A Kaminario Reference Architecture: Reference Architecture for Running SQL Server on ESXi

December 2017

TABLE OF CONTENTS

Executive Summary

The Kaminario K2 all-flash array is built to support modern workloads that require efficient use of resources, high performance and flexibility in scale. Virtualized environments, that are used to increase the utilization of servers and increase productivity, are an ideal workload for the K2's all flash architecture. One such environment is SQL Server deployed via VMware vSphere using ESXi. Kaminario has a unique value proposition for this implementation, and it is recommended that best practices and reference guides are followed in order to maximize capacity and performance efficiencies.

This reference architecture describes in detail all integration points of the K2 all-flash array, SQL Server, VMware ESXi, Kaminario best practices for ESX environments, and Kaminario SQL Server configuration best practices.

The document also includes examples and screenshots that facilitate easy understanding of the different settings and clearly demonstrate the tradeoffs between different options of configurations. Content in this document is quite technical and is intended to be used by pre-sales engineers, system and storage administrators and customers who want to deploy SQL Server on the Kaminario K2 with VMware ESXi environments.

Introduction to Kaminario K2

The K2 all-flash array is an appliance which is a combination of tested and packaged hardware, software and services. K2's Gen6 hardware platform is based on leading off-the-shelf enterprise components that enable K2's software-defined architecture and software stack. The K2 runs Kaminario VisionOS™, the next-generation flash operating system stack, that provides the core software and advanced data services framework.

The value proposition for VisionsOS is its ability to enable the following modular components and services across a real scale-out storage platform:

- **DataShrink** – Data reduction features and capabilities are mandatory for economics of flash storage. With differentiating inline, global, adaptive and selective deduplication, together with inline byte aligned compression, thin provisioning and zero detection, Kaminario is able to establish itself as a cost-efficiency leader of flash storage.

- **DataProtect** – Kaminario values its customers' data more than anything. Native array based snapshots and replication allow for returning to any point in time at any site. Data-at-rest AES256 encryption makes sure that data is kept private and safe at all times. A highly resilient design to eliminate single points of failure, non-disruptive upgrades (NDU) and a robust RAID scheme facilitate 99.999% of data availability.

- **DataManage** – The K2 can be managed by various means. Internal management includes an intuitive web-based GUI, a scriptable CLI and a fully programmable RESTful API platform.

- **DataConnect** – K2's RESTful API allows for external applications of the IT eco-system to easily integrate and seamlessly manage the K2. This eco-system is constantly growing and includes: VMware vSphere, Microsoft VSS, OpenStack, Flocker (containers) and Cisco UCS director.

Microsoft SQL Server

Microsoft SQL Server is one of the most widely used databases in the world. Most organizations often run both production and revenue-generating applications on redundant instances of SQL Server. SQL Server is often preferred by developers for its simplicity and ease of use, enabling the use of a single database to perform tasks such as Online Transaction Processing workloads (OLTP), data mining, and Online Analytical Processing workloads (OLAP). One limiting factor about SQL Server is its host utilization – needing to utilize all of the available CPU, memory and disk. This often leads to either underutilized or a lack of resources based on host configuration. The growth of virtual environments has helped overcome this hurdle. Virtualizing SQL Server allows for the optimization of resources and the ability to add resources as needed and migrate to newer hardware without incurring expensive costs. Storage is another critical component in running SQL Server, with different workloads requiring different I/O patterns and levels. Supporting VMware hosts with scalable storage is just as important for achieving the desired performance from the SQL Server environment.

Deploying SQL Server on VMware ESXi over K2 Best Practices

ESXI Host Configuration for SQL Server Installation

It is important to remember that ESXi host configuration on VMware impacts the performance of SQL Server. Specific host configurations are recommended in order to achieve the desired performance levels from SQL Server environments. However, not all settings are applicable to every host.

