

HOW MANY IOPS IS ENOUGH?



**Coughlin
Associates**

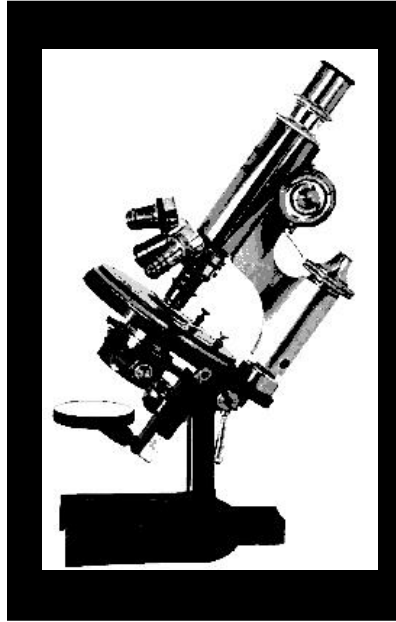
Data Storage Consulting

**OBJECTIVE ANALYSIS
& COUGHLIN ASSOCIATES**

October 2016

HOW MANY IOPS IS ENOUGH?

A JOINT REPORT



*Coughlin
Associates*

JIM HANDY, OBJECTIVE ANALYSIS
TOM COUGHLIN, COUGHLIN ASSOCIATES

October 2016

Entire contents copyright 2016 Objective Analysis and Coughlin Associates.
All rights reserved.

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form, except as permitted under Sections 107 or 108 of the 1976 United States Copyright Act, without the prior written permission from Objective Analysis. Requests for permission should be addressed to Objective Analysis, PO Box 440, Los Gatos, CA, 95031-0440, USA.

CONTENTS

Contents	iv
Executive Summary	1
Survey Objective.....	3
Why Did We Do This?	3
Methodology	3
Margin of Error	5
Anticipated Outcome	6
Storage Devices and their Characteristics.....	7
SSDs and Flash Memory vs. DRAM	14
Caching/Tiering vs. Determinism.....	14
Raw Data Analysis.....	17
IOPS Requirements.....	19
Capacity Requirements	21
Impact of Other System Bottlenecks	24
Minimum Useful Latency	35
Comparing the Results.....	39
IOPS vs. Capacity	39
Latency vs. Capacity	41
IOPS vs. Latency	44
Application Sensitivities	48
Database Systems.....	48
Metadata Servers.....	48
Data Warehousing and Search	49
Internet Server Caching	49
On Line Transaction Processing (OLTP)	68
Charge Card Processing.....	69
Reservations Systems.....	70
Algorithmic Trading	70
Currency Exchange and Arbitrage.....	71
Inter-Bank Transfers	71
Other On-Line Transaction Processing Applications	71
Cloud or Storage Services.....	91
Real Time Data/Feed Processing	93
Contextual Web Advertising.....	94
Scientific or Engineering Applications.....	110
Electronic Design Automation & Modeling.....	111
Weather/Life Sciences	112
Aerodynamics Design	112
Nuclear Fission Models	113
Software Development.....	113
IOPS Requirements by Application.....	130
The Monetary Value of Higher IOPS Figures	133
How the Market Should Develop	135

Appendix A: The Online Survey	136
Appendix B: Data Inconsistencies	138
Possible Reasons for Inconsistencies.....	140
Figures.....	144
Tables.....	149
Authors.....	150
Jim Handy.....	150
Tom Coughlin.....	150

EXECUTIVE SUMMARY

This report details the results of a months-long storage performance survey performed by Coughlin Associates and Objective Analysis through the first half of 2016 compared to earlier responses to the same set of survey questions in 2012.

The survey asked over two hundred sixty respondents to share their high-speed storage capacity and performance needs, IOPS and maximum latency, along with the speed of the rest of their system and for the application type being run on that system.

Data is presented in three graphical formats:

- A pie chart to display the survey population
- Column charts indicating the percentage of respondents falling into various categories (Pareto Charts)
- Scatter charts to help identify the relationship between the different parameters that were measured.

