

SATA HDD Rankings Based on Long Latency Reads

ULINK Analytics Reveal the Impact of Long Latency Reads on Drive Performance and Health

SANTA CLARA, CALIFORNIA, UNITED STATES, August 24, 2023

/EINPresswire.com/ -- Have you ever opened a local file on your computer system that seemed to take a long time to open? If a file takes too long to open, you might quit the opening application and try to open the file again. And if this happens on a regular basis, you might even consider getting a new computer or new drives. This may happen even if your read IOPS, a common indicator of drive speed, is relatively high.

However, chances are that this kind of frustrating user experience will be reflected in something called long latency read count. Behind the scenes, whenever your system sends a request to read data from your drive, it keeps track of the amount of time that the drive is taking to return that information to the system. If the read request takes longer than a certain amount of time, the system will increment the long latency read count. For HDDs, the criterion for incrementing long latency read count is typically when a read command takes longer than 1000ms + (the number of sectors being read / 256)*2ms.

Long latency reads on an HDD are usually caused by one of two reasons: bad sectors, and badly written data. Bad sectors, or damaged disk media, can be caused by defects that happen during the manufacturing process, or after the manufacturing process by loose particles or head crashes. Data may be badly written if the write was weak, written off-track, or overwritten by data belonging to an adjacent track, which may be caused by vibration. These two reasons usually cause the drive to retry the read several times, possibly with slightly altered head positions or electrical strength. And because each retry requires the head to make another rotation, the read may require several rotations, and thus a longer time, to succeed.



Table 1: Disk Drive Long Latency Read (LLR) Rankings (for 4 year old drives)

Rank	Mfg	Model Number	Cap(TB)	Drive Count	Avg POY	Total LLR Count	Total Read Cmts (Million)	LLR Ratio
1	Seagate	st1000vxx0004-1ze101	10	250	4.43	329	9,215	0.04
2	Seagate	st1000vnn0004-1zd101	10	2,765	4.43	4,961	85,782	0.06
3	Seagate	st1000ne0004-1zf101	10	819	4.41	2,840	19,938	0.14
4	Seagate	st1000dm010-2ep102	1	442	4.43	488	2,968	0.16
5	Seagate	st1000ne0004-2gt111	10	176	4.57	725	3,868	0.19
6	WDC	wdc wd2003fzxx-00srla0	2	103	4.47	255	545	0.47
7	HGST	hgst hus722t2ala604	2	315	4.35	1,516	2,850	0.53
8	Seagate	st800ne0004-1zf11g	8	800	4.43	8,265	14,946	0.55
9	WDC	wdc wd8003fzxx-01jddb1	8	661	4.59	55,795	77,148	0.72
10	HGST	hgst huh721212ale604	12	192	4.36	5,498	6,107	0.90

Disk Drive Long Latency Read (LLR) Rankings (for 4 year old drives)

The Importance of Long Latency Read Data in Drive Health Prediction

Furthermore, during the training of our own machine learning algorithms for [ULINK DA Drive Analyzer](#) to predict the remaining useful life of drives, we noticed that long latency read data was given high importance scores by the algorithms. This meant that the algorithms determined that long latency read data was useful in predicting drives' remaining useful life, especially in conjunction with other predictors.

Due to the likely reflection of a slow user experience on this metric and the relationship between this metric and the lifespan of drives, we thought it might be interesting to compare some drive models by their long latency read data to see which ones fared better or worse.

The Data Collection Process

The data we used to rank drive models was SATA HDD health data collected from NAS users in May of 2023. We used drives that had at least two adjacent days of data within that month. For each drive, read command counts were originally reported as lifetime running totals, so we converted these to daily values by calculating daily differences. Long latency read counts were originally reported as daily values and were kept as such.

Furthermore, we excluded drives with no model info. We excluded drives with non-ASCII model info. We excluded drives that did not report long latency read info or read command count info. We excluded drives that were not issued any read commands. We retained drives that had a power-on year value between 4 and 5 years, because we were interested in ranking drives after they had been used for some time, and because we wanted to control for the possible confound of drive age on rankings.

For each drive, we calculated the ratio of long latency reads to read commands, and excluded drives whose ratios were outliers (i.e., 1.5 IQR above Q3). This was done so that we could compare the typical user experience between drive models. We retained drive models with at least 100 drives. For each drive model, we then calculated the ratio of total long latency reads to total read commands, and then multiplied the resulting figure by 1 million. This ratio, the "long latency read ratio," was used to rank the drive models. This left us with 120 drive models and 135,501 drives for ranking.

[Find The Rankings Here.](#)

The drive models with low rank numbers (e.g., rank 1-10) have low long latency read ratios (i.e., only a few read commands result in long latency reads) and are generally expected to lag infrequently on reads. The drive models with high rank numbers (e.g., 111-120) have high long latency read ratios and are expected to lag more frequently on reads than drive models with low rank numbers. The drive models with the 10 highest (worst) rank numbers had exceptionally large long latency read ratios (1.5 IQR above Q3).

Neither the average power-on years, total number of read commands, nor capacity TB were significantly correlated ($p > 0.05$) with long latency read ratio. This means that these variables' confounding effects on the drive model rankings were minimal.

Limitations

Before we conclude, we will acknowledge some limitations of the above ranking. First, we could not control for file size per command, which is a user-specific factor that may have influenced the drive ranking results, as we did not have the data to control for such a potential confound. Second, the drive rankings were based on drives with power-on years equivalent to 4-5 years, so we cannot generalize the rankings to drives older or younger than this.

Summary

To recap, we compared several drive models and ranked them according to their long latency reads, normalized by the number of read commands issued to them. The rankings may offer some insight into how much lag may be felt by users when using certain drive models. The rankings may also be an indicator of potential drive model longevity, although we cannot say that long latency reads are by themselves an indicator of drive failure. Limitations were discussed.

ULINK Technology

DA Drive Analyzer

[email us here](#)

Visit us on social media:

[LinkedIn](#)

This press release can be viewed online at: <https://www.einpresswire.com/article/651552374>

EIN Presswire's priority is source transparency. We do not allow opaque clients, and our editors try to be careful about weeding out false and misleading content. As a user, if you see something we have missed, please do bring it to our attention. Your help is welcome. EIN Presswire, Everyone's Internet News Presswire™, tries to define some of the boundaries that are reasonable in today's world. Please see our Editorial Guidelines for more information.

© 1995-2023 Newsmatics Inc. All Right Reserved.