Panther V Storage Accelerator: Taking Storage to the Next Level

Focus on Data Integrity

WHITE PAPER

Authors:

Mike Ham
Technical Director of Applications Engineering, Storage Accelerators

Pinaki Chanda
Director of Software Engineering

KEY BENEFITS

- **Real-time verification (RTV):** Ensures automatic verification of encoded data against the original to prevent silent data corruption.
- **Comprehensive protection:** Safeguards data integrity through internal data path, memory, and PCIe packet protection.
- **Robust error detection:** Automatically detects and reports errors, facilitating real-time recovery and command reprocessing.
- **Industry-standard support:** Implements T10-DIF, T10-DIX, and NVMe Protection Information (PI) features to extend data integrity checks beyond the storage accelerator.

Introduction

The MaxLinear's Panther V storage accelerator (Panther V) has been developed to optimize the performance and efficiency of storage appliances by offloading complex tasks that require significant CPU cycles to execute. The key features of Panther V are as follows:

- **Data reduction**
  - High-quality compression to minimize the size of data before storage.
  - Hash generation to drive deduplication to further reduce the size of data before storage.
- **Data integrity**—Panther V provides the most robust protection against data corruption available.
  - Real-time verification (RTV)—All encode operations are automatically followed by a decode operation and verification that the fully decoded data matches the original data.

    Note: Transform commands are not completed until all requested data protection/verification operations are completed successfully.
  - Panther V provides full support for the T10-DIF, T10-DIX, and NVMe Protection Information (PI) data protection functions.
- **Data security**—Panther V supports NIST-certified Advanced Encryption Standard (AES) security standards used for SSL and IPsec transforms. They are optimal for securing data at rest (on the storage device), being transferred to/from the remote client or to/from a backup/replication storage appliance.
- **Highest performance**—Support for 450Gbps throughput with the lowest latencies available.

This document focuses on how Panther V provides the highest level of data integrity available.

Data Integrity

What is data integrity and why is it important?

Data integrity is the ability to ensure that data being transformed is not corrupted during the process. This includes protections for all data residing on the device, data moving through the device, and data being actively transformed. Without robust data integrity capabilities, data to be transformed could be silently corrupted (undetected) and stored. The corruption may go undetected until the data is decoded, by which time it may be too late to recover it since the original data is no longer available.

By implementing industry-leading data integrity features, Panther V can protect against silent data corruption for both encode and decode operations. Should data corruption occur, it can be detected by the Panther V device and reported. Panther V has the most robust data integrity capabilities of any solution on the market.

Panther V's data integrity features are as follows:

- Real-time verification
- Internal data path protection
- Memory protection
- PCIe packet protection
- NVMe PI data protection fields
- Robust compression algorithm support
- PCIe Advanced Error Reporting (AER)

Real-Time Verification (RTV)

Panther V supports the use of RTV, which MaxLinear highly recommends. When enabled, all encode operations are automatically and transparently fully decoded. The fully decoded data is then verified against the original data to ensure that all transform operations have been completed successfully. Only when this transform verification is complete is the command completion notification sent. In this way, Panther V can self-detect any errors and report them if they occur in real time, thus avoiding any risk of silent data corruption, with no impact on performance and minimal overhead due to Panther V's heavily pipelined architecture.

