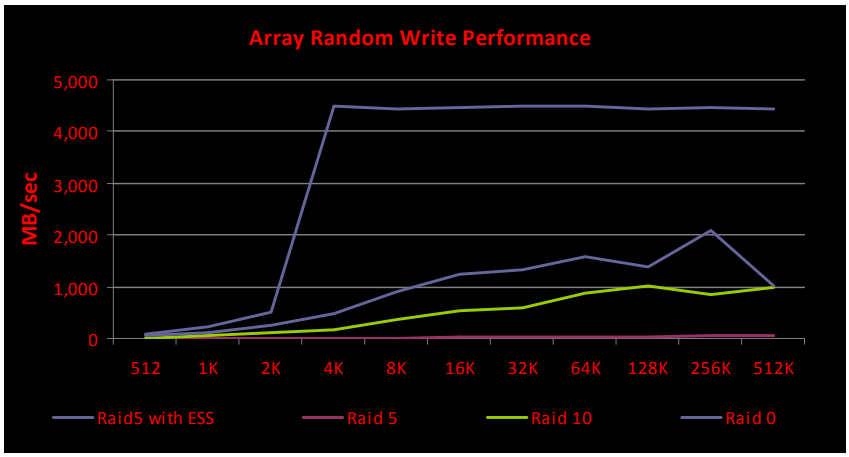

# 16-SSD Array Benchmark Discussion

*Prepared by Doug Dumitru on 8/17/2012*

In the interest of full disclosure, EasyCo tries to be very open about benchmark techniques and what they mean. With arrays of SSDs, benchmarks can produce some very confusing results. This paper is designed to answer questions for those who want to understand the performance of our Enterprise Storage Stack (ESS) in detail.




This chart comes from test configuration with 16 Samsung 830 128GB SSDs. Here is the full test system configuration:

| | |
|---|---|
| CPU: | Intel Xeon E3-1230 v2 – socket 1155, 3.3GHz, Quad Core |
| RAM: | 16 GB DDR3-1600 ECC UDIMM (4x4GB sticks) |
| Motherboard: | Tyan S5512WGM2NR – on-board LSI controller disabled |
| Host Bus Adapters: | 2 x LSI 9207-8i in PCI-e v3 x8 slots |
| Operating System: | Ubuntu 12.04 LTS 64-bit |

This chart shows the random write performance of the above hardware with the drives variously configured RAID-5 using ESS, normal RAID-5, RAID-10, and RAID-0. It incorporates normal optimizations for SSDs (noop scheduler). In addition, the raid5.c module has been patched to optimize linear, aligned writes. This patch is licensed under the GPL and is available upon request from EasyCo.

The charted test was performed at a queue depth of 10. Additional tests at queue depths of 1 and 100 were also performed and are shown in the tabular results below. The benchmark program is an in-house C application that does O_DIRECT IO from multiple threads (using fork()). The source code for bm-flash.c is also available upon request from EasyCo.

The raw results of these benchmarks are included in tabular form at the end of this paper.

**Random Read Performance:**

This is a random write benchmark. Random read results for arrays like these are generally just the performance of the drives. You will note from the raw benchmark tables that ESS does degrade the read performance of very high Q depths moderately. Our experience at EasyCo is that it is very hard to sustain Q depths of 100 or more with real world applications, so we consider this an acceptable tradeoff.

**Random Write Performance:**

Random write performance with ESS is significantly higher than either random read or normal random write measures, and maxes out at 1.15 million IOPS with a Q depth of 100

**Array Size:**

By some standards, this is a very wide array (16 drives). For EasyCo, this is a mid-sized array. The largest we have ever tested on was 84 drives.

**Big Block Size Tests:**

The bm-flash program sets Q depth without regard to block sizes. This means that 512K tests with Q=100 actually produced Q=800 at the drives (Linux cuts up large IO requests). This can swamp the scheduler and result in low benchmark numbers at very big block sizes.

