Survey Update: Users Share Their 2017 Storage Performance Needs

Jim Handy, *Objective Analysis*

Thomas Coughlin, *Coughlin Associates*

Table of Contents

List of Figures

Survey Update: Users Share their 2017 Storage Performance Needs

The following white paper is extracted from the report: "How Many IOPS Is Enough?" co-authored by Coughlin Associates and Objective Analysis. The report details the results of two end-user surveys performed by these companies in 2012 and again in 2016 in which IT managers shared their storage performance requirements across a range of applications. The report not only shows the performance requirements of a number of applications, but it also shows how these performance requirements have changed over the 4-year survey window.

What level of performance do different applications require? This is a question that was addressed by an on-line survey of digital storage end users in which IT managers in both 2012 and 2016 shared their requirements for IOPS (Inputs/Outputs per Second) and latency. The survey was designed to determine the performance requirements for several popular enterprise applications and help both users and storage providers understand how storage performance relates to other factors, including storage capacity and limitations elsewhere in the system, and to see how these requirements change over time.

The ongoing survey is still live and can be found at: http://TinyURL.com/IOPSsurvey. The updated results, compiled in a recent report, help to answer a number of important questions:

- Does a certain application really need the performance of an SSD?
- How much should a performance SSD cost?
- What have other IT managers found to be the right balance of performance and cost?

The Problem

Modern storage systems offer a dizzying range of IOPS (from hundreds to millions), as well as different latencies and storage capacities. Many IT managers find it difficult to determine which SSD or flash array to buy for their needs, or even whether they can get the speed they need from standard HDDs.

These are critical decisions. IT professionals must determine which SSD or flash array to purchase, but these same professionals may not know how to attain the bandwidth they require. Today's extraordinarily wide selection of IOPS, latencies, and capacities can only confuse those with a weak understanding of their system's requirements. Our IOPS survey was conducted to help these users to tap into their peers' experience to make well-informed choices, while also helping storage system vendors to better understand their customers' requirements.

The results of this survey provide considerable insight into these issues. The 2012 survey gathered responses from nearly 120 participants and the 2016 update added another 82 respondents to provide a respectable 6.9% margin of error.

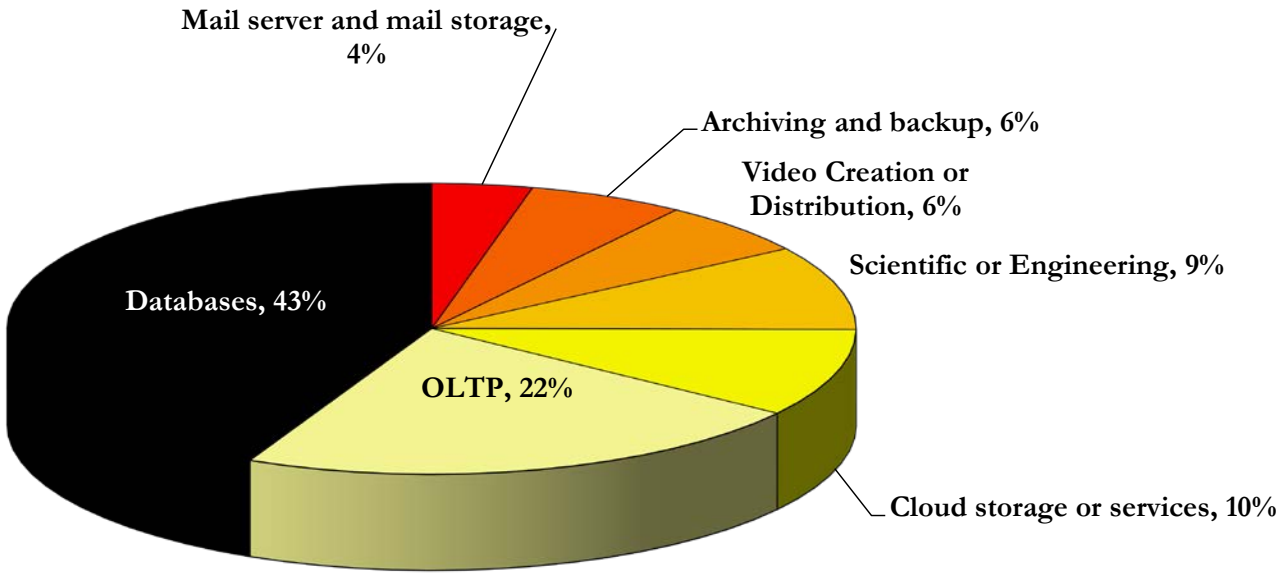
The resulting data gives a solid understanding of the performance, capacity, and cost requirements of various applications including IOPS, storage capacity and latency. We found that some applications have more rigid requirements than others, with needs varying according to the class of application.