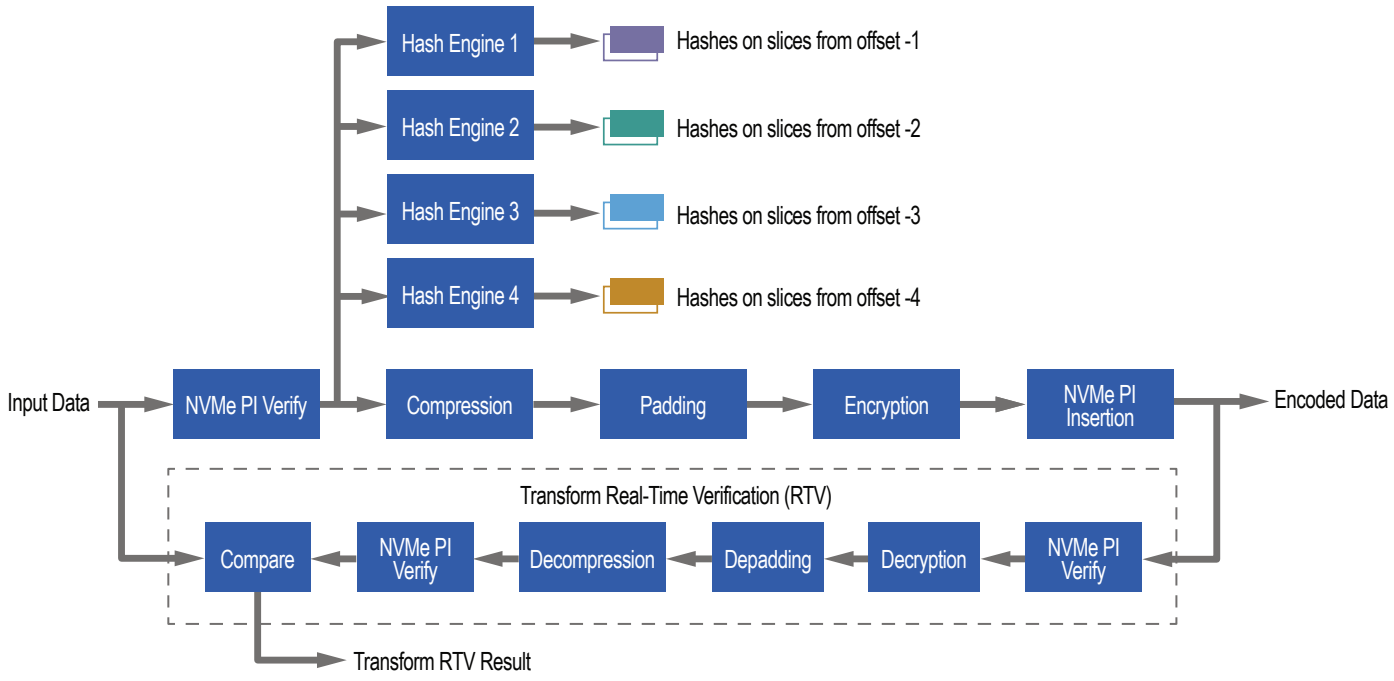

Figure 1: RTV Flow for Transform Engines

The RTV function for hash operations is implemented as two parallel engines whose two outputs are verified byte by byte since hash operations have no decode capability.

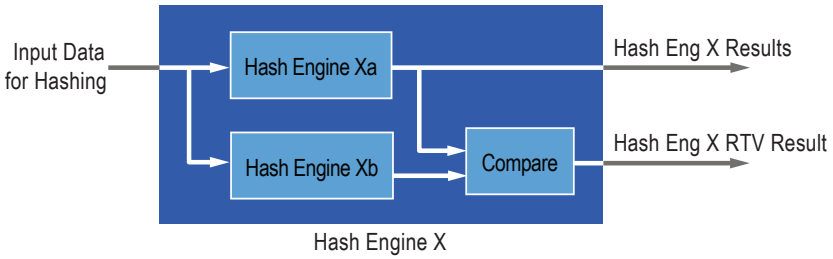

Figure 2: RTV Function for Hash Operations

One potential cause of an error on a Panther V device would be an alpha particle hit resulting in memory or logic disruption. Panther V goes beyond detecting and reporting every error occurrence. When an error is detected and reported by Panther V, the driver automatically reprocesses the command using the algorithm software library that implements all the transforms supported by the device. If successful, the command is completed with a notification that an error has been reported by the Panther V device and that the command has been successfully processed by the software. If unsuccessful using the software, the command is completed reporting a failure due to a command submission error. This helps identify the cause of the failure, while facilitating the integration of Panther V during the development process.

