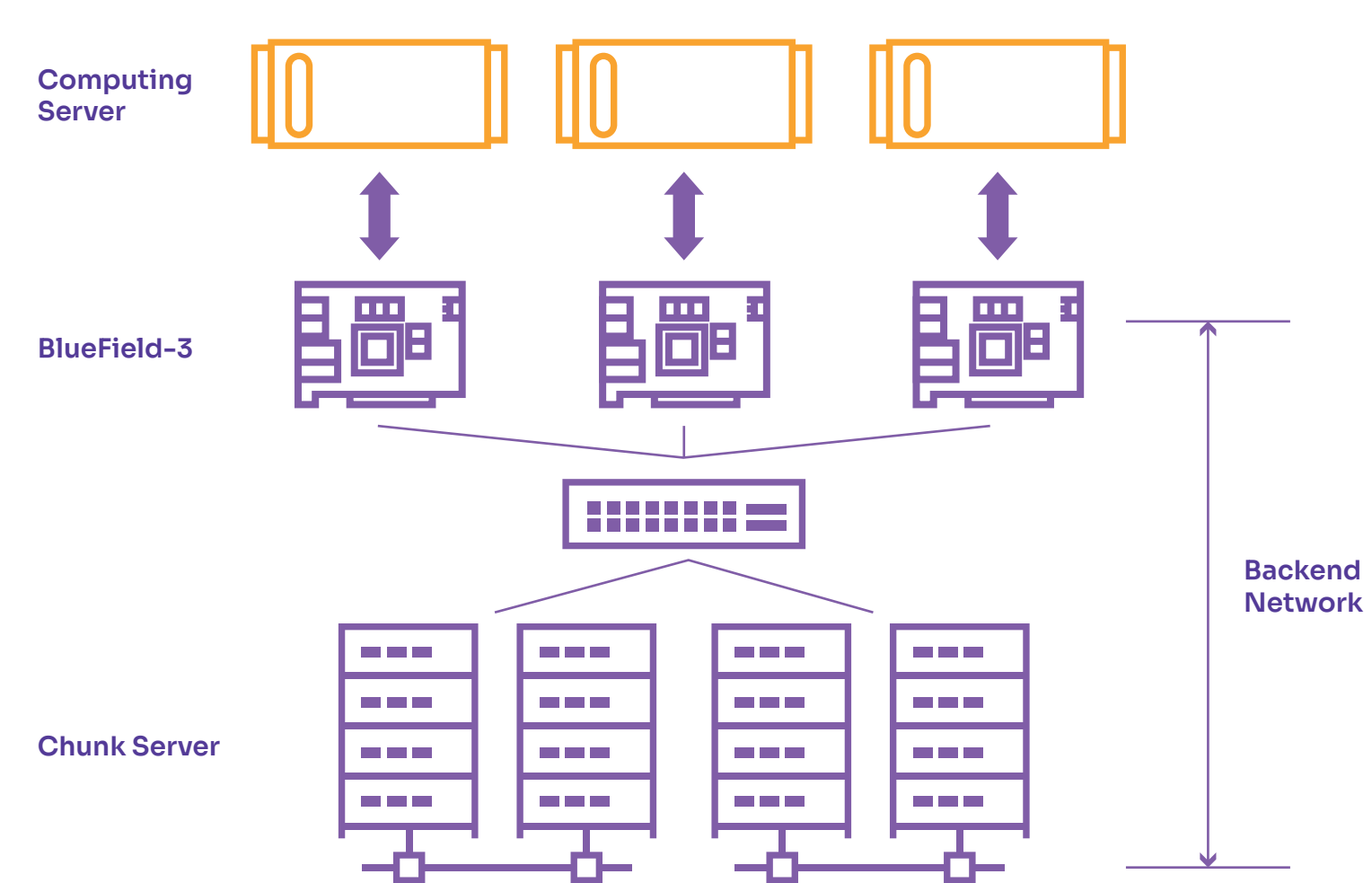
