vexata

SYSTEM ARCHITECTURE WHITE PAPER

AUGUST // 2017

SYSTEM ARCHITECTURE WHITE PAPER

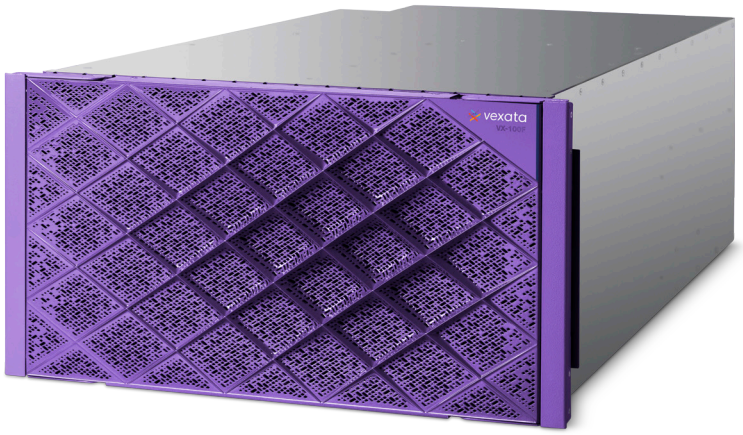
1: INTRODUCTION

1.1 Abstract

This white paper presents the Vexata System Architecture and the various components that make up the architecture. The details provide an understanding of how the Vexata system enables enterprise class features at massive performance and scale that is not achievable by legacy all-flash arrays (AFAs).

1.2 Audience

This document is intended for IT architects, and database, system and storage administrators responsible for managing enterprise applications and infrastructure.



VX-100F

1.3 Need for a new architecture for high performance data infrastructure

As enterprises increasingly shift to online business models, applications that drive the backbone of the digital economy such as Oracle, SQL Server, SAP, SAS are expected to run faster and at a larger scale. Applications are expected to handle more users, work on much larger data sets and enable real-time responses and decision making.

While compute and networks infrastructures have significantly advanced in performance to enable application acceleration, storage infrastructures have not kept pace even with the advent of high performance solid-state drives (SSD). Incumbent All Flash Arrays are unable to realize the full potential of the SSDs due to architectural limitations in existing AFA architectures.

Peak performance demands, uncompromising resiliency and a focus on maximum application utilization are key demands of a modern storage architecture. The Vexata system architecture is a groundbreaking data architecture that enables performance at scale with enterprise class services.

vexata

SYSTEM ARCHITECTURE WHITE PAPER

2: VEXATA SYSTEM ARCHITECTURE

The Vexata System Architecture is designed with the most demanding enterprise environments and meets performance and resiliency requirements at scale. The main components of this architecture are shown below

