



White Paper
Intel® I/O Acceleration
Technology

Accelerating High-Speed Networking with Intel® I/O Acceleration Technology

The emergence of multi-Gigabit Ethernet allows data centers to adapt to the increasing bandwidth requirements of enterprise IT. To take full advantage of this network capacity, data center managers must consider the impact of high-traffic volume on server resources. A compelling way of efficiently translating high bandwidth into increased throughput and enhanced quality of service is to take advantage of Intel® I/O Acceleration Technology (Intel® I/OAT), now available in the new Dual-Core and Quad-Core Intel® Xeon® processor-based server platforms. Intel I/OAT moves data more efficiently through these servers for fast, scalable, and reliable networking. Additionally, it provides network acceleration that scales seamlessly across multiple Ethernet ports, while providing a safe and flexible choice for IT managers due to its tight integration into popular operating systems.



Introduction

Business success is becoming increasingly dependent on the rapid transfer, processing, compilation, and storage of data. To improve data transfers to and from applications, IT managers are investing in new networking and storage infrastructure to achieve higher performance. However, network I/O bottlenecks have emerged as the key IT challenge in realizing full value from server investments.

Until recently, the real nature and extent of I/O bottlenecks was not thoroughly understood. Most network issues could be resolved with faster servers, higher bandwidth network interface cards (NICs), and various networking techniques such as network segmentation. However, the volume of network traffic begun to outpace server capacity to manage that data. This increase in network traffic is due to the confluence of a variety of market trends. These trends include:

- Convergence of data fabrics to Ethernet
- Server consolidation and virtualization, which is increasing server port densities and maximizing per-port bandwidth requirements
- Clusters and grids—rather than large servers—are becoming the leading way of creating high-performance computing resources
- Increased adoption of enterprise audio and video resources (VOIP, video conferencing) and real-time data acquisition are contributing to network congestion
- Increased adoption of network-attached storage, versus direct-attach storage, to meet enterprise back-up and recovery needs
- Deployment of high-density blade servers to accommodate enterprise-computing requirements

The increased reliance on networked data and networked compute resources results in a need to manage an ever-increasing data load. This increased load threatens to outpace server-processing capabilities, as demonstrated by the marginal gains from traditional performance enhancements (for example, increasing network bandwidth and adding servers). As a result, IT managers are now asking, “After investing in a 10X improvement in network bandwidth, why aren’t we seeing comparable improvements in application response time and reliability?”

To answer this question, Intel research and development teams evaluated the entire server architecture to identify the I/O bottlenecks and determine their nature and impact on network performance. Out of this investigation came a new technology called Intel® I/O Acceleration Technology* (Intel® I/OAT). This white paper discusses the platform-wide I/O bottlenecks found through Intel research and how Intel I/OAT resolves those issues for accelerating high-speed networking.

Finding the Real I/O Bottlenecks

Intel research and development teams found the real I/O bottlenecks when they examined the entire flow of a data packet as it is received, processed, and sent out by the server. Figure 1 illustrates this flow. The following numbered descriptions correspond to the circled numbers in the illustration:

1. A client sends a request in the form of TCP/IP data packets that the server receives through the network interface card (NIC). The data packet contains TCP header information that includes packet identification and routing information as well as the actual data payload relating to the client request.
2. The server processes the TCP/IP packets and routes the payload to the designated application. This processing includes protocol computations involving the TCP/IP stack, multiple server memory accesses for packet descriptors and payload moves, and various other system overhead activities (for example, interrupt handling and buffer management).
3. The application acknowledges the client request and recognizes that it needs data from storage to respond to the request.

