Intel® RSD v2.1 Sold State Drive (SSD)

Technical Advisory

Summary

One of the new features introduced in Intel® Rack Scale Design (Intel® RSD) v2.1 is the management and control of pooled Non-Volatile Memory Express (NVMe) storage resources by the Fabric Pooled System Management Engine (PSME) software. As part of the normal usage of pooled NVMe, it is strongly recommended to erase the NVMe device between compositions and this action is supported by the Fabric PSME. However, some NVMe devices are limited in the number of secure erase cycles they can perform. If the NVMe device encounters the limit on the number of secure erase cycles, the usability of the storage device may be impacted and this condition may not be covered under the manufacturer's warranty. As such, it is recommended that appropriate due diligence be exercised in the selection of NVMe drives, erase type and erase frequency to ensure that any potential format limitations are not exceeded.

The Intel® SSD DC P36XX and P37XX series (formerly Fultondale), Intel® SSD DC P35XX series (formerly Pleasantdale), and Intel® SSD DC D36XX and D37XX series (formerly Elkdale) are limited in this respect. These drives will only support approximately 100 device erases of either Secure Erase Setting (SES) type SES=1 or SES=2. An SES=0 erase may be used as a substitute (but is considered less secure).

Subsequent Intel® SSD releases, specifically upcoming Intel® SSD DC P45XX/P46XX (formerly Cliffdale) and Intel® SSD OptaneTM P4800X, do not have this design limitation. These devices can support virtually unlimited SES=2 erase cycles and this is the recommended erase type. SES=1 erase cycles are also supported, but users should be aware each SES=1 erase cycle is equivalent to one full drive write, and will impact the drive media life span.

Intel has not evaluated potential similar limitations in non-Intel SSDs.

Q&As:

Is this a SSD FW bug?

No, this is only a design limitation. The Fultondale, Pleasantdale, and Elkdale generations were designed to a use case where drive erase was expected to be an infrequent action. A limit was imposed on the number of times that a drive can be erased using SES=1 and SES=2 NVMe format command options. Once the limit is reached, the SSD cannot be erased via NVMe format command again. When the SSD is to be decommissioned, a secure erase is no longer possible.

Why is it a design limitation?

The format limitation of Fultondale, Elkdale, and Pleasantdale Intel drives is to protect the integrity of the drive. At the time of Intel® SSD design requirements definition, unlimited erase cycles were not considered. Intel® RSD introduces new innovative composability use cases which necessitate increases in the secure erasure limits for future drives.

How long does it take to reach the secure erase limit on the affected drives?

The time to reach the secure erase limit depends on the Intel® RSD usage model and the frequency of node de-composition. Applications which compose and decompose workloads frequently could reach the limit quickly. Roughly 100 secure erase cycles are supported.

What happens to the drive when the secure erase limit is reached?

Once the drive reaches its secure erase limit, further attempts to secure erase the drive with either option SES=1 or SES=2 will return an error and the drive will not be reformatted. All existing data on the drive remains on the drive and is not erased. Using the SES = 0 option will not return an error as the limit for secure erase is not encountered on SES=0. However, SES=0 erase leaves data in place on the drive, and is considered a less secure erase option.

For the Intel drives, if the secure erase limit is reached, is the SSD covered under warranty?

No, exceeding the secure erase limit is not covered under the warranty.

Was the Intel® RSD use case defined when the Intel® SSDs were developed?

The requirements as defined for (Fultondale, Elkdale, and Pleasantdale) were established to address the vast majority of the use cases understood in the market. Intel® RSD is bringing to market innovative new use cases around composable infrastructure. These use cases were defined after the drive design requirements were set.

Why would the SES options be needed?

In a composable infrastructure like Intel® RSD, the erasure of an NVMe storage device upon decomposition is done as not to expose user data from one composed system to next. This feature can be enabled or disabled based on customer requirements. The Secure Erase Setting (SES) option is a mandatory parameter on the Format NVM administrative command which is defined by the NVM Express Specification (nvmexpress.org). Three options are defined:

A) SES=0 is a non-secure erase operation which marks the media unused, but leaves data in place.

B) SES=1 is a User Data Erase which overwrites and erases all data on the drive. The default NVMe format option used by the Fabric PSME is "Secure Erase Setting (SES) option 1 – Erase User Data".

C) SES=2 is a Cryptographic Erase which makes the existing data inaccessible by deleting the current cryptographic key. The NVMe specification also allows that a drive may do a Cryptographic erase (SES=2) when a User Data Erase (SES=1) is requested if the drive is encrypted.

The Intel® RSD Fabric PSME utilizes the SES=1 option after node decomposition to ensure that data from one composed node is not exposed to the next when an SSD is re-used.

Can the SES feature be disabled?

Yes, refer to the Intel® RSD documentation on www.intel.com/intelrsd

Why does the Intel® RSD use case use/require so many formats?

In an Intel® RSD use case, there may be multiple tenants and applications. Reformatting the drive between node compositions is done so that data from the application or tenant "A" is not exposed to the application or tenant "B" when a node is composed and an SSD re-assigned. This is in many ways similar to a drive being physically decommissioned from a server and wiped before disposal or re-use in another server or application.

What happens to the data when the secure erase limit is reached?

If the secure erase limit is reached on an SES = 1 or SES =2 operation, the data on the drive remains intact and can still be accessed. The drive can still be non-securely erased using the SES=0 option.

How does this impact an Intel® RSD compliant solution?

The CTS does not require SES operational parameters for conformance.

Is there a similar limitation with non-Intel SSDs?

Unknown. Implementations vary across vendors. It is recommended to verify with the SSD vendor whether or not there is a limitation with secure erase.

How can current non-Intel SSDs be validated within the current known secure erase limits without impacting the SSD functionality in RSD use cases?

It is recommended to verify with the SSD vendor whether or not there is a limitation with secure erase. It is also recommended to test a small number of drives to confirm the secure erase operations with the Intel® RSD use case prior to deployment at scale.

§