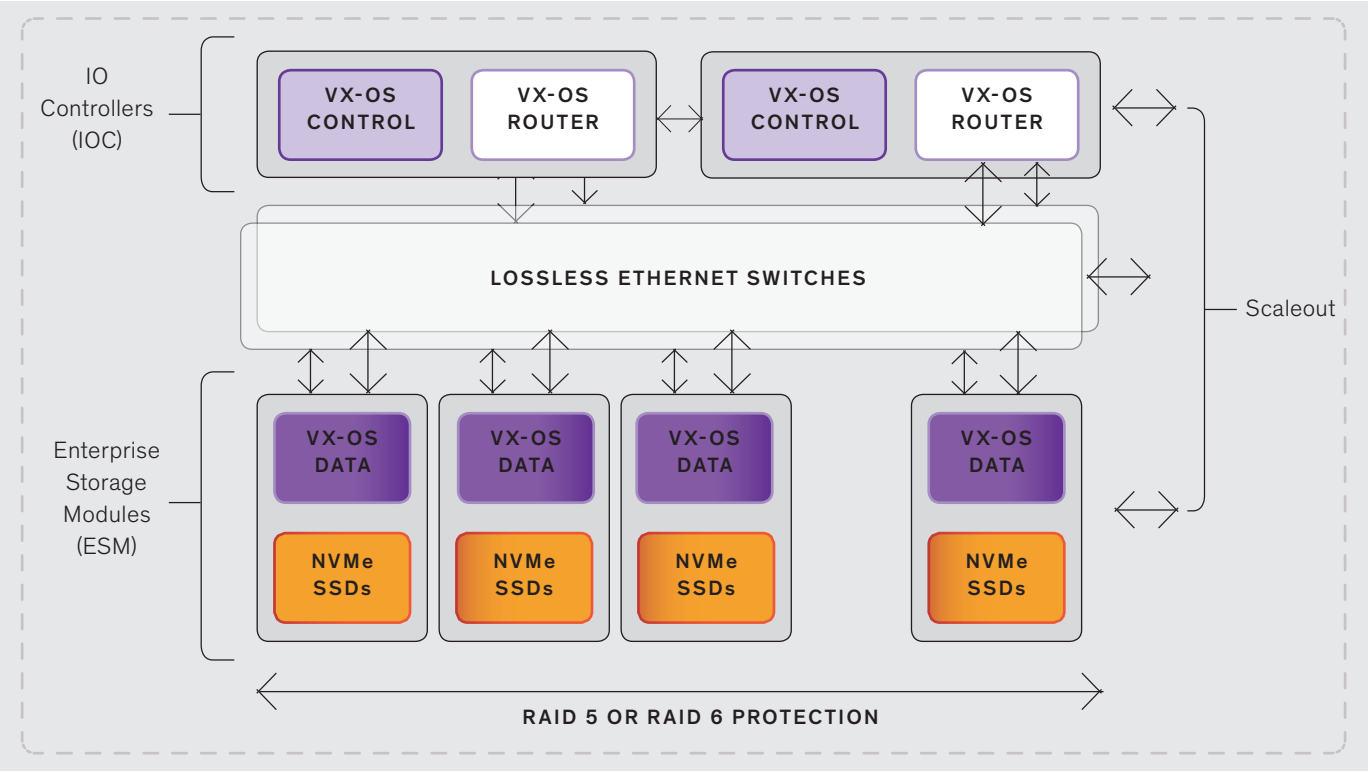


Figure 1. Vexata System Architecture

2.1 IO Controllers

The Vexata system features two active-active IO controllers which handle host IO traffic. Unlike traditional storage architectures, the IOC is split into two parts:

- VX-OS Control which implements the control path software for enterprise class storage services

- VX-OS Router which implements hardware acceleration assists for delivering ultra high performance storage services

2.2 Enterprise Storage Modules (ESM)

Unlike traditional architectures that use SSDs that are accessed in a single-threaded fashion, the Vexata system features intelligent storage modules that allow for multiple SSDs to be accessed in parallel by the front-end IOCs. The ESMs consist of a software called VX-OS Data and 4 PCIe NVMe SSDs.

2.3 Lossless Ethernet Network

Existing architectures use PCIe or SAS switches to access SSDs resulting in low access bandwidths and high latencies. These switches also don't allow easy performance and capacity scaling. The Vexata system uses low-latency lossless Ethernet switching to enable the IOCs achieve high performance parallel access to the storage modules.

The architecture enables scale out of IOC, Switch and storage modules to achieve massively parallel access to petabytes of data at ultra high performance.

