Survey: Users Share Their Storage Performance Needs

Jim Handy, Objective Analysis
Thomas Coughlin, PhD, Coughlin Associates

Table of Contents

List of Figures

Survey: Users Share their Storage Performance Needs

The following white paper is extracted from the report: How Many IOPS Do You Really Need? coauthored by Coughlin Associates and Objective Analysis. The report details the results of a survey performed by these companies in which IT managers shared their storage performance requirements across a range of applications.

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 shared their requirements for IOPS (Inputs/Outputs per Second) and latency. The survey aimed to determine the IOPS requirement for several popular enterprise applications to understand how it relates to other factors, including storage capacity and latency.

The ongoing survey can be found at: http://TinyURL.com/IOPSsurvey. The 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 survey gathered responses from over 180 participants to provide a respectable 7.1% margin of error.

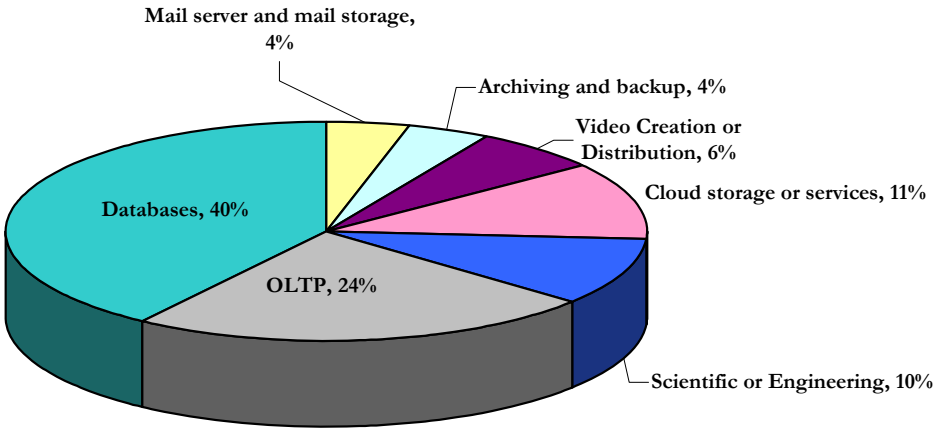
The resulting data gives a solid understanding of the performance, capacity, and cost requirements of various applications including IOPS, storage capacity and latency. We found that some applications have more rigid requirements than others, with needs varying according to the class of application.

Application Classes

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

- Databases – 40%
- On-Line Transaction Processing (OLTP) – 24%
- Cloud and Storage Services – 11%
- Scientific and Engineering Computing – 10%

Figure 1. Application Class Breakdown

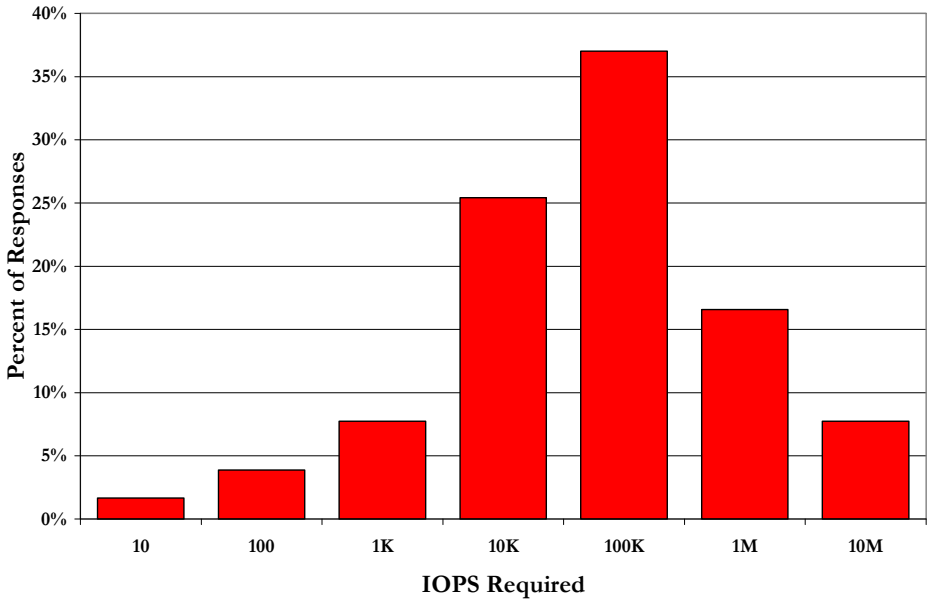


IOPS Needs

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 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 58 detailed figures.)