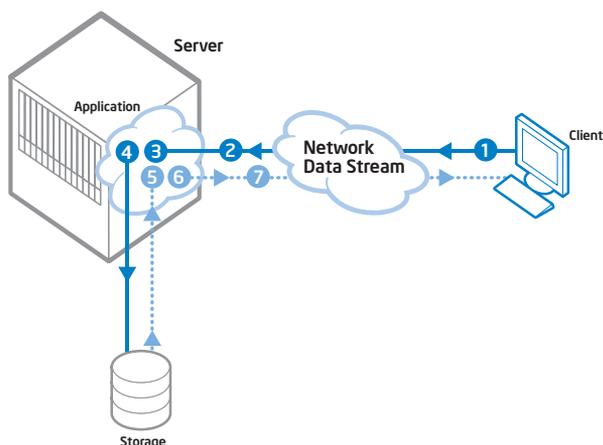
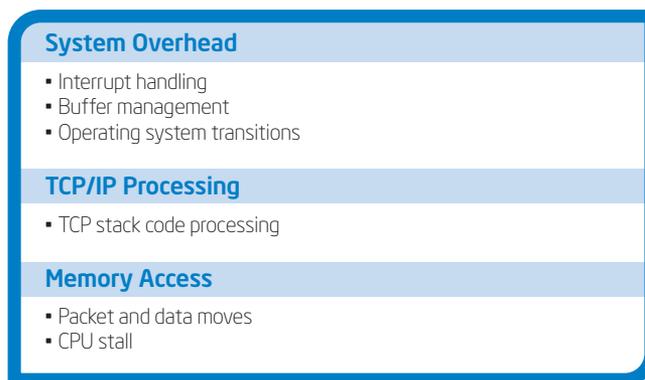


Figure 1. Data paths to and from application. Server overhead and response latency occurs throughout the request-response data path. These overheads and latencies include processing incoming request TCP/IP packets, routing packet payload data to the application, fetching stored information, and reprocessing responses into TCP/IP packets for routing back to the requesting client.

4. The application accesses storage to obtain the necessary data to satisfy the client request.
5. Storage returns the requested data to the application.
6. The application completes processing of the client request using the additional data received from storage.
7. The server routes the response back through the network connection to be sent as TCP/IP packets to the client.

The above packet data flow has remained largely unchanged for more than a decade. For the networking requirements of today, the result is significant system overhead, excessive memory accesses, and inefficient TCP/IP processing.

Figure 2 represents the entire network I/O overhead picture. It is important to note, however, that the system overhead, TCP/IP processing, and memory access categories of overhead are not proportional. In fact, as discussed later, the amount of CPU usage for each category varies according to application I/O packet size.



Source: Intel research on the Linux* operating system

Figure 2. Network I/O processing tasks. Server network I/O processing tasks fall into three major overhead categories, each varying as a percentage of total overhead according to TCP/IP packet size.

It is also important to note that some technologies already exist for mitigating some CPU overhead issues. For example, all Intel® PRO Server Adapters and Intel® PRO Network Connections (LAN on motherboard or LOM) include advanced features designed to reduce CPU usage.

These features include:

- Interrupt moderation
- TCP checksum offload
- TCP segmentation
- Large send offload

Their implementation does not impair compatibility with other network components or require special management or modification of operating systems or applications.

Other approaches exist that claim to address system overhead by offloading even more processing to the NIC. These approaches include the TCP offload engine (TOE) and remote direct memory access (RDMA).

TOE uses a specialized and dedicated processor on the NIC to handle some of the packet protocol processing. It does not fully address the other performance bottlenecks shown in Figure 2. In fact, for small packet sizes and short-duration connections, TOE may be of very limited benefit in addressing the overall I/O performance problem. Additionally, given its offloading (that is, copying) of the network stack to a fixed-speed microcontroller (the offload engine), not only is performance limited to the speed of the micro-controller, but there is also a risk that key network capabilities, such as adapter teaming and failover, may not function in a TOE environment.

As for RDMA-enabled NICs (RNICs), the RDMA protocol supports direct placement of packet payload data into an application's memory space. This addresses the memory access bottleneck category by reducing data movement overhead for the CPU. However, the RDMA protocol requires significant changes to the network stack and can require changes to application software that utilizes the network.

Defining the Worst I/O Bottleneck

The limitations of existing I/O acceleration solutions became even clearer as Intel research and development teams began to quantify each category under different conditions. Figure 3 summarizes these results for various application I/O sizes and their impact by task category on percent of CPU utilization.

Notice in Figure 3 that CPU usage by TCP/IP processing is nearly constant across I/O sizes ranging from 2KB to 64KB. Although TCP/IP processing is a significant bottleneck, it is not the most significant bottleneck. Memory accesses account for more CPU usage in all cases than TCP/IP processing, and system overhead is the worst I/O bottleneck for application I/O sizes below 8KB.

As stated earlier and indicated by the data in Figure 3, TOE and RDMA do not address the entire I/O bottleneck issue. What is needed is a system-wide solution that can fit anywhere in the enterprise computing hierarchy without requiring modification of application software and which provides acceleration benefits for all three network bottlenecks. Intel I/OAT, now available on the new Dual-Core and Quad-Core Intel® Xeon® processor-based platforms, is exactly that kind of solution.