Application Classes

Our first question asked what type is the dominant or most important application the storage would be used for. As the pie chart in **Figure 1** illustrates, most respondents wanted storage for the following application types:

- Databases – 43%
- On-Line Transaction Processing (OLTP) – 22%
- Cloud and Storage Services – 10%
- Scientific and Engineering Computing – 9%

Figure 1. Application Class Breakdown



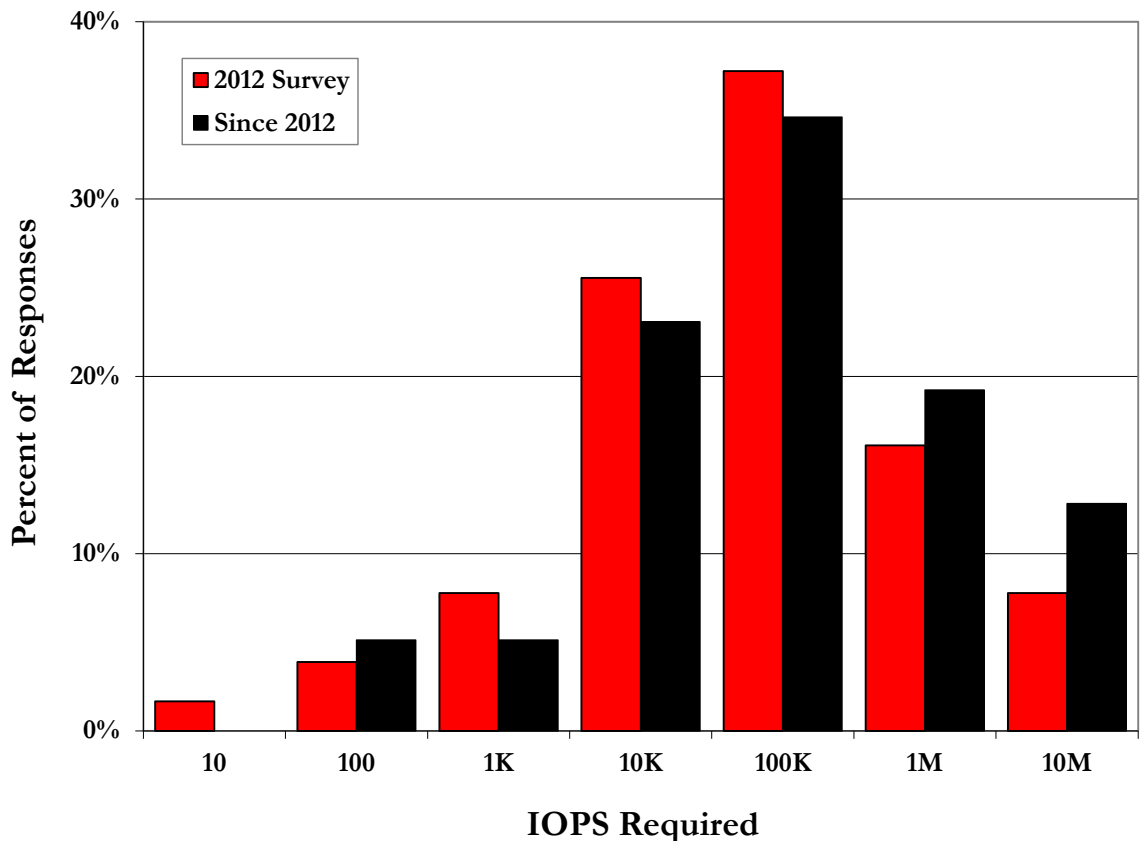
Source: Objective Analysis & Coughlin Associates, 2016

IOPS Needs

Next, we asked how many IOPS were required for this application. In order to make it as simple as possible for the respondents to answer our quantitative questions, we provided multiple-choices of simple orders of magnitude: The respondents could choose one of a number of bands, from 10-100 IOPS, or 101-1,000 IOPS, etc.