SYSTEM ARCHITECTURE WHITE PAPER

3: VEXATA ARCHITECTURE CORE DIFFERENTIATORS

Core architectural innovations separate the Vexata architecture from existing storage array architectures. They are :

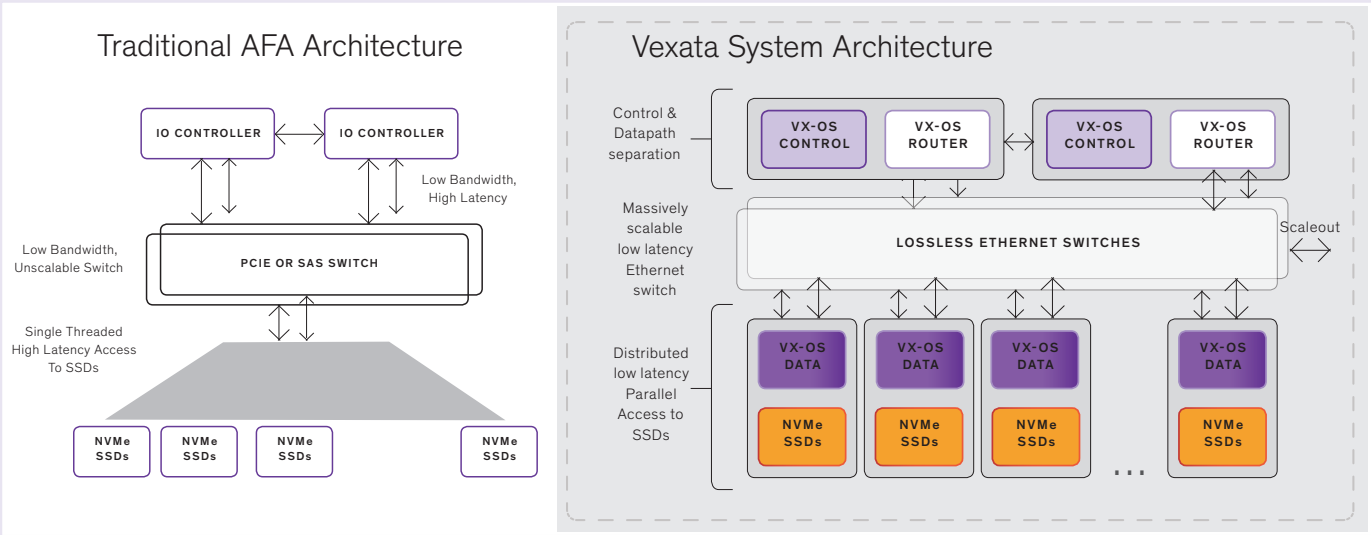


Figure 2. Vexata vs Traditional AFA differentiators

3.1 Control and Data Path Separation

Enterprise storage processing places significant demands on compute for RAID, metadata search, data reduction, data security, data movement and other services. Traditional AFA architectures that use CPUs alone run into significant bottlenecks and system performance suffers as a result. SSDs are significantly underutilized in the system and overall throughput and latency are compromised.

Vexata architecture addresses this problem by separating the control and data paths for storage processing. Control packets are processed by VX-OS Control, while all compute intensive tasks are offloaded into a separate data processor named “Router” which performs RAID, encryption, metadata search and data movement. This core innovation allows at enormous bandwidths and ultra-low latencies.

3.2 Highbandwidth parallel IO access to drives

Traditional array architectures access SSDs one drive at a time leading to wasted performance and much worse latencies. One of the key elements of the Vexata architecture is a high bandwidth parallel IO access to all the drives in the array. A ultra low latency lockless distributed operating system enables this parallel IO access resulting in almost multiplicative performance across all the drives in the array while maintaining the ultra low latencies of Flash SSDs and that of newer media such as Intel 3D Xpoint (Optane) Technology. As an example, a Vexata Flash Array with 64 drives achieves over 7M IOPS at less than 200us latency.

3.3 Optimizations for Ultra Low Latency

Minimizing the storage latency is critical to eliminating the IO bottlenecks at the host. The Vexata architecture implements high bandwidth low latency lossless Ethernet switching with 256 10GE ports to move data between the IOCs and the ESMs. The Ethernet MACs were designed very carefully to minimize receive and transmit latencies on the network. Similarly, the entire software stack has been designed in user space with the goal of reducing latency and minimizing latency variation. These optimizations combined with the other innovations already discussed make the Vexata architecture maximally utilize solid state media (Flash and Optane SSDs) and provide best in class results.

