White Paper

New Object-Oriented Paradigm Accelerates Storage Network Shift to Ethernet

Since its inception as a local area network technology forty years ago, Ethernet has made its way into more and more forms of communications. Technology developers using Ethernet for communications can build on forty years of standards developments, a well-established knowledge base, and leverage large economies of scale.

Enterprise Local Area Networks (LANs) completed the shift from semi-proprietary protocols to standards-based Ethernet many, many years ago. More recently, telecommunications service providers have shifted from TDM-based SONET/SDH to packet-based Carrier Ethernet on both the wired and wireless infrastructure side. As a result, enterprises can now shift to cloud-based resources that use Ethernet carrier services for data transmission. In addition, the industrial Internet of Things (IoT) is also undergoing a transition to Ethernet networking.

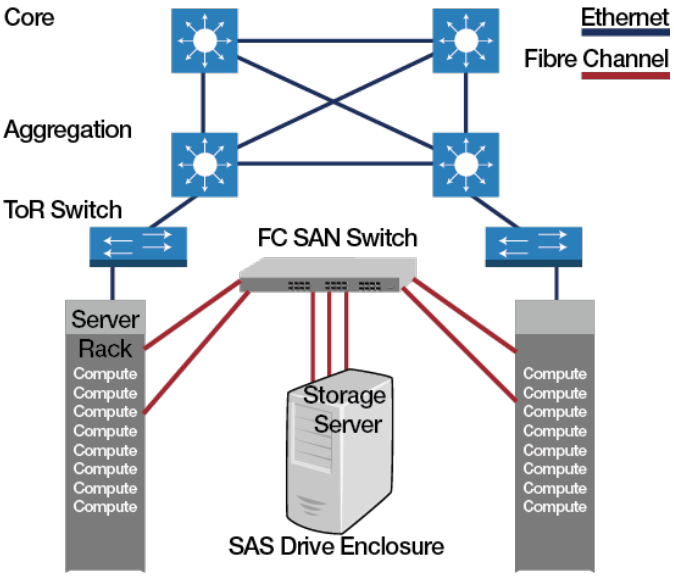
Perhaps the last bastion of specialty communications protocols is storage networking. Storage networks have continued to rely on protocols such as SAS/SATA and Fibre Channel (FC) with their own protocols and interfaces that are not well suited for networking outside a small deployment. FCoE (Fibre Channel over Ethernet) and iSCSI have promoted storage networking over IP/Ethernet networks, but specialized networks and adapters are still needed because the underlying storage is based on the venerable SCSI protocol, encapsulated many times before finally mapped to Ethernet. This reality has changed, however, with the arrival of a new Ethernet-based approach to storage networking driven by a paradigm shift in the Data Center towards object-oriented storage. Adopting Ethernet for storage networks will not only reduce costs and simplify operations, it also has the potential to accelerate the move of information technology (IT) resources to the data center cloud.

An example of such an approach is the Seagate Kinetic Open Storage Platform, developed by storage equipment manufacturer Seagate Technology with the goal of encouraging other storage manufacturers to also adopt the architecture in order to create a vibrant ecosystem. To support this goal, semiconductor technology conforming to the architecture is now available from Microsemi Corporation. This white paper explains the new architecture, how Microsemi Ethernet switch and PHY solutions support it, and the benefits provided by the total solution.

The Traditional Data Center Network

This section describes the communications protocols and connectivity that underlie traditional data center networking. The majority of large data centers are based on the Storage Area Networking (SAN) model. The following illustration shows how both Ethernet and FC are typically used within a data center to support SAN communications.

