kaminario.

WHITE PAPER

Cisco UCS and Kaminario K2

A Reference Architecture

October 2016

TABLE OF CONTENTS

2 Audience and Scope

2 Executive Summary

3 Cisco Unified Computing System (UCS)

4 Kaminario K2 All-Flash Storage Arrays

15 Joint Solution Overview and Benefits

19 Cisco UCS Manager Configuration Overview

32 Basic UCS Networking Implementations

35 Kaminario K2 Storage Configuration

39 VMware ESXi

41 Validation

43 Conclusion

44 About Kaminario

Audience and Scope

This document highlights the value of Kaminario K2 all-flash storage arrays combined with the Cisco Unified Computing System (UCS) architecture for enterprise class workloads. The combined innovative technologies of the Kaminario storage arrays and the Cisco UCS compute and networking platforms provide an ideal infrastructure platform for databases, virtual environments, and big data as this document outlines.

This document provides the following:

- Reference architecture for a solution with Kaminario all-flash arrays (AFAs) and Cisco UCS Servers
- Validated performance test results
- Outline describing why Kaminario combined with Cisco is an ideal infrastructure for all workloads

The audience should have a basic knowledge of servers, storage and at least an introductory level knowledge of Kaminario storage arrays and Cisco UCS.

Executive Summary

Customers who are deploying enterprise critical applications sometimes are forced to make difficult and limiting decisions for their infrastructure. Most application workloads are not consistent in their performance and capacity requirements, further compounding the difficulty of the infrastructure decision. A series of misleading choices often leads to over-engineering and thus overspending, or running systems that cannot meet the application demands. This paper illustrates an approach for designing and configuring the compute, storage and networking infrastructure.

An ideal compute, network, and storage infrastructure for today's enterprise would have great agility and more than meet all of the technical requirements. It would have a variety of easy to change compute and network personalities, and a storage layer that would support multiple storage protocols, contain multiple inline space saving technologies, and optimize the use of Flash technology from a performance, capacity, and cost perspective.

The combination of Kaminario K2 storage and the Cisco UCS compute platform is an ideal collective infrastructure to meet these ever-changing requirements.

- High bandwidth converged compute and networking layer from Cisco to optimize various workloads.
- Storage optimized along both the performance and capacity vectors of flash drive technologies. Providing great performance as well as large capacities.
- Ability to easily scale on both the compute and the storage level with Cisco UCS architecture and the Kaminario K2 scale up and scale out capabilities.
- Complete DR solution using Kaminario asynchronous replication at the array level along with UCS service profile exports and imports between different domains.
- Fully programmable infrastructure to reach automation goals using APIs for both UCS and Kaminario.

Cisco Unified Computing System (UCS)

The Cisco Unified Computing System addresses many of the challenges faced by database administrators and their IT departments, making it an ideal platform for mission critical implementations.

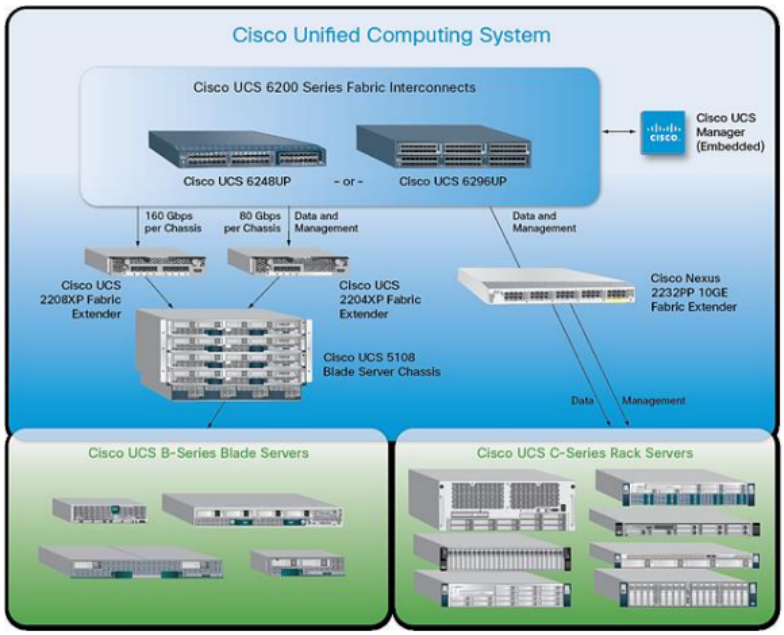


Figure 1: Cisco Unified Computing System Configuration

Comprehensive Management

The system uses an embedded, end-to-end management system with a high-availability active-standby configuration. Cisco UCS Manager uses role and policy-based management that allows IT departments to continue to use subject matter experts to define server, network, and storage access policy. After a server and its identity, firmware, configuration, and connectivity are defined, the server, or a number of servers like it, can be deployed in minutes, rather than the hours or days that it typically takes to move a server from the loading dock to production use. This capability relieves database administrators from tedious, manual assembly of individual components and makes scaling a database environment a straightforward process.

Radical Simplification

The Cisco Unified Computing System represents a radical simplification compared to the way that servers and networks are deployed today. It reduces network access-layer fragmentation by eliminating switching inside the blade server chassis. It integrates compute resources on a unified I/O fabric that supports standard IP protocols as well as Fibre Channel through FCoE encapsulation. The system eliminates the limitations of fixed I/O configurations with an I/O architecture that can be changed through software on a per-server basis to provide needed connectivity using a just-in-time deployment model. The result of this radical simplification is fewer switches, cables, adapters, and management points, helping reduce cost, complexity, power needs, and cooling overhead.

High Performance

The system's blade servers are based on the Intel Xeon 5670 and 7500 series processors. These processors adapt performance to application demands, increasing the clock rate on specific processor cores as workload and thermal conditions permit. These processors, combined with patented Cisco Extended Memory Technology, deliver database performance along with the memory footprint needed to support large in-server caches. The system is integrated within a 10 Gigabit Ethernet-based unified fabric that delivers the throughput and low-latency characteristics needed to support the demands of the cluster's public network, storage traffic, and high-volume cluster messaging traffic.

Kaminario K2 All-Flash Storage Arrays

Flash as a storage media is a game changer. It finally allows storage to match the evolution, progress and performance of CPU power and networking. However, to use flash in a storage array requires inventing a whole new storage architecture, since the legacy architectures are tightly designed to match the characteristics of spinning disks. SPEAR, the Kaminario K2 operating system, is designed from the ground up to utilize flash with the insight of preparing for the next generation of flash, CPU and networking. As an example, the K2 uses SSDs in its array, and SPEAR's metadata design can accommodate any SSD capacity size.

Kaminario K2's architecture is a facilitator for driving Flash as primary storage to the next level with the following tangible business benefits:

- **Cost efficiency** – Flash storage is considered expensive, but with the right storage efficiency features such as global inline selective deduplication, inline compression, thin-provisioning and Kaminario K-RAID™, the CAPEX of the K2 AFA is lower than legacy hard disk drive (HDD) storage. The K2 AFA is also much more economic on power, cooling and footprint so the OPEX is considerably lower. Native capacity-efficient snapshots can be utilized to create multiple production-like environments for even better ROI.
- **Enterprise resiliency** – SPEAR's high availability combined with non-disruptive upgrades results in continuous availability with no planned down time. However, availability without consistent performance is not worth much. SPEAR also facilitates excellent performance during a failure, so the productivity of the environment is not impacted. Native snapshot and replication deliver data protection with quick restore capabilities. The Kaminario HealthShield™ component monitors the array and offers proactive and preventive serviceability.
- **Business agility** – Being able to adapt to the customer's datacenter growth means the storage array needs to grow in capacity and/or performance. The K2 has the unique ability to scale out and/or scale up, non-disruptively, to accommodate such growth. The K2 scales, but it is still a single pool of storage, with a single pane of management and with automatic load balancing. Last but not least, the K2 accelerates production with Flash performance that serves mixed workloads via a global variable block size algorithm. It delivers better user-experience in VDI environments, removes the I/O-blender effect from virtual servers and allows the customer to receive real-time reports from analytics environments (OLAP) and faster database queries in OLTP environments.

System Overview

General

The K2 operating software is the secret sauce that binds best-of-breed enterprise hardware components to an All-Flash Array. K2 is software-defined, meaning it is agnostic to any hardware or technology. The K2 can easily and economically adapt to any advancements made in CPU power, networking and Flash media. Scalability combined with non-disruptive upgrades guarantees continuity and progress of both hardware and software within the array.

K-Block

The K-Block is the building block of the K2 All-Flash Array, and encapsulates the following hardware components:

- Two K-Nodes, each a 1U storage controller
- Each K-Node provides full redundancy for its components:
  - Two 8Gbps FC ports and two 10GbE iSCSI ports for external host connectivity
  - Two 40Gbps InfiniBand ports for K-Node backend interconnectivity
  - Two 6Gbps SAS ports for internal connectivity with the K-Block's SSD shelves
  - Two hot-swappable PSUs
  - Two hot-swappable Battery Back Units (BBU)
- Base SSD shelf, 2U in size:
  - 24 hot-swappable SSDs; Enterprise grade SSD with MLC or TLC Flash
  - Two hot-swappable controllers
  - Two hot-swappable PSUs

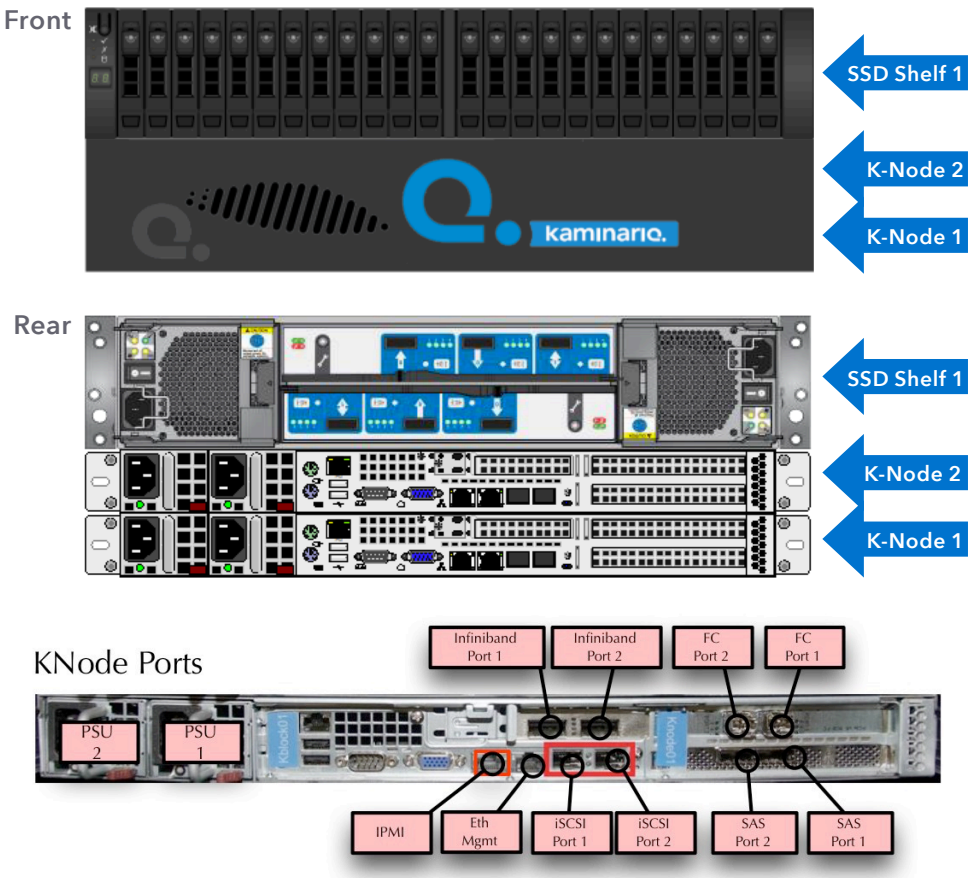


Figure 2: Basic K-Block Components K2 V5 Architecture

Scalable Architecture

SPEAR is designed to facilitate linear growth in both capacity and performance while maintaining consistently low latency, thus implementing a scale-out architecture. In addition, it has the ability to facilitate only growth in capacity with no impact on performance, thus implementing a scale-up architecture. The combination of both scale-out and scale-up architectures in a single storage array is the key feature for building a storage infrastructure that can scale in the most cost-effective way, ensuring that the exact datacenter requirements for new and existing applications are met. Any increase in capacity, results in an automated rebalancing of data within the array, with no intervention or human management.

The starting point for any K2 All-Flash Array configuration is a single K-Block, as shown in Figure 3, below:

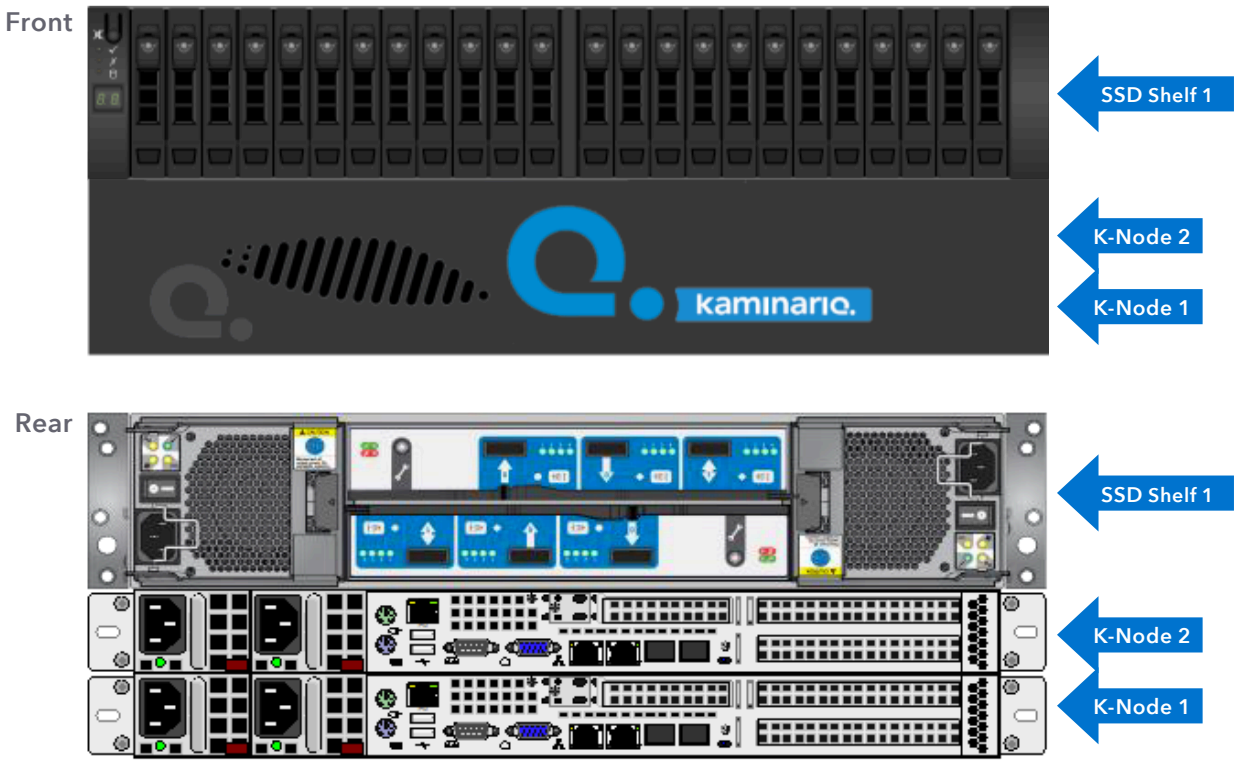


Figure 3: Single K-Block array

- The K-Nodes are connected redundantly using point-to-point IB connectivity.
- Each K-Node is connected redundantly to the SSD shelf using SAS connectivity.
- The K-Nodes have Active/Active connectivity either directly to the host(s) or through a switch, via FC or iSCSI.
- Volumes and metadata are automatically distributed between all SSDs in the array and can be accessed from every K-Node in the array.
- From this basic building block, the customer has the flexibility to scale the K2 All-Flash Array according to the datacenter needs, using a scale-up and/or a scale-out approach.

Scale-Up

Scaling up means adding more capacity by adding Expansion Shelves without adding K-Blocks.