BIOS/UEFI SETTINGS

- Enable Turbo Boost
- Enable hyper-threading
- Enable NUMA settings
- Enable advanced CPU features as VT-x/AMD-V, EPT and RVI
- Set Power Management to "OS controlled"

ESXI SETTINGS

Power Settings

Power management provides the ability to choose between performance and efficiency. It is recommended to set the Power Policy Settings to High Performance when there are critical SQL Server VMs. By default, ESXi uses Balanced Power Policy. See the Figure below for the different Power Policy Settings.
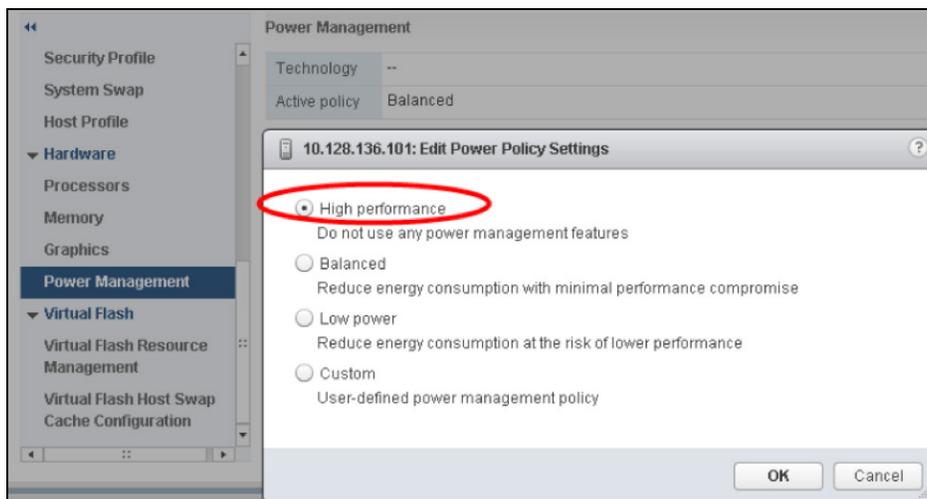
Figure 1: ESXi Power Policy Settings

K2 Best Practices for VMware ESXI Environment

ESXi has many settings that affect the K2's performance. This list details the major parameter settings and Kaminario's recommendation for each parameter.

Parameter	Scope	Default Value	Recommended Setting	Importance
Path Selection Policy	Per Volume, per ESXi host	Fixed	Round-Robin	Critical. Better utilization of K2 Ports
Number of Outstanding IOs With Competing Worlds	Per Volume, per ESXi host	32	256	Critical. Allows "pushing" more load to the K2.
Disk.SchedQuantum	Per ESXi host	8	64	Low. Less significant for All-Flash Arrays
VAAI Primitives	Per ESXi host	Enabled	Enabled	Critical. Offloads heavy workloads from the ESXi host to the K2
QLogic HBA Settings	Per ESXi host	Queue Depth=64, ZIO=1, IDT=100	Queue Depth=400, ZIO=6, IDT=1	Medium. Can increase performance significantly but may not always fit the whole eco-system
Emulex HBA Settings	Per ESXi host	lpfc_lun_queue_ depth = 30-128	lpfc_lun_queue_ depth = 30-128	Medium. Can increase performance significantly but may not always fit the whole eco-system

For additional information of the parameters and different test results refer to *K2 VMware vSphere Reference Architecture in Kaminario Support Portal*.

VM CONFIGURATION

This section summarizes Kaminario's recommendations of ESX VM configuration for SQL Server VM. Detailed information on each configuration appears after the table.

