



Maximizing Server Performance with Mellanox Socket Direct[®] Adapter

Introduction

With the exponential growth of data, enterprises and cloud providers demand higher performance from servers and compute resources in order to perform real-time analysis on vast amounts of data.

Data center servers are typically equipped with a multi-socket CPU board and a single high-speed network adapter. This paper explains how Mellanox's innovative Socket Direct technology can maximize data center return on investment by delivering much higher performance for multi-socket servers, reaching up to 25% more throughput, while reducing latency by up to 80% and reducing CPU utilization by up to 60%.

Multi-Socket Servers

To maximize server compute power and performance, data centers frequently deploy multi-socket servers in which two or more CPU sockets are mounted on the server's motherboard. Communication between the CPUs occurs via a dedicated inter-processor communication bus.

When a network interface card is plugged into a server, the traffic between the network and the various CPUs also traverses over this bus, adding overhead and latency.

Inter-CPU Data Transmission Challenge

Figure 1 shows two dual-socket servers communicating over a network. On each server, the NIC is connected to one CPU (socket). If the other socket needs to access the network, it must traverse the inter-processor communication bus, via the first socket to the NIC, and then continue on towards the network.

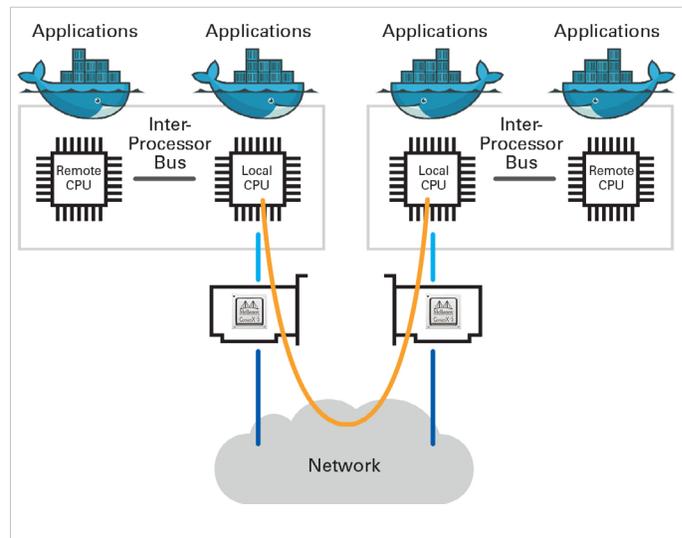


Figure 1. Dual-Socket Server with Standard Adapter

The inter-processor communication bus is capable of rapidly transmitting high quantities of data between the two CPUs. However, when data-intensive applications are running on different CPUs, while competing for access to a single network device, bottlenecks may arise. While it is possible to mitigate some bottlenecks using smart operating system resource allocation, most are unavoidable due to the nature of applications.

The Solution: Mellanox Socket Direct®

Mellanox Socket Direct® is a unique form-factor network adapter offered as two PCIe cards, wherein the PCIe lanes are split between the two cards. A key benefit that this adapter card brings to multi-socket servers is in eliminating the network traffic traversing over the internal bus between the sockets, significantly reducing overhead and latency.

Figure 2 shows a photograph of the Socket Direct adapter, which cost-effectively integrates a single network adapter silicon on a primary board, together with an auxiliary PCIe connection card and Slim Line SAS harness connecting the two.



Figure 2. Socket-Direct Adapter (Front and Back Angles)[†]

How Socket Direct Works

Socket Direct enables direct access from each CPU to the network through its dedicated PCIe interface (Figure 3).

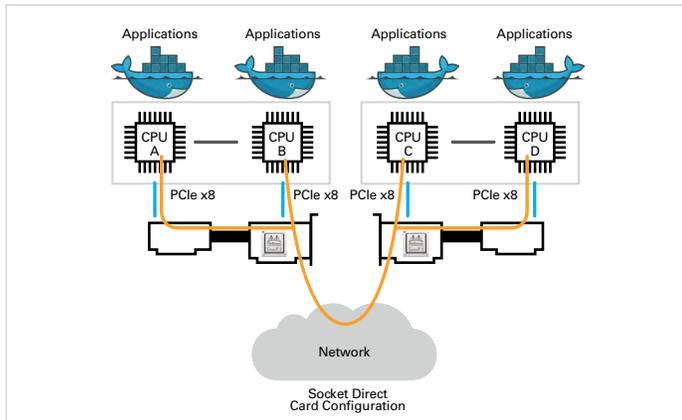


Figure 3. Configuration enabled by Socket Direct when directly connected to both CPU sockets, using two PCIe slots

Performance Comparison

Overview

Testing compared the performance of a ConnectX®-based Socket Direct adapter (installed in a dual-socket server) and a standard PCIe x16 100Gb/s adapter card, connected to only one of the CPU sockets. The measurements include TCP throughput, latency and CPU utilization, as well as RDMA benchmark tests.

