Whitepaper

How to Balance Memory on 2nd Generation Intel® Xeon™ Scalable Processors

Optimizing Memory Performance

Matt Ogle, Technical Product Marketing, Dell EMC

Trent Bates, Product Management, Dell EMC

Bruce Wagner, Senior Principal Systems Development Engineer, Dell EMC

Rene Franco, Senior Principal Systems Development Engineer, Dell EMC

Abstract

Properly configuring a server with balanced memory is critical to ensure memory bandwidth is maximized and latency is minimized. When server memory is configured incorrectly, unwanted variables are introduced into the memory controllers' algorithm, which inadvertently slows down overall system performance. To mitigate this risk of reducing or even bottlenecking system performance, it is important to understand what constitutes balanced, near balanced and unbalanced memory configurations.

Dell EMC has published this whitepaper to educate PowerEdge customers on what balanced memory means, why it is important and how to properly populate memory for a balanced configuration.

Table of Contents

1. Introduction

Understanding the relationship between a server processor (CPU) and its memory subsystem is critical when optimizing overall server performance. Every processor generation has a unique architecture, with volatile controllers, channels and slot population guidelines, that must be satisfied to attain high memory bandwidth and low memory access latency.

Memory that has been incorrectly populated is referred to as an unbalanced or near balanced configuration. From a functionality standpoint, these configurations will operate adequately, but introduce additional overhead that will slow down data transfer speeds. Conversely, memory that has been correctly populated is referred to as a balanced configuration, which will secure optimal functionality and data transfer speeds.

Intel® Xeon™ scalable processors offer a total of six memory channels with up to two memory slots per channel; a total of twelve memory slots per processor.[1] This presents numerous possible permutations for configuring the memory subsystem with traditional Dual In-Line Memory Modules (DIMMs) and Optane™ DC Persistent Memory Modules (DCPMMs), yet there are only two balanced configurations that will achieve peak memory performance for Dell EMC PowerEdge servers.

This whitepaper explains how to balance memory configured for Intel® Cascade Lake™ processors within Dell EMC PowerEdge servers, and why this is so critical.

2. Memory Topography and Terminology

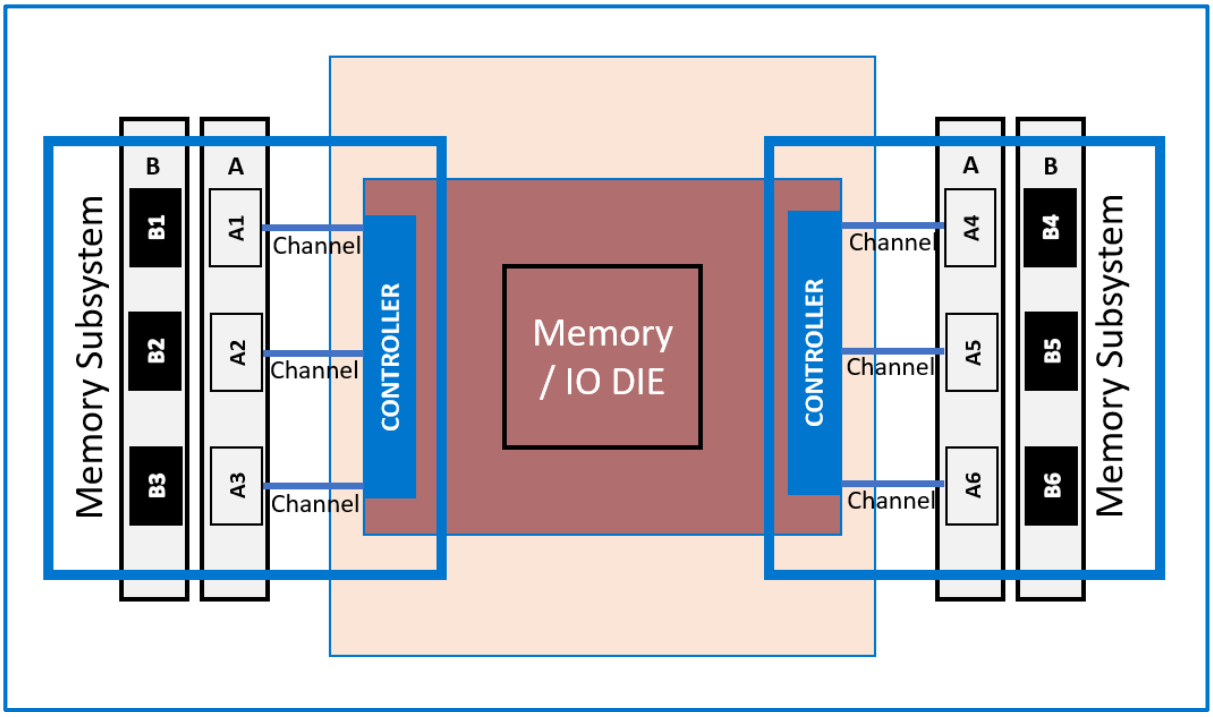

Figure 1: PowerEdge R740 CPU-to-memory subsystem connectivity for Intel® Cascade Lake™

