RAID Maintenance Free Backup Basic Functionality

June 2013
Rev 1.0

A raid controller or raid module with cache memory support will allow a raid volume's cache policy to be set to write back mode. Setting a raid volume property to write back mode can greatly enhance performance by allowing the operating system and application to go on to the next operation without having to wait for a signal from the raid controller that the data has been written to disk. However, this is inherently dangerous to critical data should the power fail during such an operation because data sitting in cache waiting to be written to disk (dirty cache) can be lost. raid controllers have relied on batteries in the past in the event of a power loss to provide enough energy to keep the raid controller's cache memory alive until the power is restored. A recent innovation to address the short life span of batteries is the adoption of using a combination of NAND flash memory and super-capacitors to protect dirty cache by providing power long enough to the cache memory to copy data from dirty cache to flash memory, to be held there until power is restored when the data can be moved back into controller cache memory. This document describes the function of the Intel RAID® Maintence Free Backup implementation of this NAND flash cache protection option.

The Intel® RAID controllers and modules that support on-controller cache memory also support cache data backup using either a raid battery (BBU) or a Maintenance Free Backup Solution (RMFBU). The RMFBU is based on using capacitors to hold the memory circuit alive until cache data can be written to flash memory. This solution is designed to protect data in-flight through controller cache in the event of a power interruption. The Intel Remote Maintenance Free Backup Unit (RMFBU) solution is designed to be remotely mounted and cannot be mounted directly on the raid controller. This allows the solution to be located in an area within the system chassis with a thermal environment conducive to longer life of the capacitors being used. Please review the RMFBU user guide for information related to environmental specifications.

Benefits of the RMFBU option include:

1. Super-capacitors have a longer usable life span than batteries which reduces maintenance and replacement costs.
2. NAND Flash memory is non-volatile which can hold data for up to 10 years without refreshing it.
3. The super-capacitors do not have the chemical volatility of batteries which carry a risk of exothermic events.
4. The RMFBU solutions do not require the comprehensive certifications required of batteries in order to import and export them.
5. The RMFBU can constantly maintain the raid volume in write back mode even when doing learn cycles.
6. The recharge rate of the Super-Capacitor is minutes as compared to hours for a battery.

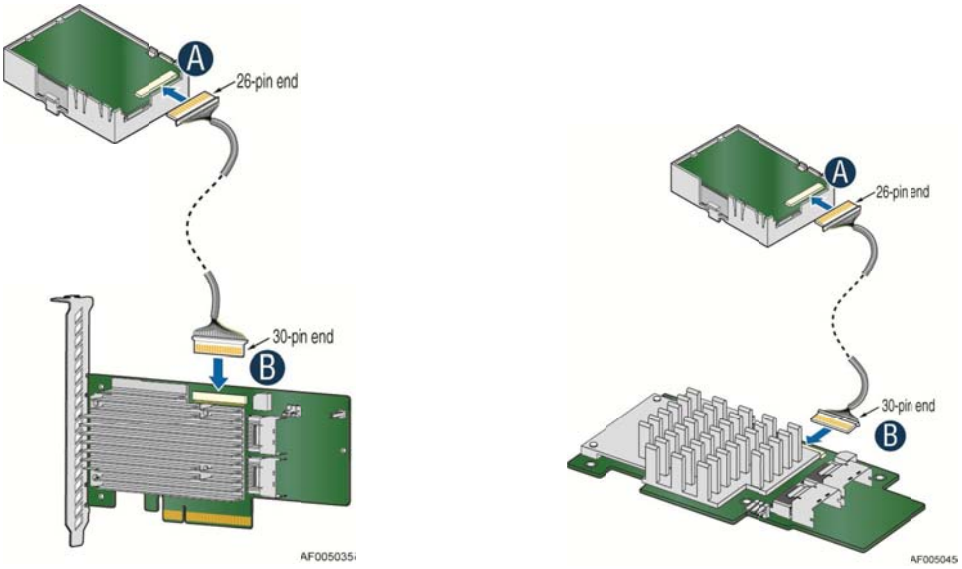


Figure 1 - RMFBU Connection to a Controller or Module

RMFBU (A) and RAID controller (B) **RMFBU(A) and RAID Module (B)**

The RMFBU solution includes a Transportable Flash Module (TFM) connected to a super-capacitor package, both of which are housed in a plastic retention bracket that allows attachment to a mounting point within the server chassis. The TFM includes circuitry that monitors capacitor voltage and charge level, and the TFM includes on-board NAND flash memory that acts as a repository for controller cache in the event of a power loss situation.

