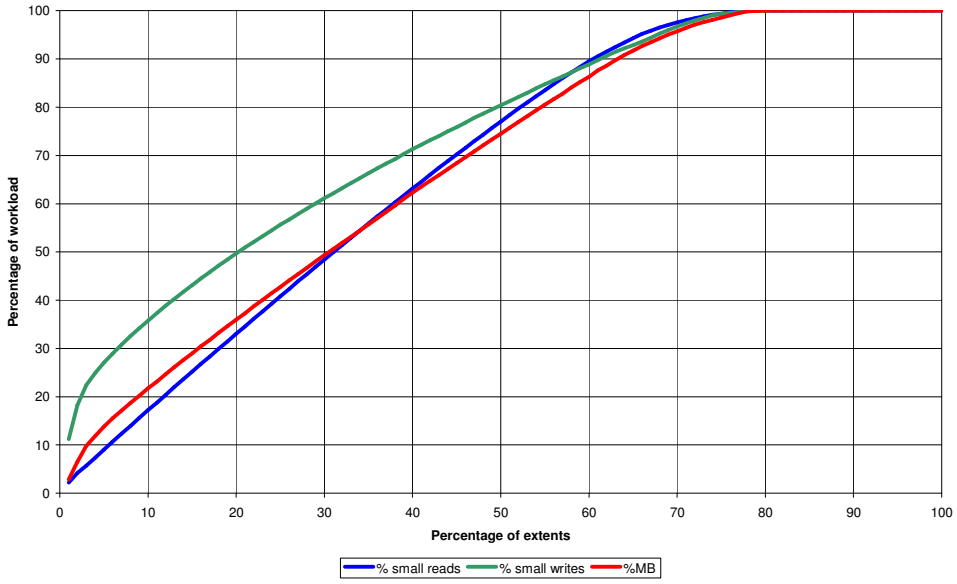

Figure 5. Skew for System i workload

3.3 Choosing an optimized configuration based on skew

Once we understand the skew of workload we can look to see how moving a percentage of data to SSDs can enable a use of larger or slower drives for the remainder of the workload based on the relative performance and capacity of the different drive types.

Figure 6 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 4.

Single-tier drive size driven by performance requirements	2–5% SSDs	5–10% SSDs
900 GB	1TB	2TB
600 GB	900 GB / 1TB	1TB / 2TB
450 GB	900 GB	1TB
300 GB	600 GB	900 GB
146 GB	300 GB	600 GB

Figure 6 Effect of different percentage of SSDs in a 2-tier configuration

The values in the table above are somewhat conservative and if the skew of a workload is known then it is possible to take a more accurate view of the percentage of SSDs which would allow a particular drive size.

Figure 7 shows, for a particular skew curve, how much SSD capacity would be required to allow different sizes of enterprise drive to be used in the configuration. If we assume 146GB drives would be required for a single tier solution then the intersection between the skew curve and the lines representing the other drives shows what capacity of SSDs would be needed to use this particular drive size.

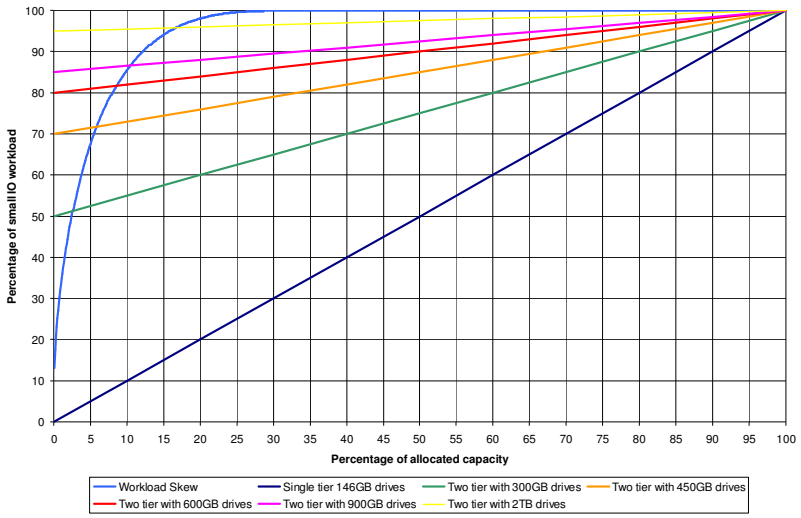

Figure 7 Drive Selection with a specific skew curve

This shows that only 3% SSDs would be required in this configuration to move to 300GB drives. 450GB drives would be possible with 6% SSDs and 10% SSDs would allow 900GB drives.

