Enhancing the Write Performance of Intel SSDs through Over-Provisioning

Application Note

Abstract

This application note aims to help users decide whether to adopt over-provisioning and choose its level in their environments.

Table of Contents

Product Models Covered by This Document

This application note applies to the following product models:

- **EonStor DS storage systems**

Note: SSD is available on selected Infortrend RAID subsystems. For detailed information on availability, please refer to individual product pages on www.infortrend.com.

Summary

This application note aims to help users decide whether to adopt over-provisioning and choose its level in their environments. Over-provisioning is a technique to improve the IOPS of an SSD by increasing the size of its reserved space. Over-provisioning is particularly beneficial for systems with small transfer size (< 16KB), random access behavior, and read/write mixed I/O. The following sections describe the concept of over-provisioning, its benefit for RAID systems, configuration steps for Intel SSD, and a performance improvement analysis.

Note: SSD is available on selected Infortrend RAID subsystems. For detailed information on availability, please refer to individual product pages.

What is Over-Provisioning?

Over-provisioning (sometimes spelled as “OP”, “over provisioning”, or “overprovisioning”) is the difference between the physical capacity of the flash memory and the logical capacity presented through the operating system (OS) as available for the user. [1]

It can also improve performance by giving the flash controller additional buffer space for managing program/erase (P/E) cycles and improving the probability that a write operation will have immediate access to a pre-erased block. [2]

During the garbage collection, wear-leveling, and bad block mapping operations on the SSD, the additional space from over-provisioning helps lower the write amplification when the controller writes to the flash memory. [1]

Note: NOT all drives can be configured by endusers for Over-Provisioning. [2]

Why is Over-Provisioning Important?

Random read and write IOPS require SSD- and CPU-intensive operations. Random writes in particular depend on multiple SSD-specific and host-specific parameters. When reads and writes are mixed, total IOPS drop significantly due to the slower writes.

In addition, IOPS degrade depending on the amount of random accesses. With higher random write access, NAND pages and blocks are written to more frequently and reclaimed in the background through erasure. Without the reclaim activity, the SSD will not be able to service host IOPS efficiently. Therefore, the percentage of random accesses impacts reclaims activity, thereby impacting the host IOPS rate.

Therefore, a higher available spare area helps to free up CPU cycles from reclaim activity, making them available for host IOPS. Therefore, available spare area can directly impact random write and mixed read/write IOPS. You can increase the spare area by over-provisioning the SSD. [3]

Configuring Over-Provisioning for Intel SSDs

The Over-Provisioning setting of an SSD can be modified by a user. It must be performed on an SSD that is in a completely clean state. This can be an SSD that is fresh out of the box that has never been used or by secure erasing the SSD. [4]

You can use Intel® Solid-State Drive Toolbox for performing a secure erase process if the SSD was not in clean state. To secure erase an Intel SSD, download Intel® Solid-State Drive Toolbox 2.0 from http://www.intel.com/go/ssdtoolbox and follow the steps for performing a secure erase on the SSD. (Industry tools are also available to secure erase an SSD using ATA commands.)

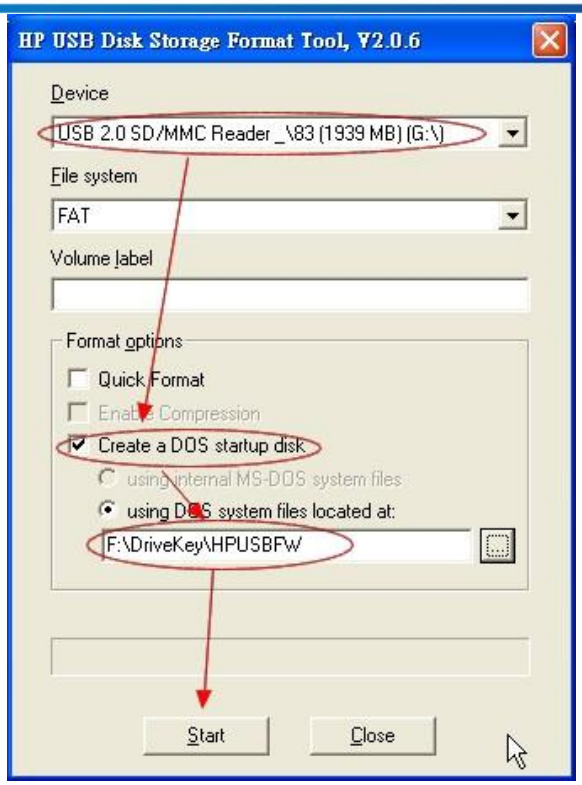
Once the SSD is in a clean state, reduce the usable capacity (which increases the spare area) using one of two methods:

- Issue the SET MAX ADDRESS command (part of the ATA specification) to set the maximum address the operating system (OS) can see. HDPARM* and HDAT2* are third-party industry tools that can be used to issue this command.
- Define a partition that is less than the maximum available capacity of the SSD. This option can be found in the OS drive configuration tools.

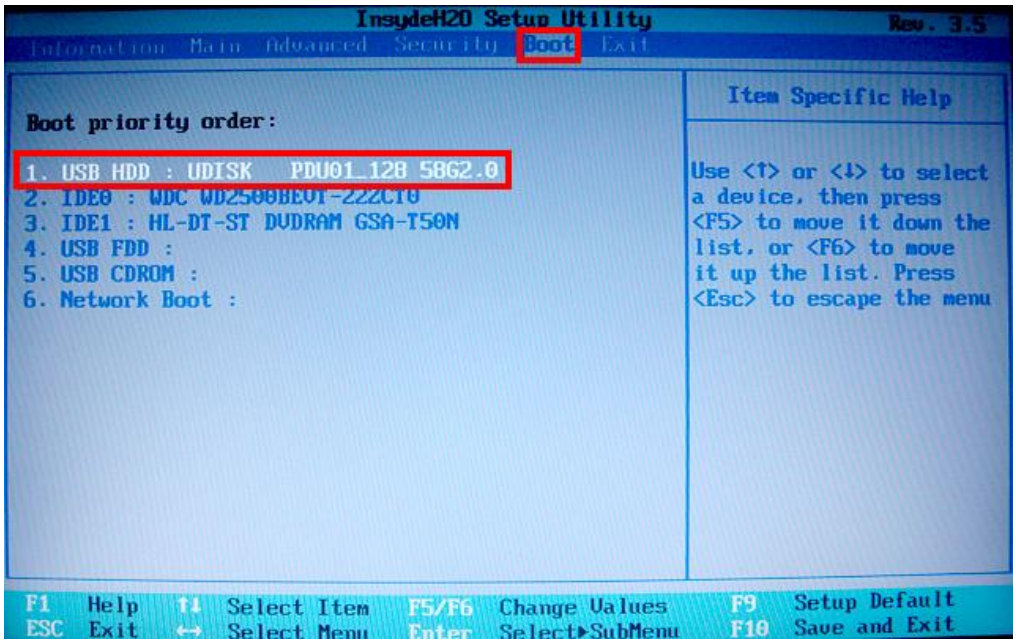
Both SET MAX ADDRESS and partitioning will reduce the user addressable space and allow the SSD to use the remaining space as part of the “ready to be written” resource pool. [5]

The steps below describe how to configure the HDAT2 program for over-provisioning in a Windows-based PC.

1. Prepare a DOS bootable USB memory stick by following these steps.
    A. Download the HPUSBFW tool from this link.
    B. Double click “SP27213.exe” and install the HPUSBFW tool into the desktop PC.
    C. Unzip HPUSBFW.exe which should be located on C:\drivekey after the previous step. DOS book image files will be unzipped into the C:\drivekey\HPUSBFW directory.
    D. Open HPUSBFW and follow the steps that are indicated in the screenshot below. *Note that after clicking “Start,” all existing data in the USB memory stick will be lost due to re-formatting.*



2. Download the HDAT2 program from this site and copy HDAT2.exe into the USB thumb drive.
3. Assign the USB thumb drive as the primary boot loader as shown in the screenshot below. A user may enter the BIOS setup menu by pressing a key (such as F2 or F10) during the boot.



4. Attach the SSD to a SATA hard drive slot and reboot the computer.
5. The computer should be booted through the USB, and a command line interface will appear on the screen.
6. Type in “HDAT2”, and then press the Enter key. The SSD model name and capacity will appear on the screen.
7. Press the Enter key and move the cursor to “set max (HPA) menu”. Press the

Enter key and the “set max address” menu will appear.

