A KAMINARIO WHITE PAPER

# Changing the Data Center Economics with Kaminario's K2 All-Flash Storage Array

November 2014

kaminario.

# Table of Contents

# Executive Summary

Solid-state storage and all-flash arrays have become synonymous with performance. However, historically they have been regarded as expensive and were only used as a solution for specific pain points where their high performance justified the added cost. Kaminario's K2 all-flash array has made this compromise unnecessary.

The K2 has the right combination of predictable performance with consistently low latencies, enterprise storage capabilities, and, at an average price of $2/GB of effective capacity, far greater cost effectiveness than traditional HDD-based or hybrid storage solutions. K2's low cost per effective GB makes it the perfect primary storage tier for new projects and an optimal replacement for aging legacy storage systems supporting key IT initiatives such as server consolidations, databases, and virtual desktops.

This white paper covers the cost efficiency and key use cases for the Kaminario K2 array, detailing the tangible advantages IT groups can realize from its enterprise features and storage efficiency capabilities including global inline selective deduplication, inline compression, high utilization K-RAID™, thin-provisioning and snapshots.

Each of these capabilities and technologies contribute to K2's cost effectiveness in both acquisition (CAPEX) and day-to-day operations (OPEX). In addition, K2 has unique advantages in three key IT use cases:

- **SQL Databases.** K2's consistent high throughput/IOPS and low latency keep performance of databases like Microsoft SQL Server and Oracle at their peak for OLTP, OLAP, and mixed workloads. K2's native inline compression provides up to 2/3rds capacity savings equaling best-case native database compression levels but without any added load placed on the database host server CPUs, eliminating any trade-offs between capacity savings and production performance SLAs.
- **Virtual Desktops**. K2's consistent throughput and IOPS performance, and low latency even under peak loads of provisioning, booting, and powering down thousands of virtual desktops simultaneously, ensure that critical VDI management tasks can continue without reducing desktop responsiveness, and without risk to user satisfaction and productivity. K2's native inline deduplication can reduce the physical capacity required for virtual desktops by over 95%, or 20:1, driving the effective cost per GB to well below $1 or less than $20 per desktop.
- **Server Consolidation.** K2's ability to cost effectively scale-out and scale-up enables consolidation of virtualized servers using a shared high-performance all-flash storage infrastructure. K2's tight integration with VMware's tools and APIs makes management of consolidated virtual servers simple. Even under demanding mixed server workloads, its performance scales linearly making it practical to use the K2 all-flash array as the primary storage tier for consolidating servers.

For details on Kaminario K2's technical architecture, benchmark reports and other information, please visit our website at [kaminario.com/resources/whitepapers.php](kaminario.com/resources/whitepapers.php)

# Cost Efficiency of the K2 All-Flash Array

One of the most important benefits of the Kaminario K2 all-flash array is the significant improvement in cost efficiency. Earlier generations of flash-based storage provided performance benefits such as low latency, but typically at a far higher cost per GB than the traditional HDD-based or hybrid storage. This earlier generation of all-flash storage solutions benefited narrow, niche use cases where extremely high storage performance was required.

The K2 array extends the no-compromise approach of its technical design to the economics of deploying SSD-based storage. Based on lower acquisition and operational costs, simpler management, and advanced capacity-saving capabilities, K2 is able to deliver substantial savings across its entire lifecycle.

## K2 Reduces Acquisition Costs (CAPEX)

K2's lifecycle expenses, including CAPEX and OPEX, can be up to 50% lower than the costs of legacy HDD-based or hybrid storage arrays, making deployment of Kaminario's storage economically feasible. K2's ability to start small and easily scale lets infrastructure investments be matched directly to current requirements. Throughout its lifecycle, K2's price averages $2/GB of effective storage capacity.