To understand the relationship between the CPU and memory, terminology illustrated in Figure 1 must first be defined:

- The memory controllers are digital circuits that manage the flow of data going from the computer's main memory to the corresponding memory channels.[2] Intel® Xeon™ scalable processors have the controller integrated into the CPU.
- The memory channels control reading and writing bandwidth operations between the CPU and memory modules.[3] Intel® Xeon™ scalable processors have six memory channels, labeled one to six, which allow for increased data transfer rates compared to previous generations.
- The memory slots host individual memory modules, such as DIMMs or DCPMMs.[4] Intel® Xeon™ scalable processors have two slots per channel, shown as columns A and B, so there are a total of twelve slots per CPU for memory module population.
- The memory subsystem is the combination of all the independent memory functions listed above.

3. Memory Interleaving

Memory interleaving allows a CPU to efficiently spread memory accesses across multiple DIMMs. When memory is put in the same interleave set, contiguous memory accesses go to different memory banks. Memory accesses no longer must wait until the prior access is completed before initiating the next memory operation. To maximize performance, all DIMMs should be in one interleaved set creating a single uniform memory region that is spread across as many DIMMs as possible.[5] Multiple interleaved sets create disjointed memory regions. Software performance will vary depending on what memory region is being used and performance will be non-deterministic.

Memory configurations can either be balanced, near balanced or unbalanced based partly on the total interleave sets created. Balanced configurations require the memory channels to be fully populated with uniform DIMMs so that only one interleave set is created, therefore maximizing performance. Near balanced configurations also have one interleave set, but do not fully populate memory channels so that the distribution of memory accesses is sub-optimal. Unbalanced configurations have more than one interleave set, which creates non-uniform memory regions. Testing performed on the PowerEdge R740 has demonstrated **unbalanced and near balanced population can reduce memory bandwidth by up to 33% from its maximum potential,** which is why it is critical to know how to properly populate a balanced configuration.

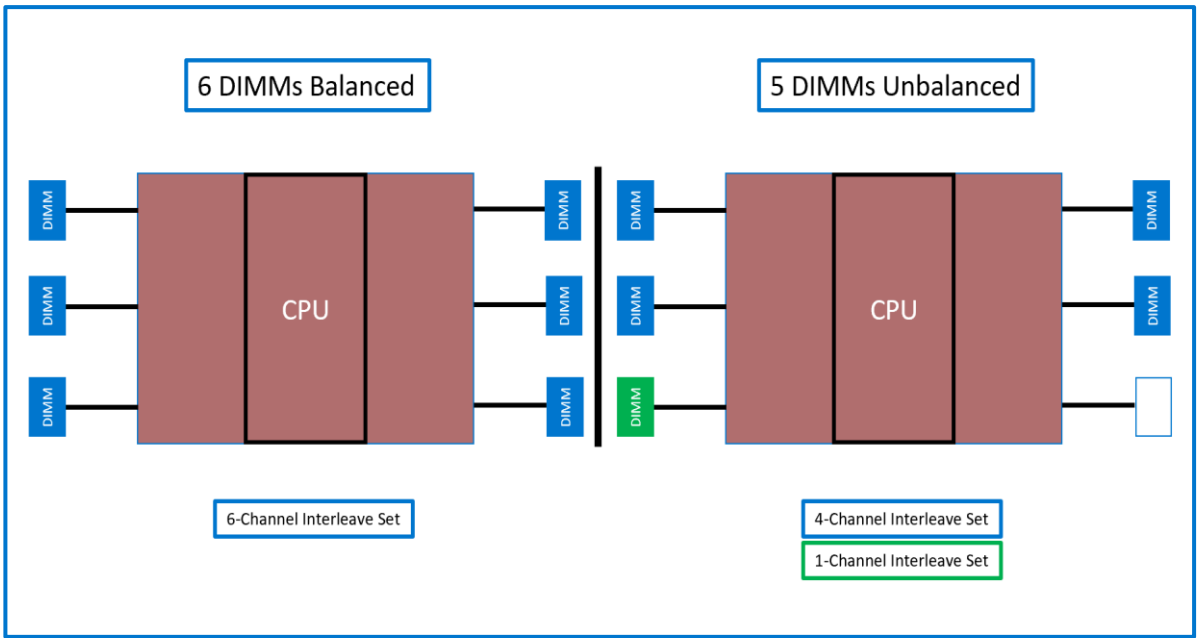

Figure 2: Side by side comparison of balanced and unbalanced memory configuration. In this case, once the balanced configuration removed one of the six DIMMs a second interleave set was introduced for the isolated DIMM, therefore making it an unbalanced configuration