8. Press the Enter key. The three areas, described below, will appear:
   - Native area = the maximum capacity of the drive at the default factory state
   - User area = the maximum capacity after enabling over-provisioning (a.k.a. manually set max LBA)
   - Hidden area = Native area – User area. The size of the hidden area should be indicated on the screen.
9. Multiple the sector number with the over-provisioning level (percentage). For example, if the over-provisioning level is 20% on a 160GB SSD device, then the user area after over-provisioning will be 312602976 x (1-0.2) = 250082380.
10. Press Insert and enter the new value. Press “S,” then select “Y (Yes).” Then press any key to complete the process.
11. Reboot the computer (this time, use the local OS disk to boot up).
12. Check the new capacity through the Disk Manager in the Windows operating system.

Note: if you want to use a different drive configuration, you can do so by typing in different values in Step 9 (for example, using a 96G drive instead of a 160G drive).

Preparing an SSD for Performance Measurement

The NAND Flash Memory in an SSD is blank at factory default settings. If you measure the drive performance in that state, the result can be exceptionally high because executing write commands does not require internal physical defragging. [6]

To accurately measure performance, you should establish conditions that represent the target workload. You may connect SSDs directly to the server or install them into a storage subsystem.. In this example, we have installed SSDs to an Infortrend RAID enclosure (EonStor DS S16F-G2850) to simulate the workloads of a real-life system:

The following tools are used in this example:

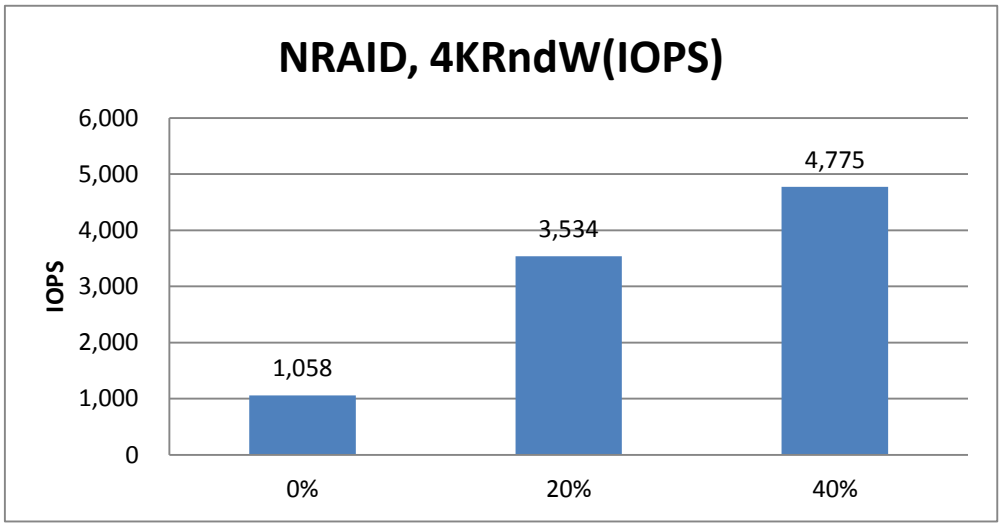
- EonStor DS S16F-G2850 x 1 (Firmware version: 3.91B02 or later)
- Intel 320 Series SSD 160GB x 1
- Drive Tray (compatible with 2.5” SSD) x 1
- IOMeter (a measurement program that is downloadable from this site)

1. Fill a hard drive slot of the RAID enclosure with an SSD.
2. Create “NRAID” for this SSD, and then map it to the host (create LUN mapping).
3. After the host has recognized the LUN mappings run IOMeter with the following parameters:
    a. Configure the target workload as follows: 4K 100% random write test, 256 outstanding.
    b. Align the access specification sector boundaries with the transfer request size.
    c. Set the first run time to 120 minutes to randomize the SSD access.
    d. Run the test again for 5 minutes to gather accurate sustained performance data.