Setting	Default Value	Recommended Setting	Comments
Power Settings	Balanced	High-Performance	
vCPU Allocation	1	Exact number of vCPU that is required at the installation time	
NUMA Consideration	–	Equal or less vCPUs than the number of physical cores of a NUMA node	
Cores Per Socket	1	Depends on ESXi version.	Different in 6.5 and previous versions
CPU Affinity	No Affinity	No Affinity	
Memory Allocation	4GB	According to what needed by SQL Server with ESX and SQL Server buffers	Reserve All guest memory
Storage VMFS or RDM	VMFS	VMFS or RDM. No Performance Impact	Use of K2 Snapshots Requires RDM Disks
VMDK File Layout	Thick Provision Lazy Zeroed	K2 Supports all VMDK File Layouts	
SCSI Controller	LSI Logic SAS	Paravirtualized	

POWER SETTINGS

Windows allows users to select different Power Plans for its servers. Power Plans enable the option to select between a high-performance environment or a high-efficient environment (performance per watt). The default windows setting is Balanced, while it is recommended for critical SQL Server installations to change the setting to High Performance. The Figure below shows the different Power Plan options.
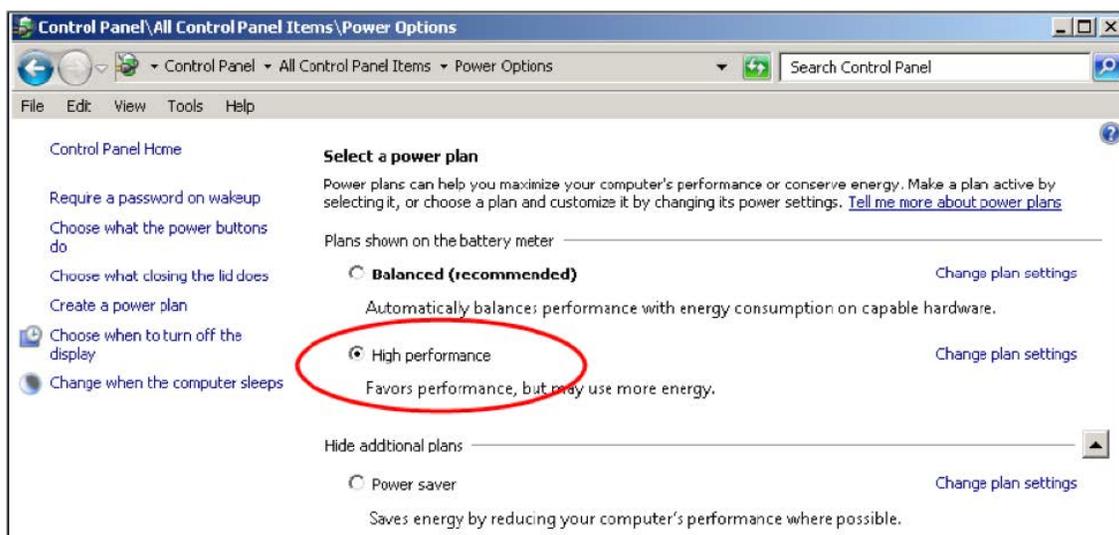

Figure 2: Windows Power Plan Options

VCPU Allocation

Deciding the right vCPU allocation for SQL Server VMs is different from the decision-making process when using physical SQL Server hardware. In a virtualized environment adding or removing a CPU from a VM is a very easy task. It is recommended to provide the VM with the exact number of vCPUs required at the installation time. It is also recommended not to give the VM more vCPUs than the number of physical cores that the ESXi has. In case the number of vCPUs is not sufficient, it is possible to add more.

NUMA Consideration

• In case of a small SQL Server installation, it is recommended to allocate the VM equal or less vCPUs than the number of physical cores of a NUMA node.

• vNUMA is enabled by default only for a VM with nine or more vCPUs.

Cores per Socket

Cores per Socket configuration enables users the option to configure the number of cores per CPU socket. This setting was created to assist users in licensing limitations and with applications that have limitations with number of cores. The settings effect performance together with NUMA considerations. There are a few factors to take into consideration:

- A VM with fewer CPU cores then a physical NUMA node will be assigned to one NUMA node. The number of sockets will not affect the vNUMA. Still, it is recommended to use the fewest number of vSockets possible.