**Linux Kernel Issues:**

The Linux Kernel is not really optimized for arrays of this size. One reason that ESS does so well is that ESS is designed to feed the underlying RAID array and the underlying drives with "perfect" IO requests. This in turn reduces the amount of work that RAID needs to do.

**Applicability to SANs:**

The performance numbers demonstrated here are much faster than any SAN link can practically deliver. Traditional Fiber Channel is only 4Gbits or 400 Megabytes per second. Newer 8Git fiber channel is still 5x slower than the tested 4K random write rate for ESS. Only 6x Infiniband is capable of actually moving this amount of data from system to system.

**Variability of Results:**

Benchmarks with ESS tend to be quick and consistent from run to run. Benchmarks with bare arrays can vary widely depending on the SSDs "preconditioning" at the time of the test.

## Read/Write IOPS Results

| Q Sz / Blk Sz | Read 1 | Read 10 | Read 100 | Write 1 | Write 10 | Write 100 |
|---|---|---|---|---|---|---|
| **Raid-5 w/ ESS** | | | | | | |
| 512 | 7039 | 65847 | 352727 | 7632 | 189507 | 235998 |
| 1K | 6941 | 64667 | 343759 | 13416 | 240528 | 241521 |
| 2K | 6435 | 59936 | 328873 | 5149 | 253990 | 251039 |
| 4K | 5689 | 53019 | 309397 | 344280 | 1146965 | 1115503 |
| 8K | 5272 | 47864 | 232264 | 279040 | 554657 | 570685 |
| 16K | 3878 | 35332 | 214633 | 200681 | 285289 | 285748 |
| 32K | 3282 | 29433 | 139786 | 124841 | 143029 | 142696 |
| 64K | 2523 | 22584 | 84244 | 70831 | 71664 | 71216 |
| 128K | 2388 | 18908 | 48547 | 35801 | 35519 | 35771 |
| 256K | 2212 | 10864 | 21392 | 17845 | 17822 | 17891 |
| 512K | 1911 | 8595 | 9715 | 8934 | 8863 | 8939 |
| 1M | 1287 | 2953 | 4446 | 4464 | 4496 | 4491 |
| **Raid-5 Bare** | | | | | | |
| 512 | 2183 | 15235 | 16404 | 981 | 3259 | 1923 |
| 1K | 2616 | 63868 | 431901 | 1008 | 1752 | 1785 |
| 2K | 6743 | 62820 | 431235 | 999 | 1716 | 1744 |
| 4K | 5845 | 54833 | 431284 | 1003 | 1726 | 1825 |
| 8K | 4651 | 43918 | 351337 | 826 | 1389 | 1384 |
| 16K | 4246 | 39606 | 264558 | 608 | 927 | 428 |
| 32K | 3536 | 32409 | 164261 | 438 | 629 | 217 |
| 64K | 2593 | 23803 | 91366 | 283 | 392 | 114 |
| 128K | 2502 | 20549 | 50597 | 227 | 285 | 68 |
| 256K | 2395 | 15742 | 25691 | 149 | 187 | 44 |
| 512K | 2226 | 10488 | 8878 | 84 | 105 | 34 |
| 1M | 590 | 4063 | 4307 | 58 | 72 | 26 |
| **Raid-10 Bare** | | | | | | |
| 512 | 7296 | 68584 | 395585 | 17205 | 112187 | 114629 |
| 1K | 7207 | 67767 | 395000 | 16860 | 65629 | 58216 |
| 2K | 6705 | 63393 | 395521 | 16514 | 52262 | 51156 |
| 4K | 5826 | 54806 | 397025 | 15326 | 45229 | 47316 |
| 8K | 4630 | 44189 | 346695 | 13978 | 46995 | 66450 |
| 16K | 4257 | 40355 | 263836 | 11133 | 33805 | 50400 |
| 32K | 3546 | 33365 | 162614 | 7653 | 18652 | 24765 |
| 64K | 2591 | 24595 | 89725 | 4942 | 14055 | 14505 |
| 128K | 1604 | 14612 | 47430 | 2712 | 8116 | 9813 |
| 256K | 1455 | 9390 | 25410 | 1571 | 3375 | 4488 |
| 512K | 656 | 4857 | 14228 | 781 | 1993 | 2459 |
| 1M | 619 | 3891 | 7581 | 678 | 1281 | 1469 |
| **Raid-0 Bare** | | | | | | |
| 512 | 7375 | 68586 | 452224 | 23327 | 123449 | 136200 |
| 1K | 7282 | 67656 | 446657 | 23260 | 112626 | 144875 |
| 2K | 6774 | 63167 | 451250 | 22859 | 127851 | 186989 |
| 4K | 5869 | 55133 | 442169 | 21846 | 122237 | 168374 |
| 8K | 4664 | 44014 | 350394 | 18293 | 115245 | 171082 |
| 16K | 4286 | 39828 | 268172 | 13736 | 80444 | 105560 |
| 32K | 3578 | 32702 | 164513 | 8935 | 42511 | 69519 |
| 64K | 2614 | 23910 | 92024 | 5055 | 25139 | 30892 |
| 128K | 1590 | 14206 | 47711 | 2846 | 10974 | 21932 |
| 256K | 905 | 8090 | 25663 | 1574 | 8410 | 11771 |
| 512K | 480 | 4337 | 13670 | 809 | 3228 | 5513 |
| 1M | 382 | 1389 | 5023 | 715 | 2442 | 3197 |