Scaling up a single K-Block with an Expansion Shelf is shown in Figure 4, below:

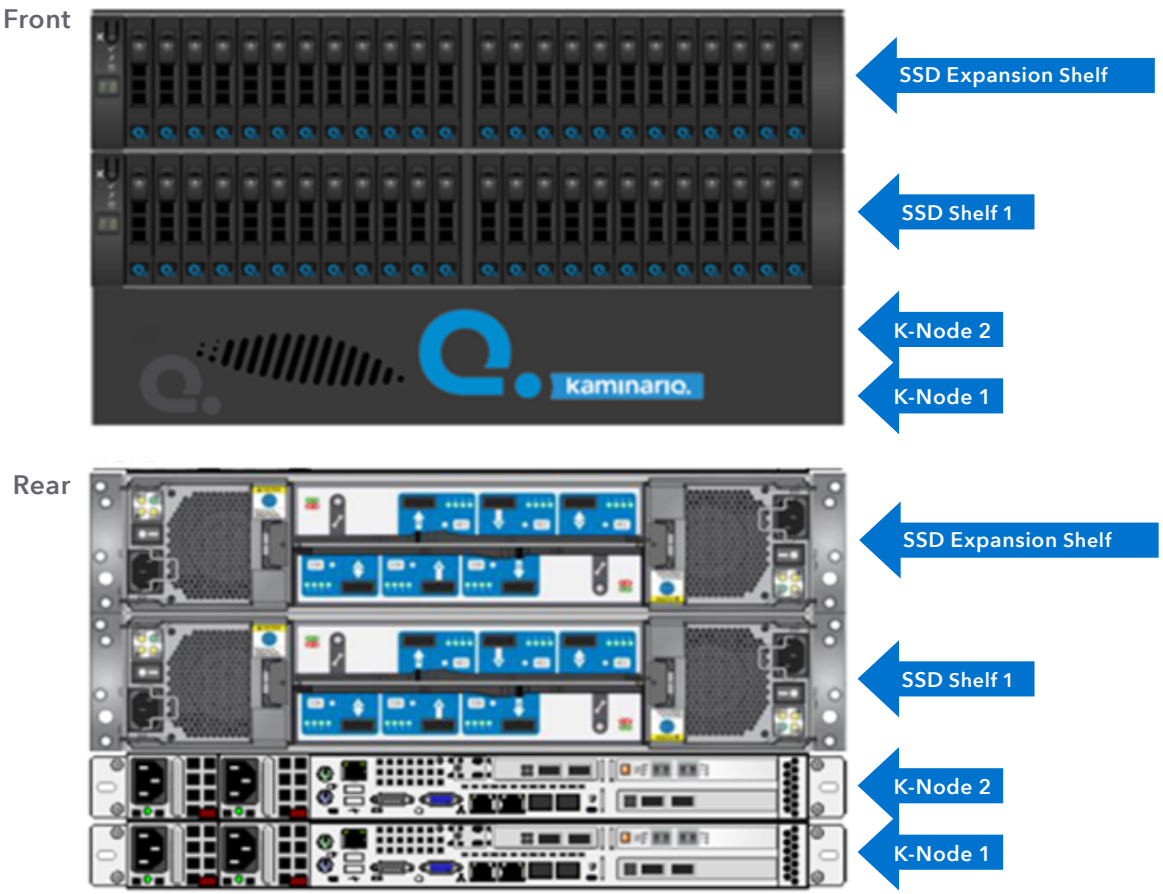


Figure 4: Scaling up a Single K-Block array with an Expansion Shelf

- The expansion increases the capacity density and reduces the cost per GB.
- The expansion shelf is connected redundantly using SAS connectivity.
- The expansion is done online with no downtime or decrease in performance.
- The new configuration has the same performance as before.
- Existing volumes are automatically redistributed between all the SSDs in the array.
- No need to change any host connections or definitions.

SPEAR is optimized for cost and performance as well as for the right balance between performance and capacity. To achieve the best optimization of cost, capacity and performance, current configurations support each K-Block to scale-up with a single expansion shelf.

Scale-Out

Scaling out means to increase the number of K-Blocks, thus adding more capacity and compute power.

Scaling out from a single K-Block to a dual K-Block array is shown in Figure 5:

- The expansion linearly increases the capacity, IOPS and throughput. The latency is kept consistently low and is indifferent to the expansion.
- Two InfiniBand switches are required to support the interconnect between all the K-Nodes in the array. Scaling beyond two K-Blocks will not require any additional networking hardware.
- The expansion is performed online with no downtime or decrease in performance.
- Existing volumes are automatically redistributed between all the SSDs in the array and can be accessed from every K-Node in the array.
- New hosts can be connected to the new K-Block, and the new and existing hosts can access all new and existing volumes.

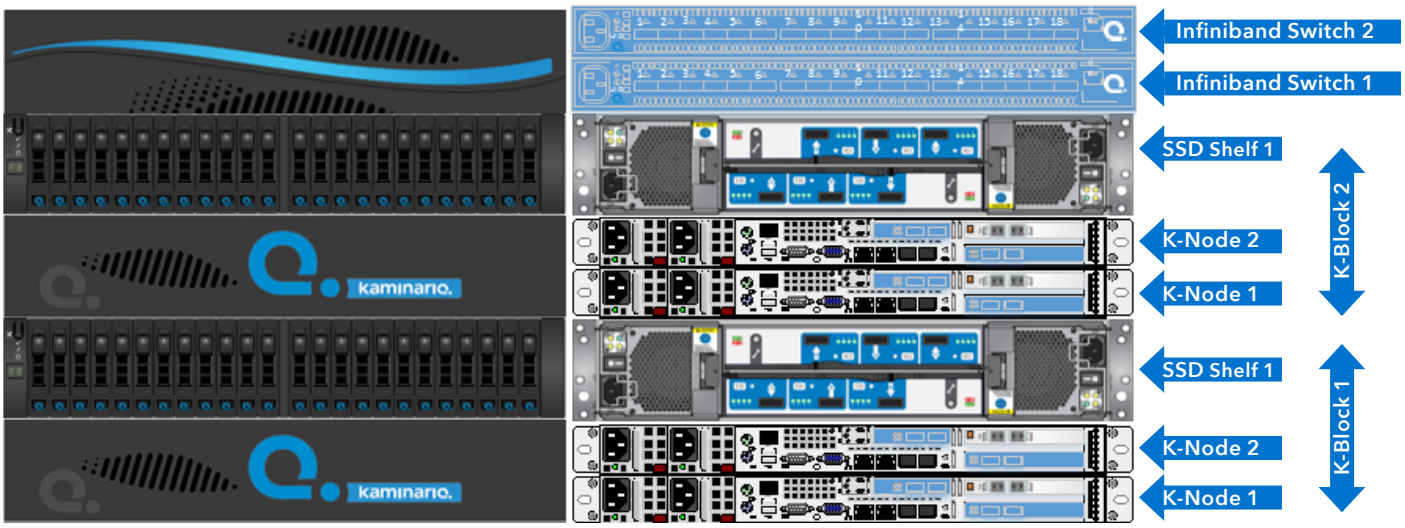


Figure 5: Scaling Out from a Single K-Block array to a Dual K-Block array

For example, scaling-up in a dual K-Block array with two Expansion Shelves is shown in Figure 6, below:

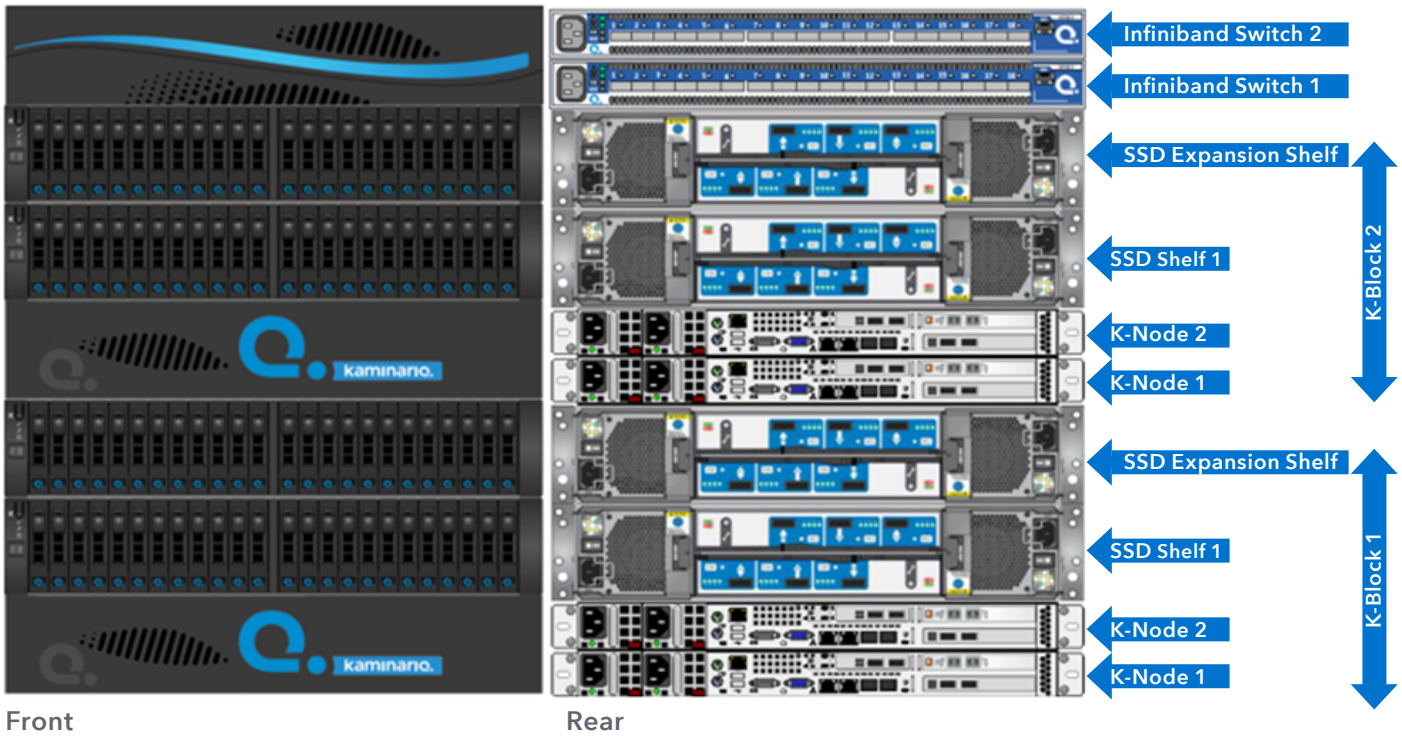


Figure 6: Scaling Up in a Dual K-Block array

Similar to the scale-up of a single K-Block, also in this case:

- The expansion increases the capacity density and reduces the $/GB factor.
- The Expansion Shelves are connected redundantly using SAS connectivity.
- The expansion is performed online with no downtime or decrease in performance .
- Existing volumes are automatically redistributed between all the SSDs in the array.
- No need to change any host connections or definitions.

The system can continue to scale out with K-Blocks and their Expansion Shelves. Continuing the example above, scaling from a dual K-Block array with Expansion Shelves to a quad K-Block array, thus adding two more K-Blocks and two more Expansion Shelves, is shown in Figure 7, below:

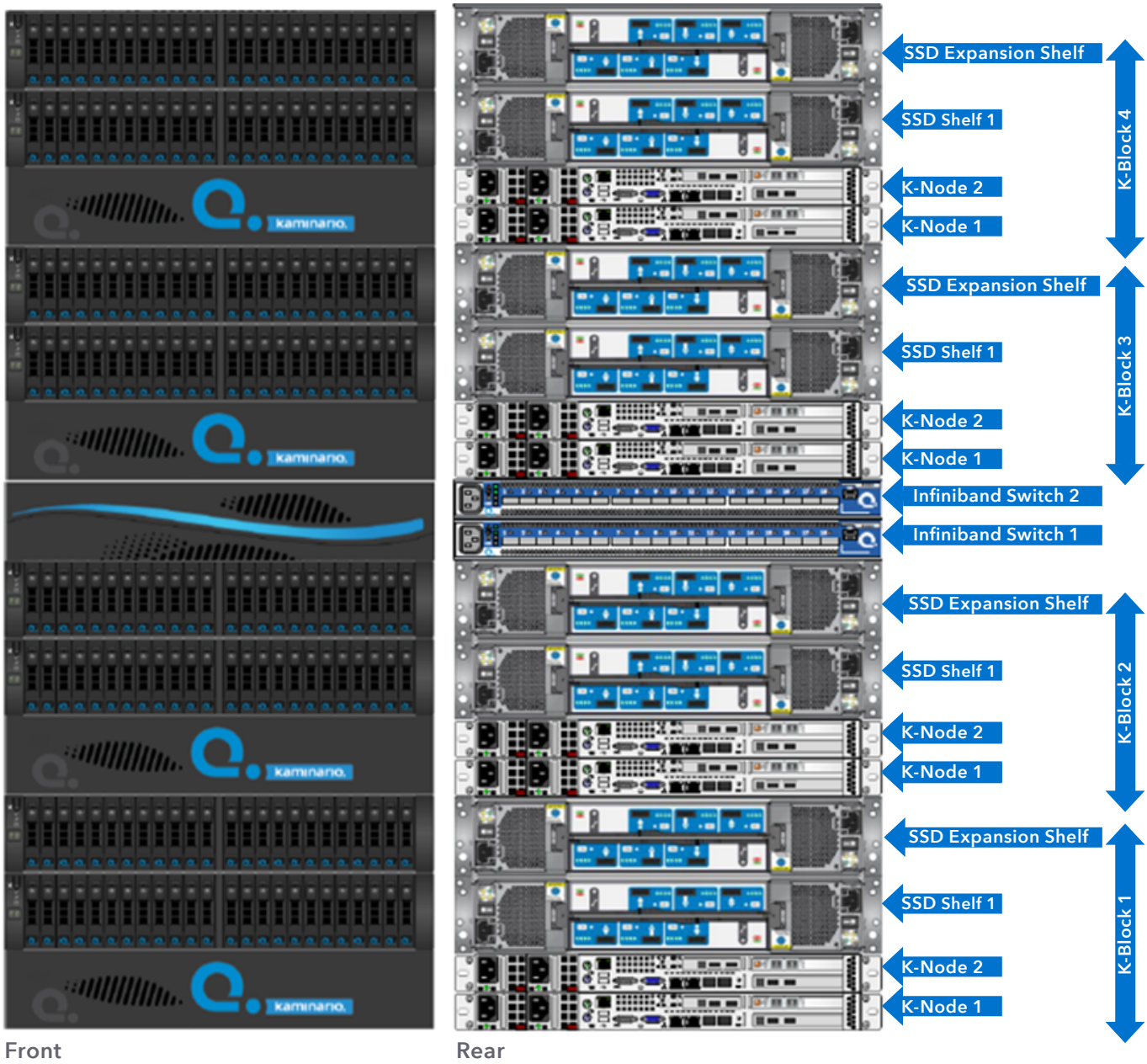


Figure 7: Scaling from a Dual K-Block array with Expansion Shelves to a Quad K-Block array

Differentiators of Kaminario Storage for Cisco UCS

The ability to build a cost-effective AFA relies almost entirely on the efficiency of the architecture or, in other words, how much effective capacity can be generated from raw physical SSD capacity. SPEAR is focused on being highly efficient, but without compromising on other features such as enterprise resiliency and consistent performance. It is therefore a combination of features that allow the efficient usage of flash media and also complement the attributes of flash. These features play a major role in the I/O processing of the UCS system.

Deduplication

Global inline selective deduplication meets the demanding requirements of eliminating redundant data so that it is stored on the array only once. The deduplication is performed globally and its processing is distributed across all the array's K-Nodes, enabling higher deduplication ratios, high performance and consistent low latency. As the array scales out, so does the deduplication. The K2 offers the unique option, amongst AFAs, of selective deduplication. It allows storing data without deduplication for applications whose data redundancy is negligible and additional performance is preferred (such as database applications like Oracle or SQL Server), as well as for security-sensitive applications where deduplication is prohibited.

Compression

The K2 uses inline real-time data compression that is optimized for low latency performance. The data reduction is gained regardless of whether the data is dedupable or not, which makes compression the de-facto method of data reduction for non-dedupable data sets that are common in database environments such as Oracle and SQL Server. SPEAR uses the LZ4 compression algorithm with the ability to compress data in the granularity of bytes. This byte-aligned compression prevents internal fragmentation and facilitates better compression ratios. The compression is performed in a 4KB granularity rather than on bigger data segments, ensuring that small reads do not result in decompression of unnecessary data.

K-RAID™

Kaminario developed K-RAID - a RAID scheme that is highly efficient with an 87.5% utilization rate. This high rate is gained without compromising on either the protection level, which is an optimized version of RAID6 protection, or on performance. This rate is achieved by deploying efficient erasure coding on each 24 SSD shelf. This erasure coding consists of two logical RAID groups, each one with a parity (P1 and P2), and an additional parity for the two groups (Q), as show in Figure 8, below:

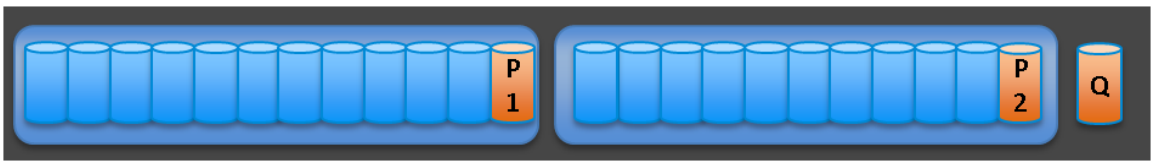


Figure 8: K-RAID logical layout of two RAID groups and a Q parity for both groups

Thin-Provisioning

Thin-provisioning allows maximum utilization of the storage array with the ability to plan storage provisioning for the long term. However, thin-provisioning can be truly utilized only with a scalable architecture that can facilitate capacity growth within the array, with no limitations. All the volumes in the K2 are thinly-provisioned, with a fine granular on-demand growth of 4KB. Un-map operations are also supported in the same granularity. SPEAR delivers the required management tools that bring the thin-provisioning feature to the next level, where the capacity management of volume provisioning is easy and hassle-free.

