Intel Xeon 5500 Memory Performance

Ganesh Balakrishnan
System x and BladeCenter Performance
# Intel Xeon 5500 Memory Performance

## ABSTRACT

---

## 1.0 INTRODUCTION

---

## 2.0 SYSTEM ARCHITECTURE

### 2.1 HS22

---

### 2.2 x3650 M2, x3550 M2, iDATAplex 2.0

---

## 3.0 MEMORY PERFORMANCE

### 3.1 MEMORY SPEED

- **3.1.1 Processor SKU**
- **3.1.2 DDR3 DIMM speed**
- **3.1.3 DIMMs per channel (DPC)**
- **3.1.4 Low-level Performance**
- **3.1.5 Application Performance**

### 3.2 MEMORY INTERLEAVING

- **3.2.1 HS22**
- **3.2.2 x3650M2/x3550M2/iDpx 2.0**

### 3.3 RANKS

### 3.4 MEMORY POPULATION ACROSS MEMORY CHANNELS

### 3.5 MEMORY POPULATION ACROSS PROCESSORS

## 4.0 BEST PRACTICES

---

---
Abstract
In this paper, we examine the architecture and performance of the new Intel Xeon 5500 processor. The new architecture poses some a number of challenges for optimizing system performance due to the differences in architecture from traditional Intel platforms and due to the variety of possible configurations. The performance analysis will cover latency to main memory, memory bandwidth and application performance. The paper addresses how the performance will vary by memory speed and by number of memory channels. This paper will also examine optimal memory configurations and best practices for the platform and make recommendations on configuring IBM platforms for optimal performance.

1.0 Introduction
The Xeon 5500 processor is the next generation Intel Quad-core processor targeted towards the two socket server space and will be the common building block across a number of IBM platforms including HS22, x3650M2, x3550M2 and iDataplex 2.0. With the Xeon 5500 processor, Intel has diverged from it’s traditional Symmetric Multiprocessing Architecture (SMP) to a Non-Uniform Memory Access Architecture (NUMA). In a two processor scenario, the Xeon 5500 processors are connected through a serial coherency link called QuickPath Interconnect (QPI). The QPI is capable of 6.4, 5.6 and 4.8 GT/s depending on the processor SKU. The Xeon 5500 processor integrates the memory controller inside the processor resulting in two memory controllers in a 2-socket system. Each memory controller has three memory channels which supports DDR3 registered memory. Depending on the type of memory, population of memory and processor SKU, the memory can be clocked at 1333MHz, 1066MHz and 800MHz. Each memory channels supports up to 3 DIMMs per channel (DPC). However, the actual number of DIMMs in the system is dependent on the system architecture.
2.0 System Architecture

In this section, we will explore the system architecture of various System x and BladeCenter Xeon 5500 based products from a memory standpoint.

2.1 HS22

HS22 is designed with 12 DIMM slots as shown in Figures 2 and 3. The 12 DIMM layout provides 6 DIMMs per socket and 2 DPC.
Intel Xeon 5500 Memory Performance

2.2 x3650 M2, x3550 M2, iDataplex 2.0
A shown in figures 4 and 5 below, the other IBM Xeon 5500 based platforms provide 16 DIMM slots. Similar to HS22, both processors have equal number of DIMM slots. However, unlike HS22 all memory channels do not have equal DPC.

Figure 3: HS22 DIMM slots board layout

Figure 4: x3650M2/x3550M2 DIMM slots architectural layout
3.0 Memory Performance

With the varied number of configurations possible in the Xeon 5500 based systems, a number of variables emerge that influence processor/memory performance. The main variables that affect memory performance are memory speed, memory interleaving, ranks and memory population across various memory channels and processors. Depending on the processor SKU and number of DIMMs, the performance of the Xeon 5500 platform will see huge variance. We will look at each of these factors more closely in the next sections.