- It is recommended to assign an even number of vCPUs in case there is a need to assign more vCPUs than physical NUMA node cores. An uneven number of vCPUs will result in sub-optimal configuration.

- Starting with vSphere 6.5 the number of cores per socket does not affect the vNUMA configuration by default.

- Prior to vSPhere 6.5, the number of cores per socket directly affected the vNUMA. It is recommended to be aware of the underlying physical NUMA topology before configuring the cores per socket.

- In case a VM is configured with more memory than of a physical NUMA node and less vCPUs than a physical NUMA node, it forces memory to be fetched remotely, degrading performance. In that case, it is recommended to assign more vCPUs to the VM forcing the hypervisor to assign vNUMA or lower the numa.vcpu.min setting to lower than nine.

CPU Affinity

It is recommended not to use CPU affinity as it is limits the hypervisor's ability to schedule vCPUs on the physical server efficiently.

MEMORY CONFIGURATION

- A VM should generally not have more memory than its ESXi physical memory. It is good practice to leave some memory for ESXi overhead and VM overhead. For more information about ESXi and VM overhead refer to the *Vsphere Resource Management.*

- To calculate how much memory is needed for a VM, use the below formula:

```
VM Memory = SQL Max Server Memory + ThreadStack + OS Mem + VM Overhead

ThreadStack = SQL Max Worker Threads * ThreadStackSize

ThreadStackSize = 1MB on x86

               = 2MB on x64

               = 4MB on IA64

OS Mem: 1GB for every 4 CPU Cores

VM Overhead: The amount of memory a VM need to power on. After a VM is running, the
amount of overhead memory might be different. Next table shows a sample of Overhead Memory
on Virtual Machines
```

Memory (MB)	1 vCPU	2 vCPU	4 vCPU	8 vCPU
256	20.29	24.28	32.23	48.16
1024	25.90	29.91	37.86	53.82
4096	48.64	52.72	60.67	76.78
16384	139.62	143.98	151.93	168.60

Table 1 Sample of Overhead Memory on Virtual Machines

More information can be found in the *Vsphere Resource Management*.

It is recommended to reserve the SQL Server VM memory to eliminate the possibility of ballooning or swapping memory. Ballooning or swapping memory can cause performance degradation. The figure below shows how to reserve guest memory.
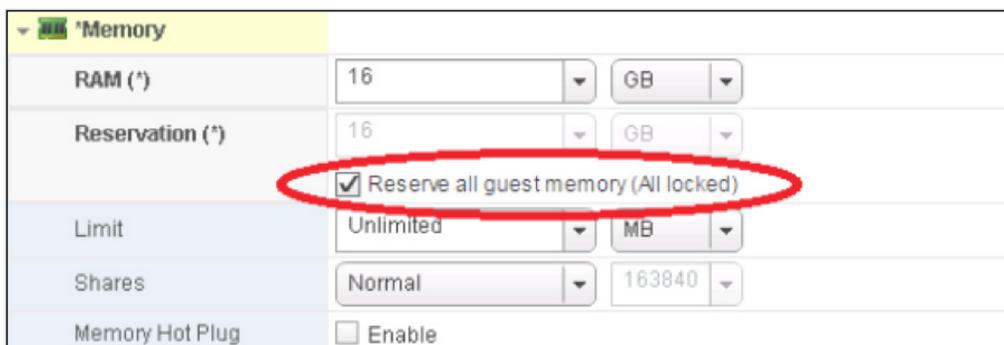

Figure 3: Guest Memory Reservation

STORAGE CONFIGURATION

VMFS or RDM

VMware Virtual Machine File System (VMFS) and RDM (Raw Device Mapping) are the most common disk options that VMware provides for its guests. K2 supports both options, creating a LUN that is used as a VMware datastore which facilitates VMDK disks or LUNs that are used as RDM disks. In terms of performance there is no difference when choosing one over the other. However, it is important to understand that there are other considerations when choosing between the options.