Enterprise Resiliency

The Kaminario K2 is architected and built to meet the requirements of the most sensitive enterprise applications. The enterprise resiliency starts by deploying only enterprise grade hardware components and continues with High Availability (HA) throughout SPEAR's design, scalability of fault domains and providing the right features for building an enterprise product.

No Single Point of Failure

SPEAR maintains a double-everything approach for K2's hardware components and all data, and metadata at rest is protected by the dual-parity K-RAID™. However, the K2 does not have passive or idle components in the array; all of its resources are being utilized at all times. There is full redundancy of every component in the system and there is not a single component that can fail and cause unplanned down time or data loss. Each K-Block is a standalone failure domain, which means that the entire array can sustain more failures as the array scales and the MTBF of the array stays constant.

Power Loss

A storage array with enterprise capabilities must have the ability to sustain a power outage in the datacenter and still keep the data intact and available for when power returns. Any metadata and/or data that were already acknowledged by the storage controllers (K-Nodes) before being stored to the K-RAID™ are saved in two distinct K-Nodes for redundancy. Each K-Node in the K2 is equipped with an internal battery that will provide sufficient power for the controller to de-stage any inflight data that has not been stored to its K-RAID. These batteries do not occupy any space within the rack since they are internal to the K-Nodes, which are 1U in size. Data that is already stored to the K-RAID is kept persistent and is sustainable through power cycles.

Healthshield™

HealthShield is Kaminario's cloud-based Call Home monitoring, analytics and reporting module that facilitates preventive, proactive and automated enterprise level support for the Kaminario K2 AFA. HealthShield continuously monitors all of the array's components and is able, through a unique decision-making algorithm, to detect in real time any error that might occur. Any error or change in the system state triggers an automatic real-time event that is sent to Kaminario Cloud Support. It is also possible for the storage admin to subscribe to the events of interest by category such as LUN configuration, FC/iSCSI connectivity, hardware errors and so on.

On an hourly basis, HealthShield collects system-wide information that is sent to the Kaminario service center data warehouse, where it is analyzed and processed to recognize issues of performance, configuration, capacity management and more. This information is reported back to Kaminario Support, which uses a sophisticated BI system to provide tailored support for each customer.

Non-Disruptive Upgrades (NDU)

The K2 can upgrade any of its hardware and software components with no impact on the availability or performance of the array. SPEAR is software-defined, so any new feature/enhancement/bug-fix/firmware can be deployed with no dependencies on maintenance windows or running workloads. In addition, hardware can be replaced/upgraded/added in the same manner. Having NDU combined with a scalable architecture, the K2 provides the best TCO of storage: No fork-lift upgrades, no need to plan down-time, new technologies can be deployed in the existing array and growing datacenter needs can be met by adding more capacity and/or performance. All of these operations are performed non-disruptively and automatically, with no human intervention.

Snapshots

K2's patented snapshot architecture follows SPEAR's guidelines of storage efficiency, performance and scalability. Snapshots are created instantly, with no performance impact and they do not take up any capacity. They track only the deltas from the volume in a 4KB granularity, using a redirect-on-write (RoW) approach. This storage-efficient design also keeps the impact on SSD endurance to a minimum. The 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 of the production volumes, without any impact on the production volumes.

The snapshots are created instantly, with no dependencies on the number or size of the volumes being snapped or how big the array is. Using the snapshot's restore functionality for recovery purposes is done without losing any of the snapshot history and is allowed at any time. The snapshots can be accessed from any of the storage controllers of the K2, without bottlenecks or load balancing of affinity to a specific controller.

Replication

Replication provides site resiliency that completes the enterprise datacenter resiliency requirements. K2 leverages its snapshot architecture to facilitate asynchronous snapshot-based replication between K2 AFAs. Since the replication is based on the snapshot architecture, there is no impact on the production's consistent high performance, and deltas of the replicated copies are captured with no dependencies on the link speed. SPEAR's deduplication and compression are also used to significantly reduce the amount of data sent between arrays.

All essential disaster recovery capabilities such as short RPO and RTO are gained natively without any third-party software components, which are costly and take their toll on the data flow.

Assured Capacity

Kaminario is the only all-flash storage vendor that offers Assured Capacity. Guesswork and anxiety are removed from storage sizing to limit capital expense, assure simplicity, streamline deployments and deliver predictable operation. Kaminario's Assured Capacity removes uncertainty from capital expenditures and storage capacity planning. It means cost-efficient assurance to expand a customer's system at Kaminario's expense to achieve promised effective capacity.* If data reduction on a system works better than predicted, customers will enjoy the benefits of additional data storage at no additional cost. K2's ultra-efficient storage architecture and superior data reduction technologies enable Assured Capacity. The system delivers high capacity with no compromise to enterprise resiliency and high performance. The Kaminario K2 storage array requires no special capacity management or tuning. Patented K-RAID™ achieves high capacity utilization of 87.5% automatically. Combined with thin provisioning, included deduplication and compression technologies maximize K2 data reduction. As a result, Kaminario drives down all-flash storage costs and drives up customer value.

	1 K-Block	2 K-Blocks	4 K-Blocks
Footprint K-Blocks Footprint +Exp.	4U 6U	10U 14U	18U 26U
Usable Capacity*	7TB-360+TB	14TB-720+TB	28TB-1.44+PB
Media (encrypted)	480GB 3D MLC SSD 960GB 3D MLC SSD 1.92TB 3D TLC SSD		
Density TB/U	Up to 60TB		
IOPS	Up to 250K	Up to 500K	Up to 1M
Bandwidth	Up to 3.2GB/s	Up to 6.4GB/s	Up to 12.8GB/s
Latency	0.35mSec		
Power (Max) Power (Typical)	1.1KW-1.4KW 0.8KW-1KW	2.5KW-3.1KW 1.7KW-2.2KW	4.8KW-6.1KW 3.3KW-4.3KW
Cooling Max Cooling Typical	3.9K-5K BTU/hr 2.7K-3.5K BTU/hr	8.6K-10.8K BTU/hr 6K-7.6K BTU/hr	16.4K-20.8K BTU/hr 11.4K-14.6K BTU/hr
Snapshots	Up to 8000 Snapshots		
Management	CLI (SSH), HTTP/HTTPS GUI, Scripting (SSH), RESTful API, VAAI, SNMP, Microsoft VSS, VMware vCenter Plug-in, VMware SRM		
Data Reduction	Global Selective Inline Deduplication, Inline Compression, Thin provisioning		
Connectivity FC/iSCSI	FC: 4x 8Gbps iSCSI: 4x 10GbE	FC: 8x 8Gbps iSCSI: 8x 10GbE	FC: 16x 8Gbps iSCSI: 16x 10GbE
Management ports	2x 1GbE		

*Capacity is subject to drive size and the application data reduction ratio. For some datasets such as VDI the range will be higher.

Joint Solution Overview and Benefits

Hardware and Software Used for Different Solutions

The following section suggests three different configurations that are commonly used by Kaminario/Cisco customers. The configurations are suitable for large/medium/small deployments. Different customers will utilize different configurations based on different performance and workload requirements. The Cisco UCS architecture combined with the Kaminario K2 array allow for an extremely diverse suite of configurations around compute power, performance, scale-up and scale-out abilities as well as extreme cost effectiveness.

The tables below display the suggested starting points of the hardware for proposed configurations.

Intro environments: Cisco Mini

Hardware	Quantity
UCS components	
Cisco UCS 6324UP Fabric Interconnect with 20-port 8GB FC	2 (configured as an active-active pair)
Cisco UCS 5108 B-Series blade chassis (6 RU)	1
Cisco UCS 2208 B-Series Blade Fabric Extender modules	2
Cisco UCS B200 M4 blade server: • 1 or 2-socket 4 - 44 core E5-2620 v4 2.4GHz CPU • 16GB – 1.5TB DDR4 memory	2
Network	
Cisco Nexus 5548UP switches	2
Storage	
KBlock Kaminario K2 V5	1 7TB – 360+TB Usable capacity*
Number of K-Blocks	1 KB
Drives – 24 per 2U shelf	24,48,72 x 480GB, 960GB, 1.92TB SSD
Physical rack space	4U – 8U
Software	**Version**
VMWare ESXi	6
Redhat Enterprise Linux 7.1 64-bit	2.6.32-279.el6.x86_64
Cisco UCS Manager	3.1(1g)
Cisco NXOS for Nexus 5548UP	5.0(3)N2(1)
Kaminario K2 SPEAR™	V5
Expected Performance (4K 80/20 read/write)	
Latency (milliseconds)	0.44
IOPS	203K
Throughput (MB/Sec)	3,100

Small environments:

Hardware	Quantity
UCS components	
Cisco UCS 6248UP Fabric Interconnect with 48-port 8GB FC	2 (configured as an active-active pair)
Cisco UCS 5108 B-Series blade chassis (6 RU)	1
Cisco UCS 2208 B-Series Blade Fabric Extender modules	2
Cisco UCS B200 M4 blade server: • 1 or 2-socket 4 - 44 core E5-2620 v4 2.4GHz CPU • 16GB – 1.5TB DDR4 memory	2
Network	
Cisco Nexus 5548UP switches	2
Storage	
KBlock Kaminario K2 V5	1 7TB – 360+TB Usable capacity*
Number of K-Blocks	1 KB
Drives – 24 per 2U shelf	24,48,72 x 480GB, 960GB, 1.92TB SSD
Physical rack space	4U – 8U
Software	**Version**
VMWare ESXi	6
Redhat Enterprise Linux 7.1 64-bit	2.6.32-279.el6.x86_64
Cisco UCS Manager	3.1(1g)
Cisco NXOS for Nexus 5548UP	5.0(3)N2(1)
Kaminario K2 SPEAR™	V5
Expected Performance (4K 80/20 read/write)	
Latency (milliseconds)	0.44
IOPS	203K
Throughput (MB/Sec)	3,100

Medium environments:

Hardware	Quantity
UCS components	
Cisco UCS 6248UP Fabric Interconnect with 48-port 8GB FC	2 (configured as an active-active pair)
Cisco UCS 5108 B-Series blade chassis (6 RU)	1
Cisco UCS 2208 B-Series Blade Fabric Extender modules	2
Cisco UCS B200 M4 blade server: • 1 or 2-socket 4 - 44 core E5-2620 v4 2.4GHz CPU • 16GB – 1.5TB DDR4 memory or Cisco UCS B460 M4 blade server: • 1 - 4-socket 4-96 core E7 v4 2.4 GHz CPU • 16GB – 6TB DDR4 memory	2 – 4 or 2
Network	
Cisco Nexus 5548UP switches	2
Storage	
KBlock Kaminario K2 V5	2 14TB – 720+TB Usable capacity*
Number of K-Blocks	2 KB
Drives – 24 per 2U shelf	48,96,144 x 480GB, 960GB, 1.92TB SSD
Physical rack space	10U – 18U
Software	**Version**
VMWare ESXi	6
Redhat Enterprise Linux 7.1 64-bit	2.6.32-279.el6.x86_64
Cisco UCS Manager	3.1(1g)
Cisco NXOS for Nexus 5548UP	5.0(3)N2(1)
Kaminario K2 SPEAR™	V5
Expected Performance (4K 80/20 read/write)	
Latency (milliseconds)	0.5
IOPS	379K
Throughput (MB/Sec)	6,600

Large environments:

Hardware	Quantity
UCS components	
Cisco UCS 6248UP Fabric Interconnect with 48-port 8GB FC	2 (configured as an active-active pair)
Cisco UCS 5108 B-Series blade chassis (6 RU)	1
Cisco UCS 2208 B-Series Blade Fabric Extender modules	2
Cisco UCS B200 M4 blade server: • 1 or 2-socket 4 - 44 core E5-2620 v4 2.4GHz CPU • 16GB – 1.5TB DDR4 memory or Cisco UCS B460 M4 blade server: • 1 - 4-socket 4-96 core E7 v4 2.4 GHz CPU • 16GB – 6TB DDR4 memory	4 – 8 or 2 – 4
Network	
Cisco Nexus 5548UP switches	2
Storage	
KBlock Kaminario K2 V5	4 28TB – 1.44+PB Usable capacity*
Number of K-Blocks	4 KB
Drives – 24 per 2U shelf	96,192,288 x 480GB, 960GB, 1.92TB SSD
Physical rack space	18U – 34U
Software	**Version**
VMWare ESXi	6
Redhat Enterprise Linux 7.1 64-bit	2.6.32-279.el6.x86_64
Cisco UCS Manager	3.1(1g)
Cisco NXOS for Nexus 5548UP	5.0(3)N2(1)
Kaminario K2 SPEAR™	V5
Expected Performance (4K 80/20 read/write)	
Latency (milliseconds)	0.48
IOPS	725K
Throughput (MB/Sec)	11,800

Cisco UCS Manager Configuration Overview

The following are the high-level steps involved for a Cisco UCS configuration:

1. Configure Fabric Interconnects for Chassis and Blade Discovery:
   - Configure Server Ports
2. Configure LAN and SAN on UCS Manager:
   - **a.** Configure VLAN
   - **b.** Configure VSAN
   - **c.** Configure and enable Ethernet LAN uplink Ports
   - **d.** Configure Port Channel
   - **e.** Configure and Enable FC SAN uplink Ports
3. Creating and Configuring UUID, MAC, WWWN and WWPN Pool:
   - **a.** Create UUID Pool
   - **b.** IP Pool and MAC Pool
   - **c.** WWNN Pool and WWPN Pool
4. Configuring vNIC and vHBA Template:
   - **a.** Create vNIC templates
   - **b.** Create Public vNIC template
   - **c.** Create Private vNIC template
   - **d.** Create Storage vNIC template
   - **e.** Create HBA templates
5. Create Service Profile Template.
   Details for each step are discussed in subsequent sections below.

Configuring Fabric Interconnects for Blade Discovery
Cisco UCS 6248UP Fabric Interconnects are configured for redundancy. They provide resiliency in case of failures. The first step is to establish connectivity between the blades and the fabric interconnects.

To configure the server ports in Cisco UCS Manager:

1. Click **Equipment.**
2. Click **Fabric Interconnects**.
3. Click **Fabric Interconnect A**.
4. Click **Fixed Module**.

5. Click **Ethernet Ports** and select the desired number of ports.
6. Right-click **Configure as Server Port**. The ports are configured as the Server Port display:

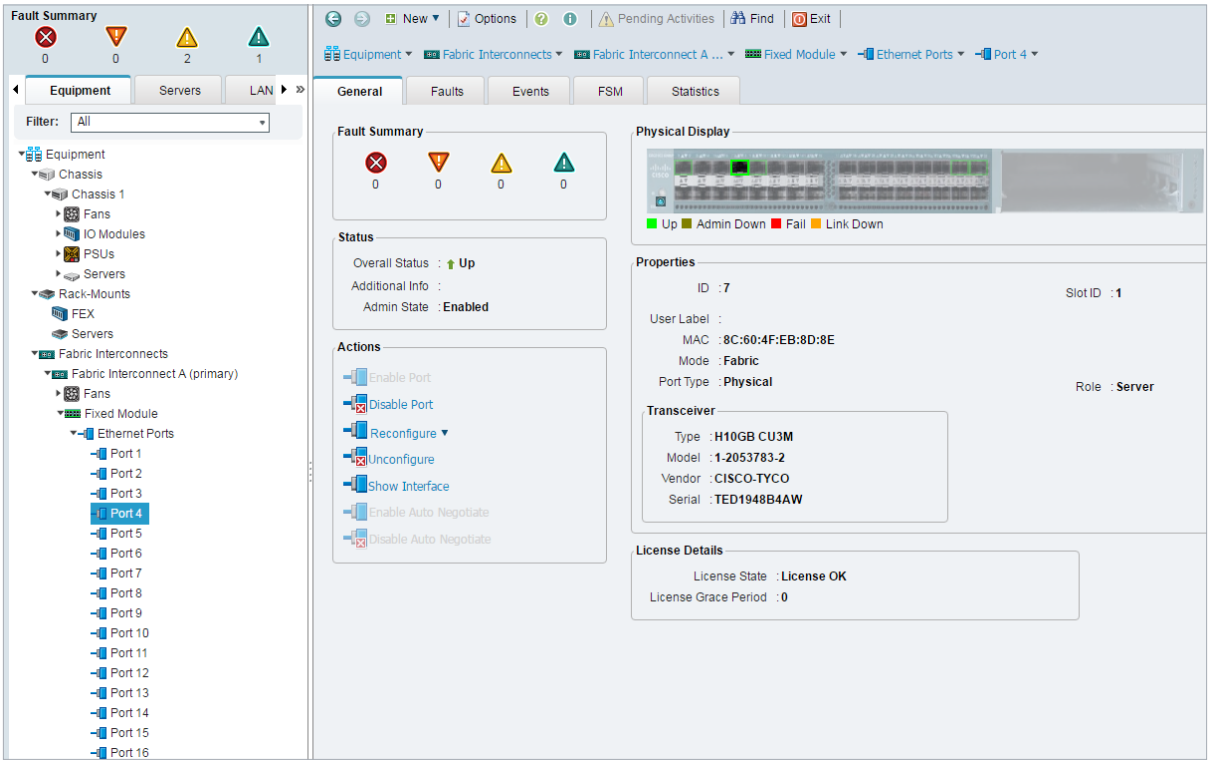


Figure 12: Configuring Ports

In the below screenshot, ports 7 and 9 are configured as Server ports on Fabric Interconnect A.

