Intel® Microarchitecture and 10 Gigabit Ethernet Transforming the Data Center

Abstract

A transformation is taking place in the data center. With more powerful compute resources available, IT managers are using virtualization to consolidate more applications onto fewer physical servers for greater efficiency and better resource utilization. The result is a more efficient, flexible data center that can better adapt to changing needs. Intelligent and powerful servers are a key element of this new data center. Another is a flexible and reliable network fabric that meets the bandwidth needs of demanding applications and simplifies the network by consolidating multiple types of data center traffic. This paper describes how new server and networking products from Intel combine to provide the foundation of this new data center model.

Table of Contents

Introduction

IT managers are turning increasingly to server virtualization and consolidation to increase data center efficiency and improve total cost of ownership (TCO). Consolidating the tasks of multiple servers onto a single server running multiple virtual machines (VMs) saves power and space costs and reduces data center sprawl, all critical issues to cost-conscious IT departments. As virtualized server deployments increase, infrastructure elements are evolving to keep pace, and these collective improvements are leading to a data center transformation that introduces new efficiencies as well as new challenges.

Today, new, more powerful processors allow servers to host more VMs than ever before, facilitating greater resource consolidation. This, however, drives the need for more network bandwidth per server, as VMs still require adequate network connectivity. Server consolidation also requires greater network storage to support the virtualized servers' data needs–including backup, live migration, and disaster recovery. Rather than allowing these elements to grow at increasing rates, however, industry leaders are developing technologies to make them more efficient and flexible. These new technologies are driving IT toward a new data center model, a model in which the various infrastructure elements, including servers, network devices, and storage devices scale performance dynamically to meet constantly changing compute needs.

The new generation of servers based on the Intel® Xeon processor 5500 series provides the processing power and adaptability needed to drive more powerful applications and support growing VM deployment without increasing power or space requirements in a dynamic data center. As the number of VMs increases, however, a server's I/O needs increase accordingly. Platforms based on this new architecture provide superb I/O scalability through higher processing power, a new local memory architecture, and the faster PCI Express* 2.0 I/O interface bus.

The Intel® 82599 10 Gigabit Ethernet Controller is designed specifically to take advantage of these platform enhancements. And with optimizations for I/O virtualization and unified networking, this controller is uniquely positioned to satisfy the needs of the dynamic data center and provide a flexible, reliable network connection capable of carrying multiple traffic types and adapting to changing network conditions.

Rising I/O Demands in the Data Center

To put the importance of the Intel Xeon processor 5500 series and the Intel 82599 10 Gigabit Ethernet Controller into perspective, it is helpful to first consider the major evolutionary stages of I/O in the data center. This begins with the traditional data center configuration shown in Figure 1(a), which follows the old model of adding a new server each time a new application is added.

This one-application-per-server approach quickly generates server sprawl and frequently leaves a significant number of servers underutilized, even idle at times. It also leads to considerable I/O infrastructure growth, with at least two Gigabit Ethernet (GbE) connections per server for LAN access and one Fibre Channel (FC) host bus adapter (HBA) per server for Storage Area Network (SAN) access. The proliferation of servers, interconnects, and cabling in the traditional data center is costly in terms of equipment expense, power usage, floor space usage, and management complexity.

Data center evolution to virtualization, as shown in Figure 1(b), reduces server sprawl and its associated costs by allowing multiple VMs to run independently on a single physical server. Each VM can run one or more different applications and can even run under different operating systems. In the case of Figure 1(b), virtualization replaces the four physical servers of Figure 1(a) with four VMs and consolidates their workloads onto one physical server. This reduction of server count saves costs, reduces data center power consumption, and conserves floor space.

As server processing power grows, the number of VMs on a single server can be increased. However, more VMs require more I/O capability, and simply adding network interface cards (NICs) a gigabit at a time quickly leads to another problem—interface, switch, and cabling sprawl. This is leading IT managers to consolidate I/O connectivity onto fewer, higher-bandwidth NICs, specifically 10GbE server adapters.