4. Guidelines for Balancing Memory

4.1 Overview

To ensure effective interleaving is created, memory must be populated into a balanced configuration. Variables such as DIMM consistency and slot population will dictate whether a configuration is balanced or unbalanced. At both the socket and server level, memory bandwidth is optimized when the guidelines below are implemented:

1. All memory modules inside the memory subsystem are identical
   - They must have the same size, speed, rank count and DIMM type
2. All populated memory channels are identical and maximized
   - Channels must be fully populated with one or two DIMMs
3. All server sockets should have identical memory subsystems

4.2 Identical Memory Modules

All memory modules being used must be the same. This means that the DIMM size (or capacity), speed, rank count and type should be consistent. For example, if all the memory modules connected to the CPU channels are 32Gb, 2133MT/s, RDIMMs with two ranks, then the configuration is following this guideline. Figure 3 demonstrates what happens when two different DIMMs are populated to the same CPU:

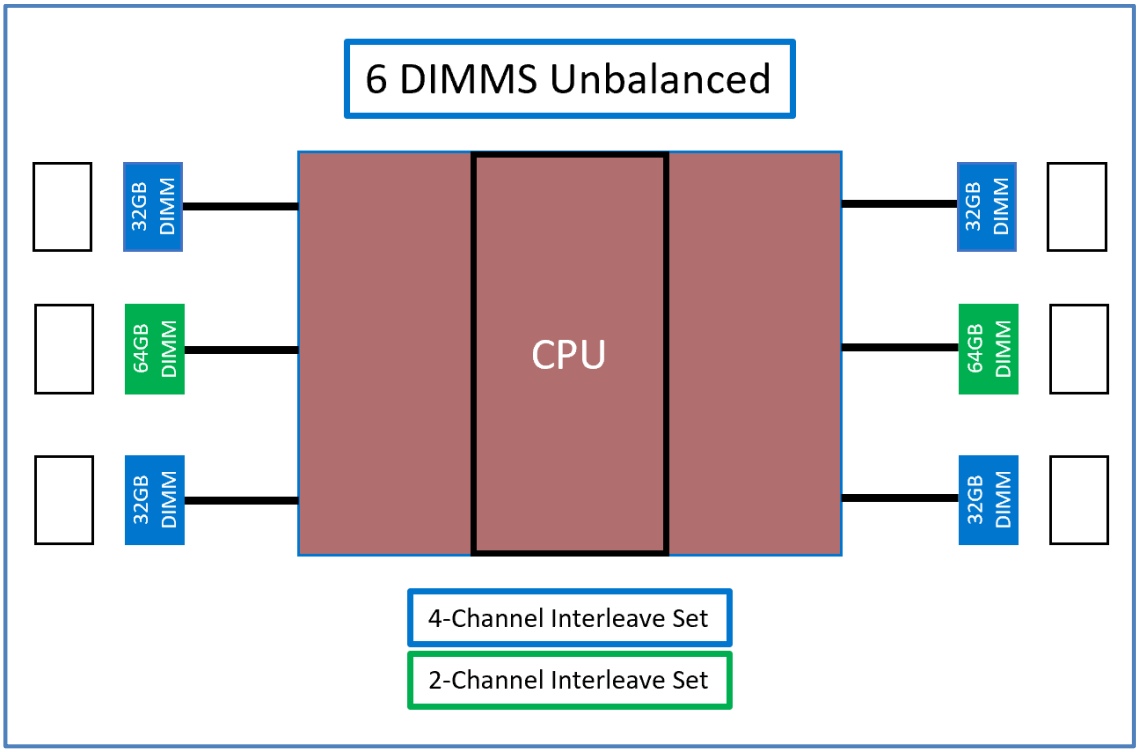

Figure 3: Two unique interleave sets were created because two unique DIMM sizes were populated