K2 sets a new standard in lowering initial deployment costs, or capital expenditures (CAPEX), including:

- **Data Reduction Technologies Lower Storage Capacity Needs.** K2's range of data reduction technologies including inline data compression, global selective inline deduplication, and writeable snapshots, dramatically reduces physical capacity requirements. Depending on the use case, K2's native deduplication and compression can lower the storage capacity needs by as much as 95% without adding management overhead or overloading the host servers.
- **Guaranteed Effective Storage Capacity**. Kaminario is the only all-flash storage array vendor on the market that offers a hard guarantee of the effective storage capacity, taking out the guesswork and anxiety from the storage sizing and deployment process. If the guaranteed effective capacity cannot be reached after the capacity reduction technologies were applied, additional hardware to expand the K2 system will be provided at Kaminario's expense to fulfill that guarantee.
- **All-Inclusive Software**. All advanced storage software including features such as deduplication, compression, snapshots, thin provisioning, adaptive block size, are bundled with the system at no additional cost.
- **Rapid Deployment and Installation.** Installation and integration of the K2 array is simple, requiring just hours rather than days. Support for VMware's vCenter allows to seamlessly integrate the system into a virtualized environment. Other 3rd party integrations are also straightforward with the support for fully scriptable CLI and RESTful APIs.
- **Consolidation Platform**. K2's consistent performance and low latency, even when faced with random I/O traffic blends, make it an ideal platform for consolidating redundant storage and server infrastructure across organization and application silos, eliminating pockets of underutilized and over-provisioned IT resources. K2's adaptive block size algorithm lets one

array support a range of workloads with different characteristics, ensuring optimal performance without requiring the deployment of separate storage or time-consuming manual tuning. As a result, required investment in hardware and software licenses, and in the associated management overhead, can be reduced.

## K2 Reduces Operational Costs (OPEX)

K2's flexible architecture, ease of management, reduced power consumption and advanced resilience and availability features ensure that on-going operating expenditures (OPEX) continue to remain well below that of legacy, hybrid or other all-flash storage solutions. Since storage OPEX costs are often four to six times higher than the CAPEX costs, they can have an enormous negative impact on the overall TCO and need to be very closely watched.

- **Flexible Scale-Up and Scale-Out Architecture**. K2's flexible and non-disruptive scale-up and scale-out design allows a K2 array to grow in capacity and performance as needed, eliminating expensive forklift upgrades. K2 is the only all-flash array on the market that supports both methods of capacity and performance scaling, eliminating the need to make trade-offs that are required with other all-flash arrays. Storage capacity is managed across the entire storage array, not by individual nodes or shelves. Scaling storage does not mean adding more management overhead as it is presented as a single pool of storage, with a single management GUI and automatic load balancing.
- **Reduced Power and Cooling Decreased Data Center Footprint**. Power and cooling and the required rack space for an all-flash array may be as much as 1/10$^{th}$ of the costs of a traditional storage array. K2's compact footprint and lower power and cooling requirements reduce energy usage and limit the use of data center resources. Typical power used by K2 arrays is as low as 170 watts per U of rack space.
- **Non-Disruptive Hardware and Software Upgrades.** Any type of system downtime in a data center is very costly, causing customer dissatisfaction, lost business or missed deadlines. But since the majority of outages are caused by planned downtime from maintenance and life-cycle management, IT managers have considerable control of how to manage system upgrades. K2's hardware and software upgrades or array expansions can be done online with no downtime, no loss of data availability, and no negative impact on performance. Upgrades and expansions retain existing configurations, and are performed automatically without any need for manual intervention or system downtime. Data recovery upon media failure is highly efficient and does not result in a loss of available capacity or significant drop in performance. Enterprise reliability and a proven design built from nonproprietary components keeps data accessible 24/7/365 year after year.
- **No Single Point of Failure.** K2 protects from data loss and a system downtime in case of unplanned failure. It has a fully redundant, no single point of failure architecture for hardware components, with all data and metadata at rest protected by dual-parity K-RAID™. K-RAID™ ensures that a K2 array can sustain up to 3 SSD failures without any data loss, with up to 2 SSD media failures simultaneously per SSD shelf. All of the K2's sub-components are hot-swappable allowing failed components to be swapped out with no downtime.

