KALRAY

White Paper

KALRAY FLASHBOX™

DPUs at the Heart of New Generations of All-Flash-Arrays

CONTENTS

DEPLOY NVMe SSDs AT MASSIVE SCALE FOR MAXIMUM PERFORMANCE

In a connected world where every millisecond counts, the speed at which an organization can access and process data determines its value. According to IDC[1] , the volume of data generated over the next 3 years will be more than the data created over the last 30 years. This data explosion is being driven by demanding AI and data analytics services and applications utilizing compute-intensive algorithms, putting massive pressure on cloud and edge data center networking and storage teams.

Datacenter technologies are continuously evolving as performance and service level requirements shift with ever-increasing application workloads. For example, remember how GPUs revolutionized compute power, or what InfiniBand and RDMA brought in terms of networking innovations.

As the third pillar of the modern data center, after compute and networking, storage has seen its fair share of innovations in an attempt to meet ever more demanding efficiency, performance, reliability, and scalability requirements. And while new NVMe SSDs have gone a long way to helping enterprise and service provider data centers meet demand, they bring their own challenges.

This white paper explains how NVMe technologies have evolved and how Kalray Flashbox™, industry-first, disaggregated NVMe storage array leverages the full potential of NVMe flash devices at a massive scale to become a powerful game-changer solution in terms of performance per Watt and dollar.

The solution leverages a new generation of processors called DPUs (Data Processing Unit), running next-generation Ethernet fabrics and featuring a scalable storage OS and an open, programmable data plane to support massively parallel workloads.

[1] *IDC's Global DataSphere Forecast Shows Continued Steady Growth in the Creation and Consumption of Data*

LEVERAGING NEW STORAGE TECHNOLOGIES

SSDs and the NVMe-oF protocol, developed to deploy NVMe over the network with the same performance as direct-attached or local storage, paved the way towards a new storage era.

KALRAY

The NVMe SSD Storage Revolution

1.1 First Revolution with SSDs

Historically, compute has consistently outpaced storage performance. As network interfaces evolved to support increasing data transfer rates, spinning, low-capacity Hard Disk Drives (HDDs) remained the bottleneck. **The advent of Solid State Drives (SSDs) revolutionized the industry** as SSD drives delivered higher performance and capacity than HDDs.

Flash memory-based SSD drives rapidly became the preferred storage medium in data centers.

However, it quickly became apparent that existing interfaces and protocols were inadequate for leveraging the full potential of SSDs. As a result, high-performance NVMe (Non-Volatile Memory Express) was developed to replace legacy storage protocols.

1.2 Maximizing Performance with NVMe SSDs

Up to 100 times faster than standard SSDs, NVMe SSDs take advantage of the NVMe protocol to unlock the bandwidth, high input/output operations per second, or IOPS[2], and low latency capabilities of solid-state drives, reducing processor overhead and streamlining operations.

In addition, using SSDs like RAM, the NVMe protocol overcomes the limitations of legacy protocols, leveraging high-speed PCI Express (PCIe) network cards to transfer data to and from SSDs over tens of thousands of channels instead of just one.

However, most NVMe SSDs are deployed as direct-attached or local (in-server) storage and can only be consumed by applications running on the server. The result is costly, highly inefficient, **siloed architectures with overprovisioned, underutilized storage pools lacking scalability.**

In addition, direct-attached NVMe increases operational costs (since servers need more storage) and decreases availability due to limited data services and extended recovery times.

1.3 Extending Performance over the Network with NVMe-oF

Overcoming these limitations, the NVMe-oF (NVMe-over-Fabric) protocol standard was developed to enable customers to deploy NVMe over the network with the same performance as local NVMe.

Extending the NVMe protocol to Ethernet and Fiber Channel, NVMe-oF leverages the full potential of NVMe SSDs, increasing the speed and efficiency of data transfer between storage and servers across the network.