(a) Traditional Data Center

(b) Server Virtualization
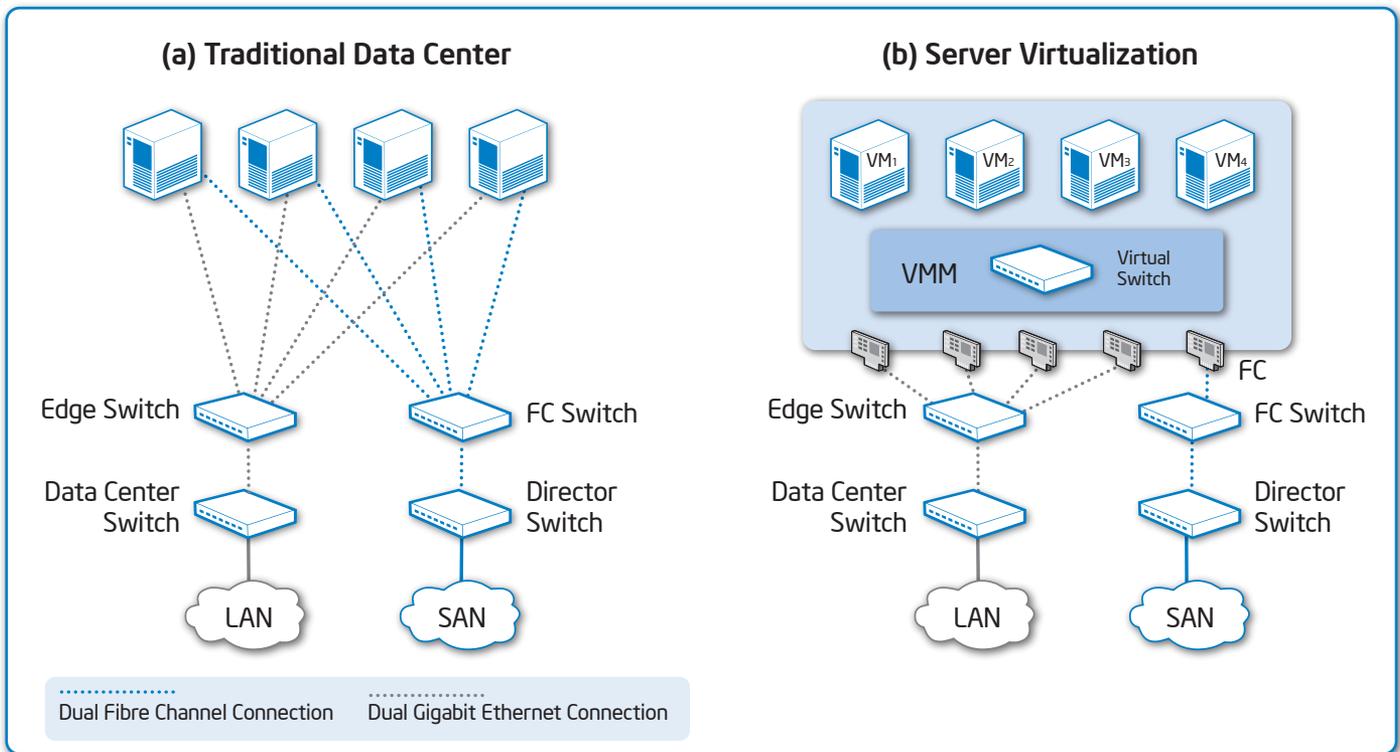
Figure 1. The applications of multiple servers in a traditional data center (a) can be consolidated onto a single physical server by virtualization, (b) where each application can run independently on a virtual machine (VM) under the control of a virtual machine monitor (VMM).

The Benefits of 10GbE

The benefit of transitioning the data center to 10GbE connectivity is apparent in Figure 2(a). Compared to the multiple GbE NICs in Figure 1(b), 10GbE connectivity provides greater bandwidth while reducing the number of server adapters and switch ports and the amount of cabling required for LAN connectivity as shown in Figure 2(a). The reduction in equipment and cabling provided by 10GbE simplifies the network and can significantly reduce TCO as compared to GbE connectivity. Also, in addition to providing the higher bandwidth demanded by many applications today, 10GbE is the foundation for the next evolutionary stage in networking, in which Ethernet will extend its reach in the data center.