The analysis method here looks at the breakdown of the average workload and concentrates on the backend random IOPS. It does not take into account the performance limits for peak workloads (e.g. sequential workloads with high MB/s) so care should be taken to understand this when moving to a much larger drive size. Disk Magic can help with this more detailed modeling.

3.4 Further Application acceleration

There are situations where an application requires 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 workload expansion requirement or an application and business benefit for significantly improved read miss performance the higher percentages of SSDs may not provide the best return on investment.

3.5 Easy Tier Data Movement

One of the other questions often asked regarding Easy Tier is how much the workload changes over time and hence how much work Easy Tier has to do on a daily basis in order to optimize the configuration.

The chart below shows the number of extents moved per 15 minute period for an Easy Tier environment once the initial promotion of data to SSDs has completed. This configuration is using a 512MB extent size so the daily movement is typically around 200GB out of a total capacity of 140TB which is less than 0.2% of data being moved. This very small percentage of data is all that needs to be moved to maintain an optimized configuration.

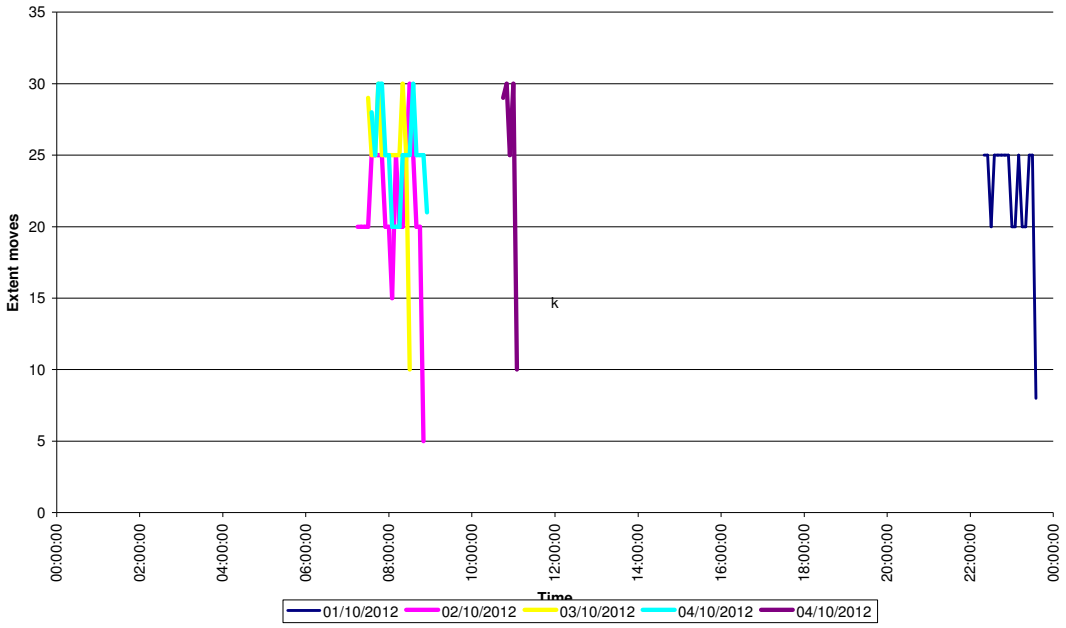

Figure 8 Example of Easy Tier data movement

The chart also shows the fact that Easy Tier is using a randomized interval for the analysis and subsequent movement of data so that it does not happen at the same time each day. In fact on the 4th November there were two periods of data movement as the Easy Tier analysis was performed twice.