3.1 Memory Speed

As mentioned earlier, the memory speed is determined by the processor SKU, memory type and DIMMs per channel.

3.1.1 Processor SKU

The initial Xeon 5500 based offerings will be categorized into 3 bins called Performance, Volume and Value. The 3 bins will have the ability to clock memory at a maximum clock speed of 1333MHz, 1066MHz and 800MHz respectively. So, the processor SKU will determine the maximum possible frequency of the memory.
3.1.2 DDR3 DIMM speed

DDR3 memory will be available in 1333MHz and 1066MHz. This represents the maximum capability at which memory can be potentially clocked. However, the memory will not be clocked faster than the capability of the processor SKU and will be clocked appropriately by the BIOS.

3.1.3 DIMMs per channel (DPC)

The number and type of DPC will also determine the speed at which memory will be clocked. The following table describes the behavior of the platform. The table below assumes a Xeon 5500 1333MHz capable SKU. If a slower SKU is used, then the Memory speed will the lower of the memory speed determined from the table and the processor SKU memory speed capability. If the DPC is not uniform across all the channels, then the system will clock to the fastest common frequency.

<table>
<thead>
<tr>
<th>DPC</th>
<th>DIMM Speed (MHz)</th>
<th>Ranks per DIMM</th>
<th>Memory Speed (MHz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1333</td>
<td>1,2</td>
<td>1333</td>
</tr>
<tr>
<td>2</td>
<td>1333</td>
<td>1,2</td>
<td>1066</td>
</tr>
<tr>
<td>3</td>
<td>1333</td>
<td>1,2</td>
<td>800</td>
</tr>
<tr>
<td>1</td>
<td>1333</td>
<td>4</td>
<td>1066</td>
</tr>
<tr>
<td>2</td>
<td>1333</td>
<td>4</td>
<td>800</td>
</tr>
<tr>
<td>1</td>
<td>1066</td>
<td>1,2</td>
<td>1066</td>
</tr>
<tr>
<td>2</td>
<td>1066</td>
<td>1,2</td>
<td>1066</td>
</tr>
<tr>
<td>3</td>
<td>1066</td>
<td>1,2</td>
<td>800</td>
</tr>
<tr>
<td>1</td>
<td>1066</td>
<td>4</td>
<td>1066</td>
</tr>
<tr>
<td>2</td>
<td>1066</td>
<td>4</td>
<td>800</td>
</tr>
</tbody>
</table>

3.1.4 Low-level Performance

It is important to understand the impact of the performance of the Xeon 5500-based platform depending on the memory speed. We will use both low level memory tools and then applications to quantify the impact of memory speed.

Two of the key low-level metrics that are used to measure memory performance are memory latency and memory throughput. We use a base Xeon 5500 2.93GHz, 1333MHz capable 2-socket system for this analysis. The memory configurations for the three memory speeds are as follows:

- 1333MHz - 6x4GB dual rank 1333MHz DIMMs
- 1066MHz - 12x2GB dual rank DIMMs for 1066MHz
- 800MHz - 12x2GB dual rank DIMMs down clocked to 800MHz in BIOS

As shown in Figure 6 below, we show the unloaded latency to local memory. The unloaded latency is measured at the application level and is designed to defeat processor prefetch mechanisms. As shown in the figure, the difference between the fastest and slowest speeds is about 10%. This represents the high watermark for latency sensitive workloads. Another important thing to note is that this is almost a 2x decrease in memory.
latency when compared to the previous generation Xeon 5400 series processor on 5000P platforms.

Figure 6: Memory latency as a function of memory speed.