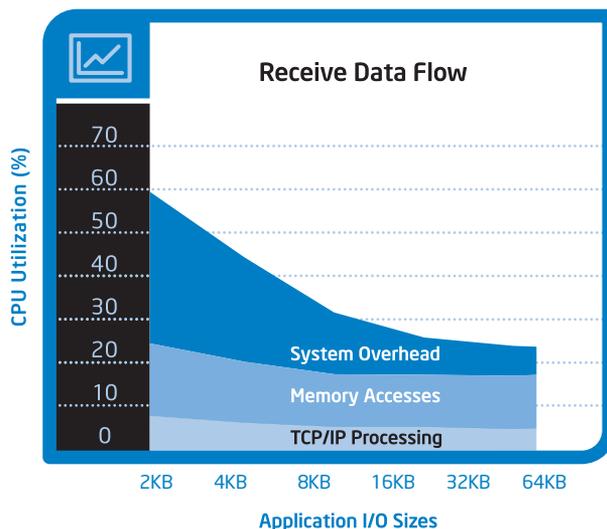


Figure 3. CPU utilization varies according to I/O size. TCP/IP processing is fairly constant and tends to be a smaller part of CPU utilization compared to system overhead.

Intel® I/OAT—The System-Wide Solution

Intel I/OAT address all three server I/O bottlenecks (illustrated in Figures 2 and 3) by providing fast, scalable, and reliable networking. In addition, it provides network acceleration that scales seamlessly across multiple Ethernet ports, and it is a safe and flexible choice for IT managers due to its tight integration into popular operating systems.

The system-wide network I/O acceleration technologies applied by Intel I/OAT are shown in Figure 4 and include:

- **Parallel Processing of TCP and Memory Functions.** Lowers system overhead and improves the efficiency of TCP stack processing by using the abilities of the CPU to execute multiple instructions per clock, pre-fetch TCP/IP header information into cache, and perform other data movement operations in parallel.
- **Affinitized Data Flows.** Partitions the Network Stack Processing dynamically across multiple physical or logical CPUs. This allows CPU cycles to be allocated to the application for faster execution.
- **Asynchronous Low-Cost Copy.** Intel® Quick Data Technology provides enhanced data movement, allowing payload data copies from the NIC buffer in system memory to the application buffer with far fewer CPU cycles, returning the saved CPU cycles to productive application workloads.

- **Improved TCP/IP Processing with Optimized TCP/IP Stack.** Implements separate packet data and control paths to optimize processing of the packet header from the packet payload. This and other stack-related enhancements reduce protocol processing cycles.

Because Intel I/OAT enhances performance while keeping all processing of the operating system's TCP stack on the Intel Xeon processor and the state information for TCP connections within the server's system memory, the technology is said to be "stateless." This is as opposed to stateful offload technologies, such as TOE. As a stateless technology, Intel I/OAT retains use of the system processors and protocols as the principal engines for handling network traffic.

Additionally, Intel I/OAT is used throughout the platform to increase CPU efficiency by reducing bottlenecks across most application I/O sizes. Because Intel I/OAT is tightly integrated into popular operating systems, it ensures full compatibility with critical network configurations such as adapter teaming and link aggregation. As a result, Intel I/OAT provides a fast, scalable, and reliable network acceleration solution with significant performance advantages over prior system implementations and technologies.