The results for all applications combined followed a standard distribution, with its peak for both the 2012 and the 2016 surveys landing in the 10K to 100K IOPS range, which is a good fit for SSDs. This is shown in **Figure 2**. (The full report breaks these down into responses by application category to produce 147 detailed figures and comparing changes from 2012 to 2016.)

Figure 2. IOPS Required Across All Applications



Source: Objective Analysis & Coughlin Associates, 2016

Most respondents (79% in 2012 and 77% in 2016) required IOPS performance of over 1,000 but fewer than 10 million IOPS over a broad range of capabilities. There is a 54% higher mean in 2016 versus 2012 and a 37% higher median value in 2016. This shows an increase in IOPS requirements since 2012.

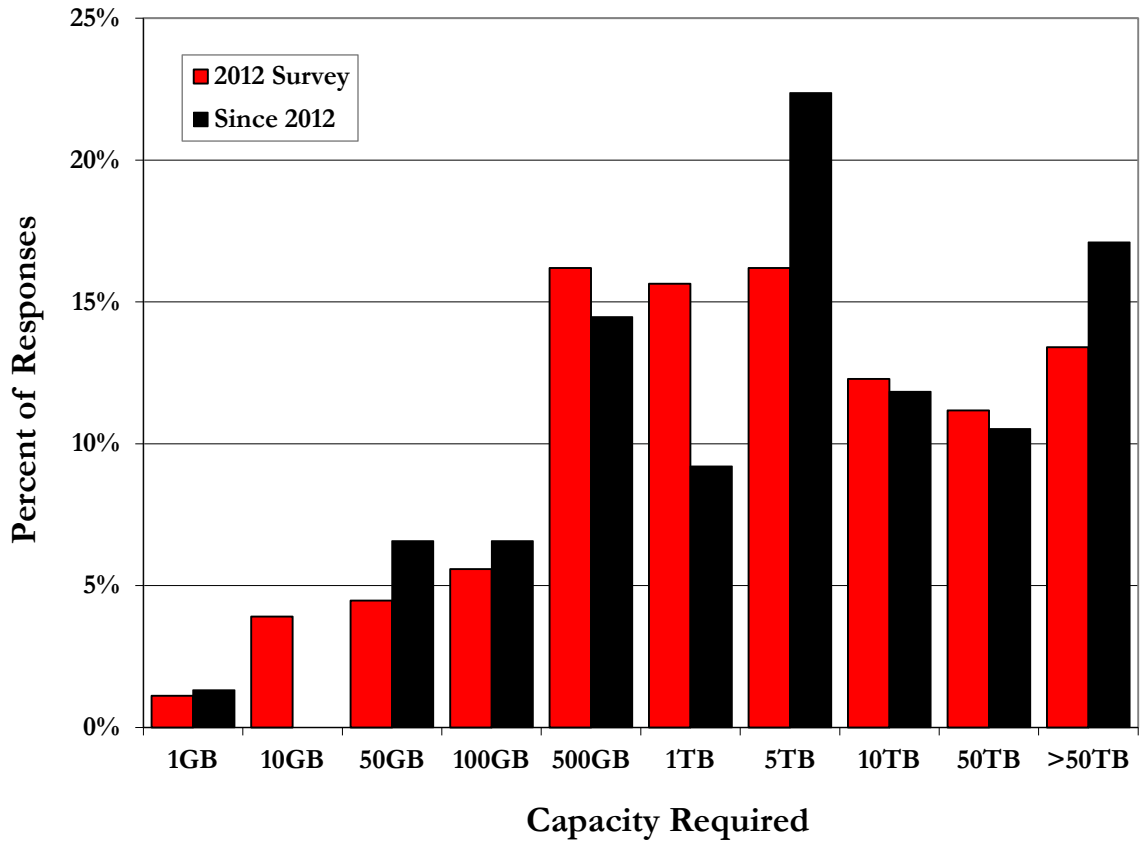
We were surprised at the large number of respondents for both surveys who claimed requirements of over one million IOPS, a level which wasn't even reached until the IBM Quicksilver demonstration at the end of 2008. The responses in the 10 IOPS and 100 IOPS ranges could easily be satisfied with less expensive hard disk drives.

These replies may be due to the fact that it is not standard practice to measure a system's IOPS, whereas data rate measurements are more common. Another survey question revealed both in 2012 and in 2016 that one third or fewer of our respondents had actually measured their systems' performance, indicating that a good number of our responses were based upon estimates rather than concrete evidence, a fact that may have led to an overestimate of IOPS needs.