- **Ongoing Lower Physical Capacity Needs**. The ability to deploy a cost-effective all-flash array relies on the efficiency of the architecture or, in other words, how much effective capacity can be generated from raw physical SSD capacity. The K2 excels in maximizing the available effective capacity using a wide range of native storage efficiency technologies:
    - **Global Deduplication.** K2's global inline selective deduplication can reduce physical capacity requirements up to 95% by eliminating redundant data so that it is stored on the array only once. Deduplication processing is performed globally across all the K2 array's K-Nodes, enabling higher deduplication ratios. As the array scales out, deduplication effectiveness increases. The unique option of selective deduplication allows this feature to be disabled for data with very low redundancy or when additional performance is preferred (such as database files used by Oracle or SQL Server), as well as for security-sensitive applications where deduplication is prohibited.
    - **Real-Time Compression.** K2's inline real-time data compression uses an LZ4 compression algorithm optimized for low latency performance. K2's native compression is highly effective with non-dedupe friendly data sets such as the database files used by Oracle and SQL Server. Its byte-aligned compression prevents internal fragmentation and is performed in a 4KB granularity, ensuring that small reads do not result in unnecessary decompression overhead.
    - **Native Writable Snapshots.** K2's native snapshots are thinly provisioned, resulting in up to a 90% capacity savings compared to full clones. Deploying a snapshot has no impact on production access to the array thus creating additional environments for QA Testing, Development, backup and more. With very little capacity overhead, K2's snapshots provide low cost environments with the same performance characteristics as production.
    - **Thin Provisioning.** All base volumes, snapshots and replicas are thin provisioned, reducing the required amount of capacity in the initial purchase and allowing it to grow non-disruptively over time.
- **HealthShield™ Active Monitoring.** HealthShield, a cloud-based Call Home monitoring, analytics and reporting module facilitates preventive, proactive and automated enterprise level support for K2. For example, it can automatically detect when media nears the end of its life cycle and will alert the customer and Kaminario's 24/7 global support team, letting the component be preemptively replaced before a failure occurs.

- **Ongoing Management.** A comprehensive and easy GUI, shown in Figure 1, combined with a seamless integration with VMware's management tools, further reduces the operational overhead for virtualized servers and desktops. Because of the built-in automation which eliminates the need for tuning or RAID management, no dedicated administrator is needed, reducing the ongoing management costs.


*Figure 1: K2 GUI Dashboard*

# K2 Helps Accelerate Business Outcomes

In addition to the acquisition and operational cost savings, the K2 all-flash array can enable enterprises to realize many "soft" benefits by accelerating business results. In applications that run highly transactional, I/O-intensive workloads, flash storage can help accelerate business by bringing new levels of real-time responsiveness.

- **Faster Time-to-Market for New Services or Products**. Certain applications or datasets require instant response times. K2 all-flash array, with its combination of predictable, consistent performance and macro-efficiency, can help drive new business opportunities in environments such as ecommerce, banking, medical services, manufacturing or web-based advertising.
- **Better End-User Experience.** K2 can boost results in transaction processing applications such as e-commerce, banking, on-line catalogs, store registers and ATMs, where the response times to user requests are critical. Flash storage can help improve the e-commerce web page responsiveness during the peak hours and shorter wait times, resulting in a higher customer satisfaction; slow responses, on the other hand, may mean a loss of business. Faster user access to product features or pricing may mean higher profitability and more web revenue through a higher number of completed and fewer abandoned sales transactions.
- **Quicker Business Insights.** Data warehousing workloads include data mining, trend analysis, business modeling and hypothesis testing that are vital in today's competitive environment for creating business strategy. To stay competitive, enterprises need to process analytical intelligence or generate reports just-in-time to make data-related decisions that provide an immediate ROI, or to formulate or adjust their business strategy based on the incoming sales results. For these workloads, the K2 all-flash array can help deliver quicker business insights by increasing the speed of completing batch jobs.