## Read/Write Bandwidth (MB/sec) Results

| Q Sz / Blk Sz | Read 1 | Read 10 | Read 100 | Write 1 | Write 10 | Write 100 |
|---|---|---|---|---|---|---|
| **Raid-5 w/ Flash SuperCharger** | | | | | | |
| 512 | 3.44 | 32.15 | 172.23 | 3.73 | 92.53 | 115.23 |
| 1K | 6.78 | 63.15 | 335.70 | 13.10 | 234.89 | 235.86 |
| 2K | 12.57 | 117.06 | 642.33 | 10.06 | 496.07 | 490.31 |
| 4K | 22.22 | 207.11 | 1,208.58 | 1,344.84 | 4,480.33 | 4,357.43 |
| 8K | 41.19 | 373.94 | 1,814.56 | 2,180.00 | 4,333.26 | 4,458.48 |
| 16K | 60.59 | 552.06 | 3,353.64 | 3,135.64 | 4,457.64 | 4,464.81 |
| 32K | 102.56 | 919.78 | 4,368.31 | 3,901.28 | 4,469.66 | 4,459.25 |
| 64K | 157.69 | 1,411.50 | 5,265.25 | 4,426.94 | 4,479.00 | 4,451.00 |
| 128K | 298.50 | 2,363.50 | 6,068.38 | 4,475.13 | 4,439.88 | 4,471.38 |
| 256K | 553.00 | 2,716.00 | 5,348.00 | 4,461.25 | 4,455.50 | 4,472.75 |
| 512K | 955.50 | 4,297.50 | 4,857.50 | 4,467.00 | 4,431.50 | 4,469.50 |
| 1M | 1,287.00 | 2,953.00 | 4,446.00 | 4,464.00 | 4,496.00 | 4,491.00 |
| **Raid-5 Bare** | | | | | | |
| 512 | 1.07 | 7.44 | 8.01 | 0.48 | 1.59 | 0.94 |
| 1K | 2.55 | 62.37 | 421.78 | 0.98 | 1.71 | 1.74 |
| 2K | 13.17 | 122.70 | 842.26 | 1.95 | 3.35 | 3.41 |
| 4K | 22.83 | 214.19 | 1,684.70 | 3.92 | 6.74 | 7.13 |
| 8K | 36.34 | 343.11 | 2,744.82 | 6.45 | 10.85 | 10.81 |
| 16K | 66.34 | 618.84 | 4,133.72 | 9.50 | 14.48 | 6.69 |
| 32K | 110.50 | 1,012.78 | 5,133.16 | 13.69 | 19.66 | 6.78 |
| 64K | 162.06 | 1,487.69 | 5,710.38 | 17.69 | 24.50 | 7.13 |
| 128K | 312.75 | 2,568.63 | 6,324.63 | 28.38 | 35.63 | 8.50 |
| 256K | 598.75 | 3,935.50 | 6,422.75 | 37.25 | 46.75 | 11.00 |
| 512K | 1,113.00 | 5,244.00 | 4,439.00 | 42.00 | 52.50 | 17.00 |
| 1M | 590.00 | 4,063.00 | 4,307.00 | 58.00 | 72.00 | 26.00 |