Increased deployment of virtualized servers also drives the need for greater connectivity to the SAN. With many VMs inhabiting a server, it is not practical for each VM to boot from a local hard drive. Instead, operating system and application data is contained in networked storage devices, typically Network Attached Storage (NAS) servers or a Fibre Channel or Internet Small Computer System Interface (iSCSI) SAN disk array. Remote storage is also critical for VM migration between physical servers and disaster recovery, where a VM needs to be brought up on a new physical server.
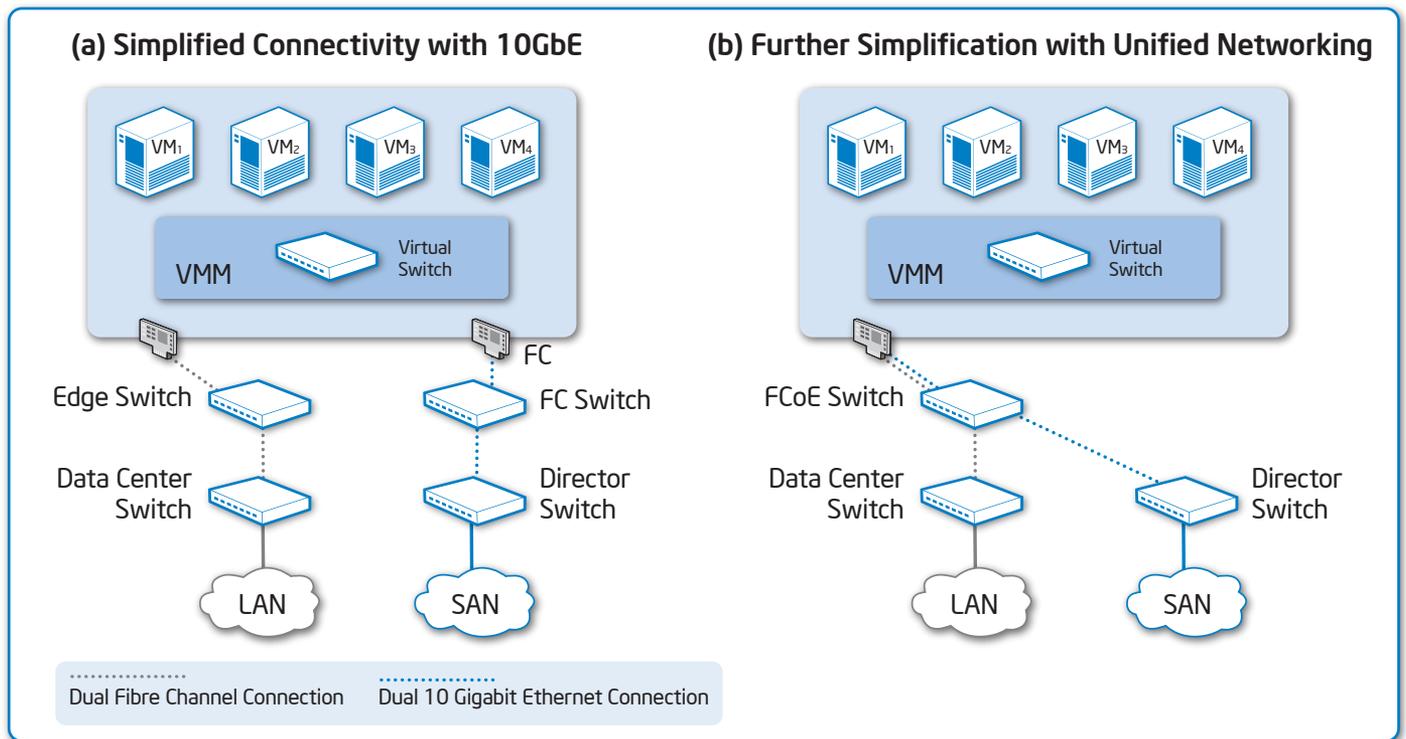
Figure 2. Migrating to 10GbE (a) eliminates the clutter of GbE connectivity, and (b) further simplification is possible by converging LAN and SAN traffic onto a single 10GbE fabric using Fibre Channel over Ethernet.

Figure 2(b) shows an example of unified networking or multiple traffic types on Ethernet. In Figure 2(b), a Fibre Channel over Ethernet (FCoE) capable 10GbE server adapter and an FCoE switch are used to combine LAN and SAN traffic onto a single 10GbE network fabric. The result is a further reduction in network infrastructure equipment and an easier-to-manage network with ample bandwidth for both storage traffic and network traffic to and from the virtualized server.