# Key Use Cases for the K2 All-Flash Array

K2's advanced enterprise storage technologies bring unique performance advantages and cost efficiencies to three key areas of IT operation: databases, virtual desktops, and server consolidation.

## SQL Databases

Getting optimal performance from SQL and other types of databases, including Microsoft SQL Server and Oracle Database servers, can be challenging for their supporting storage infrastructure. Production systems must be able to quickly process revenue-producing transactions around the clock, while simultaneously executing demanding I/O operations like backups, database replication/failover, and OLAP business analytics and reporting.

K2's ability to maintain consistent throughput and low latencies even in the face of demanding mixed database workloads, along with native technologies like writeable snapshots and inline compression, make it cost-effective to deploy SQL databases on an all-flash storage infrastructure:

- **Database Consolidation.** Consolidating databases running on a K2 array can greatly reduce the required investment in software licenses and server hardware. K2's consistent performance and writeable snapshots can allow long-running analytical and report workloads to share the same storage array with time-critical transaction processing.
- **Efficient Database Cloning**. Using K2's native writeable snapshots to clone or backup databases saves time and prevents disruption to ongoing transaction workloads. Snapshots can also save up to 90% of the capacity needed for cloned databases and let OLAP analytics run against snapshots without risk to production SLAs.
- **Capacity Savings from Compression**. K2's native compression can save up to 2/3rd of the physical capacity required for databases – equivalent to the best results from using the native compression of SQL Server and Oracle databases with no decrease in throughput and no extra load on the host server CPUs.
- **Consistent High Database Performance.** Benchmarks have demonstrated K2's lead in database performance for both Online Transaction Processing (OLTP) and Online Analytical Processing (OLAP) workloads. Benchmarks using the HammerDB OLTP-based transaction test tool showed peaks of 2.5 million transactions per minute. In OLAP benchmarks a single K2 array was able to process an average of 7.6 million row insertions in 10 seconds.
- **Backup and Recovery/Failover.** K2's native writeable snapshots enable rapid database backups to complete more quickly and with no impact on production transaction and query processing. K2 snapshots can also be used to create point-in-time database copies for analytics, development or test purposes. Benchmarking of K2 performance during database recovery and failover via a secondary replica database showed the K2 array still supported high OLTP workloads with minimal disruption during both failover and recovery processing.

## Virtual Desktops

When used to support virtual desktop infrastructure (VDI) deployments, K2 delivers both significantly improved performance and reduced costs. K2's consistent performance under load extends the scope of VDI from undemanding task workers, all the way up to performance-sensitive expert users where it was previously infeasible, keeping all of the VDI users productive and satisfied:

- **Reduced Desktop Capacity Requirements.** K2's global inline deduplication lowers virtual desktop storage physical capacity needs by up to 95%, or 20:1 in terms of data reduction ratios. K2 pricing can be below $1 per effective GB or $20 per virtual desktop.
- **Highly Responsive Virtual Desktops.** Keeping users productive and satisfied is critical to the success of any VDI deployment. K2's consistent low latency and predictable high IOPS/throughput deliver responsive virtual desktops and ensures both IT and users can get their work done. In a benchmark run with 1,250 virtual desktops on a single K2 array, the mean application response time of 0.38 seconds was well below VMware's View Planner's passing mark of 1.0 seconds
- **Rapid Provisioning of Virtual Desktops.** Benchmarks have demonstrated K2's ability to support rapid deployment of VMware View virtual desktops. During a test in which 1,000 desktops were deployed, the K2 array remained highly responsive with IOPS peaking at 70,000 while latency remained below 1.0 ms.
- **Able to Weather Boot/Shutdown Storms**. Simultaneous startup and shutdown of large groups of virtual desktops typically create storms that bring legacy storage to its knees, resulting in downtime and unhappy users. Benchmarks, run on a single K-Block K2 array holding 1,500 virtual desktops, maintained latency below 1.0 ms.
- **Simple VMware Integration**. K2's integration with VMware via a vCenter plug-in (Figure 2) and support for VMware's VAAI hardware acceleration API enables K2 arrays to be simply and easily integrated into any VMware virtualized desktop and server infrastructure.