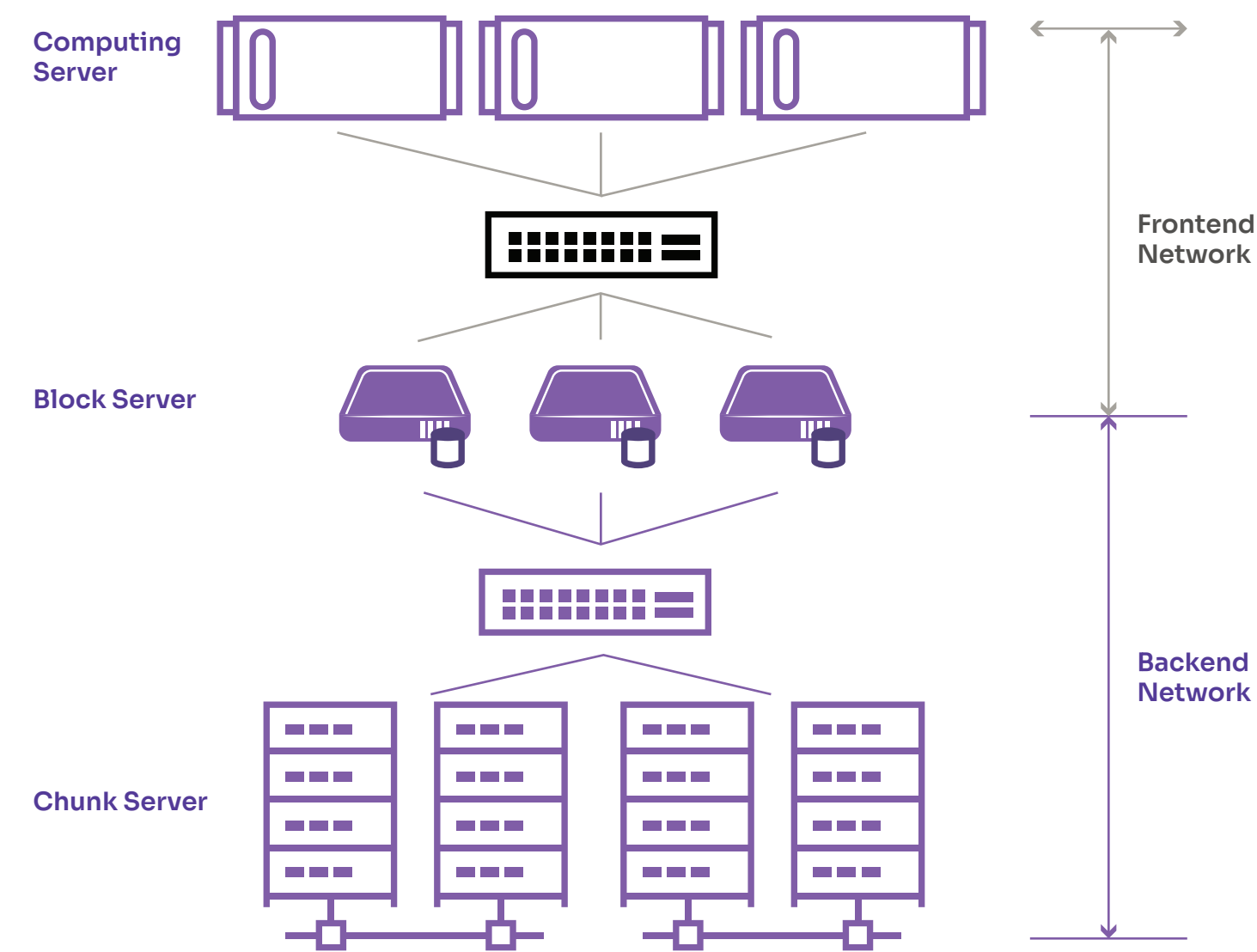