While a variety of scale-out NVMe solutions uses the NVMe-oF protocol, it still has its challenges. For example, traditional storage controllers cannot exploit NVMe capabilities, becoming a performance bottleneck when NVMe SSDs are deployed in legacy storage arrays.

Furthermore, **x86 based NVMe solutions add a substantial performance penalty when running data services such as compression, deduplication, erasure coding, and encryption.**

Taking full advantage of the performance advantages of NVMe SSDs - while minimizing tradeoffs - requires a new, disaggregated storage architecture leveraging the advanced capabilities of NVMe to seamlessly connect flash storage across the network.

[2] IOPS is the numbers of Input/Output operations per second and is a performance measurement used to characterize computer storage devices.

MAXIMIZING THE POTENTIAL OF NVMe-oF

Eliminating siloed resources and delivering performance comparable to local storage over the network.

KALRAY

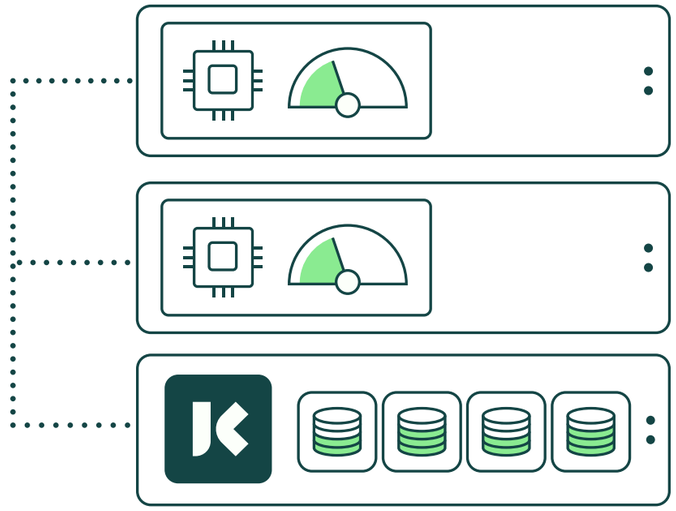
The Disaggregation Challenge

The low-latency, high bandwidth demands of new services and applications - including AI and data analytics - drive enterprises and services providers to deploy vast pools of fast NVMe SSDs. At the same time, however, compute and storage resources are still being deployed and managed as separate siloes, preventing the performance and utilization optimization of the overall environment.

Leveraging NVMe-oF protocol capabilities to deliver scalable, unified storage pools, a disaggregated architecture eliminates siloed resources and offers performance comparable to local storage.

Furthermore, NVMe-oF does not require specialized drivers or expensive network cards, utilizing industry-standard hardware and software.

Disaggregation of you server using Kalray Flashbox™



Disaggregated scale-out NVMe storage appliances enable multiple NVMe devices to function as a logical storage pool with logical volumes allocated to any application across the network. **Designed to be dynamically reconfigurable, disaggregated storage architectures enable physical resources to be reconfigured to maximize performance and minimize latency.**

However, delivering high performance and a complete range of enterprise storage services utilizing NVMe requires a new generation of processors.

THE DPU, THE ADVENT OF A NEW STORAGE REVOLUTION

A new class of data-centric programmable processors designed to eliminate bottlenecks and unleash the full potential of NVMe SSDs in disaggregated architectures.

KALRAY

MPPA® DPU

Over the past decades, chip manufacturers have innovated non-stop to produce faster, more powerful CPUs to enable modern, data-driven applications. However, **CPUs are limited when it comes to the parallel processing demands of data-centric computation**. As a result, some customers report network and storage compute overhead as high as 75%. As a result, expensive CPU cycles are wasted classifying, steering, and tracking network traffic instead of running the application workloads for which they were designed.