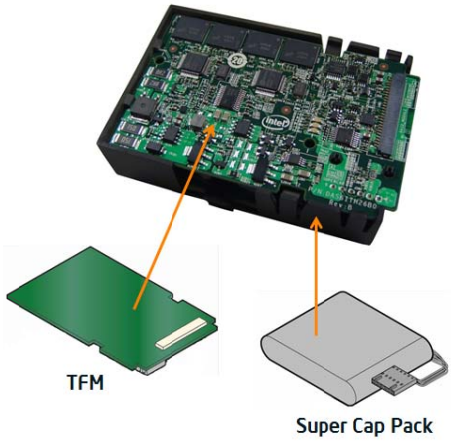


Figure 2 RMFBU Components

These raid controllers and modules include up to 1GB of DDR3 1333 memory, of which a large portion can be used as controller cache. The flash on the TFM consists of four 4Gb x 8bit width NAND flash parts yielding 2 GB of NAND flash memory that can store the cache contents in event of a power loss. The super-capacitor pack utilizes three 6.8F, 13.2V capacitors that provide the power to offload cached data from the controller DRAM to the nonvolatile flash memory on the TFM, if a power failure or outage

occurs and the arrays are set in write-back mode. Transfer of cache memory is accomplished using dual USB connections between the raid controller/module's ROC chip USB port and the USB port on the TFM's cache controller chip. The super-capacitor pack can hold the memory circuit at voltage long enough for the data to be copied from controller cache to TFM flash. The capacitors take approximately 2 minutes at power on to charge, and capacitor charge is maintained via a constant small trickle charge.

A learn cycle is initiated following the power on and charging of the RMFBU. The learn cycle allows the TFM gas gauge circuitry to calibrate the operation of the super-capacitor pack and to judge the charge level and discharge capacity. During a learn cycle the super-capacitor pack is discharged by about 30% and then recharged to slightly above 100% capacity, the gas gauge then measures discharge over time and makes an estimate of the health of the super-capacitor pack. During a learn cycle, the super-capacitor is not discharged below the level needed to maintain a cache offload operation (moving cached data from controller DDR memory to NAND flash memory on the TFM) in the event of a power outage during the learn cycle process. The controller firmware continues to monitor the TFM to manage the use of write back raid volume mode.

The firmware waits for 5 minutes after boot before the learn cycle is initiated. A new RMFBU will undergo approximately 5 minutes of charge and discharge cycles during first time use. A RMFBU previously used in the system will undergo a single learn cycle following power on. The learn cycle must complete before the raid controller's storage volumes are allowed to run in write back cache mode unless the "always write back" setting is selected in the raid volume cache properties setting in controller/module firmware. Learn cycles will be repeated on a scheduled basis. The default learn cycle schedule is once every 28 days, this setting can be adjusted by using one of the controller management utilities such as Raid Web Console 2, or the Command Tool2/StorCli utility.

Upon a loss of power, the raid controller senses the loss of power and switches the memory circuit to draw power from the super-capacitors, once the firmware determines that the transfer of power source is complete, it will copy data in cache that has not been written to disk drives (dirty cache) from the controllers DDR memory to the TFM's NAND flash memory (a cache-offload).

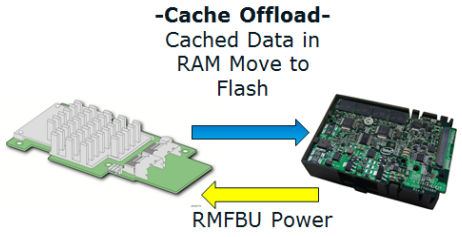


Figure 3 - Cache Offload

Once power is restored, the raid controller firmware determines if a cache offload took place by checking for a "dirty cache" bit set in a firmware register. If it finds that a cache offload took place, the raid controller will then copy the flash contents back to controller cache memory. This process happens

during Power On System Test (POST), and once the drives have been scanned and the raid volume is found to be healthy, the data in cache is flushed to the disk drives.