Figure 1 Traditional Data Center Networking Architecture



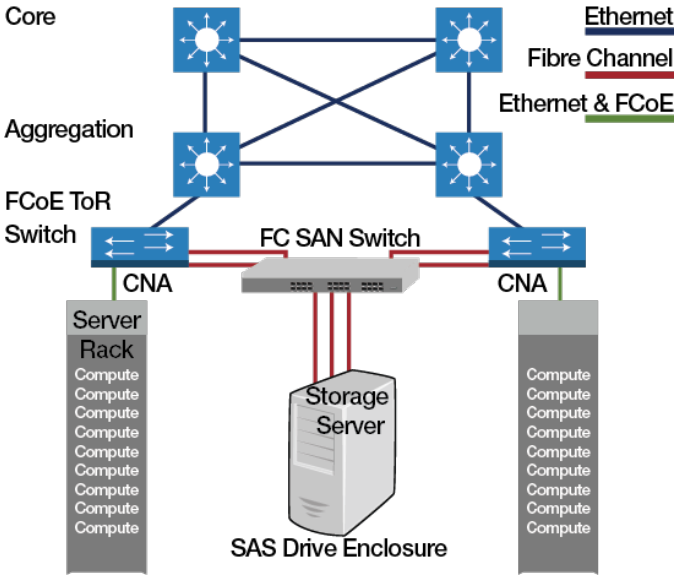
Compute servers typically have two network adapters — an Ethernet-based Network Interface Card (NIC) to connect to the Data Center LAN, and a FC-based Host Bus Adapter (HBA) to connect to the SAN. The SAN includes a FC SAN switch, which connects the HBAs on the servers to a storage server controlling numerous disk drives housed in one or multiple serial attached SCSI (SAS) drive enclosures.

Ethernet is used for communications between core and aggregation switches and between aggregation switches and top of rack (ToR) switches, which connect a stack of compute servers. Each compute server is a central processing unit (CPU) that is connected to the top of rack server through Gigabit Ethernet or 10 Gigabit Ethernet. Because the CPUs must communicate using both Ethernet and FC, each CPU has a connection to the FC SAN switch through a HBA and a separate Ethernet connection to the ToR switch using a NIC.

Fibre Channel over Ethernet (FCoE)

When FCoE is used with a SAN, the FC SAN switch communicates with the compute servers through a FCoE ToR switch. The following illustration shows this architecture in a data center networking application.

Figure 2 FCoE SAN Data Center Networking Architecture



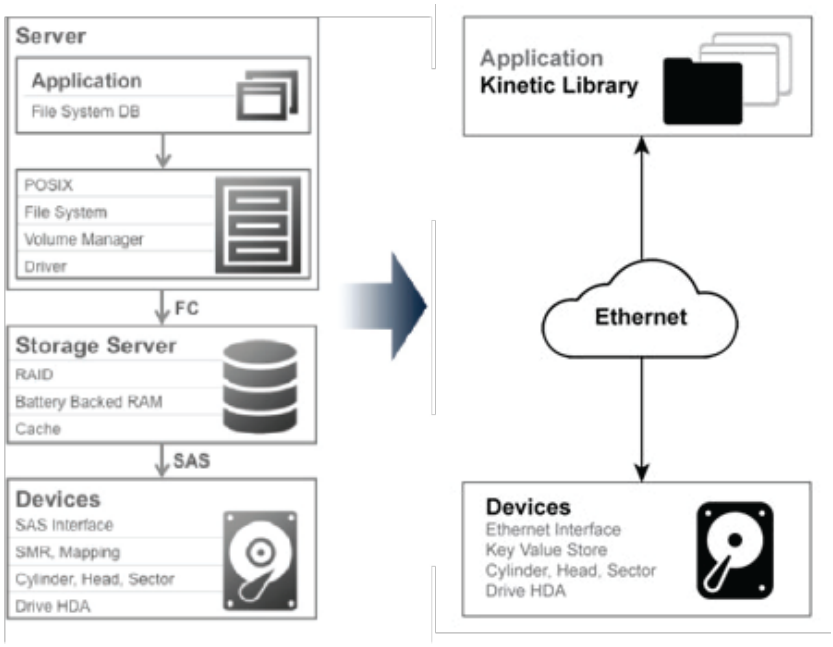
CPUs are connected to the FCoE ToR switch through a converged network adapter (CNA) that runs both FC and Ethernet over the same (typically 10 Gigabit) Ethernet connection. A key function of the ToR switch is to break out the Ethernet LAN and FC SAN traffic. Communications between the FCoE ToR switch and the FC SAN switch continue to be made via FC, as are communications between the SAS drive enclosure and the FC SAN switch. Although FCoE simplifies SAN connectivity in comparison with the traditional data center architecture by consolidating I/O from each server, it does not eliminate the use of native FC communications in the SAN.