4 Sample TCO benefits with Easy Tier

As well as providing improved performance using a multiple tier configuration, Easy Tier can also provide benefits both in terms of acquisition costs and ongoing datacentre costs for power, cooling and floor space. This section provides two different examples of storage requirements, and shows the benefits Easy Tier can provide compared to a standard single tier configuration using Storwize V7000 disk subsystems to meet the requirement. In these examples we have chosen relatively conservative assumptions regarding the skew and as can be seen in section 3.2 it is likely that Easy Tier could provide even greater savings in many environments.

4.1 Typical 50TB mixed environment

The first example is a configuration of 50TB usable capacity with performance requirements that can be met with 300GB 15K RPM drives.

Two alternative configurations are considered using Easy Tier. The first is a configuration with 5% SSDs and 95% 600GB 10K RPM drives and the second is a configuration with 10% SSDs and 90% 900GB 10K RPM drives.

Both of these alternative configurations would easily meet the performance requirements of the environment given the typical skew levels shown earlier in this paper and would generally provide a better average response time given the percentage of random workload that is moved to the SSDs by Easy Tier.

The chart below shows the comparative costs of the three configurations over three years, including hardware, software and maintenance. As can be seen the Easy Tier configuration with 5% SSD is the most cost effective configuration although the 10% SSD solution is likely to provide lower response times and has a similar cost to the base configuration with 300GB drives.

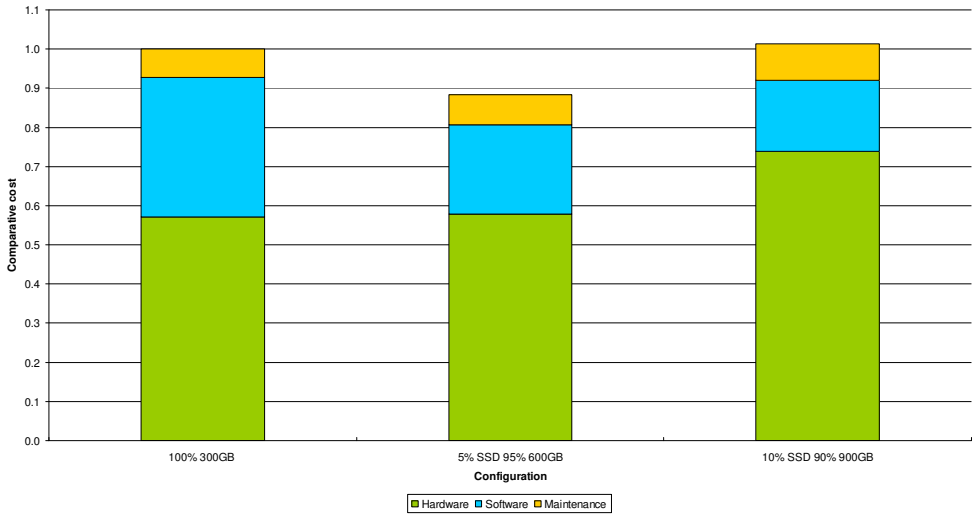

Figure 9 Three year costs for 50TB mixed environment

If we look at the datacenter costs for power, cooling and floor space for this configuration we can again see that both the Easy Tier configurations are providing a significant reduction compared to the traditional enterprise drive configuration with around a 40% saving in environmental costs.