7. Repeat the same steps to configure ports 7 and 9 as Server ports on Fabric Interconnect B

Configure LAN and SAN on UCS Manager

Perform the LAN and SAN configuration steps in the Cisco UCS:

1. In Cisco UCS manager, click **LAN**
2. Click **LAN Cloud > VLAN**
3. Right-click **Create VLANs**

Configuring VSAN

To configure the VSAN in Cisco UCS manager:

Click SAN > SAN Cloud > Fabric A > VSANs.
Right-click Create VSAN to create VSANKAM101 on Fabric A.

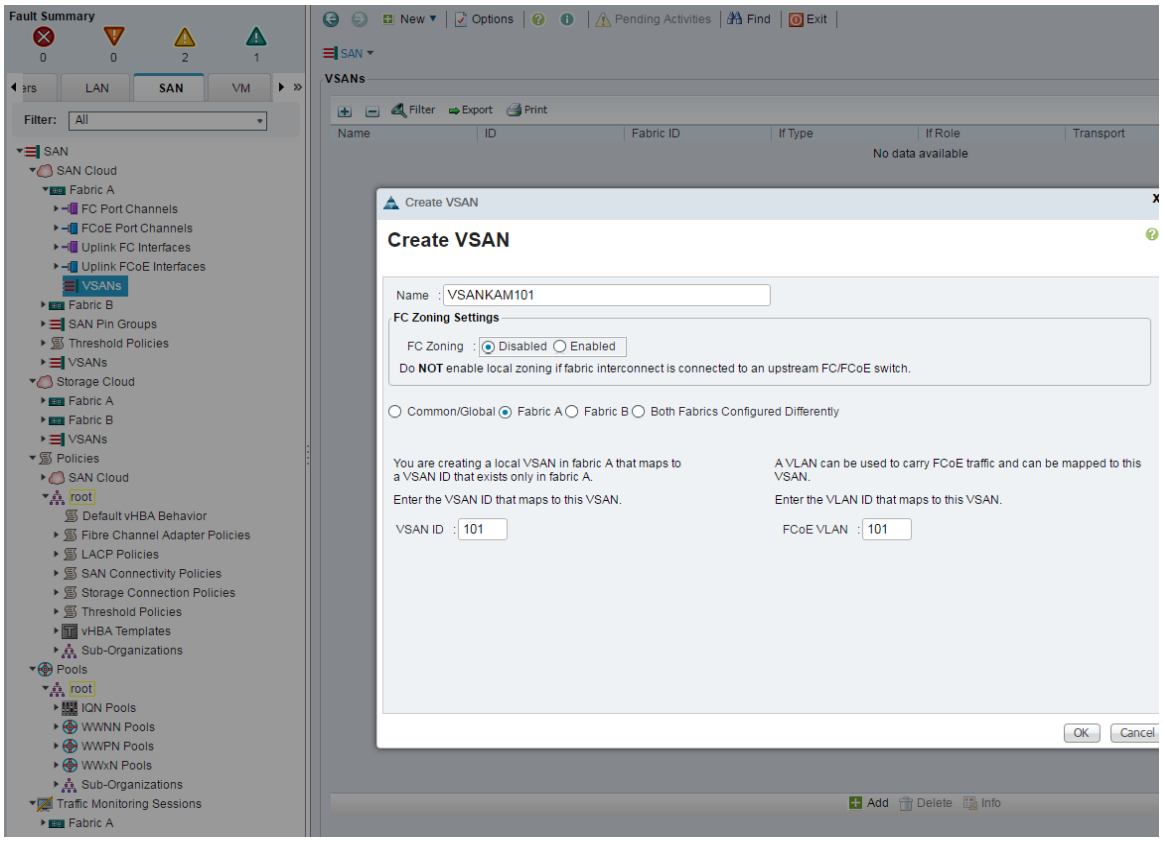


Figure 13: Configuring VSAN in Cisco UCS Manager

Note: Even if you are not currently using FCoE traffic for SAN Storage, it is mandatory to specify VLAN ID.

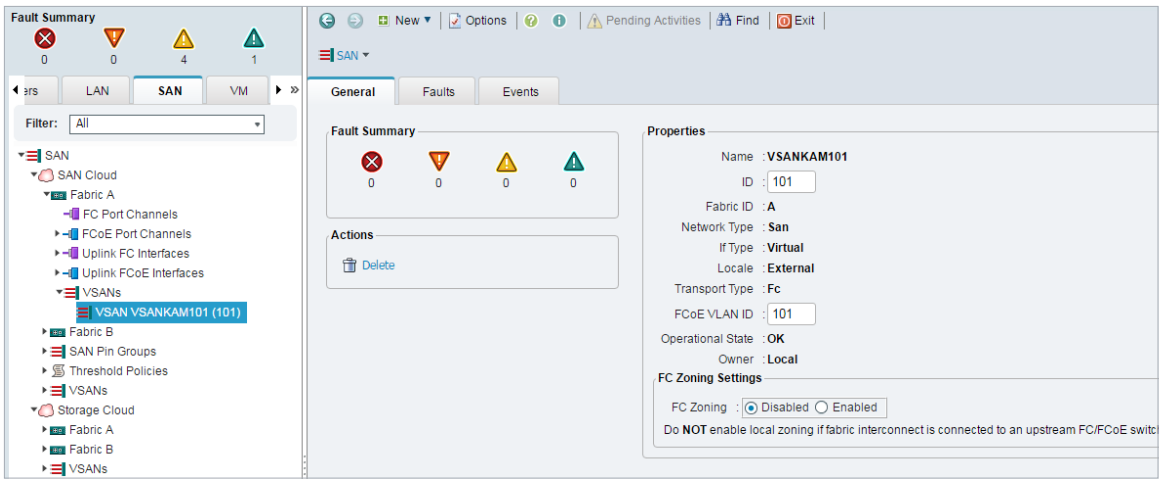


Figure 14: VSAN Summary

Repeat the steps to create VSAN 102 on Fabric B.

Configuring and Enable Ethernet LAN uplink Ports

To configure and enable Ethernet LAN uplink Ports from the Equipment tab:

1. Click Fabric Interconnects > Fabric Interconnect A > Fixed Module > Ethernet Ports.
2. Select the desired number of ports.
3. Right-click Configure as Uplink Ports.
4. In the test setup, ports 5 and 10 were configured as Network uplinks.

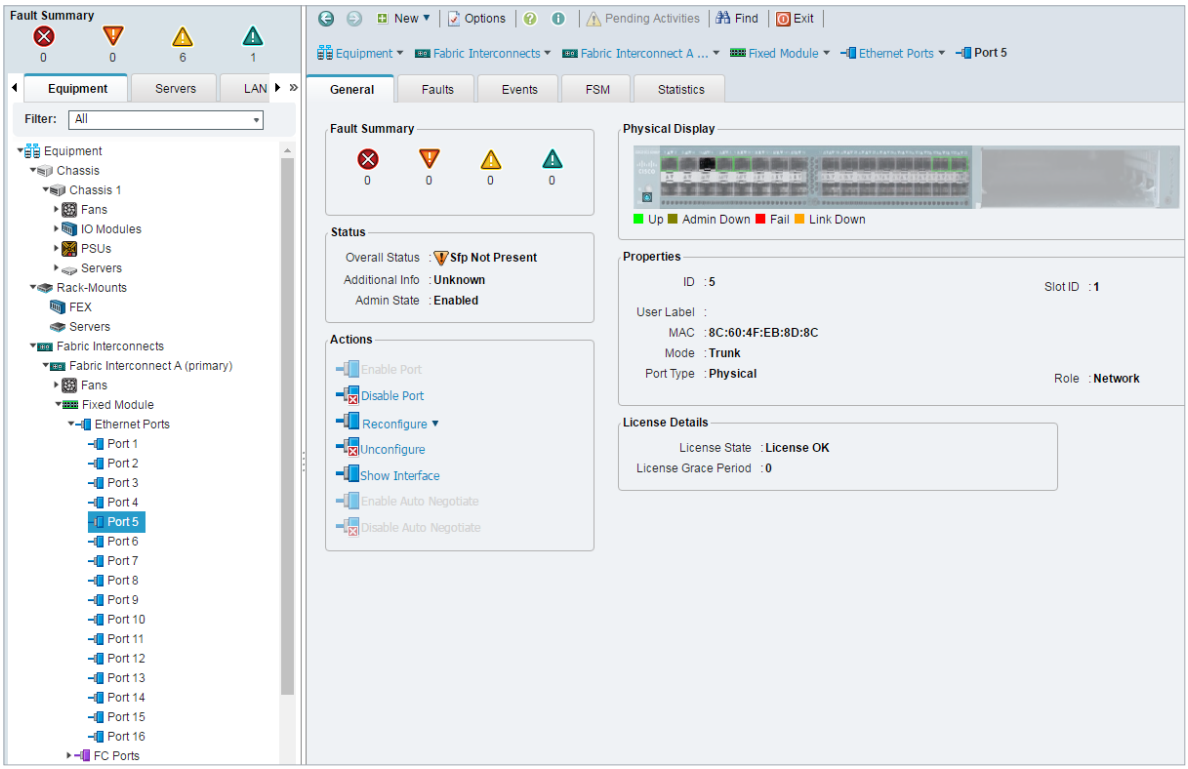


Figure 15: Network Uplink Details

Repeat the same step on Fabric interconnect B to configure ports 5 and 10 as Ethernet uplink ports

Configuring Port Channel

To configure a configure the port channel in the Cisco UCS manager:

1. Click LAN > LAN Cloud > Fabric A > Port Channels.
2. Right-click Create Port Channel.
3. In the test setup, ports 5 and 10 on Fabric A were configured as port channel 10. Similarly, ports 5 and 10 on Fabric B are configured as port channel 11.

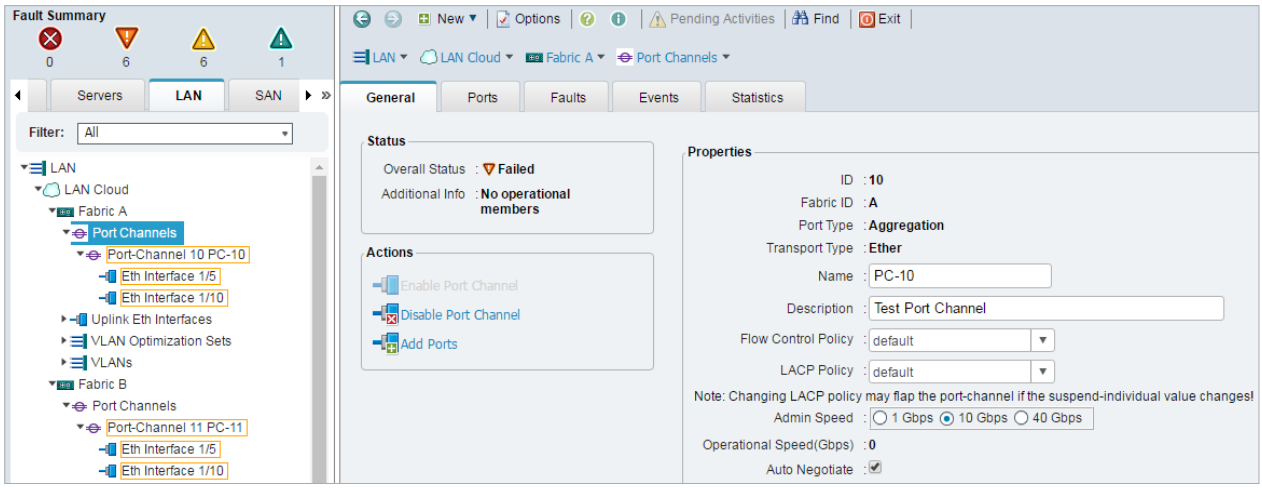


Figure 16: Port Channel 10 Details

Configuring Pools

When VLANs and VSAN are created, you need to configure pools for UUID, MAC Addresses, Management IP and WWN.

Creating UUID Pools

To configure the pools in the Cisco UCS Manager:

1. Click Servers > Pools > UUID Suffix Pools.
2. Right-click Create UUID Suffix Pool to create a new pool.

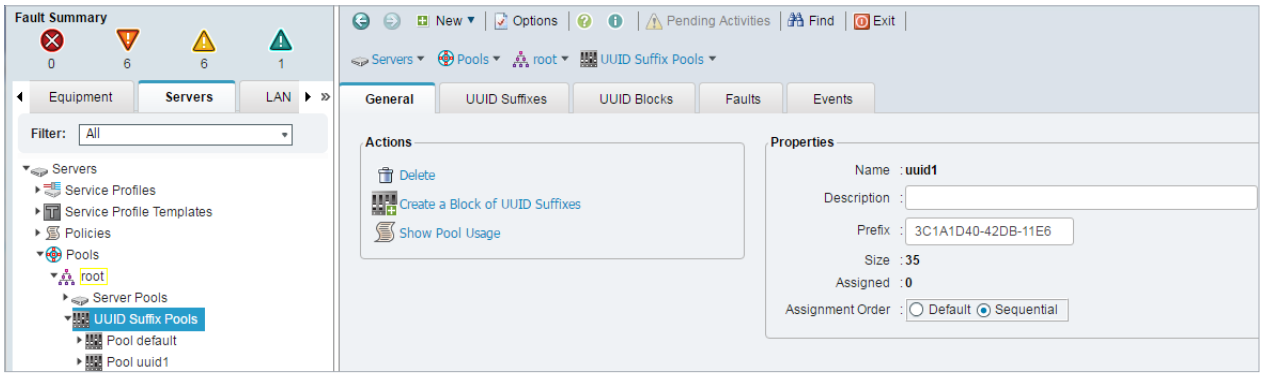


Figure 17: Creating UUID Pools

Creating IP and MAC Pools

To create an IP and MAC pool in the Cisco UCS Manager:

1. Click LAN > Pools > IP Pools.
2. Right-click Create IP Pool Ext-mgmt

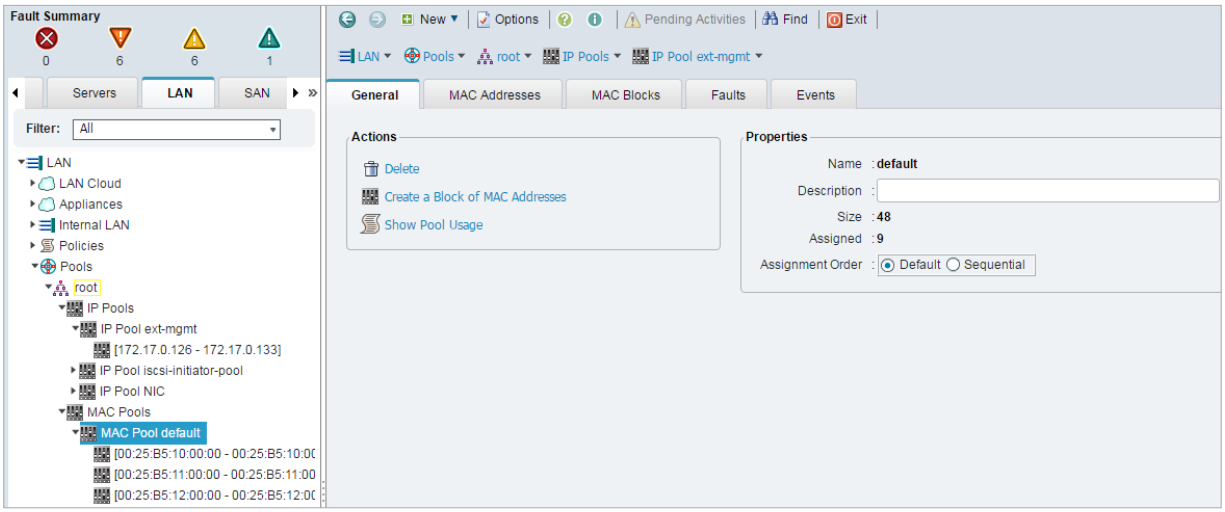


Figure 18: Creating IP Pool and MAC Pool

The IP pools are used for console management, while MAC addresses for the vNICs are used later.

Creating WWNN and WWPN Pool

To configure the pools in the Cisco UCS Manager:

1. Click SAN > Pools > WWNN Pools.
2. Right-click Create WWNN Pools.
3. Click WWPN Pools to create WWPN pools

These WWNN and WWPN entries are used for the virtual FC HBAs to access the database on the Kaminario storage.