In the past, the cost of migrating to 10GbE was a legitimate concern, but that is no longer the case. While a 10GbE server adapter does cost more than a GbE server adapter, the total bandwidth cost is lower, and as volume ramps, the cost of 10GbE ports is projected to drop at a greater than 30 percent rate over five years. Also, a single 10GbE server adapter uses only one PCI Express slot versus the multiple slots consumed by multiple GbE adapters. Adapters based on the Intel

82599 10 Gigabit Ethernet Controller include support for existing Ethernet storage applications, such as iSCSI and NAS, as well as the emerging FCoE standard, eliminating the need for separate, dedicated storage adapters.

The price of 10GbE adapters and switches is also decreasing; new standards for copper-based 10GbE networks will result in even lower infrastructure costs. These standards include:

- **SFP+ Direct Attach**. Uses a twin-ax cable assembly for a low-power, low-latency, and low-cost connection of up to 7 meters.

- **10GBASE-T**. Allows 10GbE connectivity to 55 meters over Category 6 unshielded twisted pair or to 100 meters over Augmented Category 6 cabling.

New Server Architecture Optimizes I/O Performance

10GbE as a universal network fabric provides significant advantages for creating a flexible, reliable, and agile data center. However, achieving optimum performance in the dynamic data center also requires a platform that is optimized for greater I/O scalability. Servers based on the Intel Xeon processor 5500 series provide the architectural elements that allow for new levels of 10GbE scalability.

Figure 3 illustrates some of the key differences between this new architecture and its predecessor.

Faster Memory and Caches

The most noticeable difference between the two architectures of Figure 3 is that the new architecture has an integrated memory controller and significantly higher memory bandwidth. In particular, the Intel Xeon processor 5500 series-based servers support faster DDR3 memory, as opposed to the DDR2 memory of previous generations of FBDIMMs. This enhancement provides a peak memory bandwidth available per socket of about 32 gigabytes per second (GB/s), which is much higher than what is available on previous platforms for the entire system. With 32 GB available per CPU socket, a dual-processor system has close to 64 GB/s of peak memory bandwidth, which is almost 3x more read and 6x more write bandwidth compared to the previous generation.

Other memory improvements include a fifty percent increase of the level-two cache (L2C) and the addition of a large last-level cache (LLC). The inclusive LLC enables dynamic and efficient allocation of shared cache to all four cores in each processor, increasing performance while reducing traffic to the processor cores by eliminating unnecessary core snoops to memory.

Faster Interconnect Architecture

Another key difference between the two generations shown in Figure 3 is that the Front Side Bus (FSB) has been replaced by the Intel® QuickPath Interconnect (Intel® QPI). Intel QPI provides higher bandwidth, and the new architecture also includes an Intel QPI interface from processor to processor, providing additional coherent bus bandwidth.

On previous generation platforms, all memory reads and writes from the CPUs had to traverse the FSB. With the Intel® QuickPath Architecture, local memory reads and writes do not require an Intel QPI traversal, except for snoops. This reduces the amount of bandwidth consumed on Intel QPI. The presence of an additional Intel QPI interface provides additional bandwidth for I/O-to-memory and CPU-to-CPU data transfers.