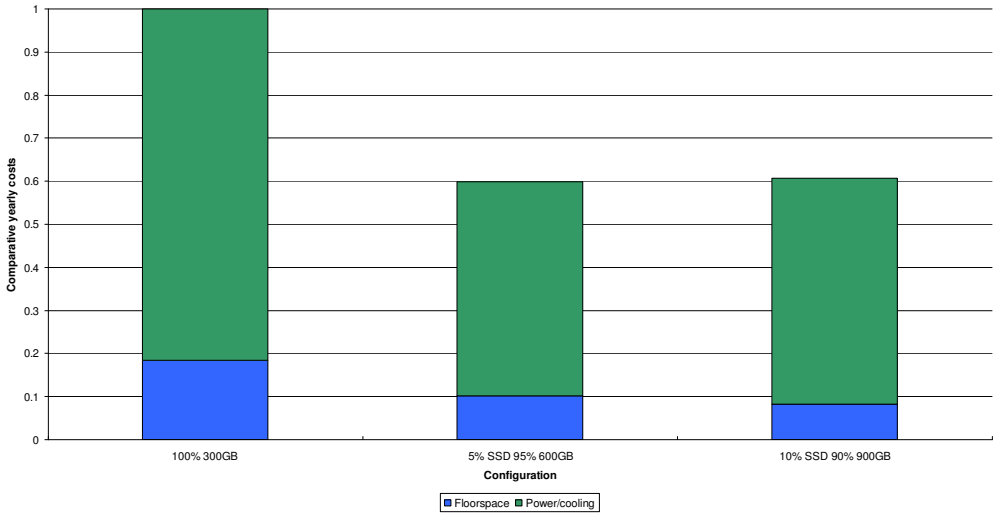

Figure 10 Yearly environmental costs for 50TB mixed environment

4.2 High performance environment

The second example is a high performance environment. The base configuration requires a full Storwize V7000 configuration of 240 146GB 15K RPM drives in a RAID10 configuration. For the sake of simplicity we have assumed that this matches the requirements both from a performance and capacity perspective.

Two alternative configurations were considered. The first is an Easy Tier configuration with 20% SSDs and 80% 600GB 10K RPM drives in a RAID5 configuration and the second is a SSD only configuration with 100% 400GB SSDs also in RAID5.

Both the SSD and Easy Tier configurations would be expected to give significantly improved performance compared to the base 146GB configuration. In the Easy Tier configuration we might expect 95% or more of the random IOs to be on the SSDs and so the performance could be very similar to the SSD only configuration.

The chart below shows the comparative costs of the three configurations over three years, including hardware, software and maintenance. As can be seen the Easy Tier configuration is the most cost effective configuration and while the SSD only configuration has a significantly higher maximum performance capability it also is more expensive than either of the other configurations.

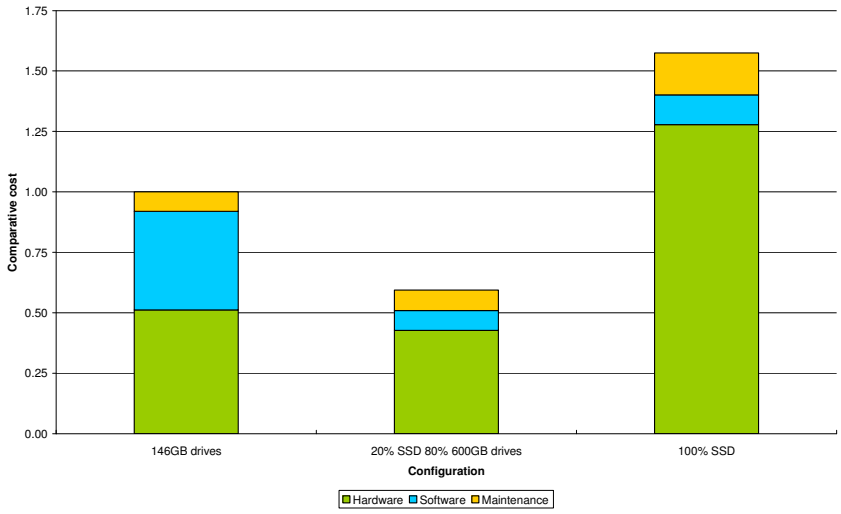

Figure 11 Three year costs for high performance environment

If we look at the datacentre costs for power, cooling and floorspace for this configuration we can again see that the Easy Tier configuration is providing a significant reduction compared to the traditional enterprise drive configuration with around 75% saving in environmental costs. What is perhaps somewhat surprising is that with this typical skew level the combination of 20% SSDs with larger enterprise drives is also more efficient from a facilities perspective than the 100% SSD solution.