T10-DIF, T10-DIX, and NVMe PI Data Protection

Panther V supports the following industry-standard data protection features:

- T10-Data Integrity Field (T10-DIF)
- T10-Data Integrity List (T10-DIX)
- NVMe Protection Information (NVMe PI)

These industry-standard data protection features extend data integrity beyond the boundary of the Panther V storage accelerator and beyond the processor root complex (RC). They are all derived from the Small Computer System Interface (SCSI) T10-DIF standard managed by the T10 Technical Committee.

The T10-DIF and T10-DIX fields are 8-byte fields that contain three sets of information:

- The 2-byte Application (App) tag, which contains customer-defined information.
- The 2-byte Reference (Ref) tag, which is also customer-defined, but typically represents a storage logical block address (LBA) and increments with each T10-DIF/T10-DIX field.
- The 4-byte Guard block, which contains a CRC generated on the preceding data that the DIF/DIX field protects.

Figure 3: T10-DIF/T-10-DIX Field Information Set Format

Each T10-DIF/T10-DIX field is generated for a data block defined as a sector and protects this data block. Each sector is traditionally 512 bytes but can be any power of 2 from 512 bytes upwards (that is, 512 bytes, 1024 bytes, 2048 bytes, and so on). The only difference between the T10-DIF and T10-DIX fields is that the T10-DIF fields are inserted directly into the data at the end of each data sector they protect, whereas the T10-DIX fields are generated and managed as a separate list from the data.

NVMe PI is an extension of T10-DIF/T10-DIX defined in the NVMe standard. The DIF and DIX configurations as described above are supported, but an additional optional customer-defined metadata block can be added. This block can be placed before or after the DIF/DIX-defined field containing up to 64KB of information. Panther V supports all configurations of T10-DIF, T10-DIX, and NVMe PI.

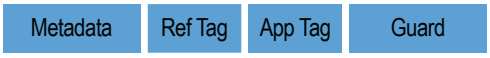

Figure 4: NVMe PI with Preceding Optional Metadata Field

Figure 5: NVMe PI with Trailing Optional Metadata Field

The advantage of these fields is that they are always in clear text (unencoded) and can therefore be checked at any time in the data flow. For example, the host can verify them, or the storage fabric, the RAID controller, the NVMe drive, among others. If they prove invalid (possible data corruption), the transaction can fail and be retried, ensuring data protection and resilience through the entire data lifecycle.

Panther V supports the following capabilities:

- Check/verify all fields of the source data submitted for encoding.
  - If submitted in DIX format, Panther V can optionally insert the fields into the data to use for subsequent decode operations.

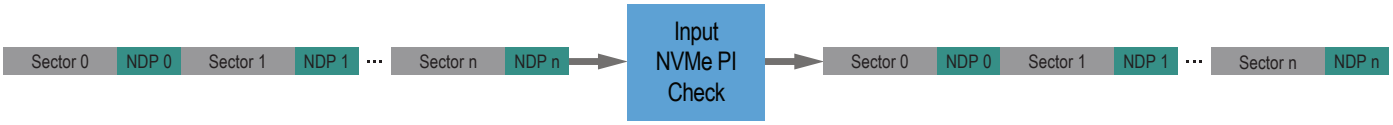

Figure 6: Source Data has Embedded DIF-Style NVMe PI to be Verified

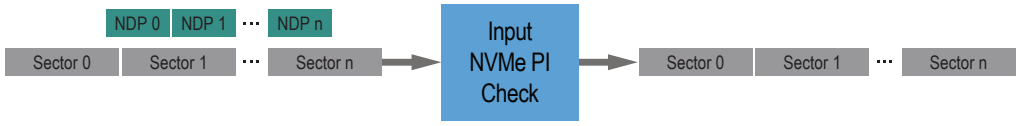

Figure 7: Source Data has DIX-Style NVMe PI to be Verified and Discarded

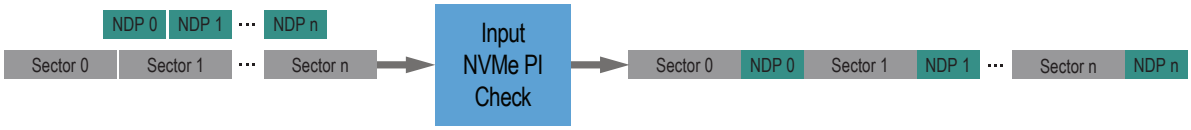

Figure 8: Source Data has DIX-Style NVMe PI to be Verified and Inserted

- Generate and insert (DIF format) or generate and return (DIX format) fields for fully encoded data.

