Power Failure Protection

Application Note AN003

February 2012

Corporate Headquarters: 39870 Eureka Dr., Newark, CA 94560, USA ♦ Tel:(510) 623-1231 ♦ Fax:(510) 623-1434 ♦ E-mail: info@smartstoragesys.com
Flash Design Center: 2 Robbins Road, Westford, MA 01886, USA ♦ Tel:(978) 303-8500 ♦ Fax:(978) 303-8757
Flash Design Center: 2600 W. Geronimo, Chandler, AZ 85244, USA ♦ Tel:(480) 792-8900 ♦ Fax:(480) 792-8901
Asia: Plot 18, Lrg Jelawat 4, Kawasan Perinudstrian Seberang Jaya 13700, Prai, Penang, Malaysia ♦ Tel:+604-3992909 ♦ Fax:+604-3992903

Table of Contents

1 Introduction

Enterprise and industrial storage system power supplies are designed with the highest reliability in mind, and subsystem designers are careful in selecting the most reliable regulation components. However, the reality is that power to data storage drives does occasionally fail. Since Enterprise- and industrial-class SSDs maintain mission-critical data, it is unacceptable for previously stored data to be lost or for data in flight to nonvolatile storage to be corrupted.

It is vital that these SSDs are designed to survive power reductions and outages without risk to the data itself.

Power loss events range from momentary loss of regulation (transient brown-out condition) to loss of all power for an extended period of time. Such events can be caused by failure of the facility's supply grid or UPS unit, failure of the system's power supply (including fusing and cabling), failure of the SSD's voltage regulation components, or mechanical failure of PCBs or connectors due to vibration, heat, or impact. Power failure risk at the SSD level depends only partly on the power delivery redundancy measures in place; power failure can cause system latency (when the drive needs to rebuild mapping tables) or permanent data loss.

This application note discusses the multiple methods of addressing power failure risk and the superior approach employed by SMART Storage Systems' XceedIOPS, XceedStor and Xcel-200 SSDs, referred to in the remainder of this application note as SMART SSDs.

2 Data Loss in Power Failure Scenarios

When programming a NAND flash page, the program operation must complete to ensure the data is stored reliably within the page. Data is at risk if flash memory cells are in the process of being programmed when power to the drive is lost. The risk is compounded for MLC NAND flash memory, which uses the same physical page of memory cells to store two logical pages of data. When power is lost during program operation of the upper page, valid data already stored in the lower page cells can be damaged. This is typically referred to as lower-page data corruption.

Solid state drives have three areas of potential data loss or corruption when system power fails:

- **Loss of data:** This can occur due to the implementation of write caching (also called "write back" or "write behind") to achieve peak performance. In this case, the host system is informed that a write operation has completed when in fact it is still in process. If power fails while the controller is "catching up" with the write operation, the data in the write buffer is not yet hardened and can be lost. When the data is requested later by the host, the controller can either report the data irrecoverable or (depending on the controller design) it can deliver a previous "stale" version of those sectors to the host. In the latter case, this translates to silent data corruption, since the host system is not informed that the data delivered is incorrect.
- **Loss of mapping information:** Every SSD controller uses mapping information to translate from the host's logical LBA addresses to physical flash memory locations. Mapping information must be created and maintained if the data is to be later retrieved from the drive, and must be updated whenever new data is

written to a previously written LBA. If the mapping information is lost when power fails, the drive may show data corruption, deliver stale (corrupted) data or may not be capable of supporting logical I/O on the next power up.

- **Lower page corruption:** MLC or E-MLC NAND flash uses each physical page to store the data of two logical pages; each memory cell represents two bits. The lower page (the logical page addressed by the lower of the two addresses) is programmed first, followed by the upper page. When programming the upper page, programming voltages are applied to the same cells already storing valid data in the lower page. If power fails while the upper page is being programmed, data in that page is lost, and already-stored data in the lower page is corrupted as well. When that lower page data is requested later by the host, the SSD will report the data irrecoverable.

