Solution Brief: 20x acceleration of RocksDB with Pliops

Problem Statement

Gartner reports that analytics and artificial intelligence are the one of the biggest growing categories for midsize, large and global enterprise with analytics adoption increasing to 50% by 2023. The unstructured data for these analytics platforms will be stored in non-relational (NoSQL) databases. Key-value stores are a popular non-relational database type known for their simplicity, flexibility and speed. It provides a dictionary like interface to easily store and retrieve data. Popular examples of key-value databases are Memcached, Facebook's RocksDB and Google's LevelDB.

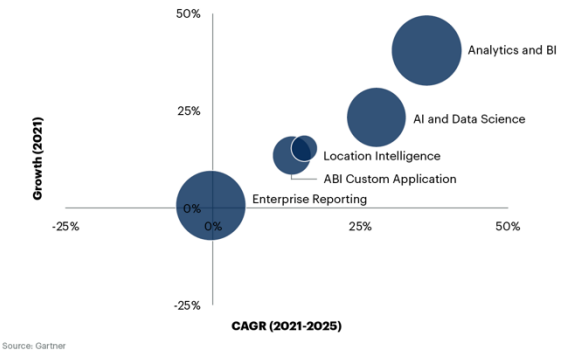


Figure 1: Gartner Analytics Platforms Market Segment Forecast

Unstructured Data Challenges

One major challenge that enterprises face with storage of unstructured data is the variety and volume of it. This means that the data is of different block sizes, not sorted and may or may not be compressible. In order to optimize for write of this data, it is typically written first to memory, and then sorted, compressed and then written to SSD disks (due to latency requirements). While this is costly from a compute perspective, this also brings up the second major challenge which is write amplification. Due to the data being sorted and consolidated into larger data structures, it causes a read-modify-write operation which causes the same data to be written repeatedly to the disk. As an example, RocksDB, which is one of the top key-value datastores has an average write amplification factor of around 7x (not including the garbage collection process of the SSDs themselves). As the SSDs get full of this unstructured data, the amplification becomes worse, and results in unpredictable latency and lowered endurance of the SSDs. The final challenge of storing this unstructured data is protection of the data itself. Enterprises are forced to choose between traditional RAID or erasure coding, the former leading to massive overprovisioning or the latter which amplifies network bandwidth requirements.

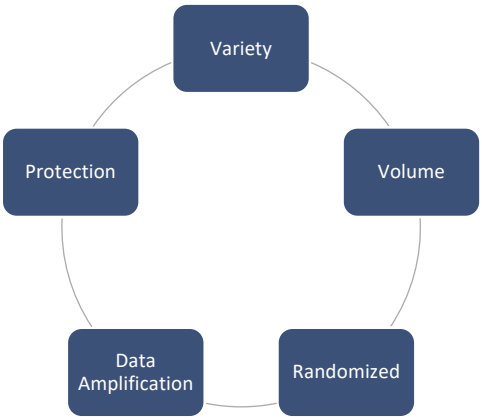


Figure 2: Unstructured Data Challenges

Pliops Solution to Data Challenges

Pliops has introduced the Extreme Data Processor (XDP) storage accelerator add-on card which overcomes all these challenges. The way it does that is, as shown in figure 3, that it intercepts the I/O from the application, sends it to the card which sort, compresses, indexes and packs the data before flushing it to the disk in a sequential manner. This frees up the CPU from using resources to sort and compress the data.

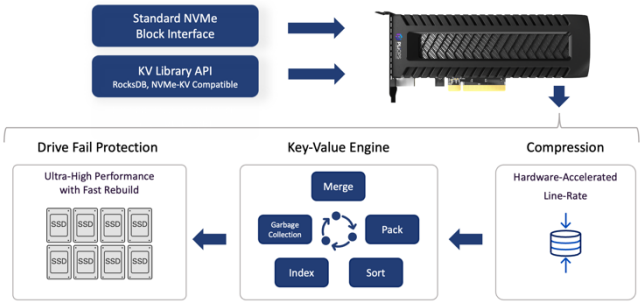


Figure 3: XDP Data flow and capabilities

Because the writes to the disk are always sequential (rather than random), the write amplification is reduced to its theoretical minimum, thus maximizing performance and increasing longevity of the disk. Finally, it also has built in drive failure protection which provides RAID5 like protection without the need for overprovisioning by allocating virtual spare capacity on all drives. The result is shown below in Chart 1. The first two figures are from our study which compares the Random read and write performance of a single RocksDB instance against XDP-Rocks at various object sizes. As seen below, that the application read performance are increased between 2x and 6x while random write performance increases between 11x and 36x.