Primarily designed for graphics processing, **graphics processing units (GPUs)** can run parallel workloads such as complex matrix operations for AI workloads. At the same time, however, **they were not designed to deal efficiently with multiple workloads and their data.**

3.1 A New Class of Processor: The Data Processing Unit (DPU)

The solution lies in a new class of processors capable of efficiently running data-centric heterogenous processing tasks. The Data Processing Unit - or DPU - is a new generation of programmable processors specially designed to run data center infrastructure services. Rather than embedding functions in the hardware, the DPU leverages software in the control plane to program the data plane. This ability provides almost infinite flexibility, giving storage administrators the power to customize the behavior of storage devices, deploy new protocols, and implement advanced data services.

In a DPU-enabled architecture, CPUs take care of the general-purpose computing workloads while the DPU acts as an intelligent traffic controller. By offloading data-related functions, the high-speed fabric can support modern workloads—faster, more efficiently, and at a lower cost than standard CPUs.

The DPU is essential for scale-out NVMe storage as it enables the peak performance of NVMe SSDs to be sustained while running data services. Instead of moving data from where it is stored to where it is processed, the data is processed where it resides. The result is higher performance and significant cost benefits compared to x86-based architectures subject to data processing bottlenecks.

3.2 Kalray's MPPA® DPU for Intelligent Cloud-to-Edge Processing

Kalray has played a vital role in developing DPUs. Designed before the industry adopted the DPU acronym, Kalray's MPPA® (Massively Parallel Processor Array) Coolidge™ processor is a market-leading solution for intelligent cloud-to-edge processing.

3.3 A Massively Parallel Architecture

Boasting 80 highly efficient cores, the Kalray MPPA® DPU is a data-centric processor offering an unprecedented level of performance and programmability, enabling all critical storage functions to run on a single chip in a disaggregated storage appliance.

The overall architecture of the Kalray 3rd generation of MPPA® aka Coolidge™ is based on a "Massively Parallel Processor Array" architecture, which is characterized by the association of computing clusters connected to each other, to the external memory and to the I/O

interfaces via two independent interconnects, suited to different types of data transfers.

The advantage of the manycore architecture is that a processor can scale to massive parallelism by replicating the computing units and extending the network on chip, whereas for a multi-core processor the replication applies to the core level.

In addition, manycore technology offers industry-leading approach to support both efficient AI computing and heterogenous multiprocessing support.

MPPA® DPU KEY BENEFITS

High Performance

Power Efficiency

Fully Programmable Data-centric

High-Speed Interfaces

To high-speed fabrics & interfaces

Security

Secure islands, encrypt/decrypt, secure boot

Heterogeneous Multi-processing

Run additional value-add services (database acceleration, analytics...)

Modular SPDK-based Software Enviroment

Enabling Efficient Disaggregation

The MPPA® DPU is integrated into the K200-LP™ acceleration cards which power the Kalray Flashbox™, bringing unique and unprecedent capabilities.

UNLEASH THE FULL POTENTIAL OF NVMe SSDS

As an industry-first disaggregated NVMe storage All-Flash-Array, Kalray Flashbox™ is a game-changer in terms of performance per Watt and per dollar.

KALRAY

Kalray Flashbox™ One Step Further

Differentiating itself from other scale-out NVMe solutions, Kalray's Flashbox™ runs all critical storage services on a single Kalray MPPA® DPU.

Kalray's Flashbox™ is the first disaggregated NVMe storage array designed from the ground up to leverage the full potential of NVMe SSDs at massive scale while ensuring the lowest storage Total Cost of Ownership (TCO).