Seagate Kinetic Open Storage Platform

Traditional storage systems use a file/directory approach similar to what is used on a personal computer, with data stored in files which, in turn, are stored in directories. However, most data stored in today's data centers is typically not stored in files and directories; a key/value, object-oriented approach is used instead. A good example of the key/value approach would be a social media website. A site user's name is the value in the key field for "name," the user's age is stored in the "age" key field and so on. With an object-oriented approach, individual data objects are referenced by a handle rather than by where they reside on a specific device. Objects are written once and read many, many times, but rarely deleted and never modified.

The Seagate Kinetic Open Storage Platform takes into consideration the object-oriented, key/field approach to storage used in today's data centers, thereby dramatically simplifying the networking architecture. The following illustration shows the Seagate Kinetic Open Storage Platform functionality in contrast with the traditional SAN approach.

Figure 3 Seagate Kinetic Open Storage Platform

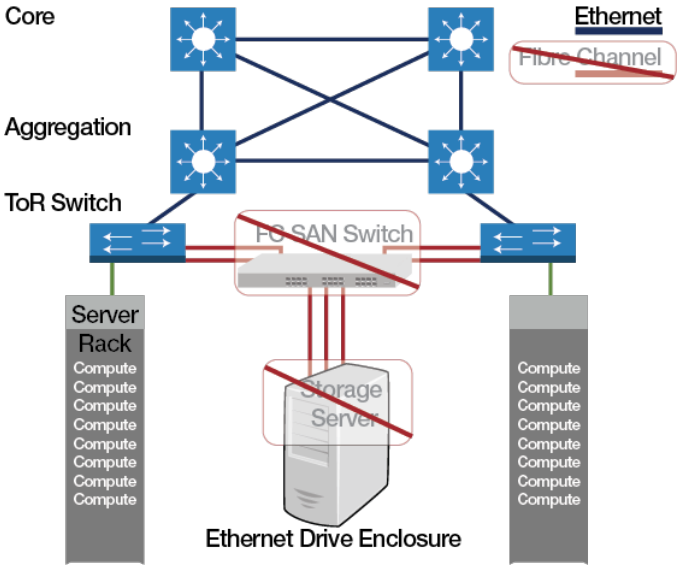


The traditional SAN architecture uses multiple layers of software and hardware to accommodate the file/directory approach. The transit path from application to storage begins with the file system database, then uses the twenty-five-year-old POSIX protocol to reach file systems, volume managers and drivers. This is followed by Fibre Channel communication to storage servers, and data finally reaches storage devices using the serial attached SCSI (SAS) protocol.

The Seagate Kinetic Open Storage Platform recognizes the changed nature of how data is stored in today's data centers. The object-oriented, key/file data stored in today's data centers doesn't need file semantics or a file system to handle space management on a storage device. Conceptually, the Kinetic storage stack eliminates multiple layers of hardware and software so that applications communicate directly with storage devices using Ethernet.

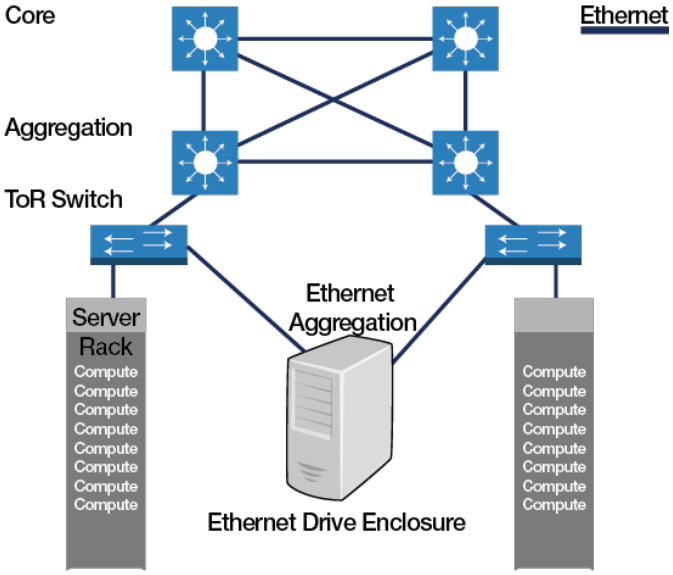
The following illustration shows the architecture of a Kinetic data center networking application.