Figure 4 shows the two network configurations, side by side.

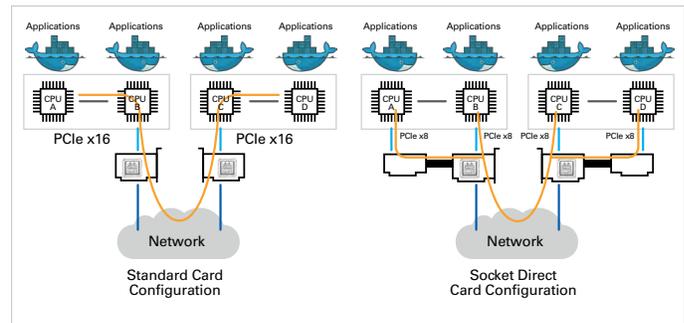


Figure 4. Standard Card Configuration versus Socket Direct Card Configuration (tests setup)

Standard Card Configuration:

The diagram in Figure 4 (left) shows a basic network of dual-socket servers, on which a single network card is installed. The gold line shows a flow path of data packets running between the servers, across the network. The flow path contains multiple data processing nodes (CPU-A, CPU-B, CPU-C and CPU-D). Traffic can pass between any two nodes on the flow path.

Socket Direct Configuration:

The diagram in Figure 4 (right) shows that CPU-A and CPU-D now have direct paths to the network controller, bypassing the inter-processor buses.

Lab Test Scenario

We tested traffic in 3 scenarios:

Test1 - Local CPU (Figure 4, left): Path started from CPU-B and ended in CPU-C.

Test2 - Remote CPU (Figure 4, left): Path started in CPU-A, went through the inter-processor bus to CPU-B, via the network to CPU-C and again through the inter-processor bus onward to CPU-D.

Test3 - Socket Direct (Figure 4, right): Path goes from both CPU-A and CPU-B through the network to CPU-C and CPU-D on the other server.

Figure 5 compares the average latency of the Socket Direct adapter against a standard network adapter. The graph shows that when using the Socket Direct adapter, latency is reduced by 80% compared to the standard adapter. This latency improvement is a result of the direct path both CPU sockets take to reach the network and the even distribution of TCP streams between the CPUs.

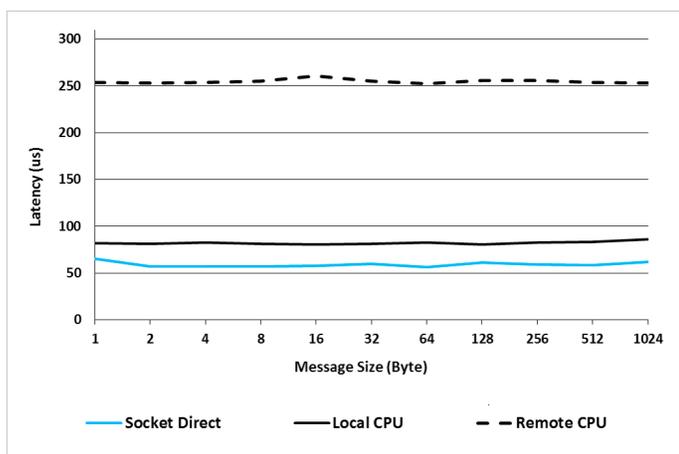


Figure 5. Average Latency

Figure 6 shows CPU utilization. It is evident that direct access to the network using Socket Direct also provides a 50% improvement in CPU utilization. Moreover, the even distribution of TCP streams reduces the average cache miss count on both CPUs versus a standard configuration server, which further improves CPU utilization.

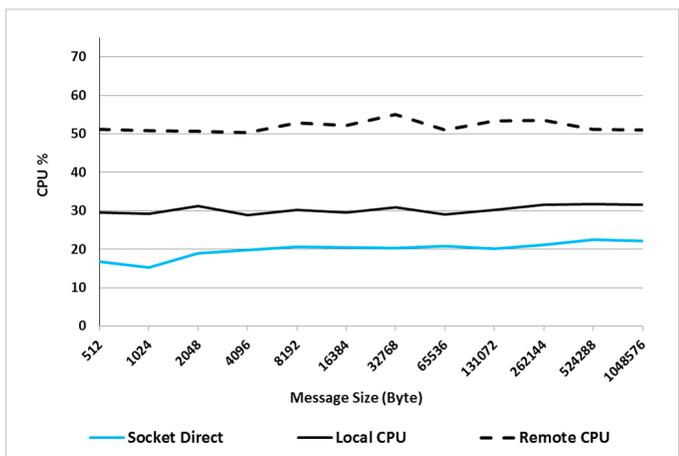


Figure 6. Host CPU Utilization (Inter-processor Load Applied)

Real Life Scenario

In a real-life scenario, applications that run on a dual-socket server generate data traversing the CPUs (over the inter-processor communication bus). To receive a more realistic measurement of the network performance, we apply an artificial load on the inter-processor bus, and then measure the effect of this load on the external data traffic of the server. We took these measurements, while comparing the two types of adapters (standard adapter and Socket Direct).