NEXT-GENERATION DPU-BASED NVMe ALL-FLASH-ARRAY

12 MIOPS
High Performance

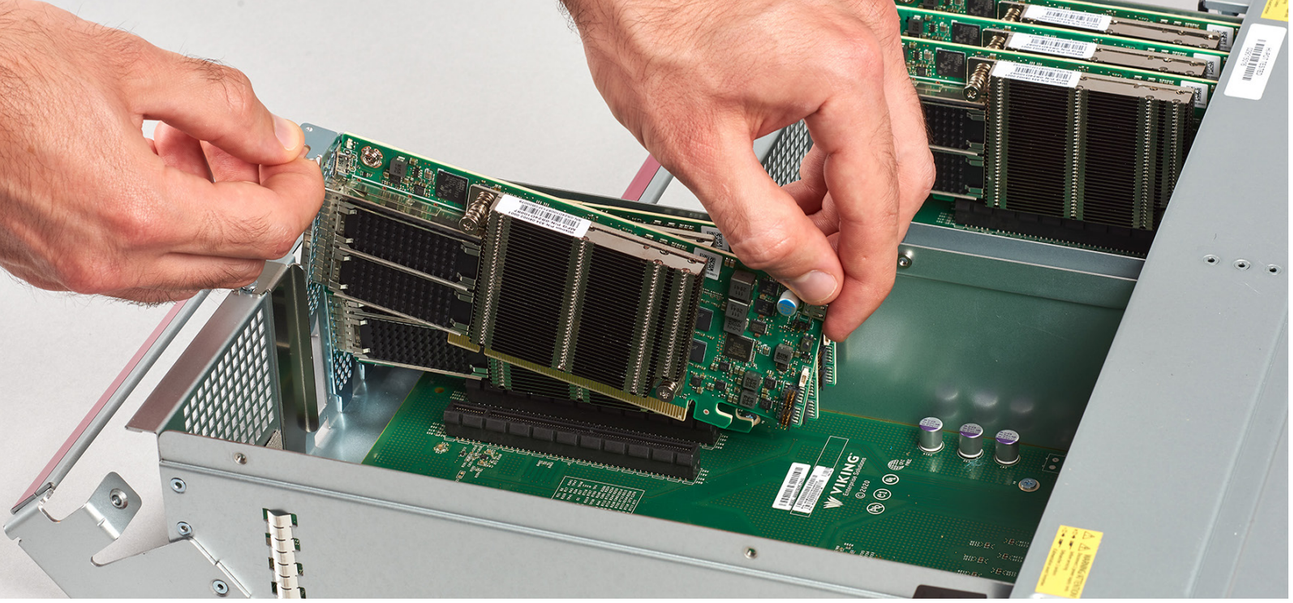
2 to 6 CARDS
High Availability Smart Storage Contollers

Up to **24 SSDs**
Hot-swappable NVMe SSDs / Array

2x100 GbE
Per card, PCIe Gen4, DDR4 @3200Mhz

30µs LATENCY
Low Latency

30 Watt
Power Consumption per Card



4.1 Featuring Kalray MPPA® DPU and Kalray K200-LP™ Smart Storage Adapters

Featuring the Kalray MPPA® DPU and Kalray K200-LP™ Smart Storage Adapters, Flashbox™ allows customers to unlock the performance of NVMe SSDs at scale. Physical devices are abstracted into one scalable pool of high-performance storage with logical volumes allocated to applications across the network as if they were locally attached.

Differentiating itself from other scale-out NVMe solutions, Kalray's Flashbox™ runs all critical storage services on a single Kalray MPPA® DPU, offering an unprecedented level of performance and programmability. In addition, the Kalray Flashbox™ is significantly more cost-effective and delivers much higher performance than x86 based architectures.
Kalray's Flashbox™ revolutionizes how customers use and store high-performance data without compromising availability or scalability.

Running on standard hardware and leveraging industry-standard protocols (NVMe/TCP, NVMe/RDMA, SPDK), Flashbox™ can easily be deployed in existing environments.

4.2 Optimized for the Most Demanding Workloads

Early benchmark results place Kalray's Flashbox™ among the best-ranking solutions on the market. **With up to 12 million IOPS, 72 GB/s throughput, and latency between 20 µs and 50 µs (RoCE & TCP protocols)**, the solution is ready for the most demanding workloads.