Figure 2. IOPS Required Across All Applications



Most respondents (87%) required IOPS performance of over 1,000 but fewer than 10 million IOPS over a broad range of capabilities.

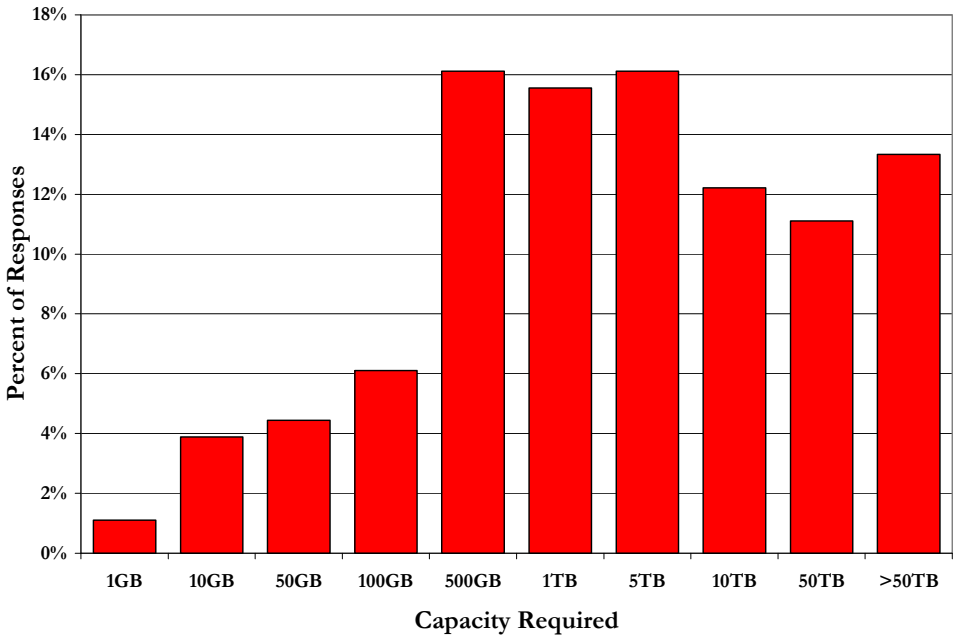
We were surprised that nearly 7% of the respondents claimed requirements of over one million IOPS, a level which wasn't even reached until the IBM Quicksilver demonstration at the end of 2008. We were also surprised to see responses in the 10 IOPS and 100 IOPS ranges, which could easily be satisfied with 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 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 lead 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, it indicates that there is a broad band of requirements with roughly similar requirements for two groups of respondents – those needing 500GB-5TB (48%, evenly divided between three adjacent bins) and those looking for 5TB and higher (37%, once again more or less evenly broken down between three adjacent bins). This second group would be reasonable candidates for flash-based arrays rather than isolated SSDs.

Figure 3. Storage Capacity Requirements for All Applications



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 $50,000 or

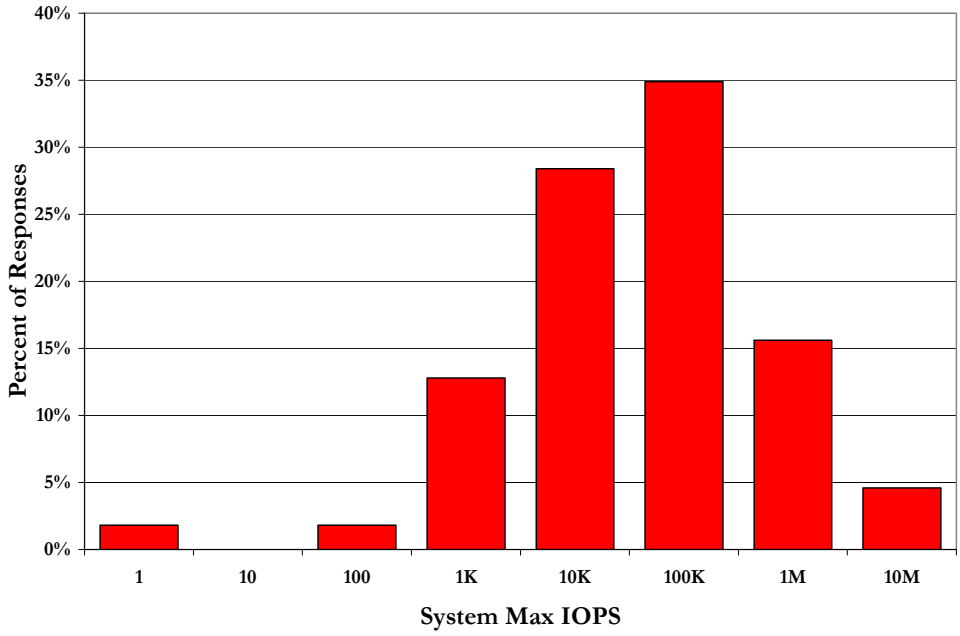
more today, while 50TB of HDD storage should cost somewhere around $2,500. Had the survey provided cost estimates we may have received more conservative estimates of capacity requirements.