Faster PCI Express Bus

To support higher-bandwidth I/O and the ability to scale across multi-port 10GbE, the new architecture in Figure 3(b) uses a PCI Express 2.0 (5.0GT/s)—PCIe2—bus from the I/O Hub to the network interface card (NIC). PCIe2 provides a greater transfer rate, to 5.0 Giga transfers per second (5.0GT/s), twice the 2.5GT/s rate of first-generation PCIe. With these new platforms, there will generally be more PCIe2 lanes made available, which can lead to more slots or wider slots. For example, a x8 PCIe2 slot is capable of a theoretical peak of 4 GB/s per direction, and with more lanes, better I/O scaling is possible, providing the ability to add more ports for superb scaling across multiple 10GbE Server Adapter ports.
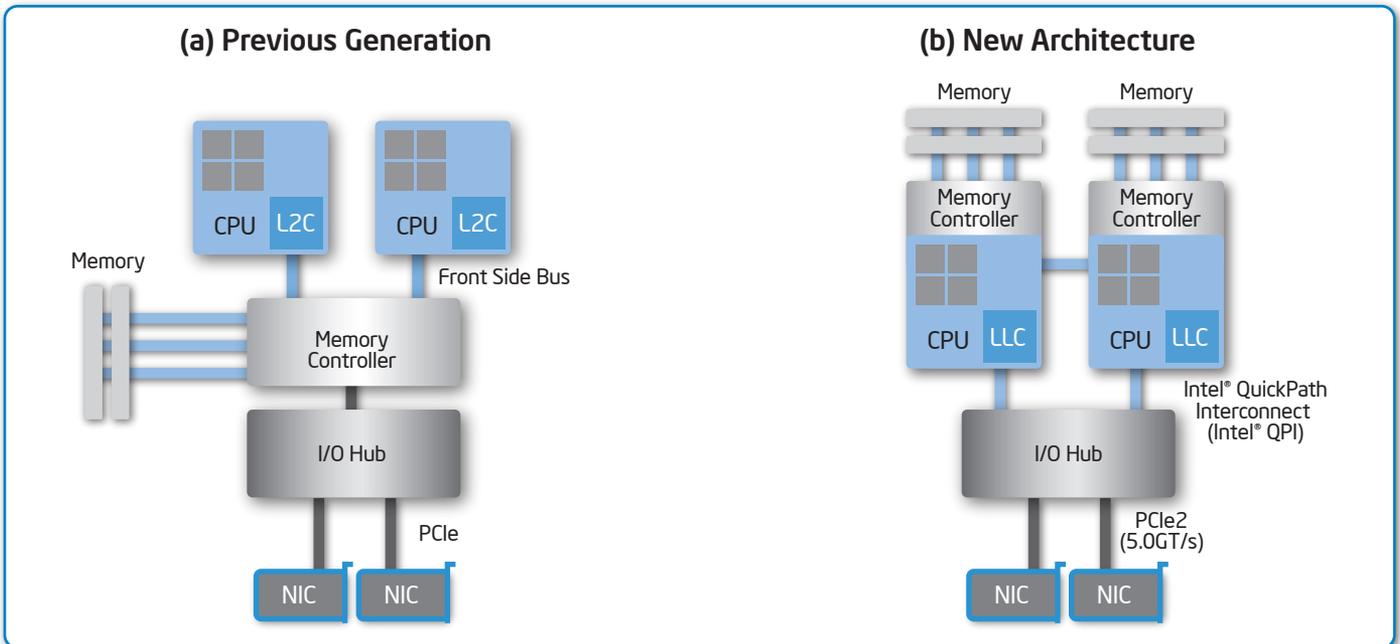

Figure 3. Comparison of previous generation architecture (a) to the new Intel® Microarchitecture, codenamed Nehalem, (b) used in the Intel® Xeon® processor 5500 series.

Figure 4 compares network throughput scalability for an Intel Xeon processor 5500 series-based server with an adapter based on the Intel 82599 10 Gigabit Ethernet Controller versus a server based on the previous generation Intel® Xeon® processor and previous generation Intel 10 Gigabit Ethernet controller.

Notice in particular that throughput on the older platform does not scale beyond two 10GbE ports and approximately 17 Gbps. This is due primarily to memory bandwidth limitations. Conversely, the Intel Xeon 5500 series processor-based server's architecture allows it to scale up to four 10GbE ports and roughly 50 Gbps. This level of throughput scalability is needed to support emerging high-bandwidth applications in today's evolving virtualized data centers.
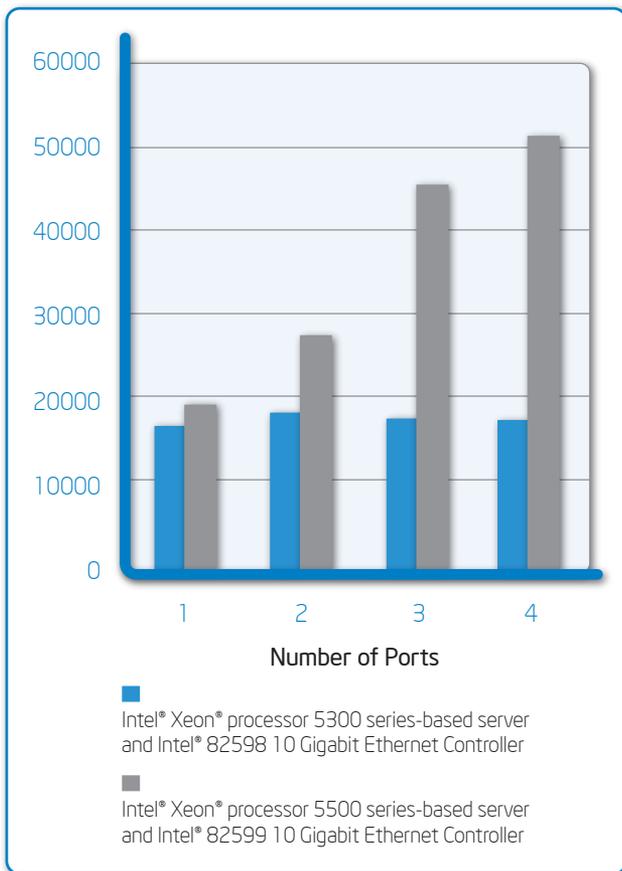
Figure 4. The Intel® Xeon® processor 5500 series-based server and Intel 82599 10 Gigabit Ethernet Controller scale superbly across multiple 10GbE ports as compared to the previous generation server and 10GbE controller.[1]

