Atomicity

Definition and OakGate Implementation

Date Prepared: 7-8-2013

OakGate Technology
3256 Penryn Road, Suite 220
Loomis, CA 95650
(916) 652-5132
www.oakgatetech.com

Definition: As it pertains to testing of SSD devices

The concept of Atomicity is germane to the behavior of a storage device performing IO write operations during periods of Power Loss or Power Failure. The expectation for this behavior is broken into three basic categories. Each of these will be explained in some detail along with the current capabilities of the OakGate Platform to verify their operation.

Case1: IO Write Operations have completed successfully during a power loss event. This is the basic operation and expectation of how a storage device should operate.

Expectation: Upon power being restored, a read of LBA's written prior to the power failure will return valid data. This will be true of IO's of any size and block format. For Queued IO Writes that access the same LBA(s), the exact data my not be deterministic but data will be valid as determined by the last write processed by the storage device.

Case2: IO Write Operations have NOT completed successfully during a power loss event. This is often referred to 'Shorn Writes'.

Expectation: Upon power being restored, a read of LBA's written prior to the power failure will return the following:

- Each LBA (independent of format) will be consistent. The data in the LBA will be either completely the new data or the prior data, but not a mixture of both. An embedded CRC of the LBA should always be valid. A failure of the CRC comparison would be considered a failure.
- The complete IO will be either totally the new data or totally prior data but not a mixture of both. If there is a mixture of data, then this will considered a failure. There is some discussion that this would only relate to 4K or 8K IO's, but for this specification, it is assumed that the IO size could range from a single LBA to the maximum supported IO size or 32MB, whichever is smaller.

Case3: Synchronous Write IO's are completed successfully during a power loss event. The objective is to ensure that the proper sequence of writes occurred and that the final data reflects the last expected write to the an LBA.

Expectation: Upon power being restored, a read of LBA's last writte prior to the power failure will return the expected data. This will be true of IO's of any size and block format. If the data is not consistent with the last successful write to that LBA, then this will be considered an error even if it matched an earlier successful write. This will support overlapping IO LBA ranges and IO sizes.

Implementation: Using OakGate Technology Platform

The OakGate Technology SVF platform has a data validation test capability that addresses much of the requirements to support the prior definition of Atomicity. There is some additional development required to provide the complete coverage. This will be highlighted as each of the cases are addressed.

Case1: IO Write Operations have completed successfully during a power loss event. This is the basic operation and expectation of how a storage device should operate.

The current OakGate data validation functionality addresses this case completely. The specific capabilities are:

- Each LBA written is individually tracked so that varying IO sizes and overlapping IO ranges can be managed.
- Each LBA includes a CRC of the data, an Index value to track the relative version of the IO and a 'magic' number to ensure that it was written by OakGate.
- Power control is integrated within the OakGate Automation environment. This is being enhanced to support PCIe slot level power control.
- Upon restoration of power, the user can select the number of writes that were successful prior to the power loss to verify. This number defaults to 1,000 but it could be 100,000's
- Writes and Reads are asynchronous events and are not related in time. Data that is writen will be read/verified at some later time, depending on the workload and IO range settings. The exception to this is the checking of completed writes following a power fail event (prior bullet).
- For Queued IO's with overlapping LBA ranges, those LBA's will have their CRC's verified but the Index value will be not be checked since it is not deterministic which order the IO's were completed.

Case2: IO Write Operations have NOT completed successfully during a power loss event. This is often referred to 'Shorn Writes'.

The current OakGate data validation functionality will need to be enhanced to address this case. Currently, only Write IO's that have completed successfully are tracked and recorded. The new capabilities are:

- Each Pending Write IO will need to be tracked separately along with each LBA within the IO. The 'index' state for each will be recorded so that both the old and expected values are known
- The CRC value will be used to ensure that each LBA is complete and consistent.
- Upon restoration of power, all incomplete Write IO's will be read. An error will be generated for

- Inconsistent LBA (mixture of old and new data with failed CRC)
- Inconsistent IO (Mixture of new and old data over the span of LBA's for that IO). This is valid only for IO's that are not overlapping or non-queue IO's

Case3: Synchronous Write IO's are completed successfully during a power loss event. The objective is to ensure that the proper sequence of writes occurred and that the final data reflects the last expected write to the an LBA.

The current OakGate data validation functionality addresses this case completely. The specific capabilities are:

- OakGate can ensure 'Synchronous Write IO's' by using either non-queued commands or with a Queue Depth = 1. This ensures that the writes are completed in the order sent to the target.
- Writes are tracked as in Case 1.
- Upon restoration of power, the user can select the number of writes that were successful prior to the power loss to verify. This number defaults to 1,000 but it could be between 1 to 100,000's