# Next-Gen Hadoop Storage: BlueField-3 and CSAL for Performance and Resilience

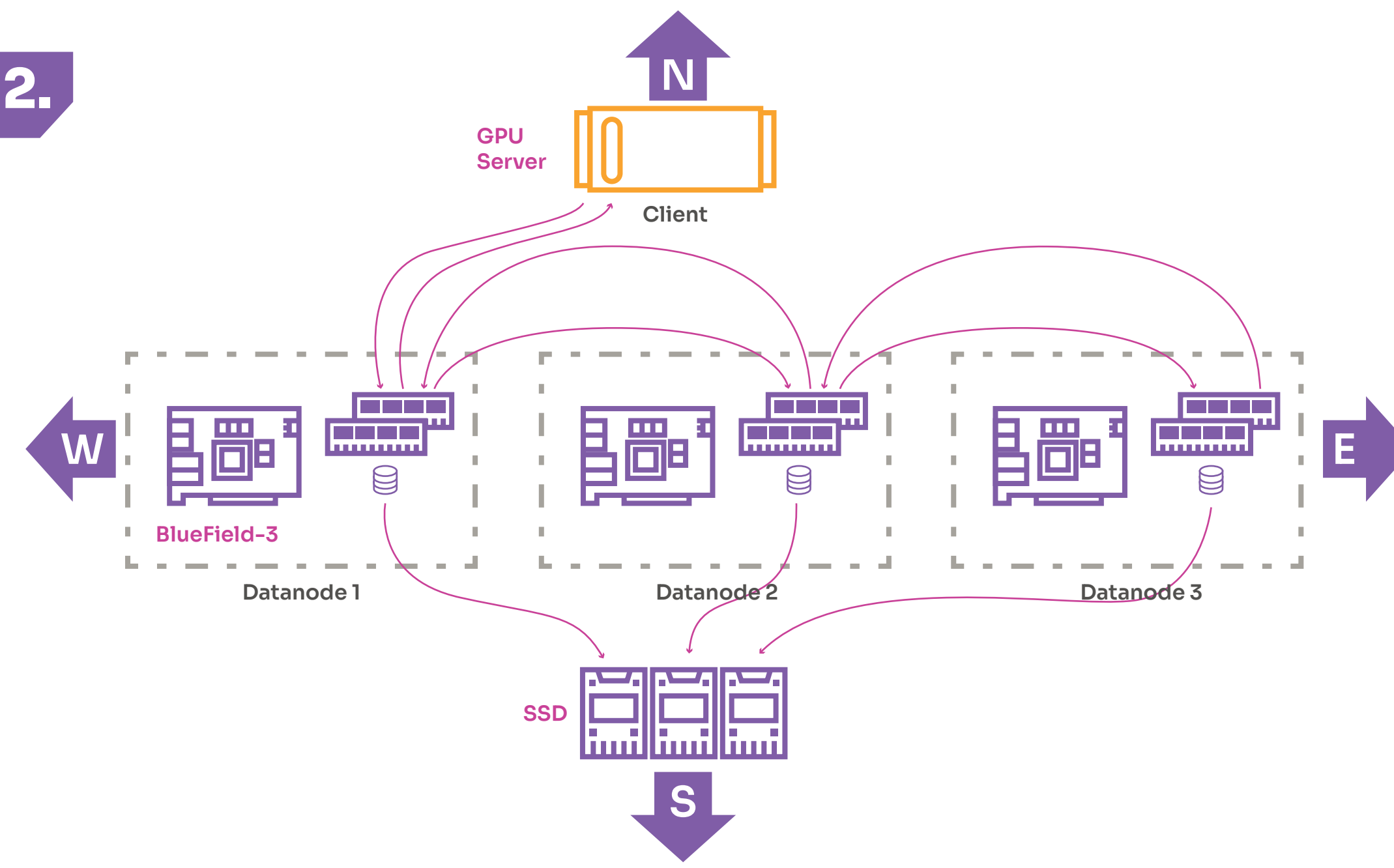
Abstract: Traditional Hadoop storage faces challenges in performance and resource efficiency for modern big data workloads. This poster introduces a solution powered by BlueField-3 and Solidigm's CSAL software, delivering superior throughput, reduced CPU overhead, and robust data protection via three-way replication, erasure coding (EC), and RAID, ensuring scalable and secure data management.