Fully resilient with no single point of failure, the Flashbox™ comes standard with independent power supplies, hot-pluggable I/O modules, and hot-swappable drives to ensure your data is always available and protected. In addition, industry-standard management application programming interfaces (APIs) allow you to seamlessly integrate and manage the appliance using existing tools.

4.3 Custom-build Your Own All-Flash-Array

While the Kalray Flashbox™ is available as a turnkey appliance, Kalray's AccessCore® Storage Software Suite and Software Development Kit (SDK) allows customers and partners to custom-build their own all-flash arrays. Based on a modular SPDK-based architecture, the AccessCore software suite can easily be enriched with additional storage and data services.

What's inside Kalray's Flashbox™?

Kalray K200-LP™

Thanks to **Kalray's K200-LP™ Smart Storage Adapter, powered by MPPA® DPU**, deploy NVMe with the ease-of-use and efficiency of traditional storage array architectures, without performance or durability trade-offs.

- 2x100GbE per card, PCIe Gen4, DDR4 @3200Mhz
- Fully programmable control & data path
- Supports Linux OS

Kalray AccessCore® Storage

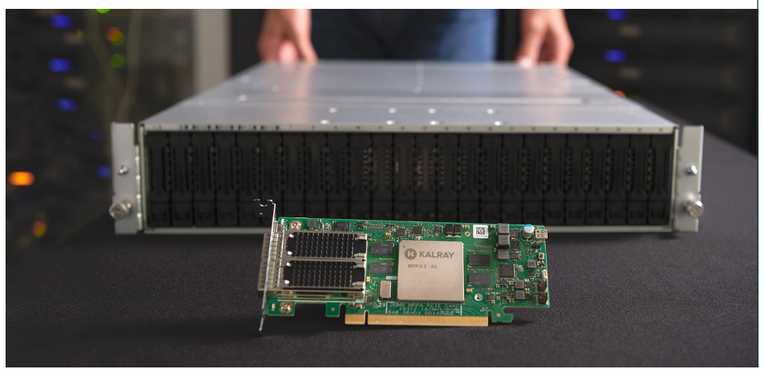
A SPDK-based open software environment with standard APIs and tool chain:

- MPPA® DPU-optimized software modules such as network (ODP) and storage (SPDK) functions.
- Based on NVMe-oF industry standard. Supports TCP and RoCE protocols
- Standard Management APIs

Viking Enterprise Chassis

A fully resilient VDS2249R system chassis:

- Versatile, 2U high, enterprise-class storage solution
- Multitude of features for customizing data storage



Technical Specifications

Hardware

2RU, 19" rack support with rails
Two I/O modules – six fans
Up to 6 adapters
Up to 12 100GbE ports
24x U2 dual-ported NVMe SSDs
Up to 364 TB in 2U with 1.9, 3.8, 7.6 and 15.2 TB SSDs
3-year built-in support

Software

SPDK-based storage suite
RAID 0, 1, 10 and 6 (based on EC algorithms)
Logical Volumes
Thin provisioning
Snapshots
Clones
NVMe-oF (RoCEv1/v2)
NVMe/TCP
Active/active or active-standby, highly available configuration
Configurable in-band or out-of-band system management

Power & Cooling

Power: 2 AC & DC 1600W power modules with 2 independent AC power inputs
Operating temperature: 0°C to 45°C
RoHS compliant
Dimensions: 87mm H x 438mm W x 697mm D 3.43 in H x 17.2in. W x 27.44 in. D
Weight: 25.5 kg

Kalray Flashbox™ NVMe All-Flash-Array

BEST PERFORMANCE PER WATT AND PER DOLLAR

KALRAY

Reduce Your Total Cost of Ownership (TCO)

Enabling efficient NVMe disaggregation, Flashbox™ delivers the most optimal utilization of NVMe resources. Because all storage services run locally on the Kalray MPPA DPU, the performance overhead is reduced to an absolute minimum.