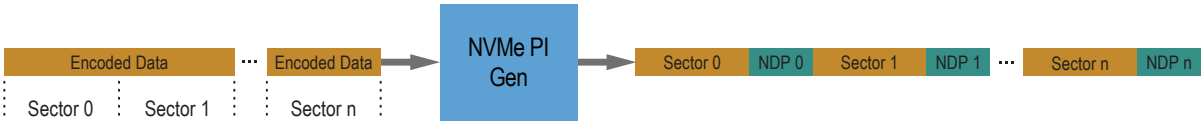

Figure 9: NVMe PI Fields Generated and Inserted into the Encoded Data

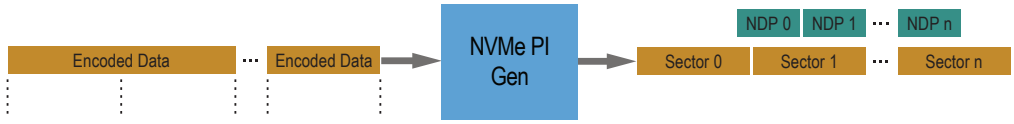

Figure 10: NVMe PI Fields Generated and Returned in Sideband Buffer as a DIX-Style NVMe PI

- Verify and discard (DIX format) or verify and remove (DIF format) fields for encoded data being decoded.

Figure 11: Encoded Data for Decode has NVMe PI Fields Supplied in Sideband Buffer in DIX Style to be Verified and Discarded

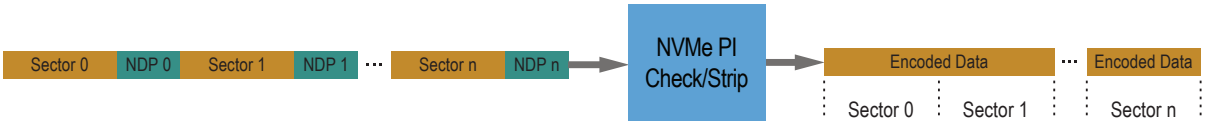

Figure 12: Encoded Data to be Decoded has Embedded NVMe PI Fields to be Verified and Stripped

- Verify the DIF format fields on fully decoded data and return them in DIF format, DIX format, or remove and discard the fields.

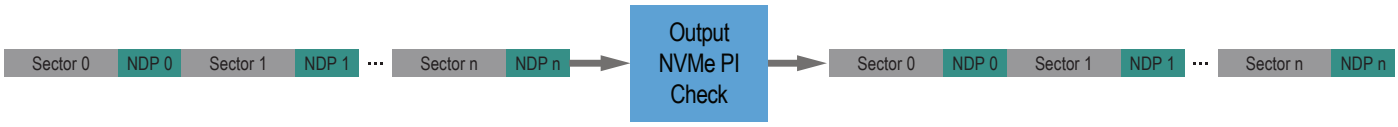

Figure 13: Decoded Data has NVMe PI Fields Verified and Returned as Embedded Fields in the Decoded Data

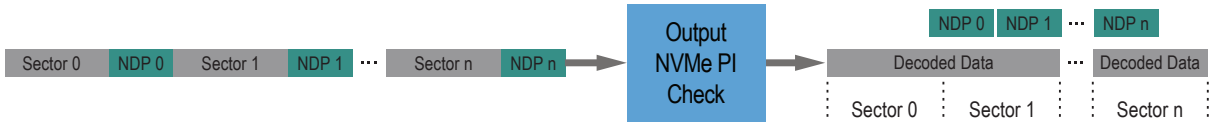

Figure 14: Decoded Data has Embedded NVMe PI Fields Verified and Returned as Sideband DIX Type List