In an effort to determine how perceptions have changed over time, this data is presented for 2012 alone (as it appeared in our original report), through 2016, and with the newer responses broken out by themselves and compared with the older data.

The following results were found:

- The majority of our survey respondents required fast storage for the following application types:
 - Database
 - On-Line Transaction Processing (OLTP)
 - Cloud and Storage Services
 - Scientific and Engineering Computing
- Those survey responses that changed between 2012 and 2016 generally represented increases in both speed and capacity and reductions in expected latency.
- Most respondents (79%) required performance of over 1,000 but fewer than ten million IOPS over a broad range of capacities
- Latency requirements had a very solid peak with 32% requiring 10 milliseconds.
- In many cases there was a reasonable correlation between two of the surveyed parameters:
 - Capacity & IOPS requirements
 - Capacity & Maximum Latency
 - IOPS Requirements and Maximum Latency
- The sample group was a respectable size, giving us a 6.0% margin of error at a 95% confidence level. We continue to collect survey data with a goal of increasing the respondent base and reducing the error margin.

- Storage pricing is loosely tied to performance. This relationship is likely to tighten as a result of this survey and as those who purchase fast storage become more sophisticated.

FIGURES

Figure 1. Graphical Depiction of the Enterprise Storage Hierarchy.....	7
Figure 2. LTO Magnetic Tape Cartridge	8
Figure 3. Storage Average Retail Price vs. Time.....	9
Figure 4. Typical Hard Disk Drive with Key Components Identified.....	10
Figure 5. Comparison of a HDD and an SSD.....	11
Figure 6. Relative Cost per Gigabyte of Flash & Other Storage Media versus Time	12
Figure 7. Typical Computer Memory/Storage Hierarchy.....	13
Figure 8. Memory/Storage Hierarchy of a Computing System Showing How NAND (SSD) Fits between DRAM and HDD.....	15
Figure 9. Distribution of Survey Dominant Application – Original Survey.....	17
Figure 10. Distribution of Survey Dominant Application – New Survey Data.....	18
Figure 11. Distribution of Survey Dominant Application – Combined Surveys.....	18
Figure 12. Overall Survey IOPS Requirements – Original Survey	19
Figure 13. Overall Survey IOPS Requirements – New Survey Data	20
Figure 14. Overall Survey IOPS Requirements – Combined Surveys	20
Figure 15. Overall Survey IOPS Requirements – Comparison of Newer and Older Survey Data.....	21
Figure 16. Overall Survey Storage Capacity Requirements – Original Survey	22
Figure 17. Overall Survey Storage Capacity Requirements – New Survey Data.....	23
Figure 18. Overall Survey Storage Capacity Requirements – Combined Surveys.....	23
Figure 19. Survey Storage Capacity Requirements – Old and Newer Survey Data on the Same Chart.....	24
Figure 20. Latency of Various Media for 4KB Random Reads	25
Figure 21. Relative Significance of System Overhead	27
Figure 22. How Many IOPS is the Most Your System Can Use? Original Survey	27
Figure 23. How Many IOPS is the Most Your System Can Use? New Survey Data.....	28
Figure 24. How Many IOPS is the Most Your System Can Use? Combined Surveys.....	29
Figure 25. How Many IOPS is the Most Your System Can Use? Old and Newer Survey Data on the Same Chart	30
Figure 26. Response Counts of IOPS Needed vs. Other System Bottlenecks – Original Survey	31
Figure 27. Response Counts of IOPS Needed vs. Other System Bottlenecks – New Survey Data.....	31
Figure 28. Response Counts of IOPS Needed vs. Other System Bottlenecks – Combined Surveys.....	32
Figure 29. Cumulative Distribution Function for Bottleneck IOPS and the IOPS Required for the Application – Original Survey	33
Figure 30. Cumulative Distribution Function for Bottleneck IOPS and the IOPS Required for the Application – New Survey Data.....	33
Figure 31. Cumulative Distribution Function for Bottleneck IOPS and the IOPS Required for the Application – Combined Surveys.....	34
Figure 32. Overall Survey Minimum Latency Requirements – Original Survey	36
Figure 33. Overall Survey Minimum Latency Requirements – New Survey Data	36