Moreover, since **Kalray's MPPA® DPU was designed to minimize power consumption, the Flashbox™ offers significant advantages in terms of performance per Watt and dollar.** Its low maximum power consumption of **30W is ten times less** than a traditional CPU and as efficient as a single SSD.

A fully loaded system with **6 x Smart Storage Adapters and 24 x NVMe SSDs** consumes less than 600W on average and 900W at peak. Moreover, the Flashbox™ provides the unique capacity of 7PB in a standard 42RU rack, delivering up to 1 TB/s throughput at less than 14kW.

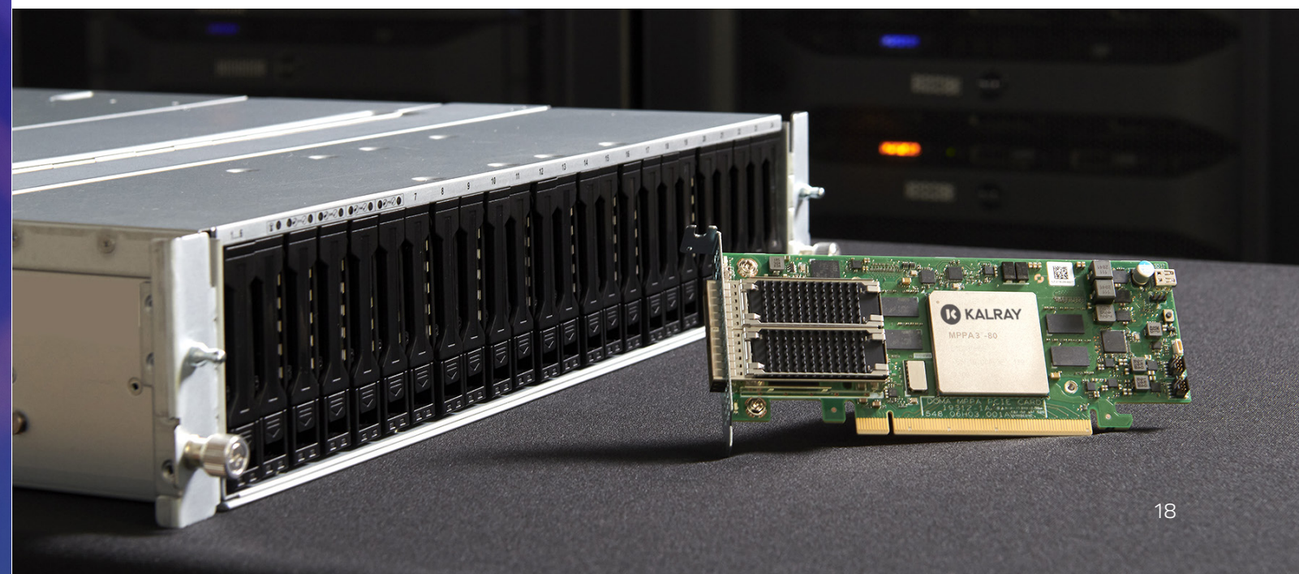
3 to 5 TIMES
Better IOPS/W

15x Better IOPS/$
vs traditional solutions

PLUG & PLAY

PERFORMANCE

COST REDUCTION



#USE CASE

Powering modern applications with next-gen, disaggregated storage.