A better indicator of application performance is memory throughput. We use the triad component of streams to compare the performance at different memory speeds. The memory throughput assumes all local memory allocation and with all 16 cores utilizing main memory. As shown in figure 7, the performance gain of using 1066MHz memory versus 800MHz memory is 28% and 1333MHz versus 1066MHz is 9%. So, the performance penalty of clocking memory at 800MHz is far greater than clocking it down to 1066MHz. However, it is possible to achieve greater memory capacity with lower memory speed. It is also possible to achieve the same memory capacity at lower cost but at a lower memory speed. So, there is a distinct trade-off of memory capacity, performance, cost and power. Regardless of memory speed, the Xeon 5500 platform represents a significant improvement in memory bandwidth over the previous Xeon 5400 platform. At 1333MHz, the improvement is almost 5x over the previous generation. This huge improvement is mainly due to dual integrated memory controllers and faster DDR3 1333MHz memory over FBD 667MHz memory. This improvement translates into improved application performance and scalability.

Figure 7: Memory throughput using Streams Triad

### 3.1.5 Application Performance

In this section, we will discuss the impact of memory speed on the performance of 3 commonly used benchmarks: SPECint2006_rate, SPECfp2006_rate and SPECjbb2005. In each case, the benchmark scores are relative to the score at 800MHz as shown in figure 8.

SPECint2006_rate tends to be more sensitive to processor frequency and less to memory bandwidth. There are very few components in SPECint2006_rate that are memory bandwidth intensive and so, the performance gain with memory speed improvements is the least for this workload. In fact, most of the difference observed is due one of the sub-benchmarks that shows a high sensitivity to memory frequency. There is an 8% improvement going from 800MHz to 1333MHz when the improvement in memory bandwidth is almost 40%. SPECint2006_rate is typically used as an indicator of performance for commercial applications.

SPECfp2006_rate tends to be memory bandwidth intensive and we expect to see significant improvements for this workload with memory frequency. As expected, a number of sub-benchmarks show improvements as high as the difference in memory bandwidth. As shown in figure 6, there is a 13% gain from 800MHz to 1066MHz and
another 6% upside with 1333MHz. SPECfp_rate captures almost 50% of the memory bandwidth improvement. SPECfp_rate is used as an indicator for HPC workloads. SPECjbb2005 is a workload that does not stress memory but keeps the data bus moderately utilized. This workload provides a middle ground and the performance gains reflect that trend.

Figure 8: Application performance as a function of memory speed

### 3.2 Memory interleaving

Memory interleaving refers to how physical memory is interleaved across the physical DIMMs. A balanced system provides the best interleaving. The Xeon 5500 system is balanced when all memory channels on a socket have the same amount of memory. The simplest way to enforce optimal interleaving is by populating 6 identical DIMMs for 1333MHz, 12 identical DIMMs for 1066MHz and 18 identical DIMMs for 800MHz.

#### 3.2.1 HS22

In the case of HS22 which has a balanced DIMM layout, it is easy to ensure that the system is also balanced for all 3 memory frequencies. Recommended DIMM population is shown in the table below assuming identical size DIMMs.

<table>
<thead>
<tr>
<th>Desired memory speed</th>
<th>DIMMs per channel</th>
<th>DIMM slots to populate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1333MHz</td>
<td>1</td>
<td>2,4,6,8,10,12</td>
</tr>
<tr>
<td>1066MHz</td>
<td>2</td>
<td>All slots</td>
</tr>
<tr>
<td>800MHz</td>
<td>2</td>
<td>All slots and clock memory speed to 800MHz in BIOS</td>
</tr>
</tbody>
</table>

#### 3.2.2 x3650M2/x3550M2/iDpx 2.0

In the case of the other systems which have 16 DIMM slots, care needs to be taken on how the DIMMs are populated especially when configuring for 800MHz. When configuring for 800MHz which would require more than 2 DPC, it would be a common error to populate all 16 DIMM slots with identical DIMMs. However, such a configuration leads to an unbalanced system where two memory channels have less memory capacity than the other 4 which leads to poor performance. Shown below in Figure 9, is the impact of poor interleaving. The 1st configuration is a balanced baseline configuration where the memory id down-clocked to 800MHz in BIOS. The 2nd configuration populates 4 channels with 50% more memory than 2 other channels causing an unbalanced configuration. The 3rd configuration balances the memory on all channels by populating the channels with fewer DIMM slot with a DIMM that is double the capacity of others. This ensures that all channels have the same capacity even though all the DIMM slots are populated. As the figure shows below, the 1st and 3rd balanced configurations significantly outperform the unbalanced configuration. Depending on the memory footprint of the application and memory access pattern, the impact could be higher or lower than the two applications cited in the figure.
3.3 Ranks