New 10 Gigabit Controller

The Intel 82599 10 Gigabit Ethernet Controller is Intel's third-generation 10GbE controller and builds on Intel's more than 25 years of experience building quality Ethernet products. This new controller takes advantage of the architectural improvements of the Intel Xeon processor 5500 series-based platforms and aids data center transformation by delivering the bandwidth and features needed to provide a reliable, unified 10GbE network connection.

Outstanding Performance

Broad operating system support, including Microsoft Windows*, Linux*, and VMware ESX*, and thorough compatibility testing with leading switch vendors, provides ease of installation and configuration. The Intel 82599 10 Gigabit Ethernet Controller also includes performance-improving features designed for multicore processors for exceptional performance. MSI-X and Receive-Side Scaling (RSS) use the multiple hardware queues in the controller to load-balance interrupts and I/O data processing across multiple cores. Intel® Ethernet Flow Director is a new feature that lowers latency and improves CPU utilization by distributing TCP flows to CPU cores.

Best for I/O Virtualization

The Intel 82599 10 Gigabit Ethernet Controller includes Intel® Virtualization Technology for Connectivity (Intel® VT-c) to deliver outstanding I/O performance in virtualized environments. Intel VT-c includes hardware optimizations that help reduce I/O bottlenecks and improve server performance. Intel VT-c consists of two components: Virtual Machine Device Queues (VMDq) and Virtual Machine Direct Connect (VMDc). VMDq improves data processing by offloading the network traffic sorting and queuing functionality from the Virtual Machine Monitor (VMM) to the Ethernet controller. VMDc provides direct connectivity to VMs to provide near-native performance and greater VM scalability. VMDc also enables VM migration between physical servers, providing greater flexibility and mobility.

Storage over Ethernet Optimizations

The Intel 82599 10 Gigabit Ethernet Controller supports iSCSI, NAS and FCoE to carry storage traffic over Ethernet. In order to meet SAN requirements for guaranteed packet delivery, the Intel 82599 10 Gigabit Ethernet Controller implements Data Center Bridging, a set of industry standards that delivers end-to-end congestion notification and quality of service throughout the network. The Intel 82599 10 Gigabit Ethernet Controller also accelerates iSCSI traffic by implementing key stateless offloads such as TCP Segmentation Offload and Receive-Side Coalescing. It also supports the trusted, native iSCSI initiators in Microsoft, Linux, and VMware operating systems and provides a robust iSCSI remote boot implementation. Further, the controller delivers a high-performance FCoE solution that offloads the main data paths for I/O read and write commands to improve throughput. The Intel 82599 10 Gigabit Ethernet Controller also greatly reduces CPU processing on FCoE receive traffic by eliminating a data copy through Direct Data placement implementation.

As the world leader in silicon innovation, Intel is refining and developing technologies, products, and initiatives that will continually transform the data center.

For More Information

Intel® 10 Gigabit Ethernet Server Adapters:
www.intel.com/go/10GbE

Intel® Virtualization Technology for Connectivity:
www.intel.com/go/vtc

Intel® Virtualization Technology:
www.intel.com/technology/virtualization/index.htm

New Generation Intel® Microarchitecture, code-named Nehalem:
www.intel.com/technology/architecture-silicon/next-gen/

Intel® Xeon® processor 5500 series:
www.intel.com/products/processor/xeon5000/index.htm