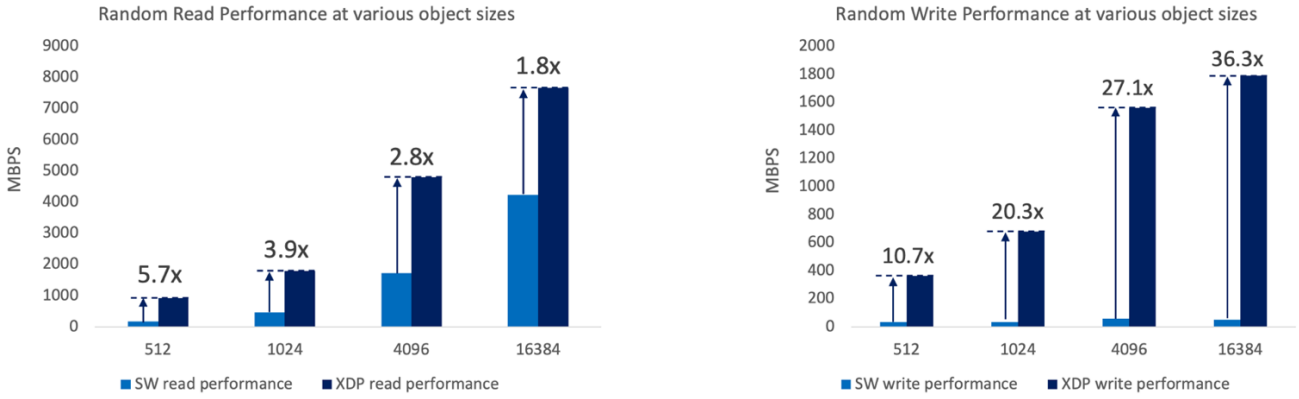


Chart 1: Random read and write performance enhancements with XDP-Rocks for various object sizes

In Chart 2, we see how XDP-Rocks can increase throughput by 5.3x while reducing 99.99% latency by 36x and reducing CPU as a percentage of kilo-ops by 3.3x in a multi database setup with 7TB data set with 4 threads, 16B key and 1KB value. Taking a closer look at this data in Chart 3, we see that while the Average throughput in kilo-ops is 6.5x higher, the standard deviation of XDP-Rocks is much lower at just 2.3% indicating a more sustained throughput.