We measured the artificial load that was applied on the inter-processor bus. The more traffic that can pass between the CPUs, the greater the potential for communication between the sockets. Figure 7 presents the amount of inter-processor traffic that the server managed to transmit. It is apparent that when using the Socket Direct adapter, additional traffic managed to pass between the sockets.

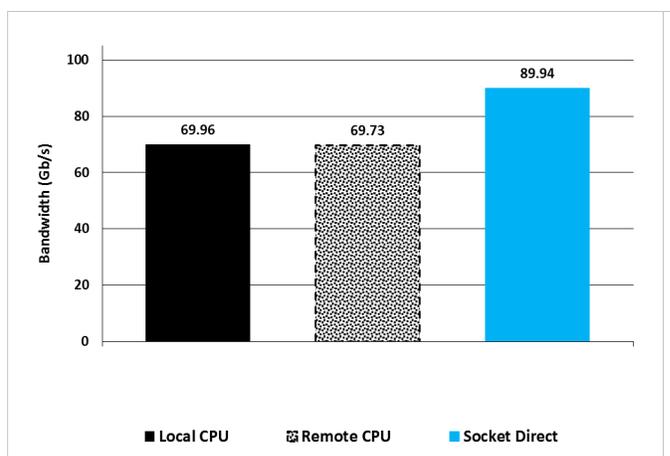


Figure 7. Load on the Inter-processor Bus

When comparing the servers' external throughput while applying the inter-processor load (Figure 8), it is evident that by implementing Socket Direct, the throughput is improved by 16%-28% compared to the standard adapter connection.

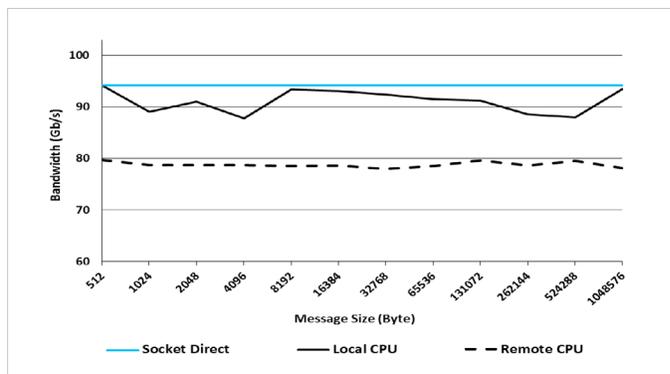


Figure 8. Ethernet Throughput (Inter-processor Load Applied)

Socket Direct and RDMA

We tested the impact of using a Socket Direct card on RDMA workloads. Figure 9 presents the results achieved, with the Socket Direct card showing 25% higher throughput compared to the throughput achieved on the remote CPU.

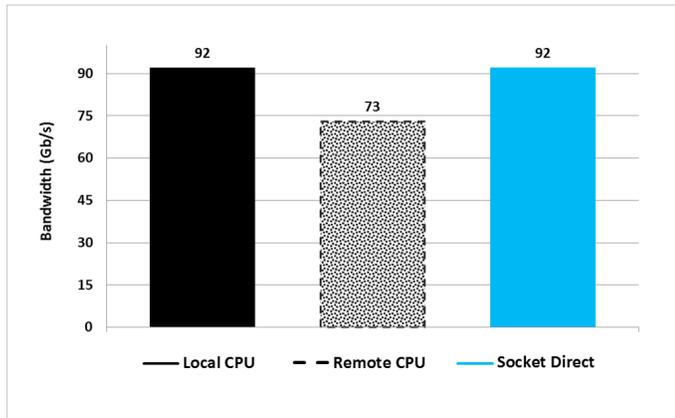


Figure 9. RoCE RDMA Throughput

Conclusion

Mellanox Socket Direct adapter provides the highest performance and most flexible solution for the most demanding applications and markets. Socket Direct extends server performance and utilization with maximum throughput connectivity. The Socket Direct adapter within a dual-socket server, enables both CPUs to connect directly to the network, delivering lower latency, lower CPU utilization and higher network throughput.

References

- From smartphones, scan to download the Mellanox Socket Direct Product Brief

