Turbo charging Spark

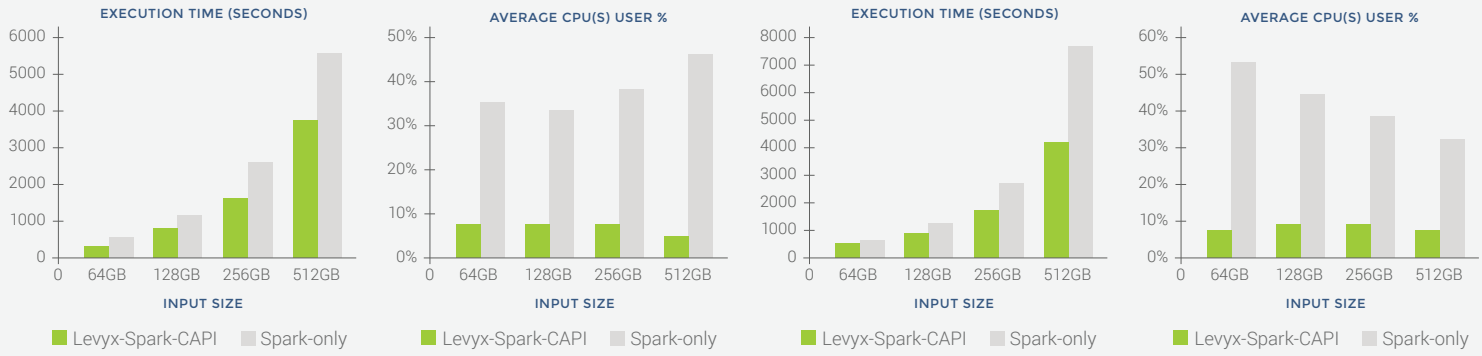
Levyx's Xenon for IBM Power 8™ CAPI™ Acceleration of Apache Spark

CHARACTERISTICS

- Core SQL functionality: filter, projection, selection, sort, join, groupby, and aggregates on structured data
- Accommodate simultaneous analytics and transactional operations on the same dataset
- Map/reduce data analytics capabilities
- Tightly integrate with the Apache Spark system
- High-performance atomic NoSQL functionality: put, get and delete on structured or unstructured data with support for point and range query
- Scale with the number of cores in the cluster and use SSD (or other high-bandwidth, low-latency persistent storage) as the storage fabric for datasets
- Provide in-memory performance
- Uses just-in-time compilation