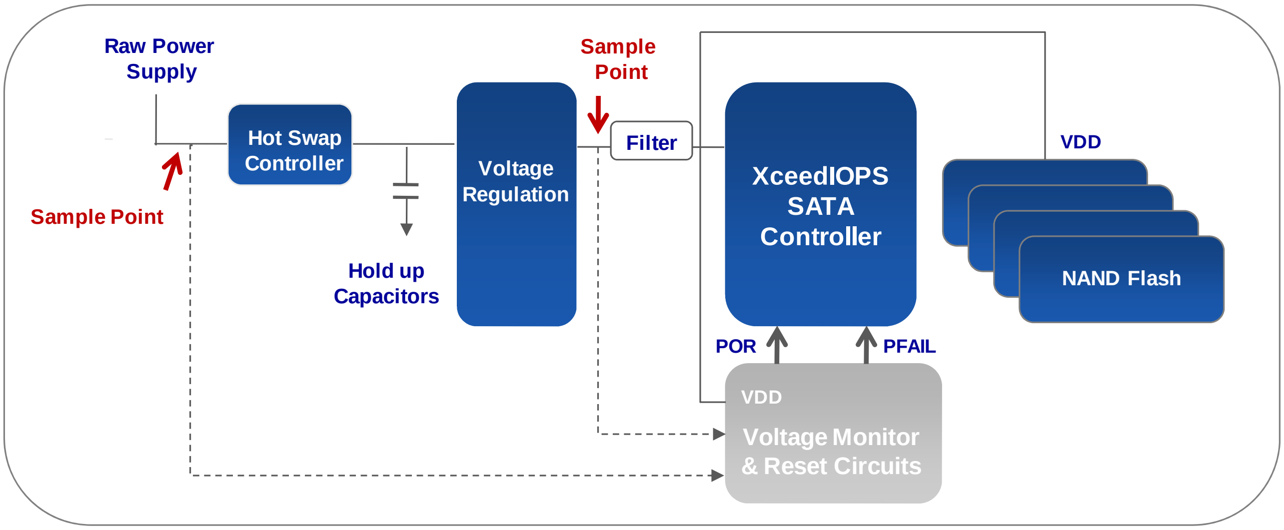
The Unrecoverable Bit Error Rate (UBER) for an enterprise-class SSD is specified to be $\leq 10^{-16}$, according to the JEDEC JESD218[1] specification. To achieve this error rate, an enterprise-class SSD must employ a power failure protection circuit to ensure that that loss of data and mapping information will not occur when power is lost.

3 Power Failure Circuitry in SSDs

Most Enterprise- and industrial-class solid state drives rely on power failure circuitry that monitors the supply voltage and generates an “early warning” signal to the SSD controller if the voltage drops below a predefined threshold. A secondary voltage hold-up-circuit is implemented to ensure the drive has power for a sufficient time to harden data whenever that warning is received.

As an example of an Enterprise-class implementation, Figure 1 below illustrates the power failure circuit block diagram of the XceedIOPS SSD.

Figure 1: Power Failure Circuit Block Diagram of XceedIOPS SSD



---

[1] JEDEC JESD218 specification is available for download http://www.jedec.org/

The secondary voltage source can be either a high capacity supercapacitor, a bank of discrete capacitors or even a battery (although no known SSD on the market uses this approach). These different implementations are not the same from a performance and reliability standpoint; some are better suited for usage in Enterprise-class SSDs than others. It should also be noted that writes are not accepted by the drive until the secondary voltage source has been sufficiently charged to protect against loss of data upon power failures.

Descriptions and relevant tradeoffs of a supercapacitor solution and a bank of discrete capacitors in an Enterprise-class SSD design are presented below.

3.1 Supercapacitor

A supercapacitor is an electrolytic capacitive charge storage device. It is capable of storing a large amount of energy in a comparatively small three-dimensional space. A generic supercapacitor-based voltage hold-up circuit is consistent with the block diagram shown in Figure 1.

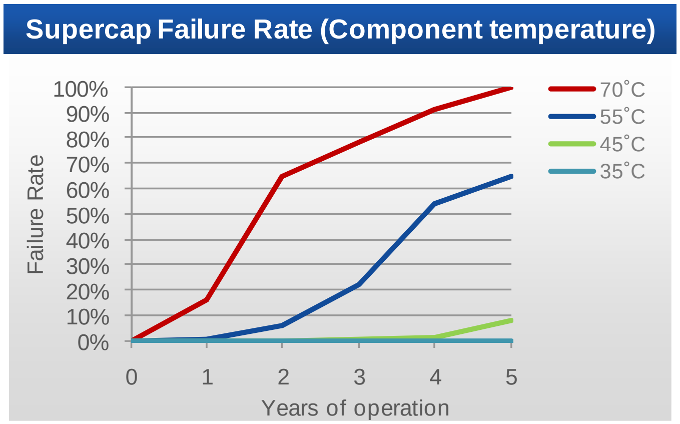
Designing a supercapacitor-based power failure protection circuit is easy to do, and many SSDs employ the approach for this reason. Unfortunately, there are a number of concerns related to long term supercapacitor reliability that makes this component type questionable for Enterprise-class SSDs.