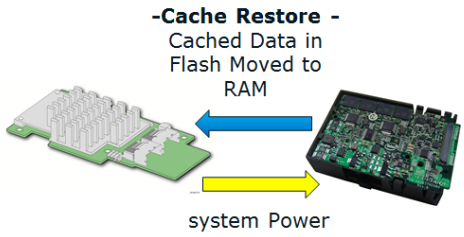
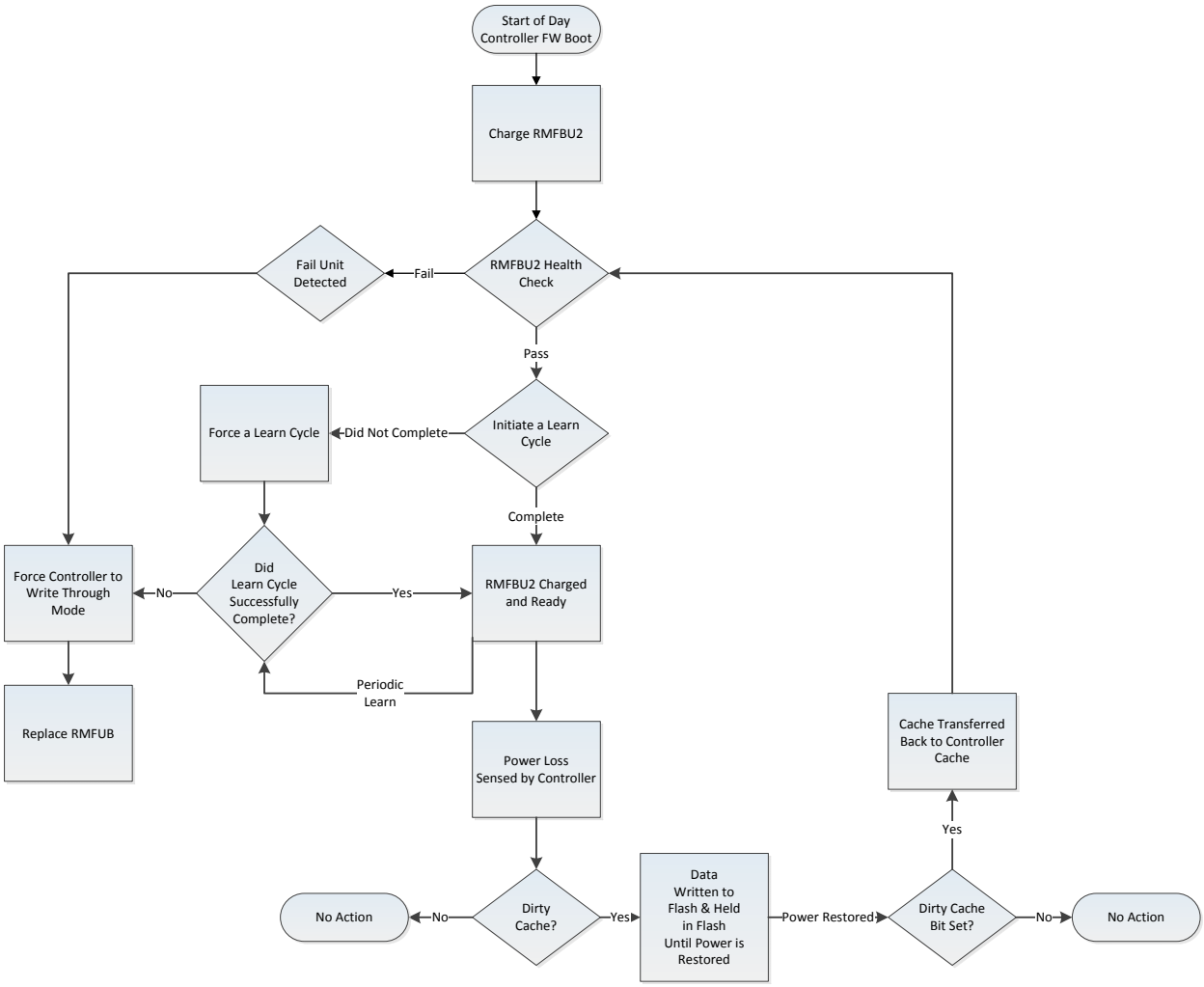


Figure 4 - Cache Restore

With cached data written to disk, the controller/module returns to a normal operational state and the Raid firmware continues to monitor the RMFBU health, and conduct periodic learn cycles.



A RAID Maintenance Free Backup Unit protects the integrity of the cached data on Intel® RAID Modules and Intel® RAID controllers that have on controller memory and support write back caching which allows offloading data stored in the RAM cache to the NAND flash if there is a complete AC power failure or a brief power outage. Furthermore, it eliminates the need for lithium ion (Li-ion) batteries traditionally used to protect DRAM cache memory on PCI RAID controllers. Therefore this is a greener and lower total cost cache protection solution. Super-Capacitor health is monitored by the TFM which is then reported to raid controller firmware. The raid controller firmware will initiate a learn cycle during power on and periodically thereafter. The learn cycle allows the TFM to measure the super-capacitor health. This document has described the function of the RMFBU and the process that allows the raid controller firmware to monitor and report the health of the super-capacitor.