It is also important to ensure that DIMMs with appropriate number of ranks are populated in each channel for optimal performance. Whenever possible, it is recommended to use Dual rank DIMMs in the system. Dual rank DIMMs offer better interleaving and hence better performance than single rank DIMMs. For instance, a system populated with 6x2GB dual rank DIMMs outperforms a system populated with 6x2GB single rank DIMMs by 7% for SPECjbb2005. Dual rank DIMMs are also better than Quad rank DIMMs because Quad rank DIMMs will cause the memory speed to be down clocked. Another important observation is to populate even ranks per channel. For instance, mixing one single rank DIMM and one dual rank DIMM in a channel should be avoided.

3.4 Memory Population across Memory Channels

It is important to ensure that all 3 memory channels in each processor are populated. The relative memory bandwidth is shown in the figure 10 and it illustrates the loss of memory bandwidth as the number of channels populated is decreased. This is intuitive because the bandwidth of all 3 memory channels is required to support the capability of the processor. So, as the channels are decreased, the burden to support the requisite bandwidth is increased on the remaining channels causing them to become a bottleneck.

3.5 Memory Population across Processors

Since the Xeon 5500 has a NUMA architecture, it is important to ensure that the memory controllers in the system are utilized by populating both sockets with memory. It is also optimal to balance the system by populating both sockets in an identical fashion to provide a balanced system. Using figure 11 as an example, socket 0 has DIMMs
populated but no DIMMs are populated on socket 1. In this case, socket 0 will have access to low latency “local” memory and high memory bandwidth as a balanced system. However, socket 1 has access to remote or far memory only. So, threads executing on socket 1 will have a long latency to access memory as compared to threads on socket 0. This is due to the latency penalty incurred to traverse the QPI links to access the data on the remote memory controller. The bandwidth to remote memory is also limited by the capability of the QPI links. The latency to access remote memory is almost 75% higher than local memory access. So, the goal should be always to populate both sockets with memory when both processors are populated.

Figure 11: Diagram showing local and remote memory access

4.0 Best Practices

In this section, we recapture the various rules to be followed for optimal memory configuration on the Xeon 5500 based platforms.

1. Always populate both processors with equal amounts of memory to ensure a balanced NUMA system
2. Always populate all 3 memory channels on each processor with equal memory capacity.
3. Ensure even ranks are populated per channel.
4. Use Dual-rank DIMMs whenever appropriate.
5. For optimal 1333MHz performance, populate 6 dual rank DIMMs (3 per processor).
6. For optimal 1066MHz performance, populate 12 dual rank DIMMs (6 per processor).
7. For optimal 800MHz performance
   a. On 12 DIMM platforms, populate 12 dual rank or quad rank DIMMs (12 per processor) and clock memory down to 800MHz in BIOS.
b. On 16 DIMM platforms,
   i. Populate 12 dual rank or quad rank DIMMs (12 per processor) and
clock memory down to 800MHz in BIOS.
   ii. Populate 14 dual rank DIMMs of one size and 2 dual rank DIMMs
of double the size as described in the interleaving section.
8. With the above rules, it is not possible to have a performance optimized system
with 4GB, 8GB, 16GB...128GB. With 3 memory channels and interleaving rules,
customers need to configure systems with 6GB, 12GB, 18GB, 24GB, 48GB,
72GB, 96GB, etc.