Figure 4 Kinetic Data Center Networking Architecture



In the Kinetic architecture, both the FC SAN switch and the storage server are eliminated. Storage drives have Ethernet rather than SAS interfaces, and Ethernet switches take the place of SAS/SATA expanders inside the storage enclosures. The following illustration shows the simplicity of the storage over Ethernet approach.

Figure 5 Storage over Ethernet: A Low Power, Lower Cost Solution

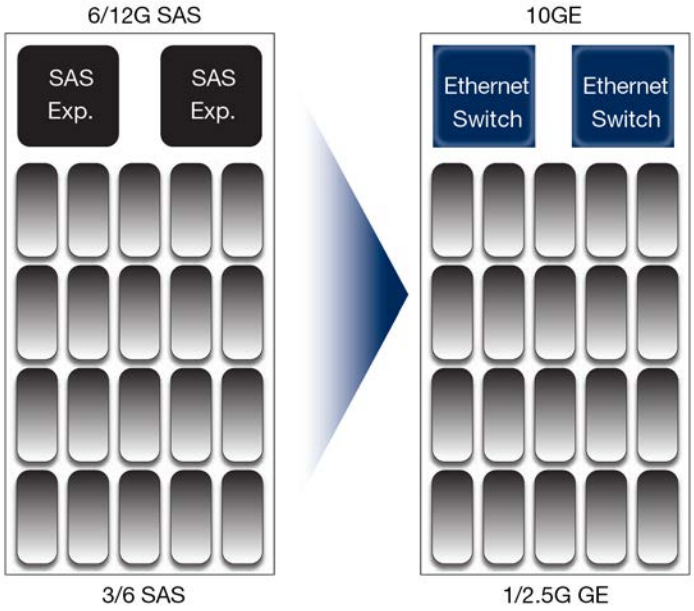


As a result, there is no need for FC or FCoE. All communication is through Ethernet. In contrast with FCoE, this is not simply input/output consolidation— it is full network integration that enables data centers to be designed differently for more efficient operations.

The Ethernet Storage Rack

The following illustration shows the key differences between a traditional storage rack and a storage rack conforming to the Seagate Kinetic Open Storage Platform.

Figure 6 The New Storage Rack

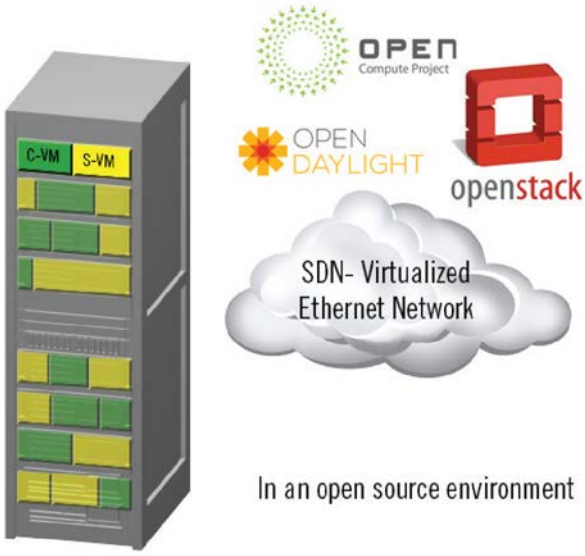


Traditionally, there is a storage server that communicates with the ToR switch using 10GE, and with SAS expanders using a 6G or 12G SAS connection. The SAS expanders communicate with individual storage drives using 3G or 6G SAS.