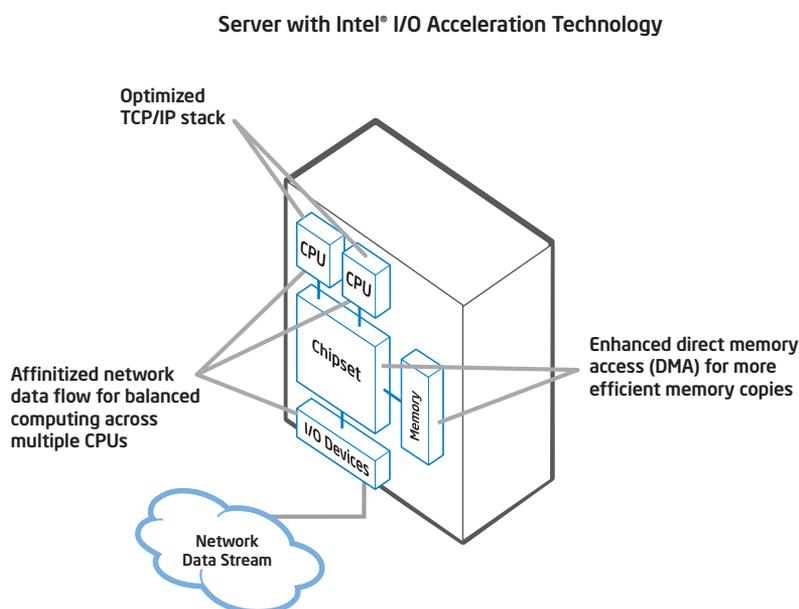


Figure 4. Intel® I/OAT performance enhancements. Intel® I/OAT implements server-wide performance enhancements in all three major server components to ensure that data gets to and from applications consistently faster and with greater reliability.

The Performance Advantage of Intel® I/OAT

The Intel I/OAT platform-level approach to improving network performance has been verified by extensive testing. Some of these results are summarized in Figure 5.

The Intel I/OAT performance tests were conducted for both Linux* and Microsoft Windows* operating systems using two Dual-Core and Quad-Core Intel Xeon processor-based servers tested across multiple Gigabit Ethernet (GbE) NIC ports (two to eight GbE ports) as represented by the X-axis. One of the servers was an Intel® E7520-based platform without the benefit of Intel I/OAT. The other was an Intel® E5000 series server using the new Dual-Core Intel Xeon processor with Intel I/OAT enabled. In the test examples of Figure 5, the graphs represent both CPU utilization percentages (the lines) and the corresponding network throughput performance (the vertical bars). Both systems underwent identical stress tests.

The left graph in Figure 5 summarizes the results of an Intel Xeon processor-based server running a Linux operating system. Notice that an Intel I/OAT-enabled platform, running Linux and using eight

GbE ports, achieved a CPU utilization improvement of over 40 percent versus a non-Intel I/OAT-enabled platform. Additionally, this same platform achieved almost twice the network throughput as the non-Intel I/OAT-enabled platform operating under the same conditions.

Similarly, with Microsoft Windows Server 2003* (right graph in Figure 5), the network throughput of the Intel I/OAT-enabled platform nearly doubled for eight GbE ports. In this test, the Intel E7520-based platform was incapable of generating CPU utilization data beyond four ports because, without the benefit of receive-side scaling, the server directs all network traffic to Processor 0, saturating the processor and limiting the system's ability to report data. However, the new Dual-Core and Quad-Core Intel Xeon processor-based platform with Intel I/OAT balanced the workload across the processors and never reached 70-percent CPU utilization, even at eight GbE ports.