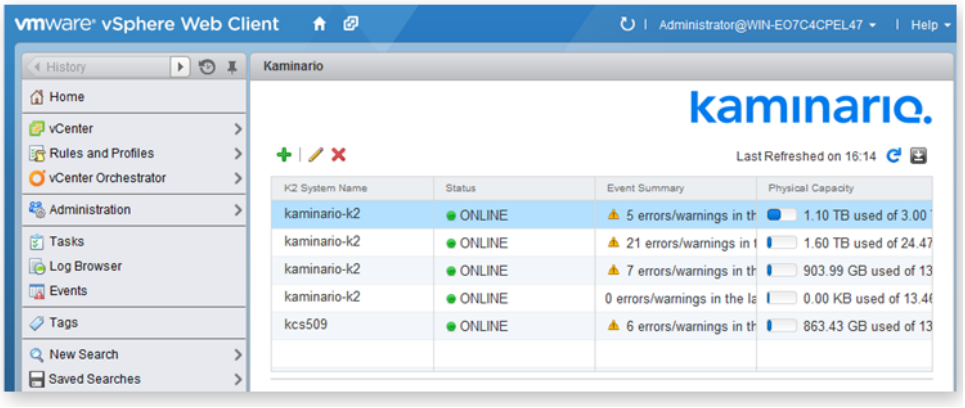


*Figure 2: VMware vSphere Web Client, displaying the Kaminario vCenter Plug-in*

### Server Consolidation

Consolidation of virtualized servers has provided IT groups with both substantial cost savings and greatly increased operational flexibility. However, consolidated servers can produce a widely varying and random blend of I/O demands that can seriously degrade the performance of most legacy and hybrid storage arrays. K2's ability to efficiently scale up and out to handle demanding mixed workloads and blended I/O make it an ideal primary storage tier for consolidating virtual servers:

- **Reliable Performance Regardless of Workload Consolidation.** K2's intelligent design provides consistent low latency and high throughput/IOPS, across even the random blends of concurrent I/O requests typical in consolidated server workloads. This reliable performance under load ensures that the K2 array does not become a bottleneck even during unanticipated periods of peak demand.
- **Highly Linear Scale-Out of Performance.** In benchmarks where host server and K2 array resources were scaled out by a factor of two, OLTP and OLAP workloads run in virtualized database servers, showed that the throughput increased by 1.7X and the amount of IOPS doubled. K2's low latency is maintained from its initial entry-level configuration to a fully scaled array, delivering predictable server and application performance.
- **Lower Licensing and Server Costs**. Consolidation of IT infrastructure silos and elimination of redundant and over-provisioned server and storage hardware by standardizing on a K2 technology supports the same number of servers, applications and users with less hardware cost and fewer software licenses.
- **Efficient Provisioning of Development, Test and Training Servers**. Like virtual desktops, the sandboxed virtual servers needed for internal development, QA testing and user training can benefit greatly from K2's inline deduplication which can reduce physical capacity requirements by up to 95%. K2's high throughput and consistently low latency allows IT and training staff to quickly clone the virtual servers without impacting the performance of production servers sharing the same array.

## Summary

The Kaminario K2 all-flash array is the right choice for primary storage tiers supporting virtualized servers and desktops or business-critical production operations like Oracle or SQL Server databases. Most importantly, the advanced technologies of K2 provide capacity savings and operational efficiencies that make the K2 all-flash array more cost effective than hybrid or legacy storage.

The K2 lets IT leverage the performance advantages of flash storage without any compromises due to high deployment costs or demanding management workloads. Any IT initiative that can benefit from the consistent performance of the K2 storage array can finally get that boost with no technical compromises or cost penalties.

For more information, please visit our website at www.kaminario.com.