Capacity Requirements

The survey also asked how much storage capacity the respondent's key application would be likely to need, to help provide an understanding of the amount of high-speed storage a system might require. Although the chart in **Figure 3** does not follow a standard distribution either during the original 2012 survey or in the 2016 follow-up, it indicates that there is a broad band of requirements. This chart illustrates the combined data for all of the application categories in **Figure 1**. The report provides similar information broken down by application, and these breakdowns show that certain applications have specific and very focused storage capacity requirements.

Figure 3. Storage Capacity Requirements for All Applications



Source: Objective Analysis & Coughlin Associates, 2016

There is still a big gap between SSD and HDD pricing, and it is likely that some of these responses didn't take flash prices into perspective. For example, 50TB of fast SSD storage would cost $10,000 or more today, while 50TB of HDD storage should cost somewhere around $1,000. Had the survey requested budget estimates we may have received more conservative responses about capacity requirements.

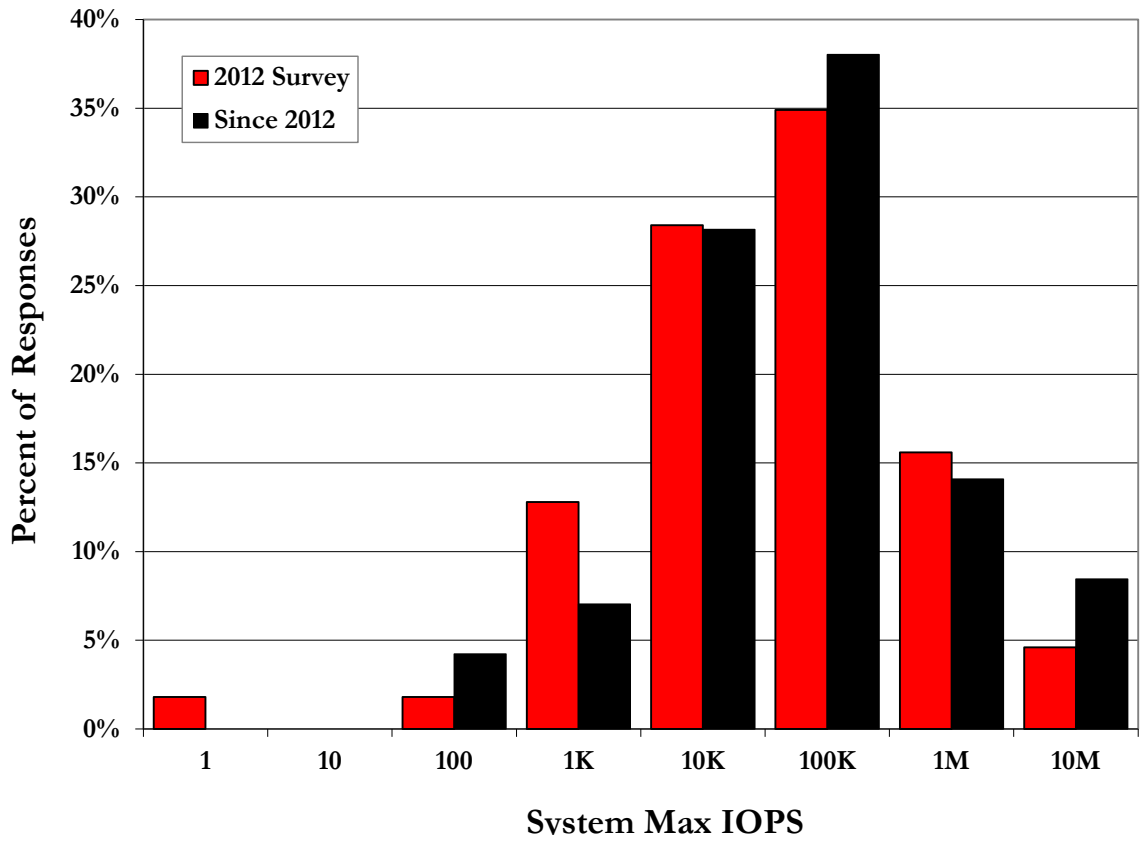
System Bottlenecks

SSD users, when they first begin to employ the technology, sometimes encounter a nasty surprise with the deployment of multiple SSDs: they might find that their system's performance doesn't scale in proportion to the number of SSDs that they add, or that an expensive 100K IOPS SSD doesn't provide any better performance than a cheaper 