- VMFS does not support FCI clustering between hosts.

- VMFS does not support Kaminario K2 snapshot. In order to utilize the K2 snapshot mechanism, RDM disks are required. Using Kaminario VSS, K2 can take an application-consistent snapshot of SQL Server databases with no performance impact and no additional capacity. Kaminario K2 snapshots can be mounted for read/write purposes, which serve to create additional working environments such as QA, Test & Dev, analytics, backup and more, all at a very low cost of storage capacity. Read/write snapshots deliver the same performance as production volumes, without any impact on production volumes. The duration of creating a snapshot has no dependencies on the number or size of the volumes being snapped or how big the array is. For more information about K2 snapshots for SQL Server refer to *SQL Server Backup and Restore using Kaminario K2 Snapshots.*

VMDK File Layout

VMFS disks can be deployed in three different formats: Thin Provisioned, Thick Provisioned Lazy Zeroed, Thick Provisioned Eager Zeroed. The difference between the options is the time to employ zeroes on the disk.

1. Thin Provisioned – Enables 100 percent storage on demand. Disk space is allocated and zeroed during writes.

2. Thick Provisioned Lazy Zeroed – Disk storage is pre-allocated. Disk is zeroed the first time the disk is written.

3. Thick Provisioned Eager Zeroed – Disk is pre-allocated. Disk is zeroed during provision time.

Kaminario K2 implements all VMware's vStorage APIs for Array Integration (VAAI). There is no performance implication when choosing either one of the file layouts. For more information regarding Kaminario K2 implementation of Vmaware VAAI refer to *K2 VMware vSphere Reference Architecture.*

PVSCSI Virtual Adapter

It is recommended to use VMware Paravirtualized SCSI Controller as the virtual SCSI Controller. The PVSCSI Controller is the optimal SCSI controller for an I/O-intensive application on vSphere. For more information refer to *Configuring disks to use VMware Paravirtual SCSI (PVSCSI) adapters.*

Kaminario SQL Server Configuration Best Practices

SQL Server has several best practices for best performance. This table summarizes those best practices and Kaminario's recommendation for each one.

Best Practice	SQL Server Recommendation	Kaminario Recommendation	Comments
Raid Configuration	RAID 10	K-RAID	No RAID configuration is needed
Separation of Files	Allocate different file types for different LUNS for best spindle performance	Separation of files is unnecessary for best performance	Separation of different database files is recommended when using K2 snapshots
Filegroups or Number of LUNs	Consider when more spindles are needed and for administration needs	No need for performance improvement	Can be considered for administration needs
Number of Files	More than one data file on different spindles for performance. TempDB files as the number of CPUs (up to eight)	TempDB files as the number of CPUs (up to eight) No need for more than a single data file for best performance	
Allocation Unit Size	64K	64K	

Contact

Contact a business development representative to answer any questions you may have.

Schedule a Demo

Schedule a demo with an engineer and learn if Kaminario's solution works for you.

Request a Quote

Request a quote for your application from our business development team.

About Kaminario

Kaminario, the leading all-flash storage company, is redefining the future of modern data centers. Its unique solution enables organizations to succeed in today's on-demand world and prepares them to seamlessly handle tomorrow's innovations. Only Kaminario K2 delivers the agility, scalability, performance and economics a data center requires to deal with today's cloud-first, dynamic world and provide real-time data access -- anywhere, anytime. Hundreds of customers rely on the Kaminario K2 all-flash array to power their mission critical applications and safeguard their digital ecosystem. Headquartered in Needham, MA, Kaminario works with an extensive network of resellers and distributors, globally.

For more information, visit www.kaminario.com

Kaminario and the Kaminario logo are registered trademarks of Kaminario, Inc. K-RAID, and Perpetual Array are trademarks of Kaminario, Inc.

Product specifications and performance are subject to change without notice.

The Kaminario ForeSight program is subject to terms and conditions.