KALRAY

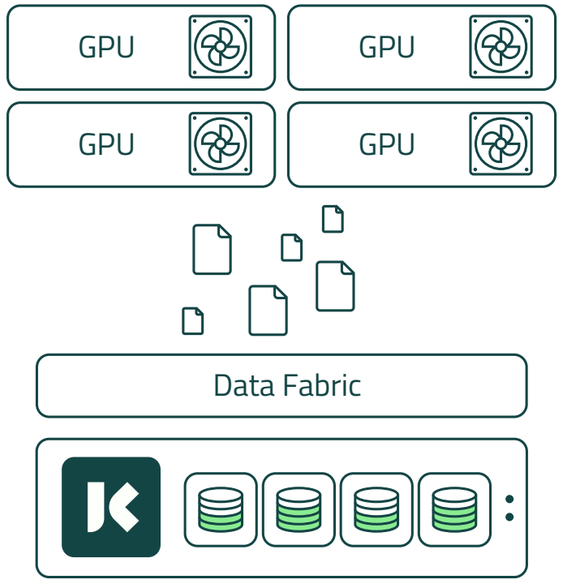
Feeding Data-hungry GPUs for AI and Machine Learning

GPUs are very efficient for Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL) matrix multiplication due to the speed at which they can load data into RAM. However, since AI datasets are usually much larger than the available RAM, **data scientists use NVMe SSDs to refresh data as fast as possible and keep the GPUs running.**

The biggest challenge of NVMe-deployments is the tradeoff between performance and scalability. However, disaggregated NVMe storage allows several NVMe devices to function as a logical storage pool. Logical volumes allocated to AI workloads across the network with full reconfigurability ensure maximize performance and minimize latency.

Leveraging all of NVMe's capabilities, **Kalray's Flashbox™ is the first disaggregated storage array delivering the performance, programmability, and scalability data scientists need to feed their data-hungry GPUs.**

Efficiently feed your GPUs with Kalray Flashbox™



#USE CASE

Boosting parallel file system performance for HPC (High Performance Computing) applications with next generation disaggregated storage.

KALRAY

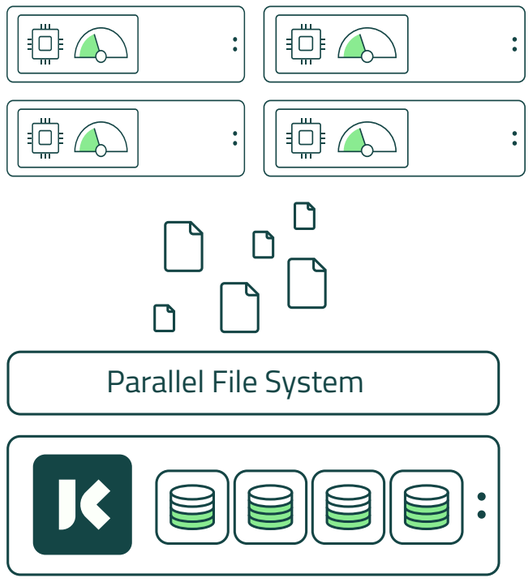
Boosting High Performance Computing

High Performance Computing is no longer the exclusive domain of university labs and government research organizations like NASA or CEA. Supercomputers have become increasingly important in our everyday lives as they enable research for a variety of fields such as weather forecasting, genomics, climate research, clean-tech, etc

HPC infrastructures significantly increase productivity for research organizations but also require higher performance networking and storage.
However, without high-performance data storage, the return on investment (ROI) of expensive infrastructure deployments can be severely affected.

Despite NVMe SSDs offering lower latency and higher performance, most are deployed as direct-attached or local storage, which lacks scalability and is highly inefficient. And even when deployed in large storage clusters under a parallel file system such as BeeGFS, Lustre, or SpectrumScale, there is the risk of stranded capacity. Furthermore, traditional storage controllers cannot handle NVMe performance, and x86 based NVMe storage solutions require expensive adapter cards to meet performance requirements when running data services, impacting TCO.

As the first disaggregated NVMe storage array designed with HPC workloads in mind, Kalray's Flashbox™ eliminates storage tradeoffs, delivering capacity, performance, and scalability without the price. In addition, **Flashbox™ integrates with all popular parallel file systems and guarantees the highest performance per Watt and dollar.**



CONCLUSION

Storage arrays are fundamental building blocks in present-day data centers. In addition, they are a convenient way for customers to deploy storage at scale.