Figure 34. Overall Survey Minimum Latency Requirements – Combined Surveys	37
Figure 35. Overall Survey Minimum Latency Requirements – Newer and Older Survey Data On the Same Chart	37
Figure 36. Latency of Four Storage Configurations	38
Figure 37. Overall Survey IOPS vs. Capacity – Original Survey.....	40
Figure 38. Overall Survey IOPS vs. Capacity – New Survey Data.....	40
Figure 39. Overall Survey IOPS vs. Capacity – Combined Surveys.....	41
Figure 40. Overall Survey Latency vs. Capacity – Original Survey	42
Figure 41. Overall Survey Latency vs. Capacity – New Survey Data.....	43
Figure 42. Overall Survey Latency vs. Capacity – Combined Surveys.....	43
Figure 43. Mix of IOPS and Latency Needs for Respondents – Original Survey	45
Figure 44. Mix of IOPS and Latency Needs for Respondents – New Survey Data	46
Figure 45. Mix of IOPS and Latency Needs for Respondents – Combined Surveys	46
Figure 46. Hotter Data is Stored in Higher Cost Devices.....	49
Figure 47. IOPS Requirement: Database Respondents – Original Survey.....	51
Figure 48. IOPS Requirement: Database Respondents – New Survey Data	51
Figure 49. IOPS Requirement: Database Respondents – Combined Surveys	52
Figure 50. IOPS Requirement: Database Respondents – Plotting Older and Newer survey Data.....	53
Figure 51. High-Speed Capacity Requirement: Database Respondents – Original Survey	54
Figure 52. High-Speed Capacity Requirement: Database Respondents – New Survey Data	55
Figure 53. High-Speed Capacity Requirement: Database Respondents – Combined Surveys.....	55
Figure 54. High-Speed Capacity Requirement: Database Respondents – Comparison of 2012 and Post-2012 Survey Data.....	56
Figure 55. High-Speed Capacity vs. IOPS Required – Database Respondents: Original Survey	57
Figure 56. High-Speed Capacity vs. IOPS Required – Database Respondents: New Survey Data.....	58
Figure 57. High-Speed Capacity vs. IOPS Required – Database Respondents: Combined Surveys.....	58
Figure 58. Latency Requirement: Database Respondents – Original Survey.....	59
Figure 59. Latency Requirement: Database Respondents – New Survey Data.....	60
Figure 60. Latency Requirement: Database Respondents – Combined Surveys.....	60
Figure 61. Latency Requirement: Database Respondents – Comparison of 2012 and Post 2012 Data.....	61
Figure 62. Latency vs. IOPS Required – Database Respondents: Original Survey	62
Figure 63. Latency vs. IOPS Required – Database Respondents: New Survey Data.....	63
Figure 64. Latency vs. IOPS Required – Database Respondents: Combined Surveys.....	63
Figure 65. Latency vs. Capacity Required – Database Respondents: Original Survey	64
Figure 66. Latency vs. Capacity Required – Database Respondents: New Survey Data ...	65
Figure 67. Latency vs. Capacity Required – Database Respondents: Combined Surveys .	65
Figure 68. Maximum IOPS System can Use – Database Respondents: Original Survey ..	66
Figure 69. Maximum IOPS System can Use – Database Respondents: New Survey Data	67