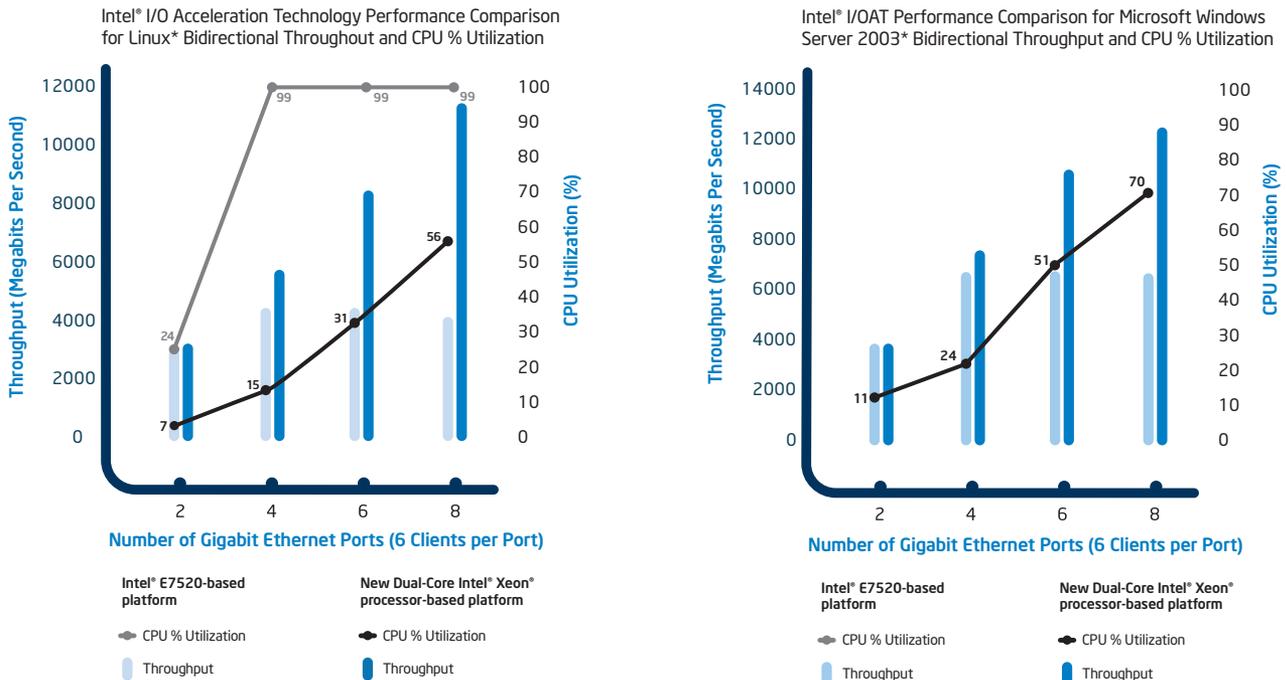


Figure 5. Network-performance comparisons for platforms with and without Intel® I/OAT. Compared to previous processors, the new Dual-Core Intel® Xeon® processor with Intel® I/OAT provides superior performance in terms of both higher throughput and reduced percentage of CPU utilization.

Conclusion

In summary, Intel I/OAT improves network application responsiveness by moving network data more efficiently through Dual-Core and Quad-Core Intel Xeon processor-based servers to provide fast, scalable, and reliable network acceleration for the majority of today's data center environments. This translates to the following significant areas of benefit for IT data centers:

- **Fast.** A primary benefit of Intel I/OAT is that it significantly reduces CPU overhead, freeing resources for more critical compute tasks. Intel I/OAT uses server processors more efficiently by leveraging architectural improvements within the CPU, chipset (Intel Quick Data Technology), network controller, and firmware to minimize performance-limiting bottlenecks.
- **Scalable.** Intel I/OAT scales seamlessly across multiple GbE ports (up to eight ports), and can scale up to 10GbE, while maintaining power and thermal characteristics similar to those of a standard gigabit network adapter.
- **Reliable.** Intel I/OAT is a safe and flexible choice because it is tightly integrated into popular operating systems such as Microsoft Windows Server 2003 and Linux, avoiding the support risks associated with relying on third-party hardware vendors for network stack updates. Intel I/OAT also preserves critical network capabilities, such as teaming and failover, by maintaining control of the network stack processing within the CPU—where it belongs. This results in reduced support risks for IT departments.

For More Information

To find out more about Intel I/O Acceleration Technology, visit www.intel.com/go/ioat.

*Intel® I/O Acceleration Technology (Intel® I/OAT) requires an operating system that supports Intel I/OAT.

INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH INTEL® PRODUCTS. EXCEPT AS PROVIDED IN INTEL'S TERMS AND CONDITIONS OF SALE FOR SUCH PRODUCTS, INTEL ASSUMES NO LIABILITY WHATSOEVER, AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY RELATING TO SALE AND/OR USE OF INTEL PRODUCTS, INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT, OR OTHER INTELLECTUAL PROPERTY RIGHT. INTEL MAY MAKE CHANGES TO SPECIFICATIONS, PRODUCT DESCRIPTIONS, AND PLANS AT ANY TIME, WITHOUT NOTICE.

Intel Corporation may have patents or pending patent applications, trademarks, copyrights, or other intellectual property rights that relate to the presented subject matter. The furnishing of documents and other materials and information does not provide any license, express or implied, by estoppel or otherwise, to any such patents, trademarks, copyrights, or other intellectual property rights. Intel products are not intended for use in medical, life saving, life sustaining, critical control or safety systems, or in nuclear facility applications. Intel® I/O Acceleration Technology may contain design defects or errors known as errata, which may cause the product to deviate from published specifications. Current characterized errata are available upon request.

Copyright © 2006 Intel Corporation. All rights reserved.

Intel, the Intel logo, Intel. Leap ahead., Intel. Leap ahead. logo, and Xeon are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

*Other names and brands may be claimed as the property of others.