System Bottlenecks

SSD users, when they first begin to use 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 more 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, system performance might increase with the addition of a single SSD, and increase 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.

Our survey respondents told us the highest number of IOPS their system could handle before some other bottleneck would get in the way, and the results appear in **Figure 4**. This distribution resembles the maximum IOPS requirement was shown in **Figure 4**, 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.

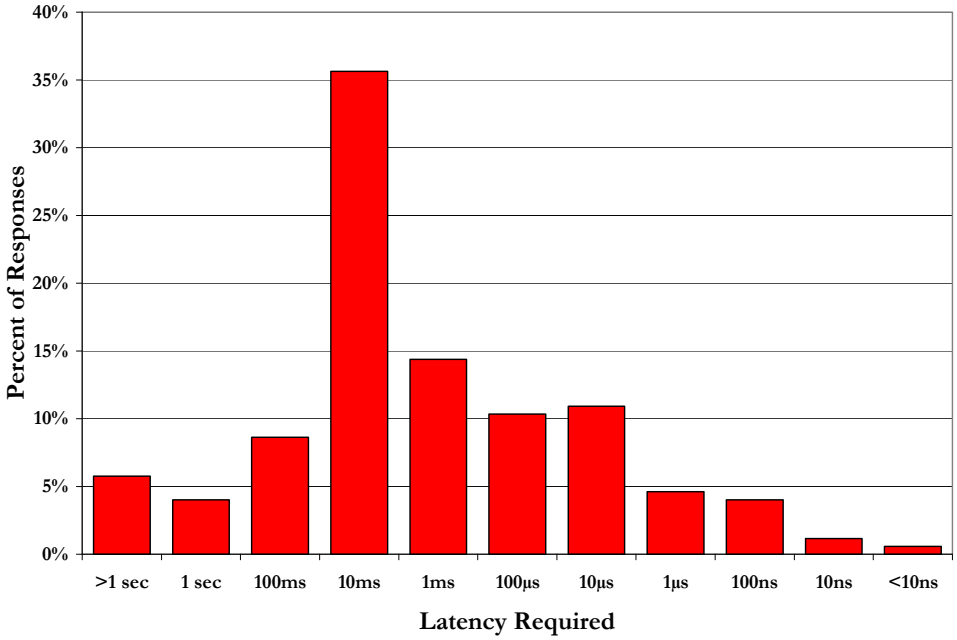
Figure 4. Highest Usable IOPS, All Applications



Respondents also reported 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 except for a pronounced 36% peak at 10 milliseconds.

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 may be because 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



The 36% peak at 10ms aligns well with the latency of networked storage, which leads us to believe that those reporting a 10ms or slower restriction (54% of respondents) was based 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 10% telling us that they were satisfied with latencies of one second or longer, and nearly 2% expressing a need for latencies of 10ns or faster, the speed of main memory or a processor cache.

Conclusion

The survey asked respondents to share a rough idea of their high-speed storage needs, expressed in capacity and performance, measured in IOPS (inputs or outputs per second) and maximum latency. Nearly 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 (87%) 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 had a very solid peak of 36% of respondents at 10 milliseconds but 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. 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 and IOPS, Capacity and Maximum Latency and IOPS and 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 some 400 member companies and 7,000 individuals spanning virtually the entire storage industry. 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 additional information, visit the SNIA web site at 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 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 solid state storage, perform market outreach that highlights the virtues of solid state storage, and collaborate with other industry associations on solid state storage 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 activities, visit www.snia.org/forums/sssi and get involved in the conversation at http://twitter.com/SNIASolidState.

Storage Networking Industry Association
4360 Arrows West Drive • Colorado Springs, CO 80907 • Phone: 719-694-1380 • Fax: 719-694-1389 • www.snia.org