SSD WRITE AMPLIFICATION

Application Note

Document #AN0035 – Viking SSD Write Amplification | Rev. A

viking TECHNOLOGY

DRAM MEMORY & FLASH STORAGE
NVDIMM, SSD, DRAM, MCP & CUSTOM
for Embedded, Industrial, Defense & Aerospace

Table of Contents

Table of Tables

1 Introduction and Definition of Write Amplification

Unlike hard disk drives, data writes to an SSD generally involve writing the same data more than once:

- Initially, when saving host data to the flash memory in the SSD for the first time
- and later again when moving valid data into a new area in the SSD during multiple garbage collection cycles (internal writes in the SSD) or during a host TRIM command that erase blocks of deleted data to free up SSD space.

As a result, it is common that more data will be written to flash memory in the SSD than was originally issued by the host system. This disparity is known as write amplification (WA), and it is generally expressed as the number of writes to the flash memory. So for instance, if 4GB of data is written to the flash in the SSD while only 2GB was sent to the SSD from the host computer, this would indicate the write amplification would be a factor of 2.0.

Since flash has a limited lifespan based on the number of writes to the flash memory, write amplification is an undesirable effect for an SSD, since it wears out the flash memory faster by lowering the SSD Endurance (measured as DWPD and TBW), and it also lowers the write performance of an SSD which is performing multiple writes or flash erases within the flash.

2 Factors Affecting Write Amplification (WA)

Factors that can contribute to write amplification are:

1. There is no WA until the SSD is written to full capacity for the 1st time
2. Sequential writes (have lower WA) vs. Random writes (higher WA)
3. Transaction size (the larger the transaction, the lower the WA)
4. Alignment of transactions determined by the host (4K aligned = lower WA)
5. Free Space from SSD over-provisioning that is set by the SSD manufacturer (the more free space = lower WA)
6. Free Space from unused user capacity on the SSD (if TRIM is supported) (more free space on the SSD = lower WA) and the organization of the free space (whether it is randomly spaced out or in sequential blocks, (similarly to hard drive fragmentation)
7. SSD GC efficiency and associated wear leveling algorithms can affect WA (more efficient = lower WA)
8. Data Entropy or randomness of the data bits/patterns in the data being written to the SSD: (lower entropy = lower WA)

9. Features of the SSD or the software managing data sent to the SSD.

Data compression or data de-duplication will lower WA and possibly lower WA to less then 1 by having the SSD writing less data to flash than was issued by the host. As the data bits/patterns randomness (data entropy) from the host goes down, less data will be written to the flash memory, leaving more free space on the SSD.

Note that an SSD WA is not a fixed value, since changing any of the variables listed above will cause a change in the write amplification. It is more appropriate to refer to WA of an application, since WA will also change depending on how the SSD is being used such as the percentage of data being written randomly (write IOPS such as writing lots of small data files to an SSD) vs. sequentially (write bandwidth in MB/sec such as streaming large compressed files to the SSD).

3 Estimation of Write Amplification (WA) for an Application

To estimate the WA for a given application it is best to perform tests with real data:

1. Use a drive at the desired over provisioning level (28%, 7% etc.)
2. Use an application (test, benchmark, real application) that can best represent the read/write, sequential/random, and data entropy the drive will typically see.
3. If the test application cannot limit the capacity accessed, use a test that will simulate a drive at full user capacity (no room for user to add more data).
4. Use a PC utility to retrieve the SMART attribute information from the SSD at start of the test and again after the end of the test.

Table 3-1: SSD S.M.A.R.T. Attributes to Determine WA

SMART Attribute # for Lifetime Writes to Flash	Indicates the total amount of data written to the Flash since the drive was deployed.
SMART Attribute # for Lifetime Writes from Host	Indicates the total amount of data written from hosts since the drive was deployed.

5. Use the calculation/equations below to determine the Write Amplification:

$$\frac{[\text{ Lifetime Writes to Flash (After Test) } - \text{ Lifetime Writes to Flash (Before Test) }]}{[\text{ Lifetime Writes from Host (After Test) } - \text{ Lifetime Writes from Host (Before Test) }]}$$

or

$$\frac{[\text{ Number of erase cycles } \times \text{ Physical capacity in GB }]}{[\text{ Amount of data written from the host in GB }]}$$

4 Free Space and Lowering Write Amplification

The NAND flash memory in solid state drives must be erased before it can store new data. In other words, data cannot be overwritten directly like on a hard disk drive. Instead, SSDs use a process called garbage collection (GC) to reclaim the space taken by previously stored data, so write demands are heavier on SSDs than HDDs when storing the same information. SSDs work better with more free space, so anything that increases free space will keep WA lower. Two of the best ways to expand free space and lower WA is to:

1) Use SSDs with more over provisioning
2) Keep more storage space free on the SSD (for operating system with TRIM)

Also, the organization of the free space on the SSD is a major contributor to WA.

When data is written randomly, the eventual replacement data will also likely come in randomly, so some pages of a block will be replaced (made invalid) and others will still be good (valid). During GC, valid data in blocks like this needs to be rewritten to new blocks. This produces another write to the flash for each valid page, causing write amplification.

With sequential writes, generally all the data in the pages of the block becomes invalid at the same time. As a result, no data needs relocating during GC since there is no valid data remaining in the block before it is erased. In this case, there is no amplification, but other things like wear leveling on blocks that don't change will still eventually produce some write amplification no matter how data is written. So, under normal conditions, an SSD will have a WA very close to 1:1 when writing data sequentially.

If data from the host is compressed and de-duplicated before being sent to the SSD, (such as PDF, ZIP and RAR files; MPEG, MP3, JPEG, and other common formats used for Audio/Visual files) the data writes to the SSD will become more sequential, thus improving free space and lowering WA.

This additional free space enables write operations to complete faster, which translates not only into a higher write speed at the host computer but also into more DWPD and higher TBW, and ultimately a longer-lasting SSD.

5 Reference Documents

- Viking SSD Product Datasheets
  http://www.vikingtechnology.com/products/ssd/ssd.html
- Wikipedia
  http://en.wikipedia.org/wiki/Write_amplification

6 About Viking Technology

Viking Technology develops and delivers innovative high-technology products that optimize the value and performance of our customers' applications. Founded in 1989, Viking Technology has been providing Original Equipment Manufacturers (OEMs) with industry leading designs, engineering, product support and customer service for 20 years. For more information visit http://www.vikingtechnology.com.

7 Revision History

11/29/17		Initial release.

Global Locations		
US Headquarters	India Office	Singapore Office
2950 Red Hill Ave. Costa Mesa, CA 92626 Main: +1 714 913 2200 Fax: +1 714 913 2202	A 3, Phase II, MEPZ-Special Economic Zone NH 45, Tambaram, Chennai-600045 India	No 2 Chai Chee Drive Singapore, 109840

For all of our global locations, visit our website under global locations. For sales information, email us at sales@vikingtechnology.com

viking TECHNOLOGY

DRAM MEMORY & FLASH STORAGE
NVDIMM, SSD, DRAM, MCP & CUSTOM
for Embedded, Industrial, Defense & Aerospace