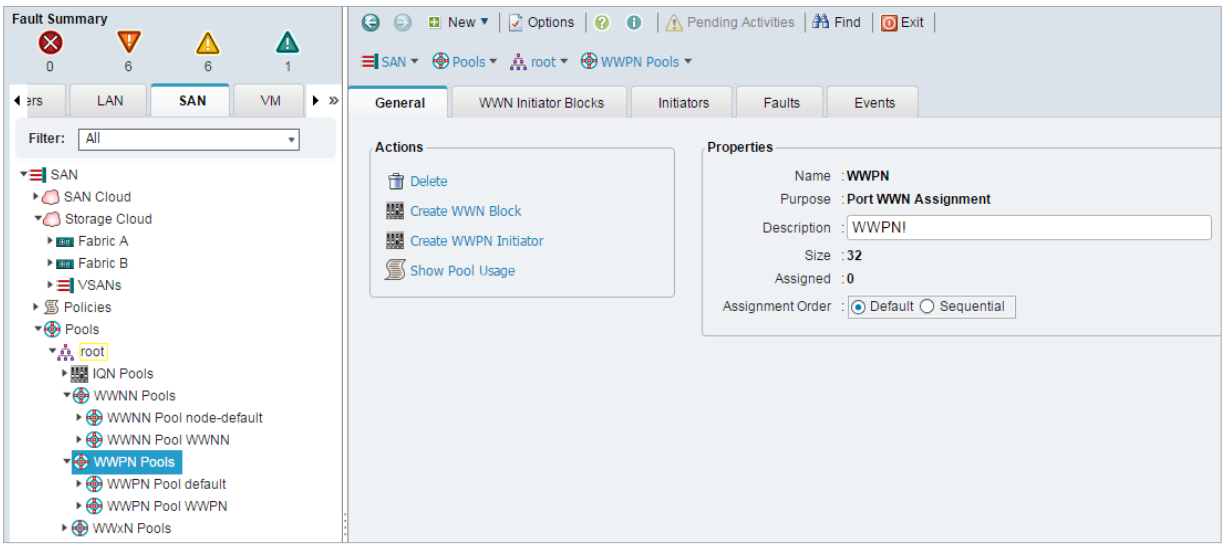


Figure 19: Create WWNN and WWPN Pool

Creating a vNIC Template

To create a vNIC template in the Cisco UCS Manager:

1. Click **LAN > Policies > vNIC** templates.
2. Right-click **Create vNIC Template**.

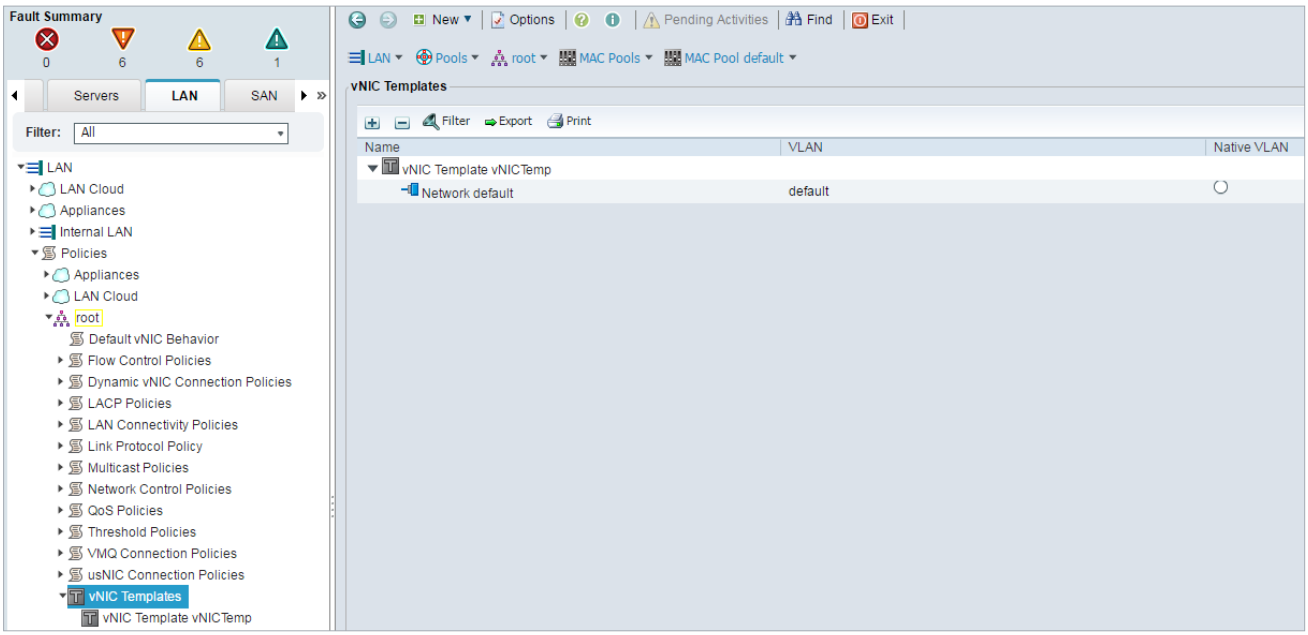


Figure 20: Creating Private vNIC Template

Creating a vHBA Template

To create HBA templates in the Cisco UCS Manager:

1. Click SAN > Policies > vHBA templates.
2. Right-click Create vHBA Template.

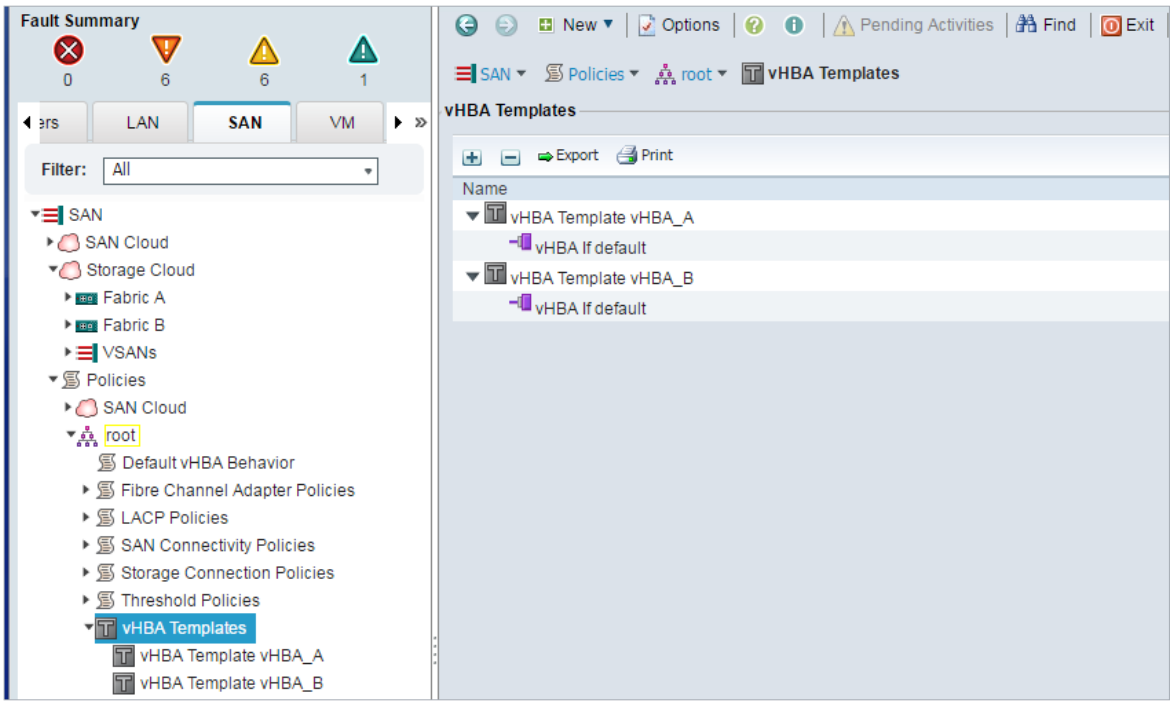


Figure 21: HBA Template Summary

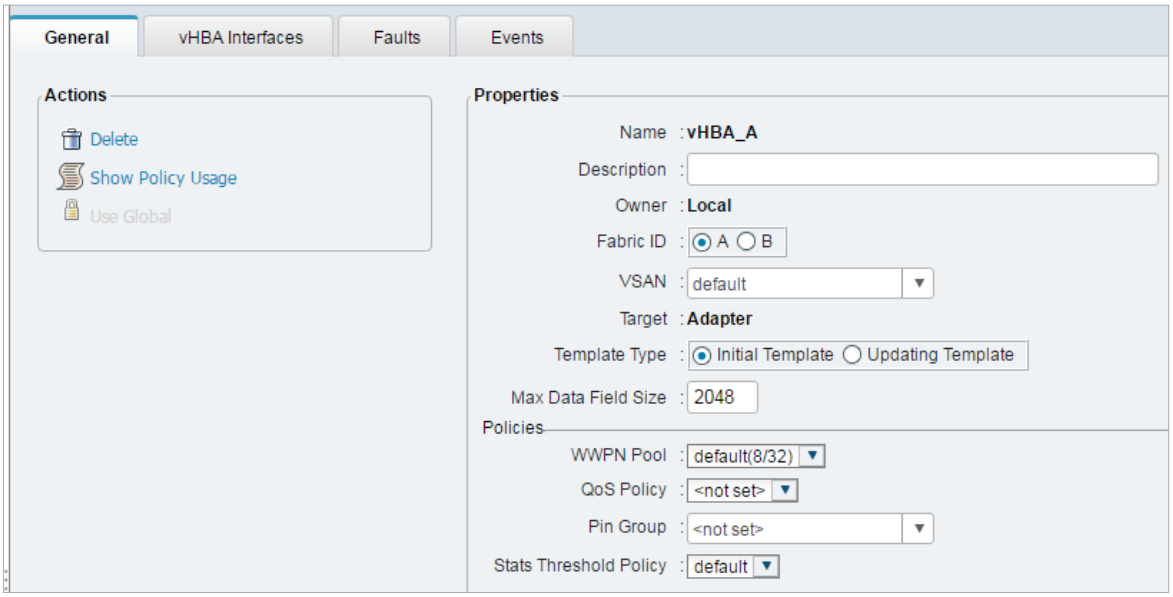


Figure 22: vHBA_A Template Properties

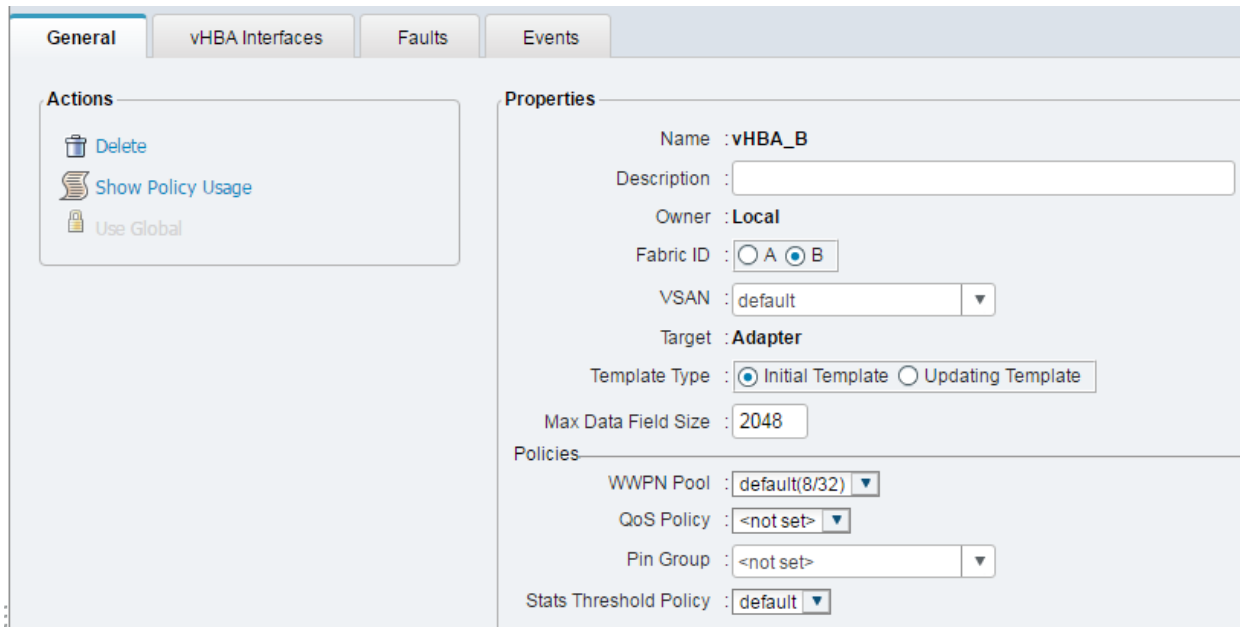


Figure 23: vHBA-B Template Properties

Service Profile Creation and Association to Cisco UCS Blades

Service profile templates enable policy-based server management that helps ensure consistent server resource provisioning suitable to meet predefined workload needs.

Creating Service Profile Templates

To create a Service Profile template in the Cisco UCS Manager:

1. Click Servers > Service Profile Templates > root.
2. Right-click root and select Create Service Profile Template.

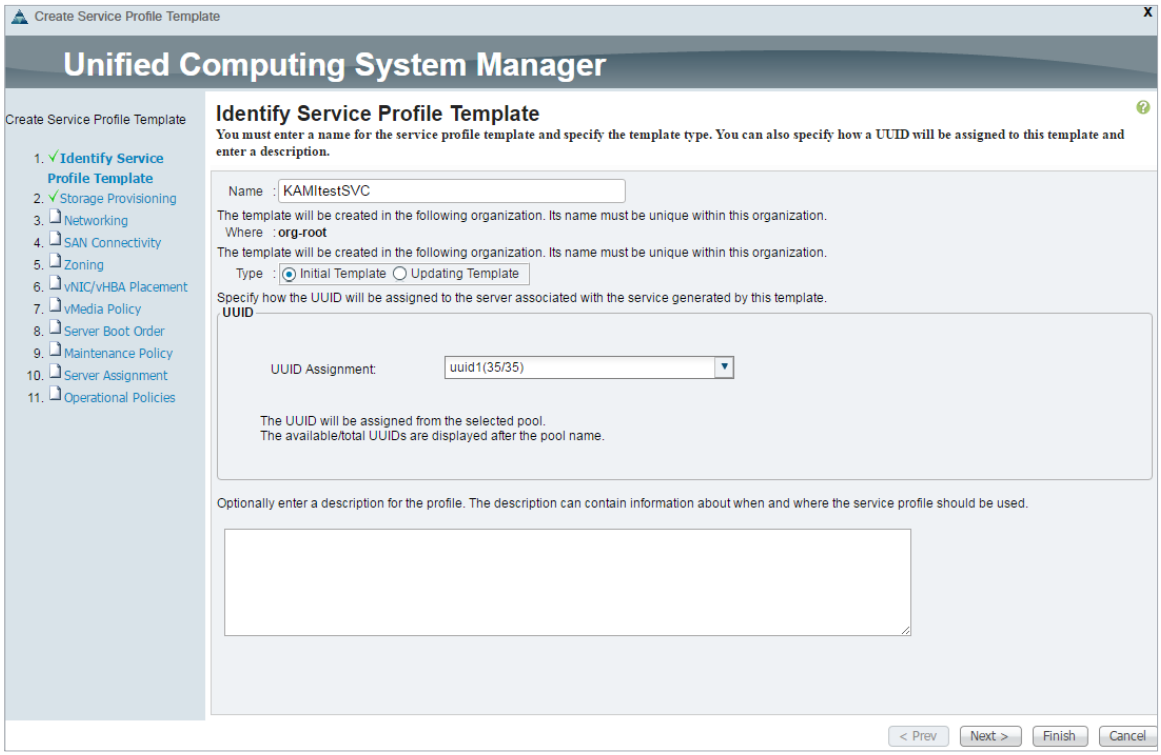


Figure 24: Creating Service Profile Template

3. Enter the template name and select the UUID Pool that was created earlier.
4. In the Networking screen, click **Expert** to enter Expert mode

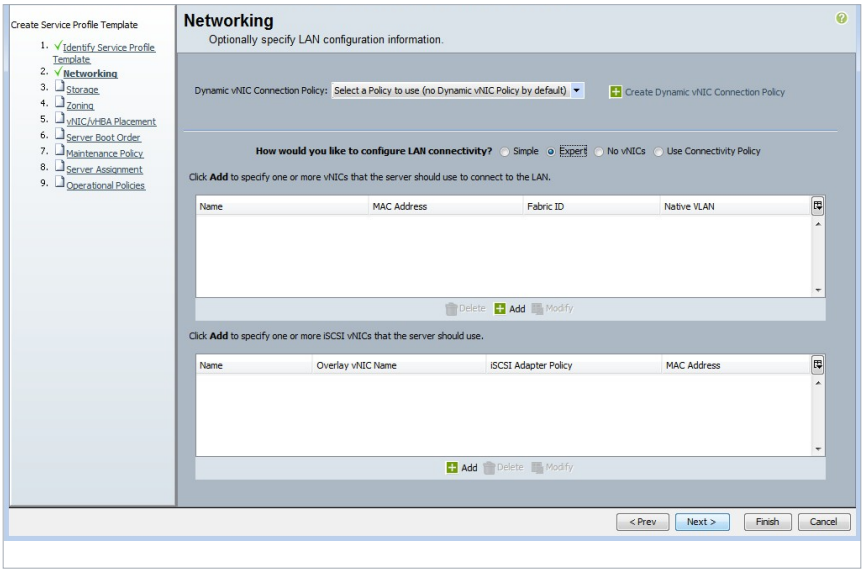


Figure 25: Expert Mode

5. Click **Add** to add a vNIC for the public network.

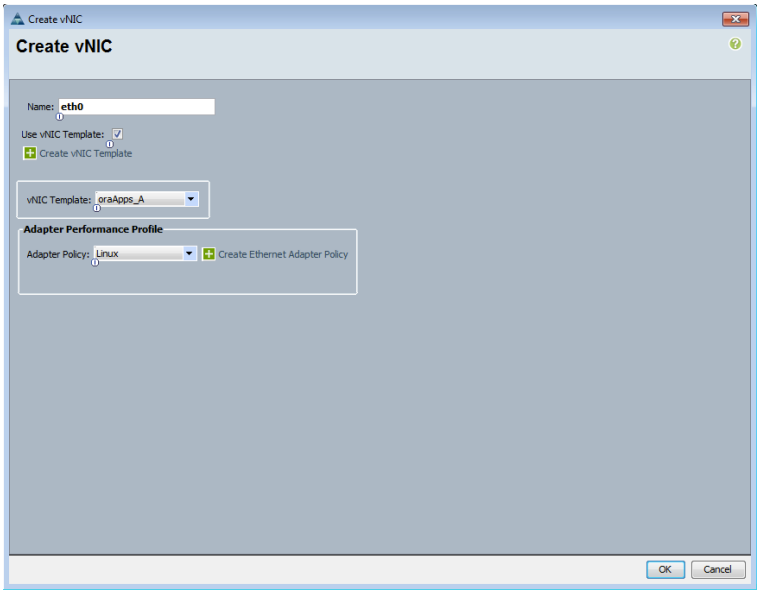


Figure 26: Creating a vNIC for a Public Network

6. Click **Next** to go to the Storage configuration screen.
7. In the Storage screen, click **Expert** to enter expert mode.
8. Click **Add** to add a vHBA for Fabric A.