**1.** By adopting the BlueField-3 + CSAL solution, we restructured the three-layer Hadoop storage architecture, bringing the following advantages:

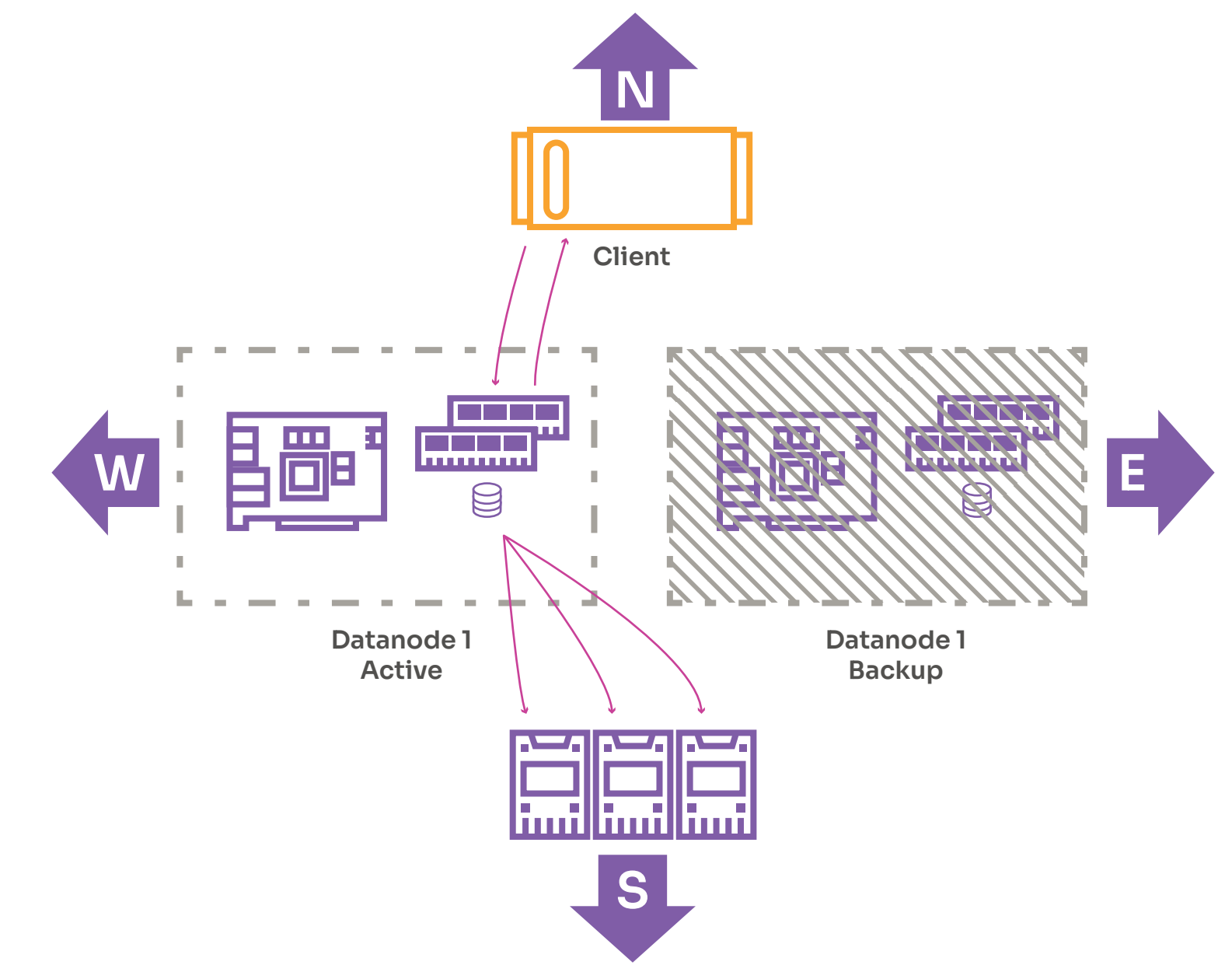
- Utilizing the hardware computing resources in BlueField-3, Solidigm's CSAL software supports various data protection backup mechanisms such as EC, RAID, and multiple copies.
- The original architecture of three layers of servers + two layers of switches has been compressed into two layers of servers + one layer of switches, significantly reducing the read/write latency of stored data.
- Offloading storage-related computing resources to the BlueField-3 has saved CPU computing resources, allowing the CPU to focus more effectively on Hadoop computations.