Figure 70. Maximum IOPS System can Use – Database Respondents: Combined Surveys	67
Figure 71. Maximum IOPS System can Use – Database Respondents: Comparison of 2012 to Post 2012 Data	68
Figure 72. IOPS Requirements – OLTP Respondents: Original Survey	73
Figure 73. IOPS Requirements – OLTP Respondents: New Survey Data	73
Figure 74. IOPS Requirements – OLTP Respondents: Combined Data	74
Figure 75. IOPS Requirements – OLTP Respondents: 2012 and Post 2012 on Same Chart	74
Figure 76. High Speed Capacity Requirement: OLTP Respondents – Original Survey	76
Figure 77. High Speed Capacity Requirement: OLTP Respondents – New Survey Data .	76
Figure 78. High Speed Capacity Requirement: OLTP Respondents – Combined Surveys	77
Figure 79. High Speed Capacity Requirement: OLTP Respondents – 2012 and Post 2012 Data Plotted.....	78
Figure 80. High-Speed Capacity vs. IOPS Required - OLTP Respondents: Original Survey	79
Figure 81. High-Speed Capacity vs. IOPS Required - OLTP Respondents: New Survey Data	80
Figure 82. High-Speed Capacity vs. IOPS Required - OLTP Respondents: Combined Surveys.....	80
Figure 83. Latency Requirement: OLTP Respondents – Original Survey	81
Figure 84. Latency Requirement: OLTP Respondents – New Survey Data.....	82
Figure 85. Latency Requirement: OLTP Respondents – Combined Surveys	83
Figure 86. Latency Requirement: OLTP Respondents – Plotting 2012 and Post 2012 Data Together	83
Figure 87. Latency vs. IOPS Required - OLTP Respondents: Original Survey	85
Figure 88. Latency vs. IOPS Required - OLTP Respondents: New Survey Data	85
Figure 89. Latency vs. IOPS Required - OLTP Respondents: Combined Surveys	86
Figure 90. Latency vs. Capacity Required - OLTP Respondents: Original Survey	87
Figure 91. Latency vs. Capacity Required - OLTP Respondents: New Survey Data	87
Figure 92. Latency vs. Capacity Required - OLTP Respondents: Combined Surveys	88
Figure 93. Maximum IOPS System can Use - OLTP Respondents: Original Survey.....	89
Figure 94. Maximum IOPS System can Use - OLTP Respondents: New Survey Data	90
Figure 95. Maximum IOPS System can Use - OLTP Respondents: Combined Surveys.	90
Figure 96. IOPS Requirement: Cloud or Storage Services Respondents – Original Survey	94
Figure 97. IOPS Requirement: Cloud or Storage Services Respondents – New Survey Data	95
Figure 98. IOPS Requirement: Cloud or Storage Services Respondents – Combined Surveys.....	96
Figure 99. High Speed Capacity Requirement: Cloud or Storage Services Respondents – Original Survey	97
Figure 100. Capacity Requirement: Cloud or Storage Services Respondents – New Survey Data	98
Figure 101. Capacity Requirement: Cloud or Storage Services Respondents – Combined Surveys.....	98

Figure 102. High-Speed Capacity vs. IOPS Required - Cloud or Storage Services Respondents: Original Survey	100
Figure 103. Capacity vs. IOPS Required - Cloud or Storage Services Respondents: New Survey Data.....	100
Figure 104. Capacity vs. IOPS Required - Cloud or Storage Services Respondents: Combined Surveys	101
Figure 105. Latency Requirement: Cloud or Storage Services Respondents – Original Survey	102
Figure 106. Latency Requirement: Cloud or Storage Services Respondents – New Survey Data.....	103
Figure 107. Latency Requirement: Cloud or Storage Services Respondents – Combined Surveys.....	103
Figure 108. Latency vs. Capacity Required - Cloud or Storage Services Respondents: Original Survey.....	104
Figure 109. Latency vs. Capacity Required - Cloud or Storage Services Respondents: New Survey Data.....	105
Figure 110. Latency vs. Capacity Required - Cloud or Storage Services Respondents: Combined Surveys	106
Figure 111. Latency vs. IOPS Required - Cloud or Storage Services Respondents: Original Survey	107
Figure 112. Latency vs. IOPS Required - Cloud or Storage Services Respondents: New Survey Data.....	107
Figure 113. Latency vs. IOPS Required - Cloud or Storage Services Respondents: Combined Surveys	108
Figure 114. Maximum IOPS System can Use - Cloud or Storage Services Respondents: Original Survey.....	109
Figure 115. Maximum IOPS System can Use - Cloud or Storage Services Respondents: New Survey Data	109
Figure 116. Maximum IOPS System can Use - Cloud or Storage Services Respondents: Combined Surveys.....	110
Figure 117. IOPS Requirement: Scientific or Engineering Respondents: Original Survey	115
Figure 118. IOPS Requirement: Scientific or Engineering Respondents: New Survey Data	115
Figure 119. IOPS Requirement: Scientific or Engineering Respondents: Combined Surveys.....	116
Figure 120. High Speed Capacity Requirement: Scientific or Engineering Respondents – Original Survey.....	117
Figure 121. High Speed Capacity Requirement: Scientific or Engineering Respondents – New Survey Data	118
Figure 122. High Speed Capacity Requirement: Scientific or Engineering Respondents – Combined Surveys	118
Figure 123. High-Speed Capacity vs. IOPS Required - Scientific or Engineering Respondents: Original Survey	119
Figure 124. High-Speed Capacity vs. IOPS Required - Scientific or Engineering Respondents: New Survey Data.....	120