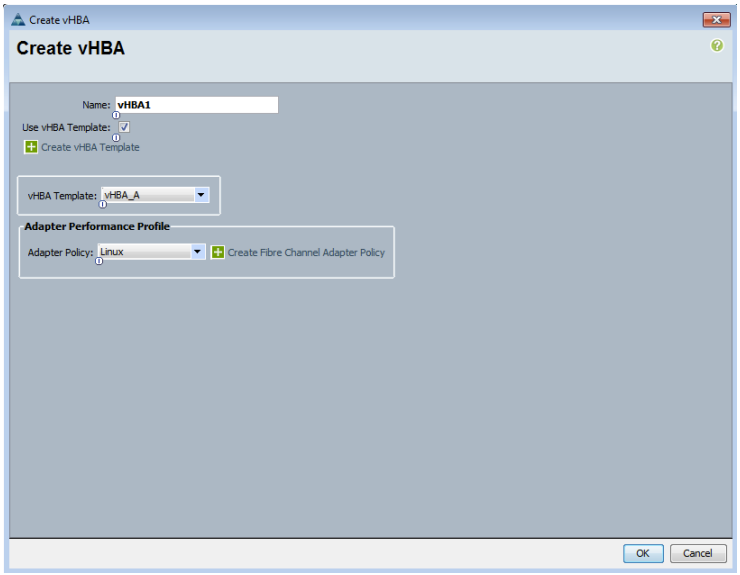


Figure 27: Creating vHBA

9. Click **Add** to add a vHBA for Fabric B

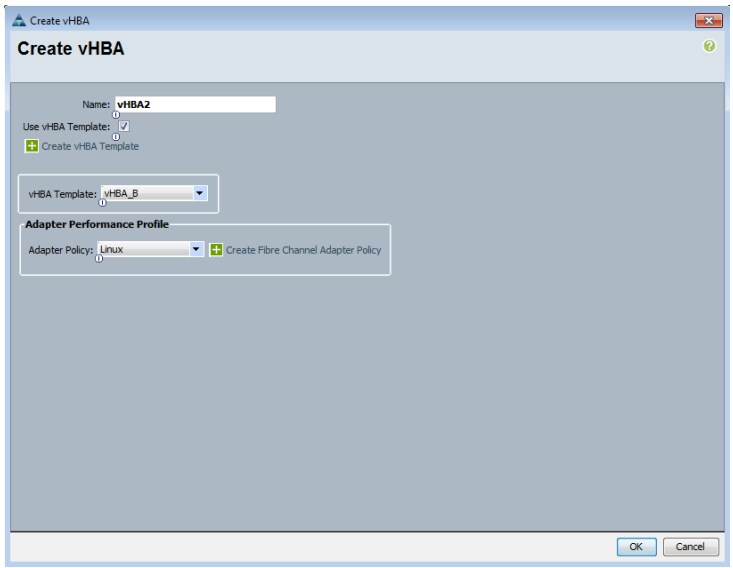


Figure 28: Creating vHBA

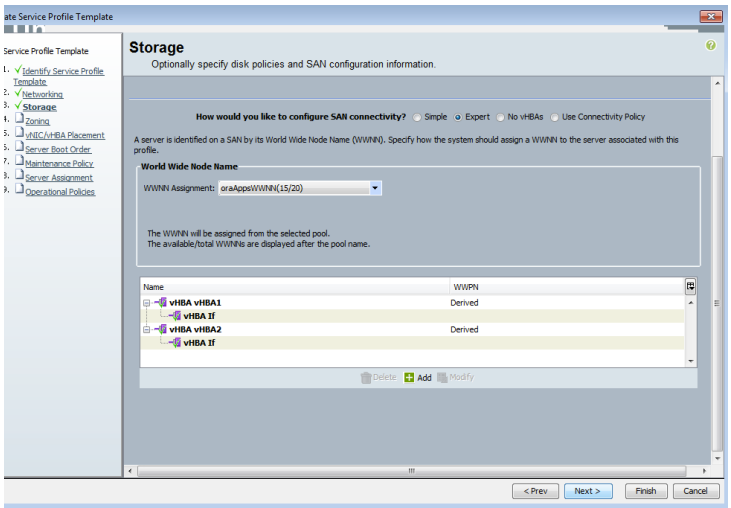


Figure 29: Storage vHBA summary

Note: Make sure to set the WWNN assignment to the WWNN pool created previously.

10. Click **Next** to proceed to the next screen.
11. Leave the configuration to default in the rest of the screen, and click Finish to complete the service profile template creation.

Creating Service Profiles from Service Profile Templates

To create a service profile from a service profile template:

1. Click Servers > Service Profile Templates.
2. Right-click Create Service Profiles from Template.

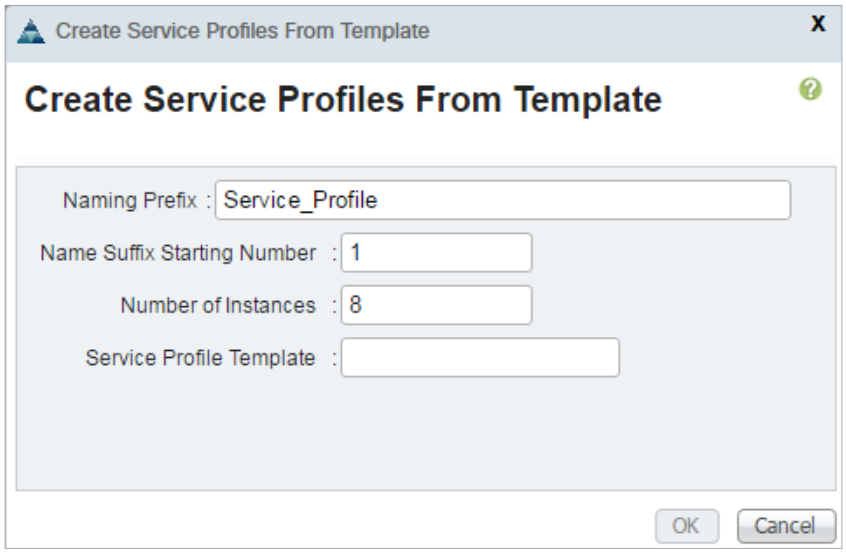


Figure 30: Creating Service Profiles

In the test setup, from the service profile template, or the AppsDB, Kaminario created eight service profiles with prefix: Service_Profile.

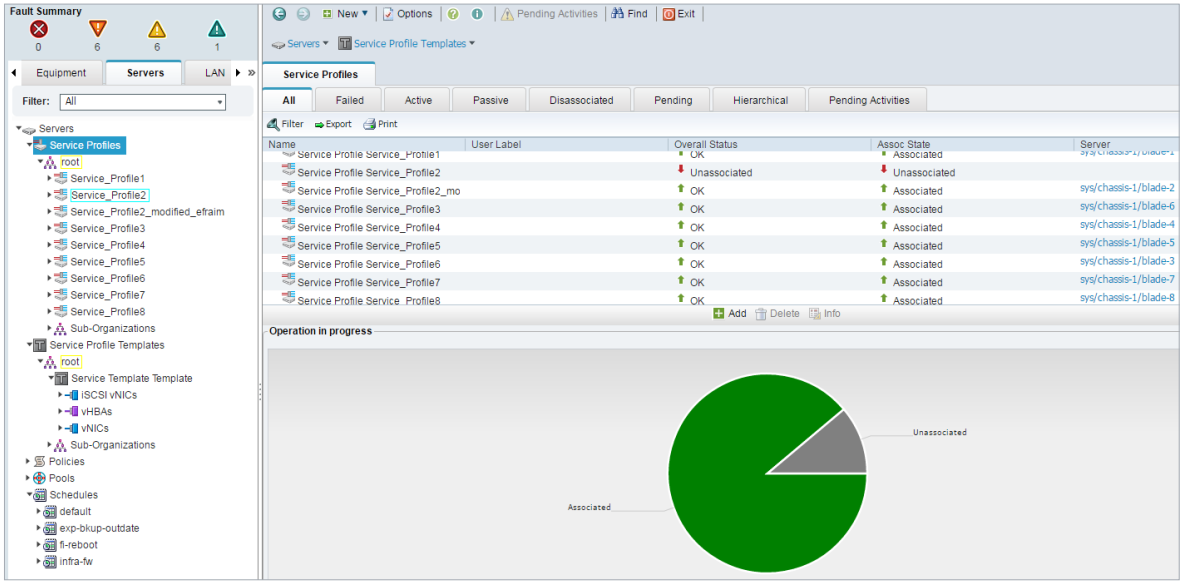


Figure 31: Service Profiles

Associating a Service Profile to Servers

When you create service profiles, you can associate them with servers.

1. In the UCS Manager, click **Servers tab > Service Profiles > Root.**
2. Select the desired service profile.
3. In the Associate Service Profile screen, set the Server Assignment to **Select Existing Server**.
4. Select the desired blade server and click **OK** associate the blade server to the service profile.
5. **Note**: Make sure the FSM (Final State Machine) Association progress status completes to 100 percent.
6. Repeat the same steps to associate the other service profiles for the respective blade servers.

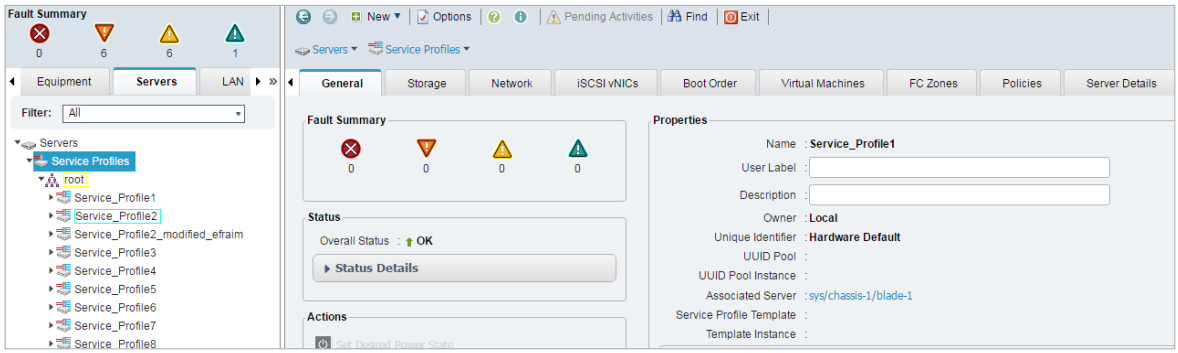


Figure 32: Associating Service Profiles to Servers

Basic UCS Networking Implementations

iSCSI Implementation

In both the Cisco UCS Mini and the full on UCS chassis, there are two ways of connecting the Kaminario K2 array. The array can either be directly connected to the Fabric Interconnect, hereinafter referred to as FI, or to another switch that the FI will be connected to.

NOTE: If the K2 array is directly connected to the FI, it will only be accessible within that UCS domain as per Cisco's design.

DIRECT FABRIC INTERCONNECT CONNECTION

When connecting the K2 array directly to the FI, please first verify that the FI is operating in end-host mode (default). This can be verified via UCS Manager by going to **Equipment > Fabric Interconnect > Fabric_Interconnect_name.**

In addition, the port personality used on the FI must be set as an appliance port as the FI will not automatically recognize the connection. The port personality can be set through UCS Manager by following the below listed steps.

Step 1: Select the **Equipment** tab

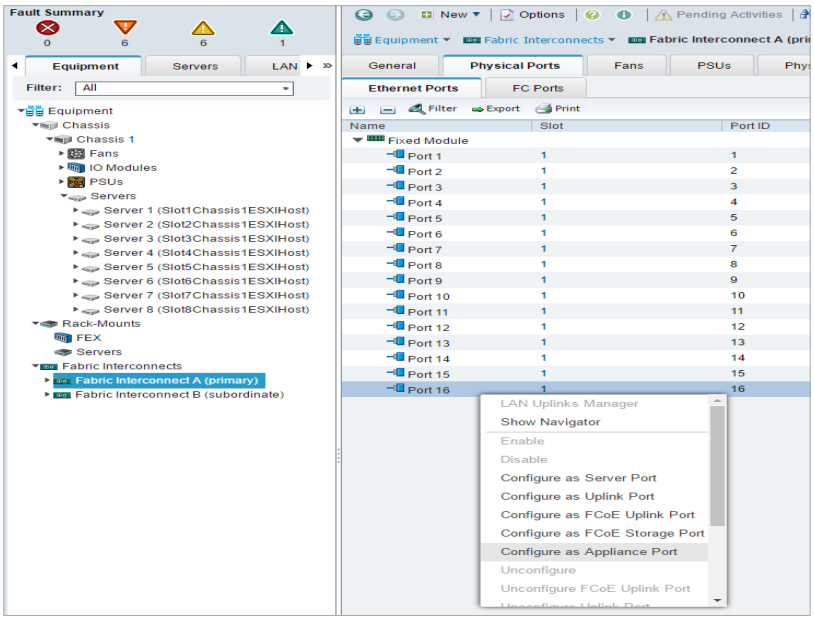
Step 2: Expand **Equipment > Fabric Interconnects > Fabric_Interconnect_Name**

Step 3: Select the port to be configured

Step 4: Select the **General** tab

Step 5: In the **Actions** area, select **Reconfigure**

Step 6: Select **Configure as Appliance Port f**rom the drop-down list



NOTE: If an expansion module is installed on the FI and changes are made to the ports on the expansion, only the expansion will need to reboot and not the FI as a whole. However if a change is made to the fixed module, the entire FI will reboot causing a loss of traffic. In cluster environments with vNICs that are configured for failover (within the Service Profile) this can be avoided.

INDIRECT FABRIC INTERCONNECT CONNECTION

Another option is to connect the K2 array to a switch higher up in the stack. In that case, the port on the FI would need to be configured as an uplink port. This can be done within UCS Manager by going to **Equipment > Fabric Interconnects > Fabric_Interconnect_name**. Within the **General** tab there will be an option to reconfigure a port from which it can be configured as an uplink port.

NOTE: Uplink ports on the FI are always trunk ports. Therefore the northbound port on the uplink switch must also be configured as a trunk port.

It is recommended to create a port channel from the FI to the uplink switch for resiliency. A port channel can be created by selecting the **LAN** tab, expand **LAN > LAN Cloud**, select the FI where the port channel will be added, right-click the **Port Channels** node and select **Create Port Channel**. There can be up to 16 Ethernet ports to a port channel.

Fiber Channel Implementation

As with the iSCSI implementation, there are two ways of connecting the Kaminario K2 array to the UCS environment. Direct connection to the FI or to a northbound switch. Like the iSCSI implementation, direct connection to the FI will make the array a dedicated array for the UCS domain.

DIRECT FABRIC INTERCONNECT CONNECTION

If the fiber channel array is directly connected to the FI, the FI must be configured to Fiber Channel Switching Mode. Please see the steps below via UCS Manager.

NOTE: When reconfiguring the FI to Fiber Channel Switching Mode both fabric interconnects will reload **simultaneously** resulting in a system-wide downtime.