The move from HDDs to SSDs inspired a new generation of faster, all-flash arrays. However, traditional SSDs use HDD interfaces that were not optimized for fast I/O. When NVMe SSDs arrived on the scene, they were up to 100 times faster than previous flash drives as the NVMe interface unlocked increased throughput, higher bandwidth, and ultra-low latency.

However, most NVMe SSDs were deployed as direct-attached or local (in-server) storage that does not scale and is highly inefficient. Moreover, traditional storage adapters cannot handle the NVMe performance, becoming performance bottlenecks when deploying NVMe in legacy storage arrays. In addition, x86-based NVMe storage solutions lack the ease of use of array architectures and require expensive adaptor cards.

Eliminating availability, performance, and scalability tradeoffs, Kalray's Flashbox™ is the first disaggregated NVMe storage array designed from the ground up to leverage the full potential of NVMe flash devices at massive scale.

Featuring the Kalray MPPA® DPU and Kalray K200-LP™ Smart Storage Adapters, Flashbox™ enables customers to deploy NVMe with the efficiency and simplicity of traditional storage array architectures while ensuring the lowest possible storage TCO.

KALRAY
Flashbox™

Revolutionizing the Way Data Is Used and Stored, With the Best Performance per Watt and per $.

THE AUTHORS

Rémy Gauguey

Senior Software Architect, Datacenter Business Line, Kalray

A Senior Software Architect in Kalray's Data Center Business Unit, Rémy Gauguey has more than 25 years of experience in the high-tech industry, specializing in SoCs, RTOS, and high-performance packet processing. Leveraging the Kalray MPPA® manycore technology, Rémy develops advanced architectures for composable disaggregated infrastructure.

Rémy previously developed his expertise at Conexant, Mindspeed Technologies, and the French Atomic Energy Commission (CEA, Commissariat à l'énergie atomique et aux énergies alternatives). He holds several patents in the fields of software architecture and packet processing.

Loic Hamon

VP Corporate Development and strategic marketing, Kalray

Before his time at Kalray, Loïc spent 10 years at Inside Secure. The last position he occupied was executive vice president of corporate development and communication. Prior to that, he was the vice president of Inside Secure's NFC business unit. Loïc also served as director of strategic marketing for the wireless business unit of Texas Instruments after holding several strategic and operational marketing positions within TI.

Loïc has a Master's Degree in Marketing Intelligence at the HEC School of Management in Paris. He has also been awarded a Master's Degree in Electrical Engineering from ESIGELEC in Rouen and a postgraduate degree in Microelectronics from Paris XI University.

ABOUT KALRAY

Kalray (Euronext Growth Paris - FR0010722819 - ALKAL) is a fabless semiconductor company, a leading provider of a new class of processors, specialized in Intelligent Data Processing from Cloud to Edge. Kalray's team have created and developed its leading edge technology and products to help its clients maximize the market possibilities presented by a world dominated by massive, disparate and pervasive data.

Thanks to Kalray's patented manycore architecture, Kalray's MPPA® Intelligent Data Processors are natively capable of managing multiple workloads with no bottlenecks to enable smarter, more efficient and energy-wise data-intensive applications. Kalray's offering includes processors, acceleration cards with associated software environment and appliances, allowing its customers to design the best solutions in fast growing sectors such as modern data centers, 5G, AI and Edge Computing, autonomous vehicles and others.

Founded in 2008 as a spin-off of CEA French lab, with investors such as Alliance Venture (Renault-Nissan-Mitsubishi), Safran, NXP Semiconductors, CEA and Bpifrance, Kalray is dedicated through technology, expertise and passion to offer more: More for a smart world, more for the planet, more for customers and developers.

kalrayinc.com
contact@kalrayinc.com

KALRAY
Flashbox™

Unleash the Full Potential of your NVMe SSDs

Kalrayinc.com