Figure 125. High-Speed Capacity vs. IOPS Required - Scientific or Engineering Respondents: Combined Surveys	120
Figure 126. Latency Requirement: Scientific or Engineering Respondents: Original Survey	121
Figure 127. Latency Requirement: Scientific or Engineering Respondents: New Survey Data	122
Figure 128. Latency Requirement: Scientific or Engineering Respondents: Combined Surveys.....	122
Figure 129. Latency vs. Capacity Required - Scientific or Engineering Respondents: Original Survey	123
Figure 130. Latency vs. Capacity Required - Scientific or Engineering Respondents: New Survey Data.....	124
Figure 131. Latency vs. Capacity Required - Scientific or Engineering Respondents: Combined Surveys	124
Figure 132. Latency vs. IOPS Required - Scientific or Engineering Respondents: Original Survey	125
Figure 133. Latency vs. IOPS Required - Scientific or Engineering Respondents: New Survey Data.....	126
Figure 134. Latency vs. IOPS Required - Scientific or Engineering Respondents: Combined Surveys	126
Figure 135. Maximum IOPS System can Use - Scientific or Engineering Respondents: Original Survey	127
Figure 136. Latency vs. IOPS Required - Scientific or Engineering Respondents: New Survey Data.....	128
Figure 137. Latency vs. IOPS Required - Scientific or Engineering Respondents: Combined Surveys	128
Figure 138. Comparison of IOPS Needs by Application: Original Survey	131
Figure 139. Comparison of IOPS Needs by Application: New Survey Data	131
Figure 140. Comparison of IOPS Needs by Application: Combined Surveys	132
Figure 141. Cost per Gigabyte as a Function of IOPS.....	133
Figure 142. Scatter Chart of Responses: IOPS Needed vs. System Bottleneck – Original Survey	139
Figure 143. Scatter Chart of Responses: IOPS Needed vs. System Bottleneck – New Survey Data.....	139
Figure 144. Scatter Chart of Responses: IOPS Needed vs. System Bottleneck – Combined Surveys.....	140
Figure 145. About Two Thirds of Respondents Estimated Answers: Original Survey	141
Figure 146. The same held true for the New Survey Data.....	142
Figure 147. Naturally, the Combined Surveys showed the same split	142

TABLES

Table 1. Mapping of IOPS Requirements to Numerical Values.....	4
Table 2. Mapping of Capacity Requirements to Numerical Values	4
Table 3. Mapping of System Bottleneck IOPS Responses to Numerical Values	5
Table 4. Mapping of Maximum Latency to Numerical Values	5
Table 5. Margin of Error for the Combined Survey and For Sub-Groups (As of August 1, 2016)	6
Table 6. Margin of Error for the Post 2012 Survey and For Sub-Groups (As of August 1, 2016)	6
Table 7. Breakdown of Latency in Different Media.....	26
Table 8. Latencies of Standard Storage Devices	44
Table 9. System IOPS Capability& Stated IOPS Need vs. How Data was Determined ..	142

AUTHORS

Jim Handy



Jim Handy, a widely recognized semiconductor analyst, comes to Objective Analysis with over 30 years in the electronics industry including 14 years as an industry analyst for Dataquest (now Gartner) and Semico Research. His background includes marketing and design positions at market-leading suppliers including Intel, National Semiconductor, and Infineon.