Supercapacitors are typically Aluminum Electrolytic Capacitors. This type of capacitor is known for a high capacitance-to-size ratio, and is an attractive choice for applications requiring large bulk capacitance like a solid state drive. However, like all electrolytic capacitors, supercapacitors suffer from a well known set of deficiencies with regard to long term reliability. Supercapacitors “wear out”, resulting in reduced capacitance over time. They use a wet electrolyte, and the packaging is subject to ongoing losses via leakage and diffusion.

The performance of the supercapacitor degrades slowly with electrolyte loss, until the onset of total failure occurs with little or no warning. In addition, loss rate increases with higher operating voltage, and in higher operating and non-operating temperature environments. For every 10°C of ambient operating temperature rise, the life expectancy of a supercapacitor can be cut approximately in half.

A thorough analysis done by SMART Storage Systems has shown that supercapacitors are not reliable enough to meet the required reliability standards for the high-performance enterprise and industrial computing markets served by SMART SSD product line.

fFigure 2 shows an example of a supercapacitor reliability projection, based on component life test data.



F

igure 2: Supercapacitor Failure Rate by Temperature

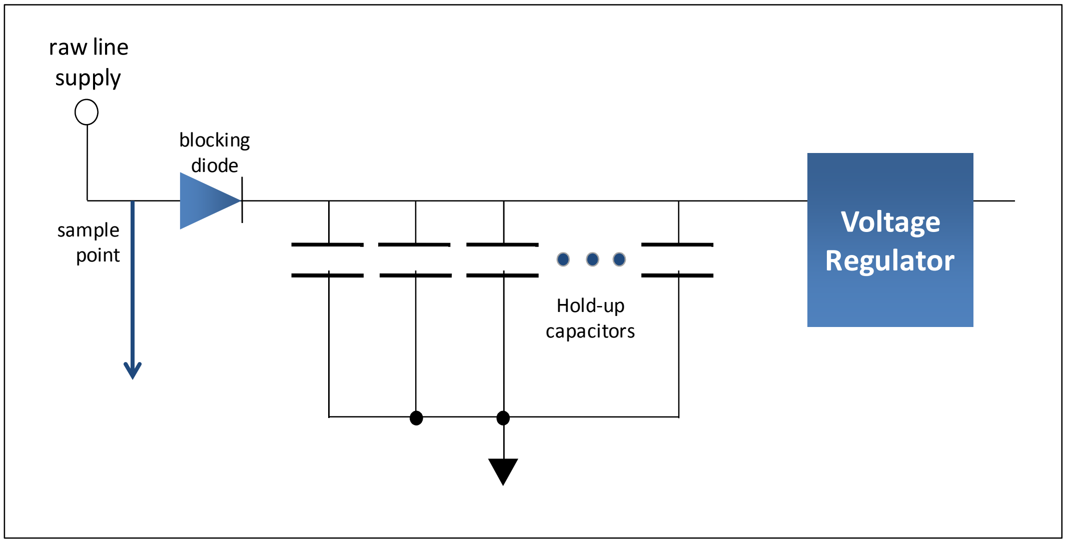
Due to the reliability concerns associated with this capacitor type, it is imperative **t**hat the SSD constantly monitor the capacitor's operating capabilities to ensure continued reliable operation as the SSD ages. This is done by periodically measuring the supercapacitor's charge/discharge ("Hold Up") time under a controlled load.

The challenges associated with performing this test seamlessly and transparently to the host system are many. Because the secondary power system is under a "live test" during hold up time measurements, the SSD must harden data prior to testing (in case the test fails). This operation and the test time itself almost always result in extended latencies (as much as 100ms or more) for host commands issued during the test interval.

3.2 Discrete Capacitors

This approach requires more design expertise, but overcomes the supercapacitor limitations. A discrete capacitor-based voltage hold-up circuit employs a bank of discrete capacitors connected in parallel, shown in Figure 3 below. This is the approach employed by SMART SSDs.

Figure 3: Discrete Capacitor Rail Hold-Up Sub-Circuit



SMART SSDs utilize either Niobium Oxide or Polymer Tantalum capacitors.