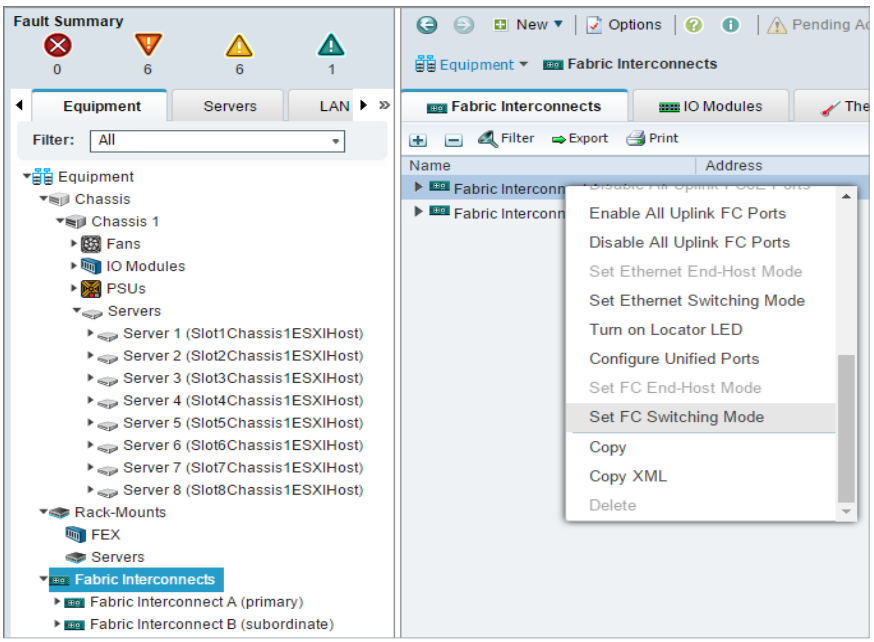
Step 1: Select the **Equipment** tab

Step 2: Exand **Equipment > Fabric Interconnect > Fabric_Interconnect_name**

Step 3: Select **General** tab

Step 4: Select Se**t Fiber Channel Switching Mode**

Step 5: Confirm the changes



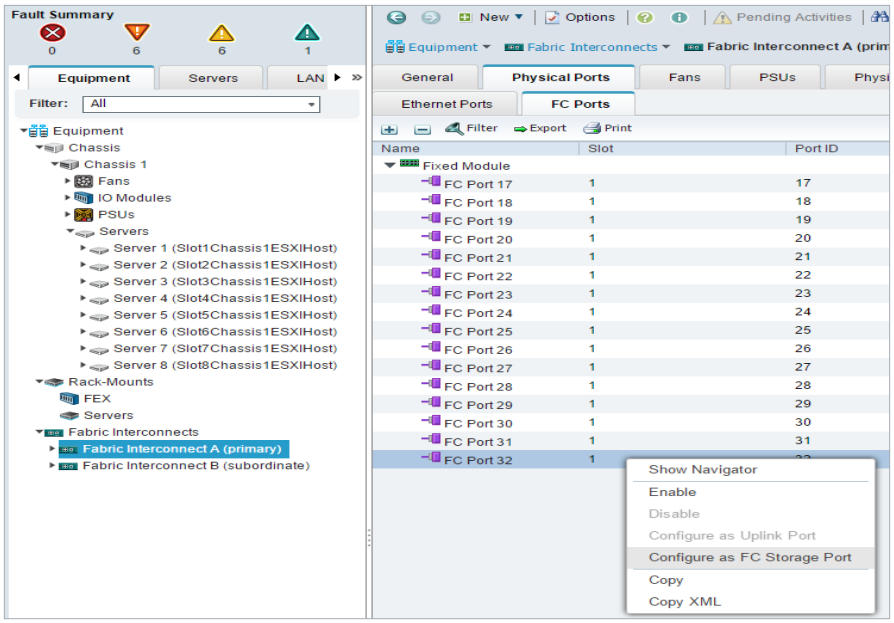
Once the FI has been reconfigured to Fiber Channel Switch Mode, it is now time to configure the port personality.

Step 1: Select the **Equipment** tab

Step 2: Expand **Equipment > Fabric Interconnect > Fabric_Interconnect_name**

Step 3: Select port(s) and right-click and select **Configure as FC Storage Port**

Step 4: Confirm changes



Additionally, in the case of direct FI connection, all zone configuration must be done on the FI.

DIRECT FABRIC INTERCONNECT CONNECTION

When the array is not directly connected to the FI, the FI must remain in end-host mode (NPV mode). Fiber channel ports are, by default, configured as Fiber Channel uplink ports. Any and all zoning configurations would need to be done on the Fiber Channel switch.

Kaminario K2 Storage Configuration

The following section discusses the Kaminario storage layout design considerations when deploying a UCS solution.

OS Installation

For the test setup, Kaminario configured a single node ESXi-6.0.0 system using a Cisco B200 M4 blade server. The OS components were configured to use the FC protocol on Kaminario storage.

Component	Details	Description
Server	1xB200 M4	2 Sockets with 12 cores with HT enabled
Memory	131 GB	Physical memory
NIC1	Public Access	Management and Public Access, MTU Size 1500

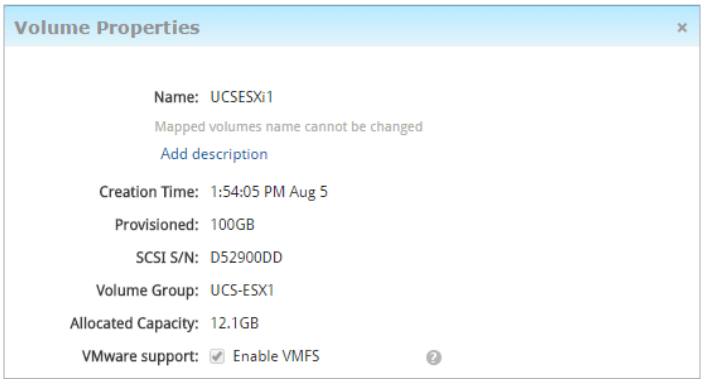
Complete the following steps to install the OS and other required packages:

1. Install 64-bit ESXi-6.0.0 following the standard VMWare installation guide

Configuring Kaminario provisioned volumes on ESXi 6.0

The following guidelines should be taken into consideration while provisioning volumes to a VMware system from the Kaminario K2 storage array.

When creating a volume that is intended for use by VMware as VMFS datastore, the VMware support flag should be enabled at the Kaminario level:



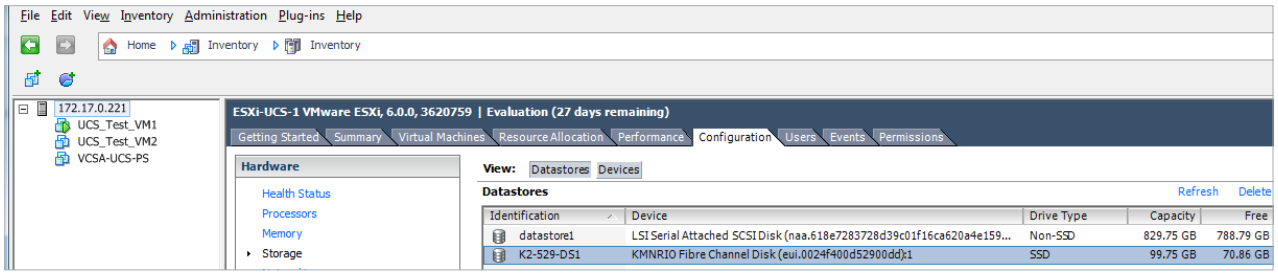
When creating a volume that is intended for use by VMware as an RDM device, the VMware support flag should be disabled at the Kaminario level.

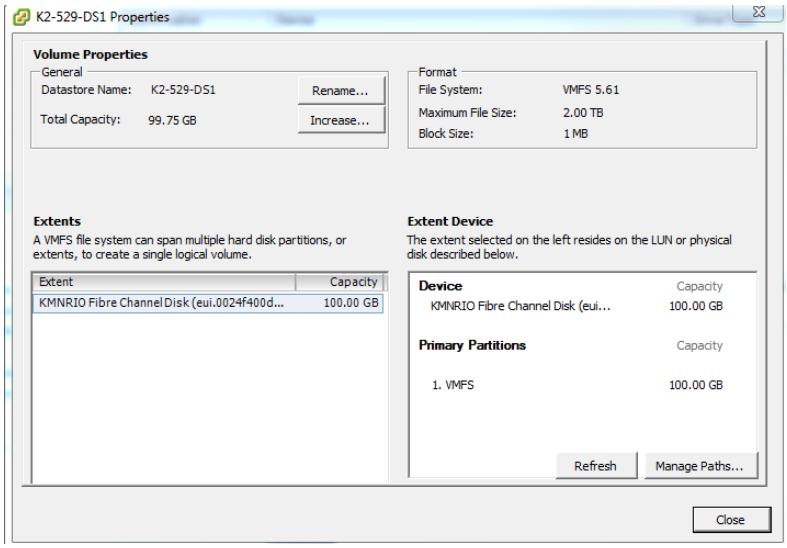
Round-Robin Path Configuration

Kaminario architecture works in an active/active fashion. It is highly recommended to use a round-robin load-balancing policy. Round-robin can be set for each LUN via the vSphere client. However, this can be a time consuming task if you are creating many LUNs. This section will show how to change the load-balance policy after a LUN has been created and mapped to the ESXi server. If you wish to change the default settings so that new LUNs will automatically be set to round-robin, that requires a CLI command with will be covered in the next section.

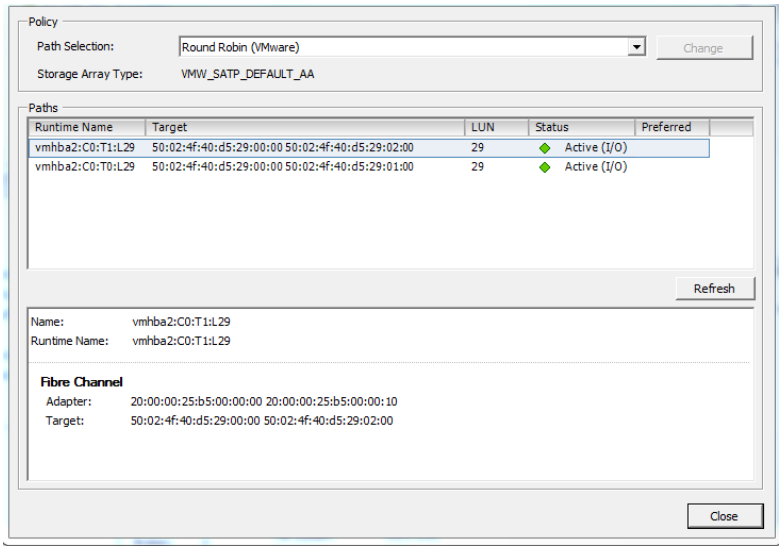
To set the load balance policy to round-robin for a LUN, select an ESXi server, click on the configuration tab and select storage in the hardware section. Click on the LUN and click "properties".

Click on the Manage Paths button.





Change the path selection to Round Robin (VMware) and click the Change button. Once the changes have taken affect, click close on the manage paths windows and click close on the LUN properties window.



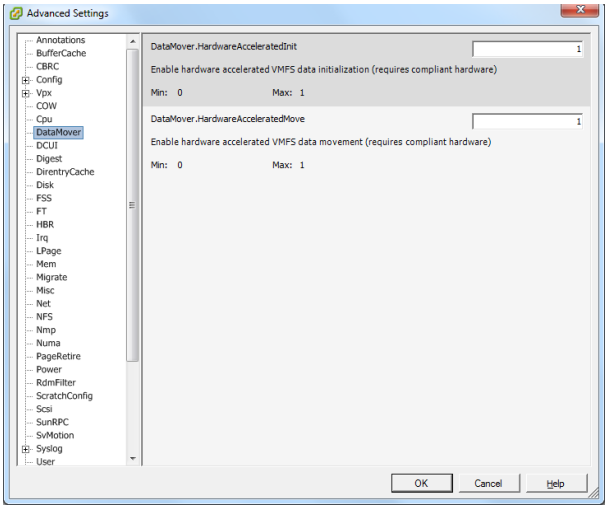
VMware API for Array Integration

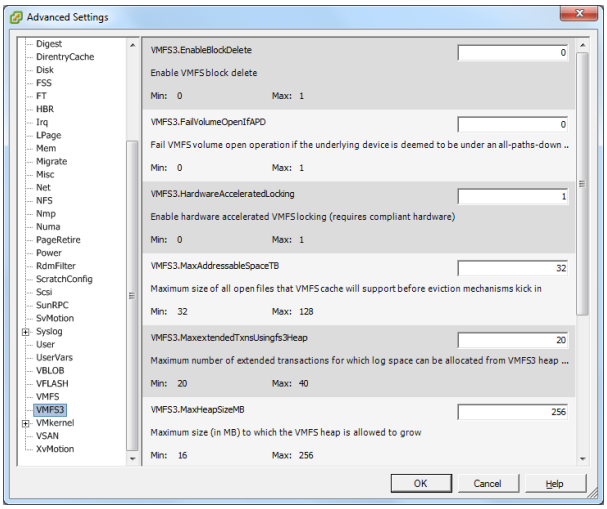
Enabling VAAI is strongly recommended in order to gain the most performance out of the K2. VAAI enables the host to offload certain storage functions to the array in order to reduce resource overhead on the host and significantly improve performance for storage intensive operations.

Make sure all servers in the cluster are VAAI-enabled. To enable VAAI (or verify that it is already enabled) run the vSphere client, select an ESXi server, click the configuration tab and click on advanced settings in the software section.

Click on DataMover and set (or verify that) HardwareAcceleratedInit and HardwareAcceleratedMove to 1.

Click on VMFS3 and click set (or verify that) HardwareAcceloratedLocking is set to 1.



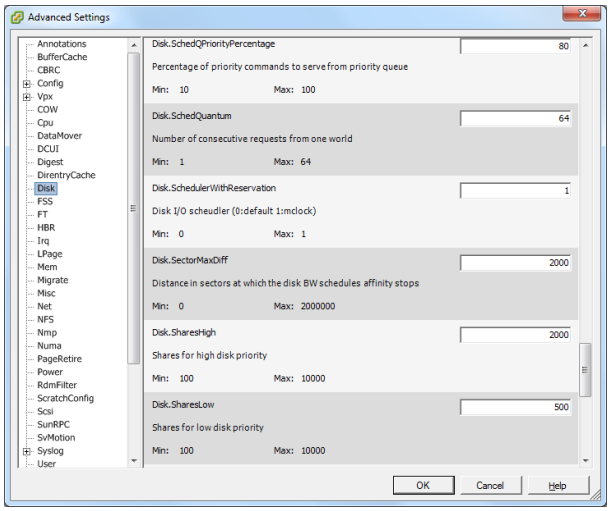


The Disk.SchedQuantum parameter defines the maximum number of consecutive commands issued from the same virtual machine, even though there might be I/O from another virtual machine. This is done in order to maintain the proper sequence of virtual machine IO streams.

By default, the Disk.SchedQuantum parameter is configured to a value of 8. Kaminario recommendation is to set the Disk.SchedQuantum paramter to 64. When Disk.SchedQuantum is configured to 64, the K2 system serves I/O requests in a rate that improves virtual machine throughput with no impact on latency.

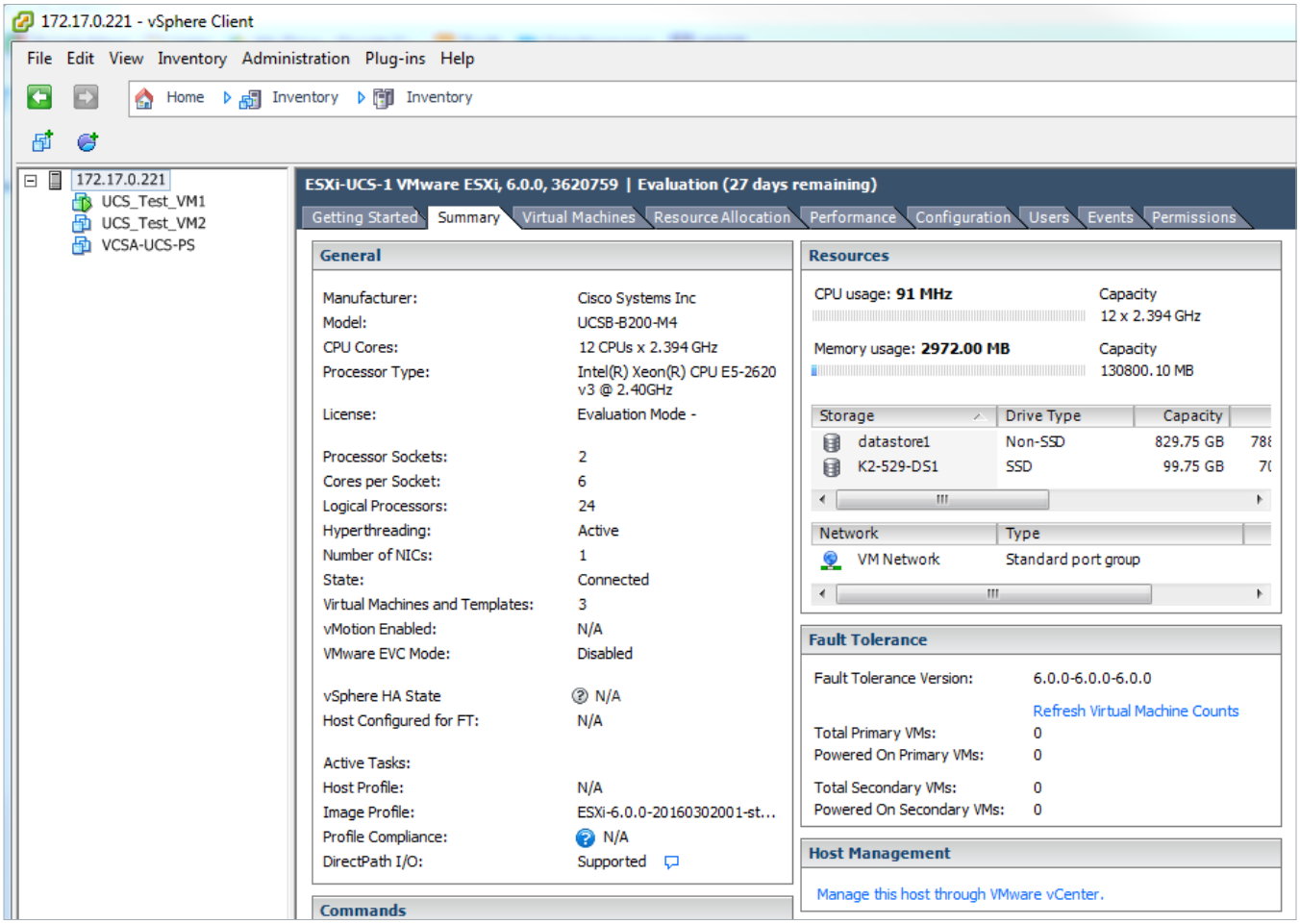
Click on Disk and set the Disk.SchedQuantum parameter to 64.