**2.**

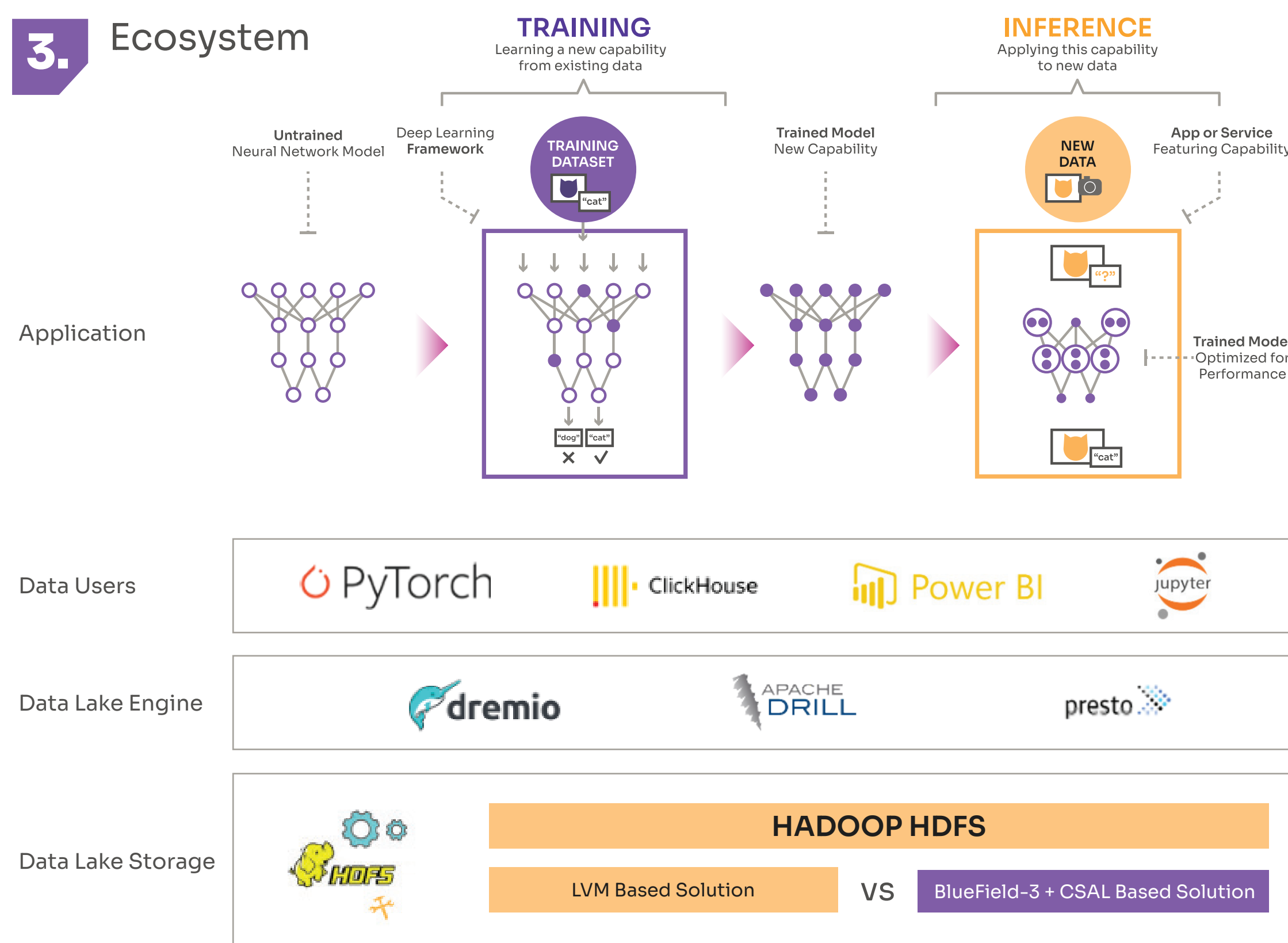


**Traditional Three-Replica:**  
**High Reliability:** Ensures fault tolerance with three copies.  
**Chained Replication:** Data flows sequentially through Datanodes, increasing latency.  
**ACK Overhead:** Multi-step acknowledgment adds complexity and latency.