These discrete capacitors do not employ a "wet" electrolyte and are not susceptible to the leakage related issues that plague supercapacitor technology. Niobium and Polymer Tantalum capacitors are rated to 85ºC, providing considerably more temperature operating range than is typical for a supercapacitor (70ºC). As a result of these factors, a discrete component based hold up circuit is better able to meet the demands of enterprise and industrial computing environments.

Another advantage of discrete capacitors over supercapacitors is that they are highly predictable and reliable. Provisioning can be selected so that it is optimal for the SSD’s needs over its lifetime. Derating and significant over provisioning is not required. SMART SSDs discrete capacitor circuit has been carefully optimized to meet the holdup requirements.

Discrete capacitor implementations require more careful design. Lacking the compactness of supercapacitors, the capacitance-to-size ratio of a discrete solution is less space-efficient. To provision sufficient capacitance with a discrete capacitor solution requires packing components as densely as possible onto the PCBA board.

Tradeoffs are necessary to achieve a balance between cost, reliability and operating margin. Not all SSD manufacturers have the knowledge and experience to navigate these tradeoffs correctly.

4 Summary

The Early Warning / Rail Hold-Up approach is the most reliable solution for a backup power circuit in an Enterprise-class or Industrial SSD design. SMART Storage Systems’ SSD design team has in excess of 1000 man-years of experience in the design of memory storage and data storage devices. SMART’s expertise has allowed us to identify the risks associated with supercapacitor reliability, and implement a discrete capacitor circuit which is superior in every regard. Discrete capacitors do not degrade over time or with elevated temperatures and can operate reliably in environments up to 85ºC.

Implementing a discrete capacitor backup power circuit into the design differentiates SMART SSD products from their competitors in a way that provides meaningful advantage to our customers.

Disclaimer:

No part of this document may be copied or reproduced in any form or by any means, or transferred to any third party, without the prior written consent of an authorized representative of SMART Storage Systems. ("SMART"). The information in this document is subject to change without notice. SMART assumes no responsibility for any errors or omissions that may appear in this document, and disclaims responsibility for any consequences resulting from the use of the information set forth herein. SMART makes no commitments to update or to keep current information contained in this document. The products listed in this document are not suitable for use in applications such as, but not limited to, aircraft control systems, aerospace equipment, submarine cables, nuclear reactor control systems and life support systems. Moreover, SMART does not recommend or approve the use of any of its products in life support devices or systems or in any application where failure could result in injury or death. If a customer wishes to use SMART products in applications not intended by SMART, said customer must contact an authorized SMART representative to determine SMART's willingness to support a given application. The information set forth in this document does not convey any license under the copyrights, patent rights, trademarks or other intellectual property rights claimed and owned by SMART. The information set forth in this document is considered to be "Proprietary" and "Confidential" property owned by SMART.

ALL PRODUCTS SOLD BY SMART ARE COVERED BY THE PROVISIONS APPEARING IN SMART'S TERMS AND CONDITIONS OF SALE ONLY, INCLUDING THE LIMITATIONS OF LIABILITY, WARRANTY AND INFRINGEMENT PROVISIONS. SMART MAKES NO WARRANTIES OF ANY KIND, EXPRESS, STATUTORY, IMPLIED OR OTHERWISE, REGARDING INFORMATION SET FORTH HEREIN OR REGARDING THE FREEDOM OF THE DESCRIBED PRODUCTS FROM INTELLECTUAL PROPERTY INFRINGEMENT, AND EXPRESSLY DISCLAIMS ANY SUCH WARRANTIES INCLUDING WITHOUT LIMITATION ANY EXPRESS, STATUTORY OR IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

©2012 SMART Storage Systems. All rights reserved.

Corporate Headquarters: 39870 Eureka Dr., Newark, CA 94560, USA ♦ Tel:(510) 623-1231 ♦ Fax:(510) 623-1434 ♦ E-mail: info@smartstoragesys.com
Flash Design Center: 2 Robbins Road, Westford, MA 01886, USA ♦ Tel:(978) 303-8500 ♦ Fax:(978) 303-8757
Flash Design Center: 2600 W. Geronimo, Chandler, AZ 85244, USA ♦ Tel:(480) 792-8900 ♦ Fax:(480) 792-8901
Asia: Plot 18, Lrg Jelawat 4, Kawasan Perinudstrian Seberang Jaya 13700, Prai, Penang, Malaysia ♦ Tel:+604-3992909 ♦ Fax:+604-3992903