5K IOPS SSD, even though the lower-IOPS SSD was found to dramatically improve performance over the original all-HDD system. In other cases, users might find that their system's performance increases with the addition of a single SSD, and increases even more with the addition of a second SSD, but that there is no benefit from adding a third SSD. This happens as a result of bottlenecks elsewhere in the system. One installation described exactly this kind of problem, which the company traced to the network. By replacing the original 1Gb/s Ethernet with a 10Gb/s Ethernet, the company once again was able to improve performance by adding SSDs.

With this in mind we included a question in the survey to uncover the impact of system bottlenecks.

Our survey respondents for both the 2012 and 2016 surveys told us the highest number of IOPS their system could handle before some other bottleneck would get in the way. The results appear in **Figure 4**. This distribution resembles the maximum IOPS requirement that was shown in **Figure 2**, which stands to reason: once the IOPS have reached the system's limits there would be no benefit in pushing the storage system's performance any higher.

Figure 4. Highest Usable IOPS, All Applications



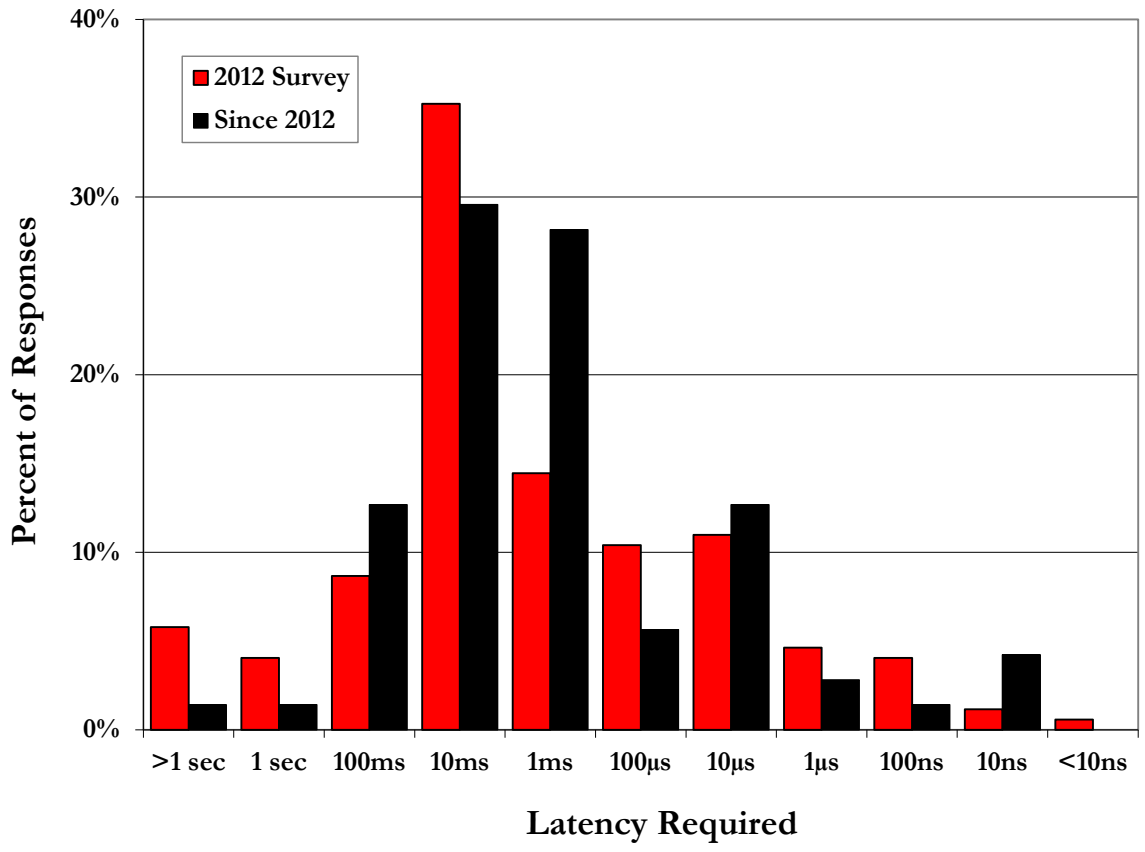
Source: Objective Analysis & Coughlin Associates, 2016

Although the responses for the 2012 survey and the 2016 survey differ less in this chart than they do in the previous charts, the data still clearly moves towards higher performance, a fact that comes as no surprise since system speeds steadily improve over time.