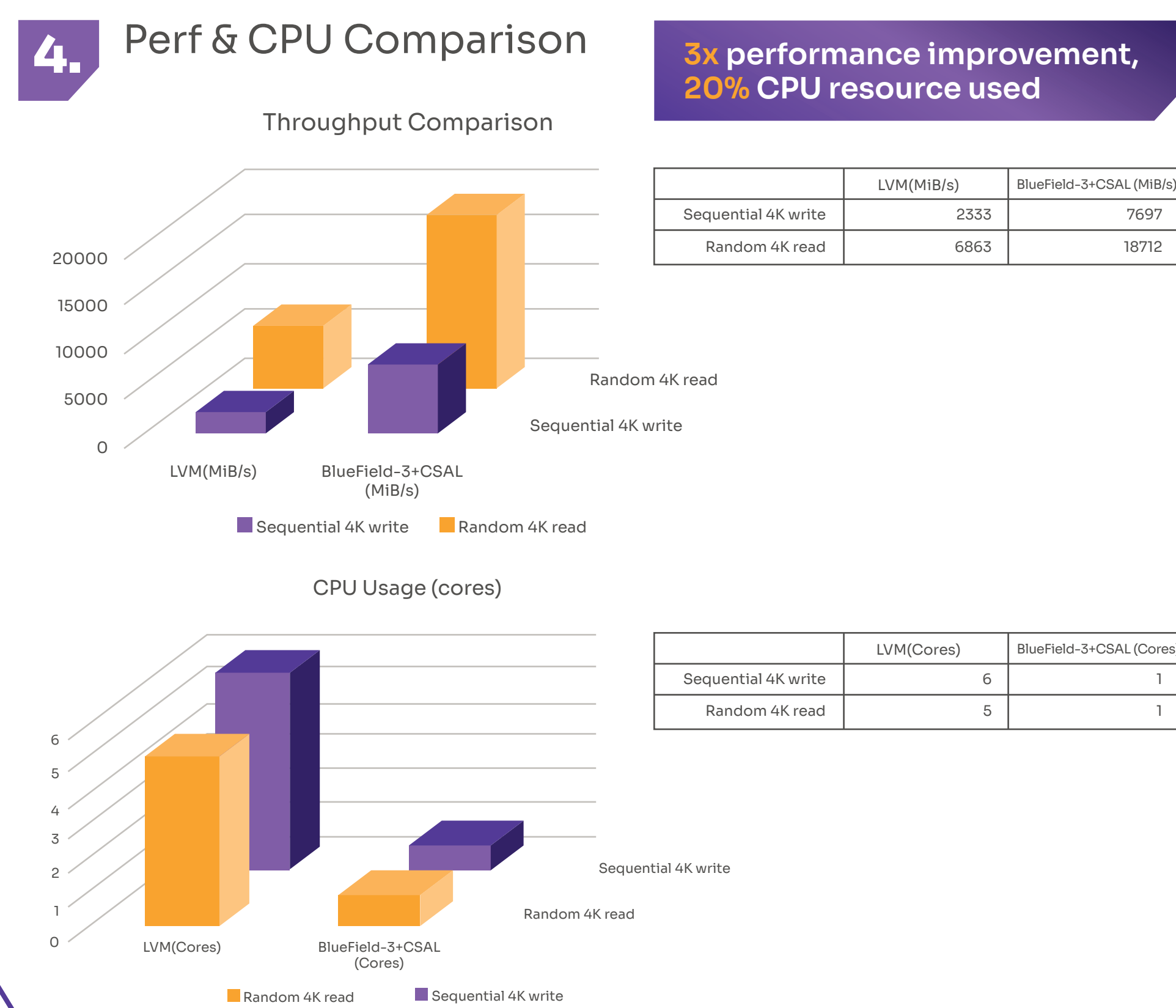


**BlueField-3+ CSAL engaged Active/Backup Architecture:**  
**Simplified Workflow:** Directly replicate data inside one BlueField-3  
**Improved Performance:** No East-West Traffic, reduced replication steps and lower latency.  
**Optimized for High-Performance Scenarios:** Ideal for latency-sensitive workloads.  
 An additional management plane mechanism is required to enable rapid failover to the backup node in the event of a Datanode failure.