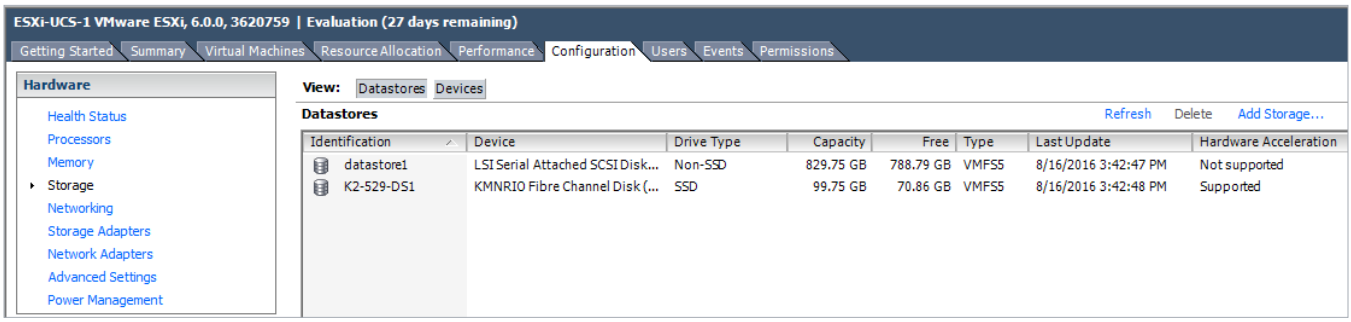
➲ The **Disk.SchedQuantum parameter** is a global parameter, meaning it has impact on all storage arrays attached to the specific ESXi host. Consult other storage array vendors before applying this setting.

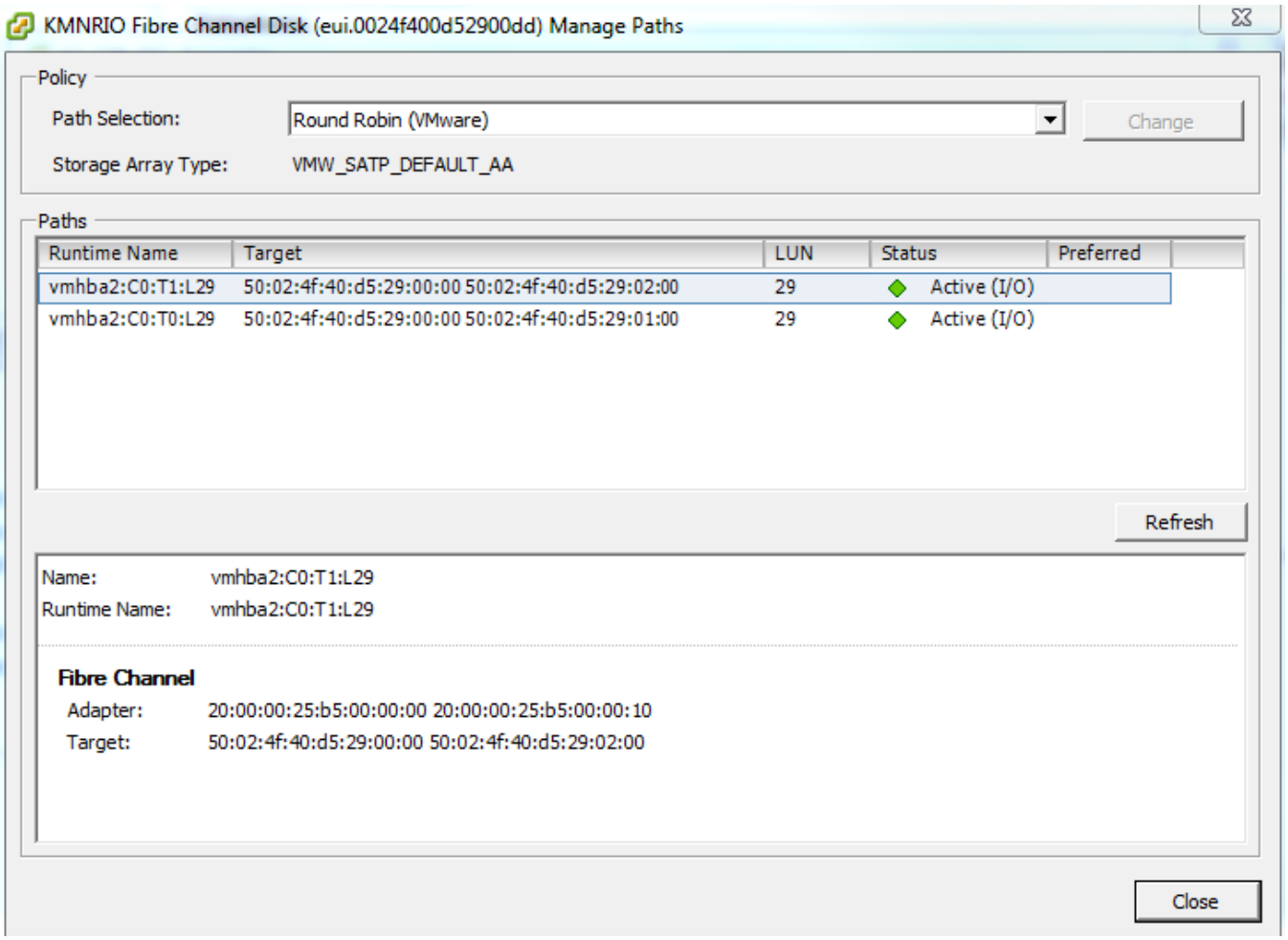


VMware ESXi

ESXi Configuration





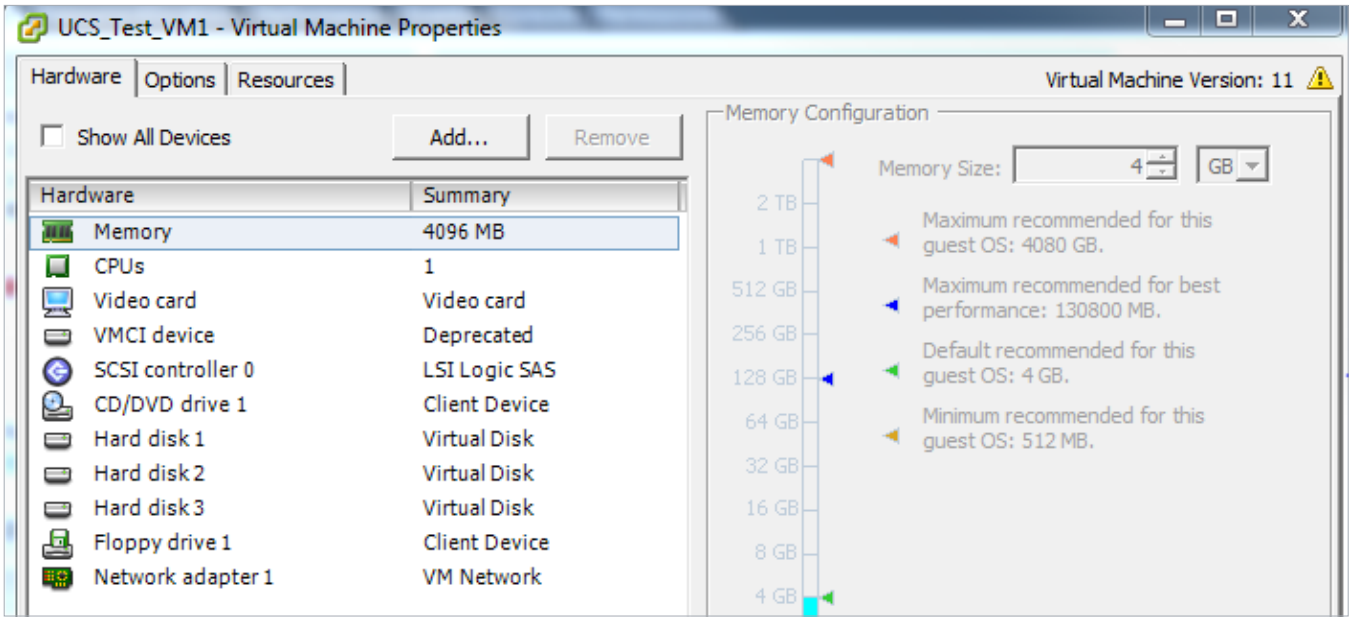


VM's configured with 1 vcpu, 4GB RAM,

/ = 20Gb

/data = 500 Gb

/log = 200 Gb



Validation

We have tested the performance of the different configurations by running three different workloads utilizing the Oracle Vdbench utility. Vdbench is a command line utility specifically created to help engineers and customers generate disk I/O workloads to be used for validating storage performance and storage data integrity. We have tested and reported the performance of each configuration by looking at both IOPS and throughput.

For IOPS, we reported the value that the Kaminario K2 could handle while keeping IO latency below 1ms.

Three different workloads were used for the IOPS test:

- 4k 100% Random Reads
- 4k 100% Random Writes
- A mixed workload of 80% random Reads and 20% Writes

For throughput, we utilized similar testing methodologies but with a much larger IO block size of 32K.

The same three workloads were then used for the throughput test:

- 32k 100% Random Reads
- 32k 100% Random Writes
- A mixed workload of 80% random Reads and 20 Writes

The table below compares the performance of the different configurations according to pre-defined tests and use-cases. Readers who would like to obtain the actual Vdbench parameter files used for each test are encouraged to contact their Kaminario representatives.

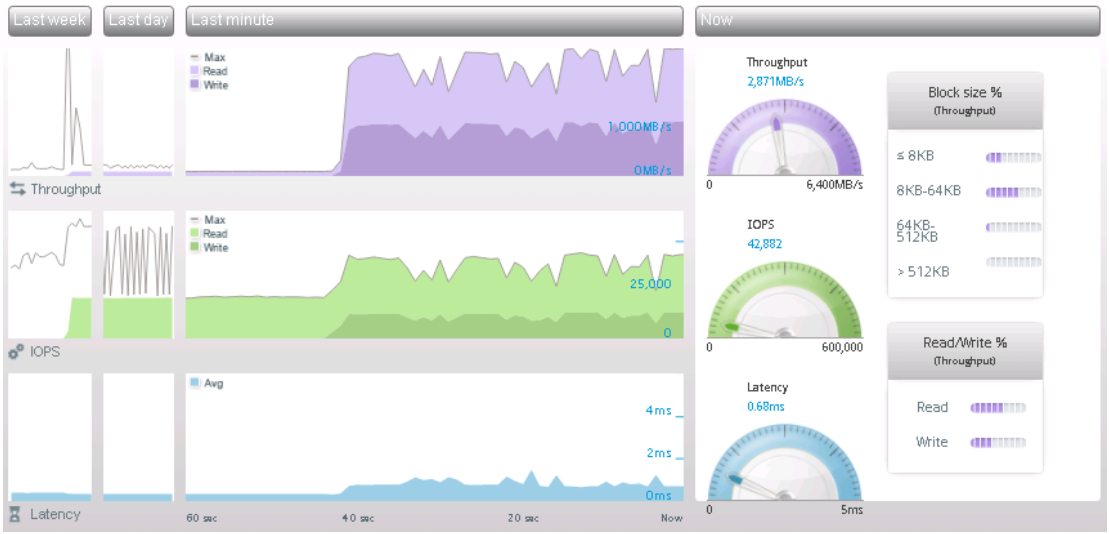
	Use Case	Small (1 KB)	Medium (2 KBs)	Large (4 KBs)
IOPS	4k 100% Random Reads	245K	470K	881K
	4k 100% Random Writes	104K	210K	420K
	4k 80% Random Reads, 20% Random Writes	203K	379K	725K
Throughput	32k 100% Random Reads	3,000 MB/s	6,000 MB/s	12,200 MB/s
	32k 100% Random Writes	1,000 MB/s	2,200 MB/s	4,100 MB/s
	32k 80% Random Reads, 20% Random Writes	3,100 MB/s	6,600 MB/s	11,800 MB/s

In addition to the Vdbench validation, we also simulated real database activity attempting to reflect real world scenarios. We executed long running batch jobs utilizing large block reads and writes to simulate an OLAP workload and single record queries to demonstrate typical OLTP characteristics.

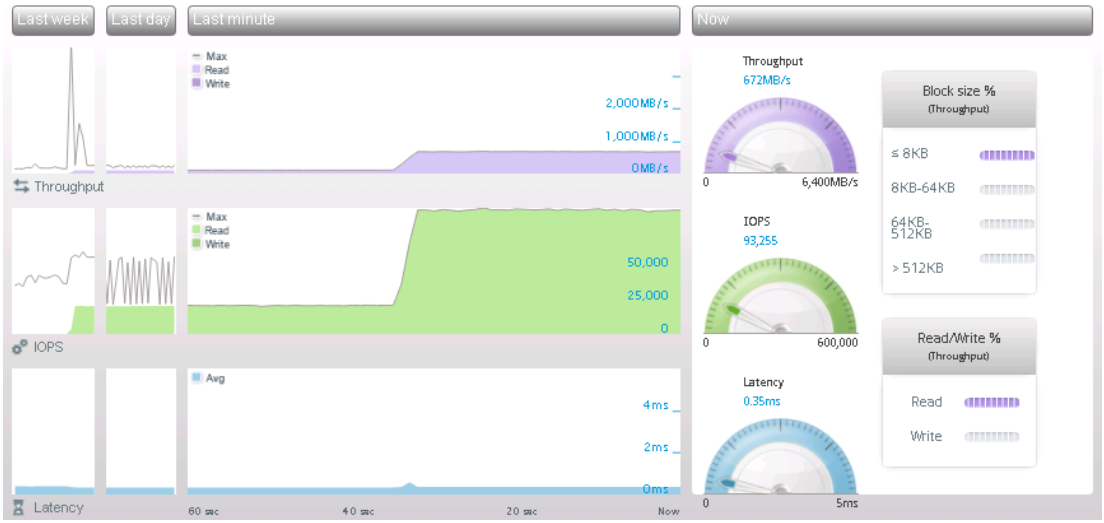
The screenshots below present the Kaminario GUI and how it displays the performance of the Kaminario UCS configuration (see "Small" configuration in Hardware and Software Used for Different Solutions) according to pre-defined tests and use-cases.

Metrics used:

- **K2 Latency:** The inner latency measured from the time an IO request arrives until the time the K2 sends an acknowledgement to the initiator for that IO. The K2 GUI shows only the inner latency.
- **IOPS**: Input/output operations per second as measured by the K2.
- **Throughput:** Accumulative size of all the input/output operations processed by the K2 per second.



OLAP



OLTP

Conclusion

Cisco UCS paired with Kaminario Storage Arrays create a highly performing and cost effective platform for enterprise level deployments.

The decidedly differentiated architectures are well aligned to bring customers value for enterprise level deployments initially, as well as with evolving technology over time.

Critical Business Driver	UCS Enabling Technology	Kaminario Enabling Technology
Performance and Capacity Optimization	Varying blade models with CPU, memory options. VIC virtual interfaces and 10G integrated fabric	All-flash offerings, enterprise SSD
Data Protection	Service Profile – stateless computing. Fully redundant HA fabrics	Snapshots, no overhead Replication, K-Raid™
Time to Market for Enterprise Applications	Service Profile template Export, Import of XML schema	Views, read/write without performance or space overhead
Consolidation	Varying workloads on same fabric using instrumented QoS and Virtual interfaces Multiprotocol via different port options	Guaranteed low latency allows for fewer server cores Denser storage means few racks in the datacenter
Reduce Infrastructure Costs	Stateless computing allows fewer spare systems Converged 10G fabric, FCoE reduces infrastructure	Thin Provisioning Inline Compression, 50% space savings

kaminario.

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, the Kaminario logo and SPEAR are registered trademarks of Kaminario Inc.

K-RAID, Perpetual Array and HealthShield are trademarks of Kaminario, Inc.

Product specifications and performance are subject to change without notice.