Respondents also told us in both 2012 and 2016 what was the minimum storage latency that their system could use before some other bottleneck would get in the way. The reported latency requirements (**Figure 5**) had a broad distribution in 2012 except for a pronounced 36% peak at 10 milliseconds. By 2016 this peak had shifted toward 1ms, but the rest of the distribution held strong similarities in 2016 to the responses we received in 2012. Primarily because of this peak at 1ms the 2016 median minimum latency is 54% lower than the 2012 median minimum latency.

It is rather odd that this trend's distribution is unlike that of the IOPS requirement, since IOPS is a function of latency and bandwidth. This could stem from the fact that latency, like data rate, may be a more commonly understood measure of system performance by many users.

Figure 5. Minimum Usable Latency across All Applications



Source: Objective Analysis & Coughlin Associates, 2016

The 10ms-1ms peak aligns well with the latency of networked storage, which leads us to believe that those reporting a 10ms or slower restriction based their responses upon the notion that all of their storage would be network-connected. As with the IOPS data shown in **Figure 2** the outliers are difficult to explain, with a few respondents telling us that they were satisfied with latencies of one second or longer (perhaps for archiving applications), and others expressing a need for latencies of 10ns or faster, the speed of main memory or a processor cache.

Conclusion

The survey asked respondents in 2012 and again in 2016 to share a rough idea of their dominant storage needs, expressed in capacity and performance, measured in IOPS (inputs or outputs per second) and maximum latency. Over two hundred respondents also shared the speed of the rest of their system and for the application type being run on that system.

Most respondents (nearly 80% for both survey periods) required performance of over 1,000 but fewer than ten million IOPS over a broad range of capacities. For the most part, their needs mirrored the maximum IOPS level their systems could support. System latency limitations for both the original and later surveys peaked near 10 milliseconds but showed a wide distribution at other latencies.

Although this white paper is too brief to share the complete details, the survey report provides the above data analyzed for each of the five leading applications and comparing the changes over time. For each application and for the combination of all applications the report also uses scatter charts to illustrate the relationship between different surveyed variables to find that some parameters correlate well with each other. These include capacity vs. IOPS, Capacity vs. Maximum Latency and IOPS vs. Maximum Latency.

The report shows that the best storage choice for a given application is a function of the value of performance vs. the costs of storage. Higher performance (IOPS) is critical for certain applications like OLTP but may be less important for others, including many cloud services. Likewise, some applications' latency requirements are more stringent than others. High-capacity applications may employ higher latency storage tiers based on HDD or magnetic tape technology to reduce total storage costs with some flash memory serving as an acceleration layer.

Complete survey results are analyzed in the report: ***"How Many IOPS is Enough?"*** which is available for purchase from www.Objective-Analysis.com or http://www.tomcoughlin.com/techpapers.htm. This study explains the need for IOPS for various applications and maps the survey's results in a way that provides deeper insight into the various tradeoffs of capacity, IOPS, and latency.

About the SNIA

The Storage Networking Industry Association (SNIA) is a not–for–profit global organization, made up of member companies spanning the global storage market. SNIA's mission is to lead the storage industry worldwide in developing and promoting standards, technologies, and educational services to empower organizations in the management of information. To this end, the SNIA is uniquely committed to delivering standards, education, and services that will propel open storage networking solutions into the broader market. For more information, visit http://www.snia.org.

About the Solid State Storage Initiative

The SNIA Solid State Storage Initiative (SSSI) fosters the growth and success of the market for solid state storage (SSS) and persistent memory (PM) in both enterprise and client environments. Members of the SSSI work together to promote the development of technical standards and tools, educate IT communities about persistent memory and solid state storage, perform market outreach that highlights the virtues of these technologies, and collaborate with other industry associations on SSS and PM technical work. SSSI member companies represent a variety of segments in the IT industry.
(See http://www.snia.org/forums/sssi/about/members)

For more information on SNIA's Solid State Storage and Persistent Memory activities, visit www.snia.org/forums/sssi and get involved in the conversation at http://twitter.com/SNIASolidState.