Performance Analysis

Write IOPS

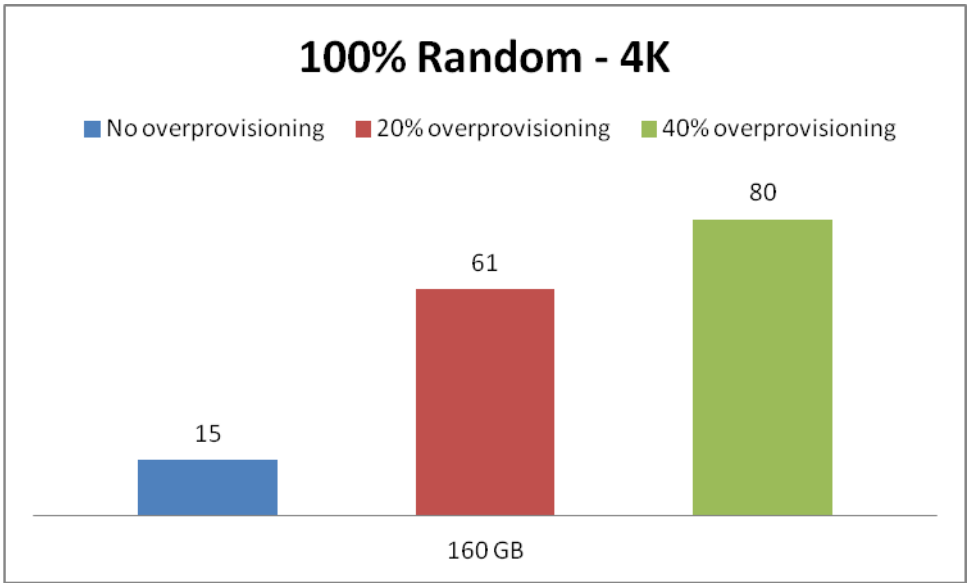
The figure below shows the write IOPS performance results with different levels of over-provisioning by a single SSD configured as NRAID. Performance improves by as much as 3x by allocating 20% of the total space to over-provisioning, and 4x by allocating 40%.



Measurements were made using the following equipments: a desktop system with Intel Xeon E5405 @ 2.0GHz (2 processors) and 16GB PC2-5300 DRAM running Microsoft Windows* 2008 R2 64bit O/S and an Infortrend ESDS S16F-G2850 Subsystem with 16 160GB Intel 320 Series SSDs. Any difference in system configurations may affect the result.*

Endurance

Although the endurance of an over-provisioned SSD is not inside our test criteria, it is also worthy to note that SSDs receive long-term benefits from over-provisioning due to lower write amplification. The figure below, taken from an official document from Intel [3], shows how the endurance levels on Intel SSD 320 Series increases by adopting over-provisioning. For example, when the transfer size of a 160GB SSD is set at 4 KB, there is a 4X endurance improvement from 0% to 20% over-provisioning (total writable capacity changes from 15TB to 61TB), and a further 1.3X improvement from 20% to 40% over-provisioning (61TB to 80TB).



Endurance number depends on the host's workload.

Summary

The table below summarizes the changes in usable capacity, random write IOPS, and cost performance of an SSD (160GB Intel 320 Series) under different over-provisioning settings.

Configuration	OP 0%	OP 20%	OP 40%
Usable Capacity(GB)	160	128	96
Random Write IOPS		3X	4X
Endurance		4X	5.3X
$ / Random Write IOPS	$0.317	$0.095	$0.070

SSD pricing may vary by vendor. Please consult Infortrend or Intel for more information.

References

1. Wikipedia entry: Write Amplification. Wikipedia. Retrieved 2012-05-30.
2. "Definition of over provisioning (SSD overprovisioning)". Search SolidState Technology. Retrieved 2012-05-30.
3. "Random workload characterization of 320 series SSD (478927_Intel_SSD_320_Series_Random_Workload_Characterization_325774_rev003_FINAL_2011-09.pdf)". Intel. Retrieved 2012-05-30.
4. Intel High Performance SATA Solid-State Drive. Intel. Retrieved 2012-05-30.
5. "Using HDAT2 to do over-provisioning (Use HDAT2 To Do Over-provision.pdf)". Retrieved 2012-05-30.
6. "Measuring the performance of SSD using IOMeter (Performance_SSD_IOMeter_Performance_Measurement_Process_320672_09Jun.pdf)". Retrieved 2012-05-30.