The Seagate Kinetic Open Storage Platform completely eliminates the storage server, and the SAS expanders are replaced by Ethernet switches supporting the Kinetic approach. The Ethernet switches communicate directly with the ToR switch using 10 GE. Storage drives look the same as those used with the traditional architecture and have the same connectors, but the connectors communicate using 1G or 2.5G Ethernet instead of 3G or 6G SAS.

Total cost of ownership is reduced dramatically by using the Kinetic approach instead of the traditional approach. The savings result from the elimination of the storage server, increased efficiencies from eliminating multiple protocol and processing layers, and by using Ethernet rather than SAS for communications. While SAS is a specialized protocol supported by a small number of manufacturers, Ethernet benefits from the economies of scale of a broad ecosystem supported by a large number of manufacturers. Another advantage of the Seagate Kinetic Open Storage Platform is that because both compute and storage devices use Ethernet for communications, they can be mixed as needed in the same rack for potential additional cost savings.

Figure 7 Virtualization Becomes Much Easier



The ability to mix compute and storage drives in the same rack also offers more flexibility. For example, it is well suited to supporting applications that require a relatively large amount of compute resources but a relatively small amount of storage (or vice versa).

Storage Virtualization

By moving to all-Ethernet communications, the Seagate Kinetic Open Storage Platform leverages the interoperability that has been a cornerstone of Ethernet development. As a result, it has the potential to enable storage virtualization in a manner that supports devices from multiple manufacturers. Currently, a few vendors offer storage virtualization, enabling a single virtual machine to span multiple devices, but products supporting that capability have not interoperated with products from other vendors until now. Properly designed, the Ethernet switch replacing the SAS expander has the ability to reassign virtual storage to different physical storage resources in real-time without interrupting communications, and to do so in a manner that interoperates with devices from other vendors. This has been common practice for compute server virtualization for a number of years, but not universally accomplished for storage. Critical capabilities for the switch include quality of service and VLAN management and the ability to move VLAN and TCP/IP addresses from one storage device to another without disrupting communications.

Enabling open environment storage virtualization would give storage the same level of flexibility that already exists for compute resources, and should provide an impetus for enterprises to move more applications to the data center cloud. To fully support the vision for on-demand cloud-based IT resources and applications, there is one other piece that needs to be put in place. The connectivity and networking capacity between the virtual compute machines and virtualized storage in the data center must be able to be adjusted on demand.

Fortunately, the move toward software defined networking (SDN) for the now completely Ethernet-based Data Center network should make that possible. SDN separates the control plane from individual network devices, enabling a centralized control plane to first create a flat interconnect Ethernet "fabric" from aggregation/leaf and core/spine switches, and then slice that fabric into multiple virtual networks interconnecting the proper virtual compute and store resources as needed. That makes it well suited to supporting on-demand changes to bandwidth, quality of service parameters, and so on. As the Seagate Kinetic Open Storage Platform and SDN take root, data center operators should be able to realize the full vision for on-demand IT in which the network, computing, and storage are fully virtualized. Achieving this will mean linking network, computing, and storage resources end-to-end through open standard Ethernet communications.

Considerations for Storage System Designers

As storage system designers prepare to implement the Seagate Kinetic Open Storage Platform, they should consider several factors:

- Power is always important in big data centers. Air flow can be a problem in the storage enclosures, and space is not always sufficient to allow for cooling fins. Designers should look for Ethernet switching integrated circuits with minimal power requirements.
- The Ethernet switch should have enough SGMII ports to support the number of drives (typically 60–90+) in a storage enclosure, and ideally those interfaces should have the ability to support double-speed SGMII for future capacity upgrades.
- The number of traces across the mid-plane between the Ethernet aggregation cards and the storage drive is sometimes limited. In that case, designers should seek a solution that gives them the ability to multiplex the interface to higher speeds (such as 2.5G, 5G, or 10G) and de-multiplex on the other side again before connecting to the Kinetic drives.
- Storage enclosures typically come in various capacities and sizes — from about twenty-four drives for smaller systems up to ninety-six drives in larger storage enclosures. Ideally, Ethernet switching capacity should scale accordingly, without requiring a different chip set and different software. Hardware stacking capability of the Ethernet switches is a good way to address this, and this approach scales power, cost, and port count linearly with the number of drives supported.
- The Ethernet switch should have an embedded CPU so that the entire operating system and protocols of the switch (including stacking software) can be run directly on the chip, without requiring another external processor.
- For advanced data center architectures, the aggregation switch on the storage enclosure could be viewed as a port aggregator into the large ToR switch. That means the aggregation switch should be manageable as an extension of the ToR switch. This is similar to how virtual hypervisor "vSwitches" are managed through open APIs like JSON-RPC from the ToR switch.
- To support virtualization, designers should look for flexible VLAN management in the switches, and should be able to move VLAN from one drive to another without significant traffic disruption.
- Future applications of Ethernet-based storage beyond object-based storage will likely involve East-West in addition to North-South traffic, and that will require the switch to support advanced QoS capabilities. Advanced packet classifications should also be supported to enable fine-grain policy control.