4: PRODUCTS AND FEATURES

The Vexata system architecture has been validated successfully with two enterprise class storage arrays:

- The VX-100F Flash Storage System with off-the shelf single-ported PCIe NVMe SSDs from multiple vendors
- The VX-100M Optane Storage System with Intel Optane drives

Both products offer ultra high performance with comprehensive storage services. The Vexata architecture ensures that performance is not compromised even under impaired conditions as described in the sections below.

4.1 High Availability

This white paper presents the Vexata System The Active-Active IOC architecture ensures that data stored in the Vexata array is available 99.9999% of the time. Host IO operations and subsequent metadata changes are always kept in sync between the two IO controllers using very low latency Ethernet links. Host reads are served by both IOCs resulting in balanced performance.

During normal operation, each IOC is capable of handling 7 million IOPS thanks to the Router offload. When an IOC is impaired, the second IOC takes over operation within milliseconds. Router offloads ensure that even a single IOC delivers phenomenal performance and application responsiveness is significantly minimized even during IOC failure.

4.2 Non-disruptive Upgrades (NDU)

The Vexata architecture supports NDU updates of the VX-OS software in the field. VX-OS Data on the ESMs can be upgraded without any disruption to normal operation. VX-OS Control on the IOCs can be updated one IOC as a time with no impact to data availability.

4.3 RAID

This white paper presents the Vexata The Vexata architecture supports both RAID5 and RAID6 options for protecting data stored in the storage modules. Since RAID computation is offloaded to the Router, enormous data rates are achieved with RAID always turned on. In the event of an ESM failure, the Router is able to read all ESMs in parallel and create the missing data. Router RAID offload provides the acceleration needed to compute and serve missing data until the failed ESM is replaced. RAID offload in the Router speeds up rebuild performance when a new ESM replaces the failed component. Rebuild rates of 2TB/hour are easily achieved on the Vexata Array.

4.4 Encryption

Data-at-Rest-Encryption (DARE) is an absolute necessity to secure mission critical data from theft. The Vexata architecture supports in-line DARE in the Router using AES-256-XTS or AES-256-CBC methods. Key management is done by the Controller with keys stored in the array, or in external frameworks accessible through REST APIs. High performance is achieved with encryption enabled in the system

4.5 Snapshots and Clones

The Vexata Array supports space efficient zero cost read-only snapshots and writeable clones of application volumes. Because of the offloaded architecture, the CPUs in the Vexata Array are lightly loaded due to data services getting offloaded into the Router. As a result, the CPU and its associated memory are able to provide much faster, scalable snaps and clones. With very little CPU memory, the Vexata Array supports up to 256 snapshots per volume and a total of 64K volumes/snaps/clones.

Snaps are created in memory with a Redirect on Write (RoW) implementation. Snap creation is therefore simple and instantaneous with snapshot configuration persisted in the array. Snap deletes require consolidation and copying of configuration across different snapshots and are compute intensive tasks. The snap delete offload function provided by the Router accelerates configuration data movement resulting in much faster deletes when compared to existing AFA implementations.

5: SUMMARY

Vexata's mission is to build high performance data infrastructure to accelerate applications. The Vexata Architecture is key to enabling solid state storage performance at scale with enterprise class services. Vexata offers the VX-100F Flash Array and VX-100M 3D Xpoint Array based on the differentiated architecture.

ABOUT VEXATA

Vexata is the leader in active data management solutions. Vexata's unique breakthrough enterprise offerings enable transformative performance and scale from database and analytics applications. With unparalleled ability to consume the latest in media like NVMe Flash and now with with Optane™ SSDs, Vexata systems deploy simply and seamlessly into existing storage environments. Learn more at www.vexata.com

WP-1008-09142017

Contact Vexata: info@vexata.com