

Sun Blade X6440 and STREAM benchmark performance

Andrew Lewis
16 June 2010

ETH Zurich has discovered an issue with benchmark performance variations on the X6440 blade. This is a AMD based 4 Socket blade using Barcelona Socket F processors. The customer is reporting that STREAM benchmark results are lower when the benchmark is run on CPU 0 or 3 compared to the results when the benchmark is run on CPU1 or CPU2. In this case, the benchmark is run on an individual core on each physical CPU.

The customer results are below.

a) Sun X6440 with Opteron 8380, 2.5 GHz:

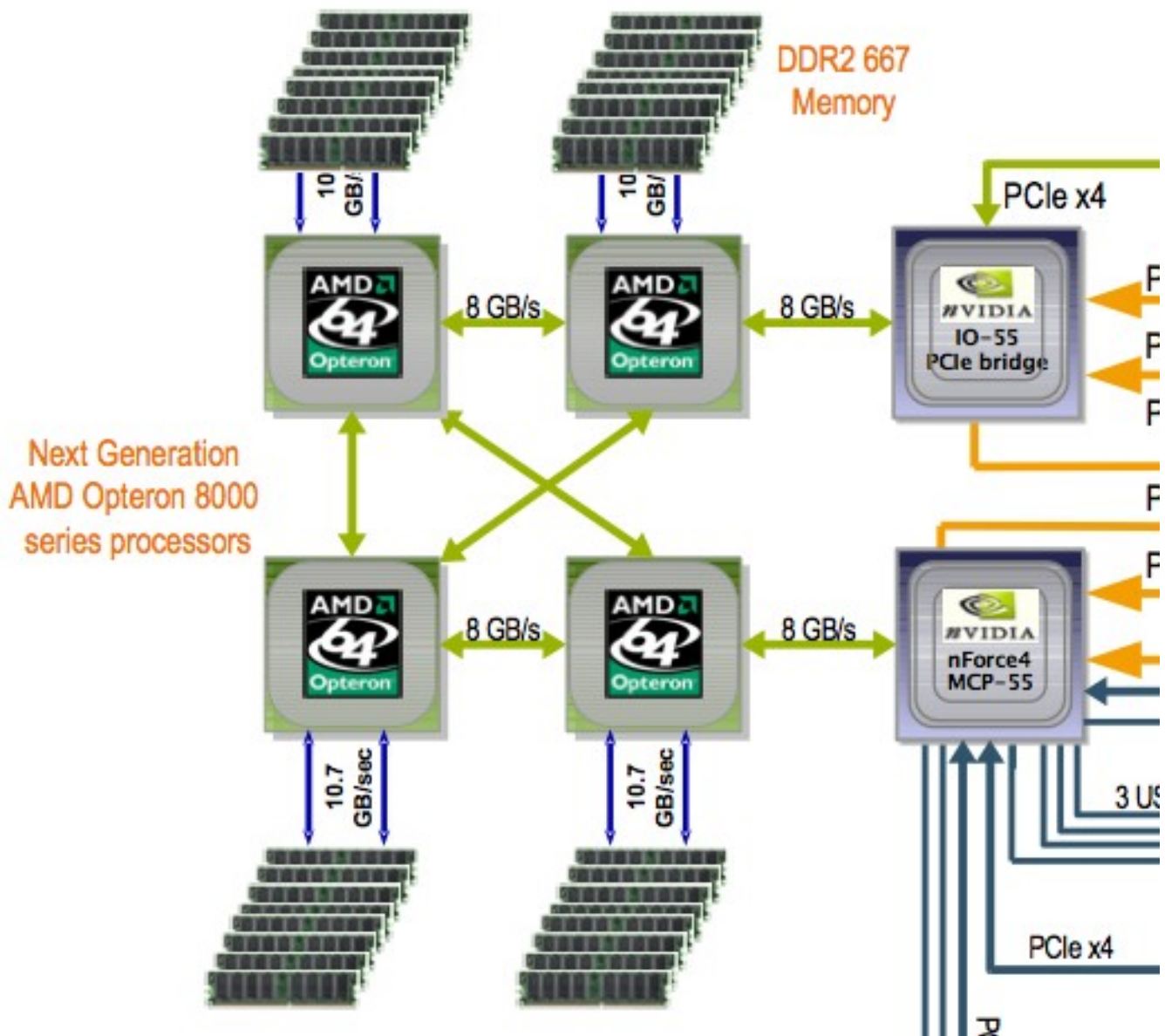
```
# for core in 0 4 8 12; do taskset -c $core ./STREAM; done | grep Triad
Triad:      3939.7779      0.0244      0.0244      0.0246
Triad:      4693.9123      0.0205      0.0205      0.0206
Triad:      4609.1780      0.0209      0.0208      0.0209
Triad:      3945.4141      0.0245      0.0243      0.0247
```

b) Sun X6440 with Opteron 8384, 2.7 GHz:

```
# for core in 0 4 8 12; do taskset -c $core ./STREAM; done | grep Triad
Triad:      4173.5666      0.0230      0.0230      0.0233
Triad:      4969.6772      0.0194      0.0193      0.0195
Triad:      4979.7569      0.0194      0.0193      0.0195
Triad:      4145.2519      0.0232      0.0232      0.0233
```

The customer has tried this on multiple X6440 blades and has observed a memory throughput drop of around 16%.

The block diagram of the X6440 blade is on the next page



This shows the Hypertransport (HT) links between each CPU, and between the CPU's and the I/O Bridges.

In the diagram above, CPU0 and CPU3 are the CPU's on the right side, connected to the I/O bridges, and CPU1 and CPU2 are on the left side.

Local memory reads generate snoop requests which are sent to each of the other sockets in the system. The remote sockets need to service these snoop requests and return a response to the requester prior to a memory read completing. Snoop requests which originate from CPU0 or CPU3 will require an additional "hop" to reach the most distant processor in the system. This means that any transaction which originates from CPU0 and targets CPU3 will need to pass through either CPU1 or CPU2. Similarly, any transaction which originates from CPU3 and targets CPU0 will also need to pass through CPU1 or CPU2. The additional HT hops required for snoop transactions originating from CPU0 or CPU3 delay read completions and result in increased local memory latency for these two

sockets. These increased read latencies directly contribute to a reduction in local memory latency for the processor.

This behavior is not unique to this blade, it is common to all of Oracle's 4-socket AMD socket F systems. Some other manufacturers chose not to make use of all the available HT links from each of the processors. These systems incur the latency increase and bandwidth reduction described above across all 4-sockets. The Oracle design significantly improves the performance of the two of the processors, but this optimization results in asymmetric performance across all of the processors installed in the system.