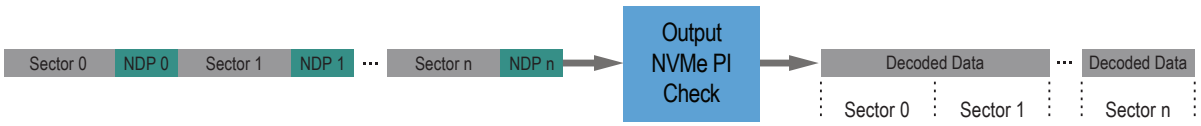

Figure 15: Decoded Data has Embedded NVMe PI Fields to be Verified and Stripped

This automation increases data reliability while offloading a transform which, when performed in software, is very CPU- and memory-intensive. Additionally, Panther V's support for this data protection feature extends data integrity to the entire data lifecycle.

Robust Data Compression Algorithm Support

Panther V supports several compression algorithms. The most popular are Deflate/zlib/gzip and XP10. They all provide some level of data integrity protection. Part of the protection is built into the encode function, but the most important feature is the use of a CRC on the unencoded data that is automatically verified upon decode. Only Deflate does not support a built-in CRC, which is why MaxLinear recommends using zlib/gzip and XP10 for data compression.

Panther V Memory Protection

All Panther V's internal memories are protected by parity or error correction codes (ECCs). ECC-protected memories can correct single-bit errors and detect two-bit errors.

The 32MB BAR4 Panther V memory that is mapped to the system memory is ECC-protected.

PCIe Advanced Error Reporting (AER)

Panther V and the supplied SDK support PCIe AER mechanisms and callback functions as defined by the PCI standard and the OS/kernel AER framework. All detected data integrity and PCIe errors will be reported and recovery attempted.

PCIe Packet Protection

By using Panther V's support for the optional endpoint cyclic redundancy check (ECRC) for PCIe packet transfers, data protection against possible data corruption in the data path between the CPU's RC and the Panther V device can also be detected and prevented.

Conclusion

Panther V has been designed and architected to provide optimal data integrity support. Many of its features offer very significant levels of CPU offload in terms of CPU cycles, latency, memory bandwidth, and power. Panther V is easy to integrate and has been designed with maximum uptime and absolute data/transform reliability in mind. There is no better solution to today's challenges for ensuring next-level data integrity.

References

- *MxL890x Software Development Kit (SDK) Getting Started User Guide* (210-CUG).
- *MxL890x Software Development Kit User Guide* (209UG).
- *MxL890x Raw Acceleration Application Program Interface (API) Reference Guide* (200AG).
- *MxL890x Storage Accelerator User Guide* (204-UG).
- *MxL890x Performance Evaluation Tool User Guide* (219UG).
- *MxL890x Linux Performance Application Note* (294AN).
- *MxL890x Linux Performance Tuning Application Note* (298AN).
- *MxL890x Software Development Kit Linux Release Notes* (202-CRN).
- *MxL890x FreeBSD Performance Application Note* (297AN).
- *MxL890x FreeBSD Performance Tuning Application Note* (299AN).
- *MxL890x Software Development Kit FreeBSD Release Notes* (209-CRN).
- *Panther Storage Acceleration SDK: Simplifying Hardware and Software Integration White Paper* (005WP).
- *Panther V Storage Accelerator: Next-Generation MaxHash™ Deduplication White Paper* (006-GWP).
- *Panther V Storage Accelerator: Taking Storage to the Next Level—Focus on Data Reduction White Paper* (007WP).

MaxLinear, Inc.
5966 La Place Court, Suite 100
Carlsbad, CA 92008
Tel.:+1 (760) 692-0711
Fax: +1 (760) 444-8598
www.maxlinear.com