Summary

The new object-oriented storage architecture migrates storage networks from overly-complex and costly communications protocols and architectures to a simpler, more flexible approach based on Ethernet. With the new architecture as exemplified by Seagate's Kinetic Open Storage Platform, the traditional FC SAN switch and the storage server are no longer required in the Data Center. Storage drives have Ethernet rather than SAS interfaces, and Ethernet switches take the place of SAS/SATA expanders. SAS/SATA, iSCSI, FC, and FCoE are all replaced with Ethernet. This enables data centers to be designed for more efficient operations and yields a substantial reduction in total cost of ownership.

The migration of storage networks to Ethernet also will facilitate the virtualization of storage in an open environment – a capability that today exists only in a single-vendor environment. Data center compute resources already use Ethernet communications. By also migrating storage to Ethernet, the new architecture completes the migration of the data center to Ethernet, which should help accelerate the further shift of IT applications and resources to the data center cloud.

Microsemi Corporate Headquarters
One Enterprise, Aliso Viejo,
CA 92656 USA
Within the USA: +1 (800) 713-4113
Outside the USA: +1 (949) 380-6100
Sales: +1 (949) 380-6136
Fax: +1 (949) 215-4996
E-mail: sales.support@microsemi.com
www.microsemi.com

©2014–2016 Microsemi Corporation. All rights reserved. Microsemi and the Microsemi logo are registered trademarks of Microsemi Corporation. All other trademarks and service marks are the property of their respective owners.

Microsemi makes no warranty, representation, or guarantee regarding the information contained herein or the suitability of its products and services for any particular purpose, nor does Microsemi assume any liability whatsoever arising out of the application or use of any product or circuit. The products sold hereunder and any other products sold by Microsemi have been subject to limited testing and should not be used in conjunction with mission-critical equipment or applications. Any performance specifications are believed to be reliable but are not verified, and Buyer must conduct and complete all performance and other testing of the products, alone and together with, or installed in, any end-products. Buyer shall not rely on any data and performance specifications or parameters provided by Microsemi. It is the Buyer's responsibility to independently determine suitability of any products and to test and verify the same. The information provided by Microsemi hereunder is provided "as is, where is" and with all faults, and the entire risk associated with such information is entirely with the Buyer. Microsemi does not grant, explicitly or implicitly, to any party any patent rights, licenses, or any other IP rights, whether with regard to such information itself or anything described by such information. Information provided in this document is proprietary to Microsemi, and Microsemi reserves the right to make any changes to the information in this document or to any products and services at any time without notice.

About Microsemi
Microsemi Corporation (Nasdaq: MSCC) offers a comprehensive portfolio of semiconductor and system solutions for aerospace & defense, communications, data center and industrial markets. Products include high-performance and radiation-hardened analog mixed-signal integrated circuits, FPGAs, SoCs and ASICs; power management products; timing and synchronization devices and precise time solutions, setting the world's standard for time; voice processing devices; RF solutions; discrete components; enterprise storage and communication solutions; security technologies and scalable anti-tamper products; Ethernet solutions; Power-over-Ethernet ICs and midspans; as well as custom design capabilities and services. Microsemi is headquartered in Aliso Viejo, Calif., and has approximately 4,800 employees globally. Learn more at www.microsemi.com.