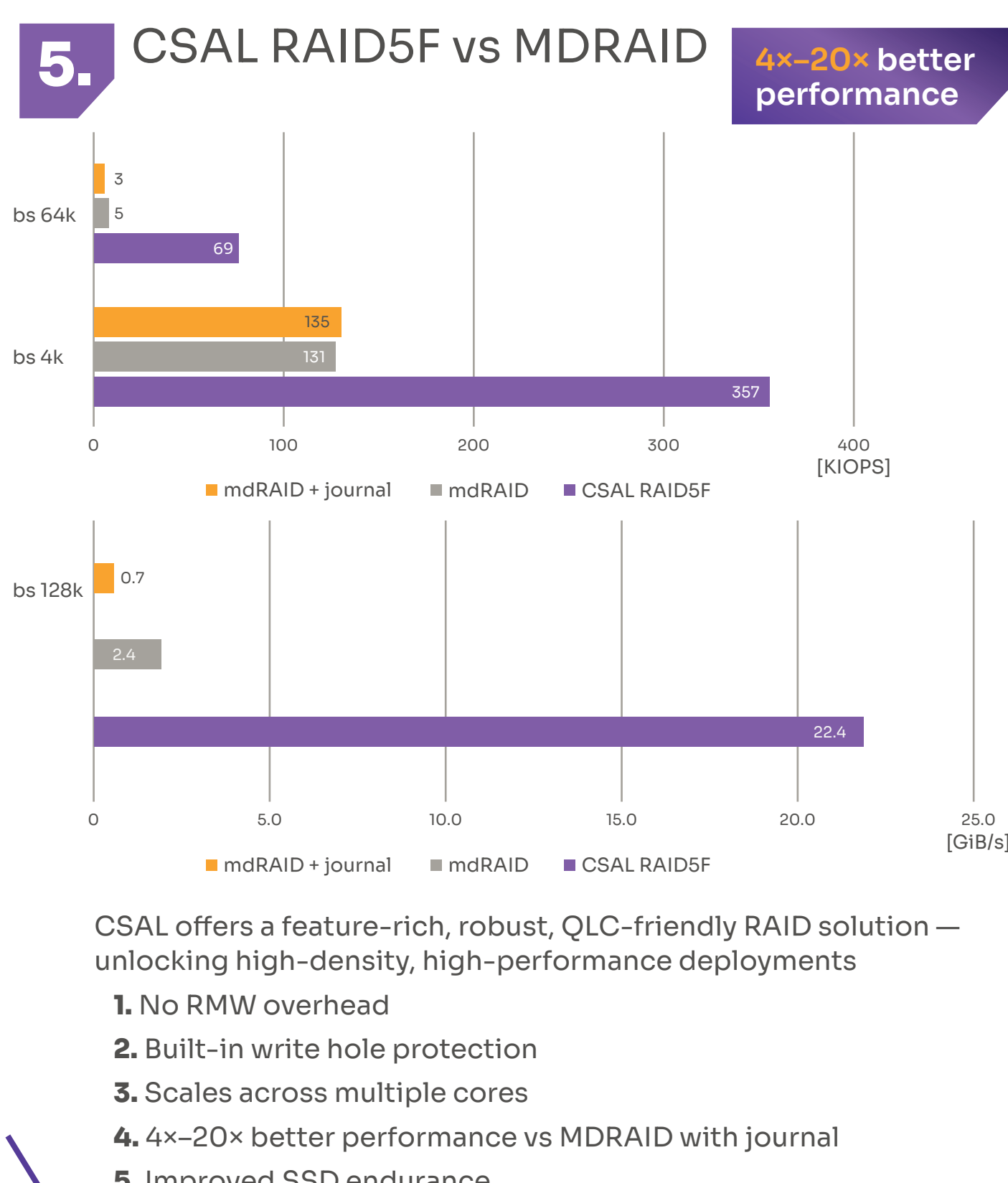
**3.** Ecosystem



**4.** Perf & CPU Comparison



**5.** CSAL RAID5F vs MDRAID



**6.** TCO Analysis