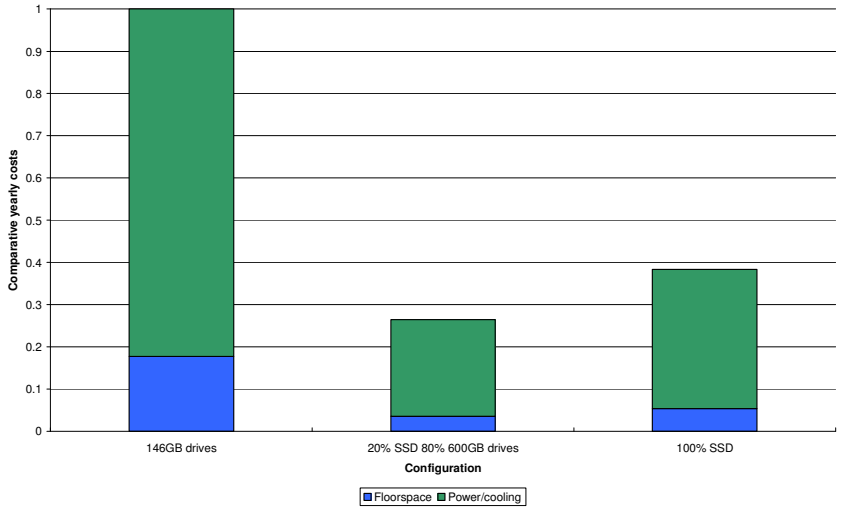

Figure 12 Yearly environmental costs for high performance environment

5 Conclusions

With Easy Tier an automated solution is made available for automatically managing data placement for performance in a variety of different configurations.

Typical real world workloads shown in this paper have a significantly uneven distribution of the IO activity over the capacity which enables significant performance benefits from relatively small percentages of SSDs.

We have also shown that Easy Tier can provide a reduction in both acquisition costs and the ongoing datacenter costs by enabling the use of larger enterprise drives along with SSDs.

This paper reflects performance and cost of storage in early 2013. The reader must remember though that the capabilities and cost of flash storage are improving markedly every quarter. The opportunities described here for flash storage will only increase as the technology develops.

In particular we expect to see an increasing opportunity for 100% SSD systems to be deployed in the near future, based on a number of factors:

- Use of technology such as IBM's Real-time compression to reduce the cost of SSD storage
- Cost savings that can be realized outside of the storage management environment through the use of SSD storage, by improving server and application performance and reducing costs their. There are also further opportunities to reduce application development and maintenance costs, through the benefits of reduced response time that SSD storage provide
- The creation of new business applications that take advantage of SSD storage and its unique capabilities to solve latent business problems in novel ways

Realizing the full benefits needs new co-ordination within organizations, since some of the benefits can only be achieved by reaching out from the storage management domain into other parts of the IT organization, and into other business units. However the trend is set in terms of increasing use of SSD storage in all data centers.

6 References

The following redbooks and whitepapers contain information about the architecture and implementation of Easy Tier on the SVC and Storwize V7000 as well as performance information with benchmark workloads.

Implementing the IBM System Storage SAN Volume Controller V6.3
http://www.redbooks.ibm.com/abstracts/sg247933.html?Open

IBM System Storage SAN Volume Controller Best Practices and Performance Guidelines
http://www.redbooks.ibm.com/abstracts/sg247521.html?Open

Using the Storage Tier Advisory Tool (STAT) and Easy Tier on the IBM Storwize V7000
http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP101852

SAP and IBM demonstrate capability of candling high billing volume in a Telecommunications Scenario
http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP101984

Microsoft SQL Server 2008 R2® with IBM System x3850 X5 and IBM Storwize® V7000 Easy Tier™
http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP101988

Performance benefits of IBM Storwize V7000 with IBM Easy Tier for Oracle ASM
http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP101990

The following whitepapers provide information about Easy Tier functionality as implemented on the on the DS8000.

IBM® System Storage™ DS8700™ and DS8800™ Performance with Easy Tier® 3nd Generation.
http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP102024

IBM® System Storage™ DS8700™ and DS8800™ Performance with Easy Tier® 2nd Generation.
http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP101961

IBM® System Storage™ DS8700™ Performance with Easy Tier®
http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP101675

The SPC-1 benchmarks referenced in this whitepaper can be found here:

http://www.storageperformance.org/results/benchmark_results_spc1/#spc1