A frequent presenter at trade shows, Mr. Handy is known for his widespread industry presence and volume of publication. He has written hundreds of articles for trade journals, Dataquest, Semico, and other prior employers, and is frequently interviewed and quoted in the electronics trade press and other media. Mr. Handy writes the [Chip Talk blog](#) for Forbes online and contributes to two Objective Analysis blogs: [The SSD Guy](#) and [The Memory Guy](#).

Mr. Handy has a strong technical leaning, with a Bachelor's degree in Electrical Engineering from Georgia Tech, and is a patent holder in the field of cache memory design. He is the author of "The Cache Memory Book" (Harcourt Brace, 1993), the leading reference in the field. Handy also holds an MBA degree from the University of Phoenix. He has performed rigorous technical analysis on the economics of memory manufacturing and sales, discrediting some widely held theories while unveiling other true motivators of market behavior.

Mr. Handy may be contacted at Jim.Handy@Objective-Analysis.com, or by telephone at +1 (408) 356-2549.

Tom Coughlin



Tom Coughlin, President, Coughlin Associates is a widely respected storage analyst and consultant. He has over 39 years in the data storage industry with multiple engineering and management positions at high profile companies.

Dr. Coughlin has many publications and six patents to his credit. Tom is also the author of *Digital Storage in Consumer Electronics: The Essential Guide*, which was published by Newnes Press. Coughlin Associates provides market and technology analysis (including reports on several digital storage technologies and applications such as professional media and entertainment, and a newsletter) as well as Data Storage Technical Consulting services.

Tom is active with SMPTE, SNIA, the IEEE Magnetics Society, IEEE CE Society, and other professional organizations. Tom is the founder and organizer of the Annual Storage Visions Conference (www.storagevisions.com) as well as the Creative Storage Conference (www.creativestorage.org). Tom is also the general chairman of the annual Flash Memory Summit. He is a Leader in the Gerson Lehrman Group Councils of Advisors and a member of the Consultants Network of Silicon Valley (CNSV). For more information go to www.tomcoughlin.com or call Tom at 408-202-5098.



A COMPELLING REPORT BY COUGHLIN ASSOCIATES & OBJECTIVE ANALYSIS

HOW MANY IOPS DO YOU REALLY NEED?

NEW REPORT DETAILS PERFORMANCE NEEDS FROM IT MANAGER SURVEY

Do you 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? Many find it difficult to determine which SSD or flash array to buy or even whether they can get the speed they need from standard HDDs.

Modern storage systems offer a dizzying range of IOPS (from hundreds to millions), latencies, and capacities. This report probes these questions to yield considerable insight into these issues. By surveying over 180 IT managers and end users, Coughlin Associates and Objective Analysis have gained a deep understanding of the performance, capacity, and cost requirements of various applications including IOPS and latency, all of which are all spelled out graphically in this thorough report.

The results of this ongoing survey are revealing. Some applications have more rigid requirements than others, with needs varying according to the type of application. Since the survey (TinyURL.com/IOPSSurvey) is ongoing, this report will be updated in the future.

Order Today!

This report can be ordered online for immediate download at www.Objective-Analysis.com, or via email or fax using the form below.

ORDER FORM

How Many IOPS Do You Really Need? \$5,000 Single-User Price

NAME: _____ TITLE: _____
COMPANY: _____ ADDRESS: _____
CITY: _____ STATE: _____ ZIP: _____
TELEPHONE: (____) _____ E-MAIL: _____

Visa Mastercard American Express

Card Number: _____ Expiration Date: _____
Signature: _____

Make checks payable to: Coughlin Associates, 1665 Willowmont Avenue, San Jose, CA 95124
Telephone: 408-202-5098 Fax: 866-374-6345 Email: Tom@TomCoughlin.com