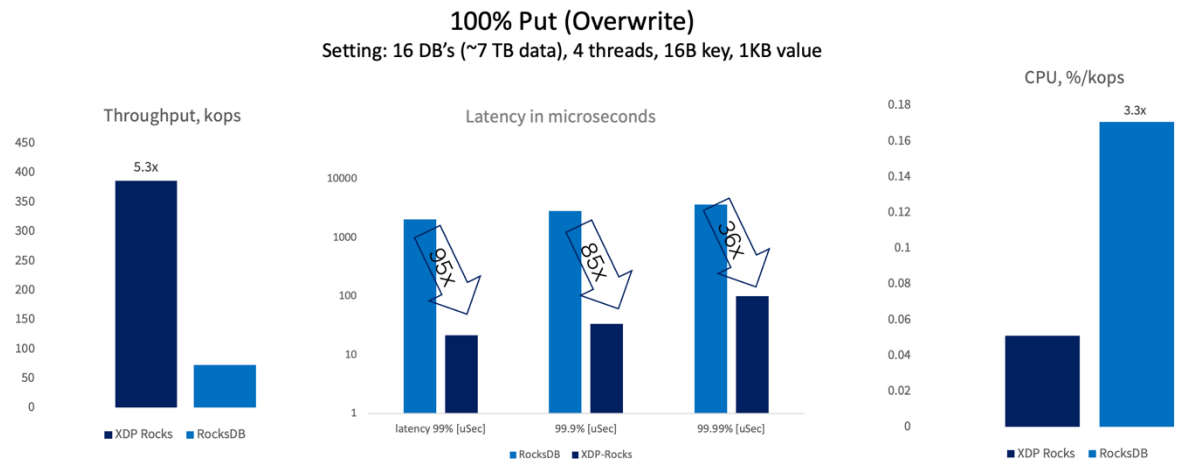


Chart 2: Multi RocksDB Workload Enhancements

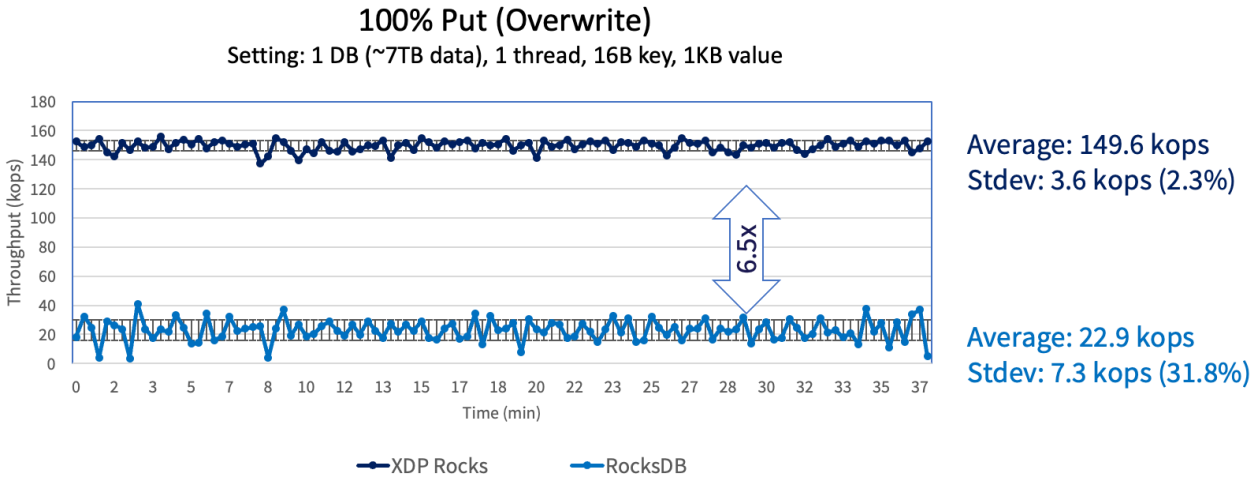


Chart 3: Write Workload Performance – a Zoom-In

Looking at read and mixed workload performance in Chart 4, we see that XDP-Rocks can achieve 6.8x throughput over traditional RocksDB in read operations which fails to scale beyond 100K Ops with a 10 ms SLA target. We're also able to get 5.9x throughput in mixed workloads over RocksDB which fails to scale beyond 100K Ops with a 50 ms SLA target.

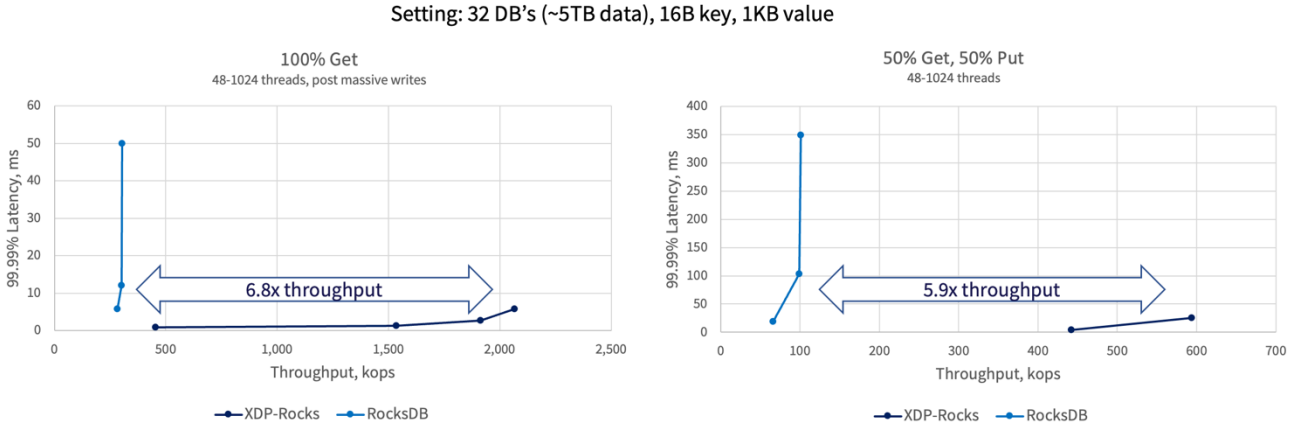
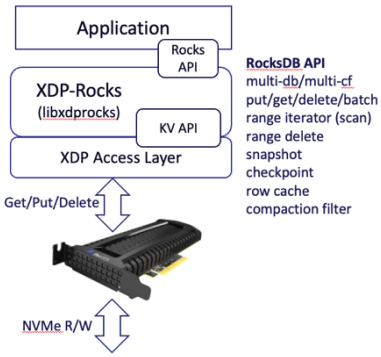


Chart 4: Read and Mixed Workload Performance

We see similar benefits when XDP-Rocks is used with applications on top of RocksDB. For example, we have another solution brief available of the value XDP-Rocks provides with KVRocks, an open-source distributed key-value NoSQL database for storing and processing large datasets, on Pliops.com/resource

RocksDB API Compatibility

All these benefits of the XDP accelerator card are provided with an API interface called XDP-Rocks which is presented as an adaptation layer to RocksDB. Thus, developers can switch to the XDP-Rocks API interface with minimal effort to take advantage of these performance, reliability and efficiency capabilities. The XDP-Rocks API is then translated to native XDP KV commands to take advantage of the hardware acceleration.



Conclusion

Check our Product page to learn more about how your key value application can be accelerated in 5 simple steps by the Pliops XDP card using our XDP-Rocks API interface.