| **Raid-10 Bare** | | | | | | |
| 512 | 3.56 | 33.49 | 193.16 | 8.40 | 54.78 | 55.97 |
| 1K | 7.04 | 66.18 | 385.74 | 16.46 | 64.09 | 56.85 |
| 2K | 13.10 | 123.81 | 772.50 | 32.25 | 102.07 | 99.91 |
| 4K | 22.76 | 214.09 | 1,550.88 | 59.87 | 176.68 | 184.83 |
| 8K | 36.17 | 345.23 | 2,708.55 | 109.20 | 367.15 | 519.14 |
| 16K | 66.52 | 630.55 | 4,122.44 | 173.95 | 528.20 | 787.50 |
| 32K | 110.81 | 1,042.66 | 5,081.69 | 239.16 | 582.88 | 773.91 |
| 64K | 161.94 | 1,537.19 | 5,607.81 | 308.88 | 878.44 | 906.56 |
| 128K | 200.50 | 1,826.50 | 5,928.75 | 339.00 | 1,014.50 | 1,226.63 |
| 256K | 363.75 | 2,347.50 | 6,352.50 | 392.75 | 843.75 | 1,122.00 |
| 512K | 328.00 | 2,428.50 | 7,114.00 | 390.50 | 996.50 | 1,229.50 |
| 1M | 619.00 | 3,891.00 | 7,581.00 | 678.00 | 1,281.00 | 1,469.00 |
| **Raid-0 Bare** | | | | | | |
| 512 | 3.60 | 33.49 | 220.81 | 11.39 | 60.28 | 66.50 |
| 1K | 7.11 | 66.07 | 436.19 | 22.71 | 109.99 | 141.48 |
| 2K | 13.23 | 123.37 | 881.35 | 44.65 | 249.71 | 365.21 |
| 4K | 22.93 | 215.36 | 1,727.22 | 85.34 | 477.49 | 657.71 |
| 8K | 36.44 | 343.86 | 2,737.45 | 142.91 | 900.35 | 1,336.58 |
| 16K | 66.97 | 622.31 | 4,190.19 | 214.63 | 1,256.94 | 1,649.38 |
| 32K | 111.81 | 1,021.94 | 5,141.03 | 279.22 | 1,328.47 | 2,172.47 |
| 64K | 163.38 | 1,494.38 | 5,751.50 | 315.94 | 1,571.19 | 1,930.75 |
| 128K | 198.75 | 1,775.75 | 5,963.88 | 355.75 | 1,371.75 | 2,741.50 |
| 256K | 226.25 | 2,022.50 | 6,415.75 | 393.50 | 2,102.50 | 2,942.75 |
| 512K | 240.00 | 2,168.50 | 6,835.00 | 404.50 | 1,614.00 | 2,756.50 |
| 1M | 382.00 | 1,389.00 | 5,023.00 | 715.00 | 2,442.00 | 3,197.00 |

**EasyCo LLC**
220 Stanford Drive
Wallingford
PA 19086 USA
Tel: (+1) 610-237-2000
888-473-7866
Email: sales@EasyCo.com
Web: http://EasyCo.com