Representative Benchmarks Comparison of IBM Power™ Spark

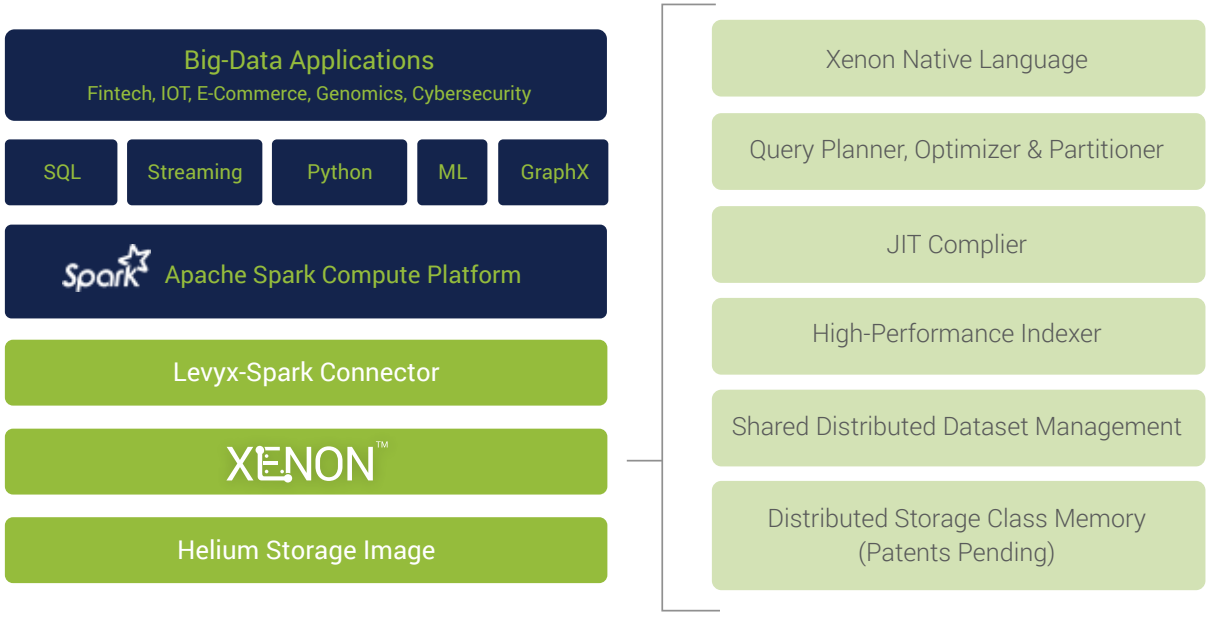
WITH AND WITHOUT LEVYX XENON CAPI OFFLOAD ACCELERATION



KEY FEATURES OF JOINT LEVYX-IBM POWER SOLUTION

- Works with standard Spark Distributions including Hortonworks™ HDP
- Ideal for Accelerating and Consolidating I/O intensive Spark workloads by up to 80+%
- Simultaneously frees up 90% CPU utilization via CAPI I/O offloading
- Enable Apache Spark DataFrames to use FlashGT or IBM FlashSystems as memory extension
- Enables concurrency across multiple Spark contexts accessing the same DataFrame
- Simple to Install/Integrate with existing Spark Applications
- Levyx Xenon plus Spark can accelerate workloads such as Trading Algorithm BackTesting simulations by over 32X over Hadoop-based implementations

(source: https://stacresearch.com/news/2017/06/26/LEVX170603)



Xenon™ was designed with the following key principles in mind:

1. Bypass file-system and kernel-buffers and directly perform flash friendly I/O to (a) reduce I/O latency -- important for transactional operations and (b) saturate the I/O bandwidth of the storage devices -- important for analytics operations. Reduce read/write amplification to absolute minimum and issue I/O to the device in large sequential write and small random read patterns. Make use of NVMe technology to increase I/O efficiency.
2. Reside on a Distributed Storage Class Memory abstraction (patent pending) providing persistent byte addressable memory at flash capacities and DRAM performance across a cluster of nodes.
3. Use just-in-time compilation to perform high-bandwidth database operations as close to the metal as possible and avoid any and all form of data interpretation.
4. Utilize modern processor ISA capabilities (e.g., memory barriers, bit-level instructions, etc.) to design lock-free and scalable data-structures and algorithms.
5. Design a custom, modern, compressed, memory-resident, serializable, and lock-free indexer (patent pending) that tracks objects residing on persistent storage. Ultimately, allow high-performance in-memory point and range-query operations performed directly on the compressed indexer while being able to retrieve or update any object on persistent storage with a single read or write issue to the device.

About Levyx

Levyx was founded 2013 in Irvine California with a mission of enabling Real-Time Persistent Computing for Big Data™. To that end Levyx has developed next-generation database Storage and Query offload engines that fully exploit the latest commodity hardware technologies including multi-core servers such as Power 8, internal and external flash systems, and IO offload engines such as CAPI.

The result is unprecedented performance and latency reductions for IO intensive workloads such financial service backtesting or streaming analytics.

Levyx's new software-based engines allow for the first time, persistent computing to be possible on Big Data platforms such as Apache Spark thru use of SSDs instead of volatile memory-only designs.

Levyx is now delivering the World's fastest key value store (Helium) and World's first distributed storage/analytics offloads engines on flash (Xenon).

Levyx is an IBM Business Partner.

49 Discovery, Suite #220
Irvine, CA 92618

(949) 502-6369
info@levyx.com