4.3 Identical Channel Population

Intel® Xeon™ scalable processors will have six channels, identified in Figure 4 as one to six, with up to two slots per channel, identified as A and B. Optimal performance is achieved when a channel is completely populated. For example, Figure 4 would have to have the inner grey slots A1, A2, A3, A4, A5 and A6 populated to fulfill this guideline. This guideline also applies to the outer black B slots, which can only be populated once the grey slots are populated. Therefore, only a total of six or twelve DIMMs can be balanced to return the best bandwidth and latency. When mirrored channels are identically, yet only partially, populated then the configuration is near balanced. For example, if A1, A2, A4 and A5 were populated while A3 and A6 were empty, this configuration would be near balanced. Near balanced and full balanced configurations are similar because they both only create one interleave set, but because near balanced does not populate every channel slot for A, the performance will degrade in comparison to balanced.

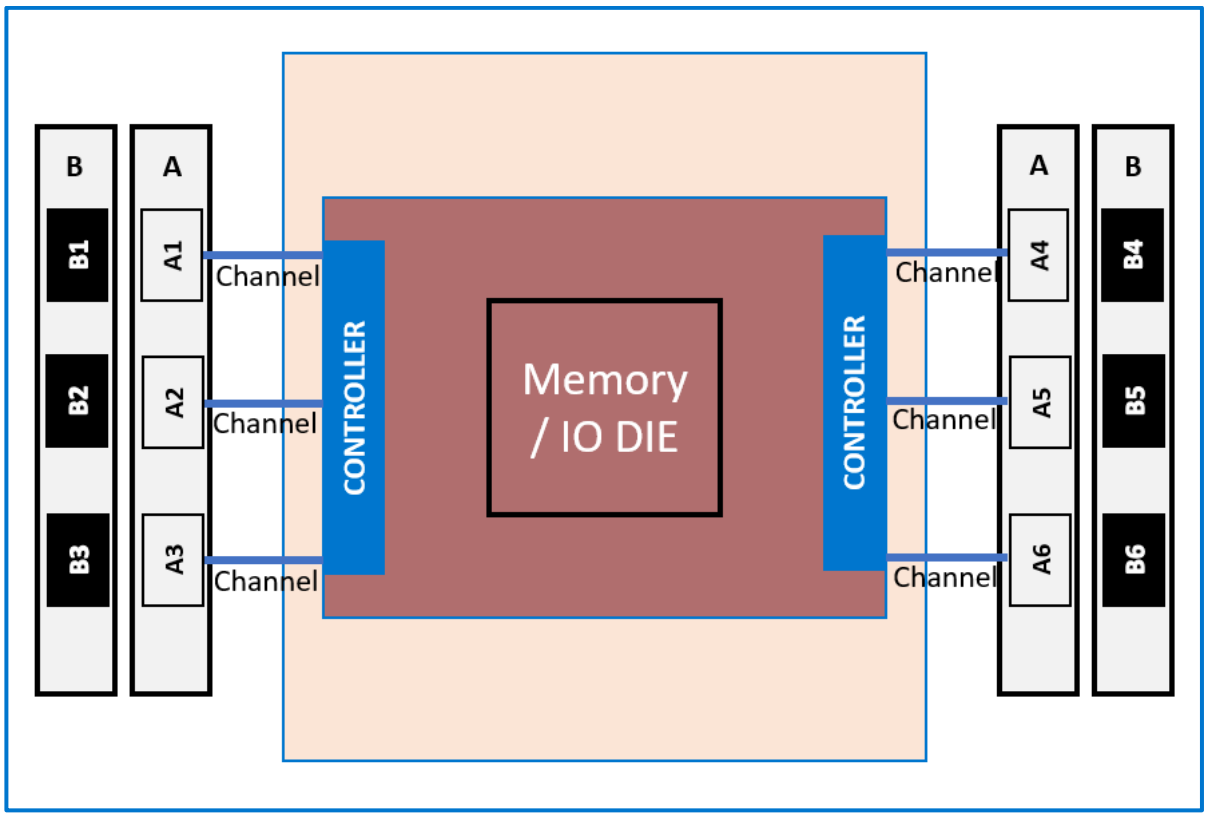

Figure 4: Memory channels are broken out into columns A and B, with six channels for each column

4.4 Identical Sockets

A CPU socket must have identical memory subsystems meeting guidelines 4.2 and 4.3. A physical server should also have identically configured CPU sockets. As seen in Figure 5, when only one unique memory configuration exists across all sockets within a server, memory bandwidth is further optimized because the CPU and memory controllers can command information to be distributed across many more channels. This principle is consistent when applied to even broader scopes; data centers that have identical servers with identical balanced memory subsystems will inherit increased memory performance.

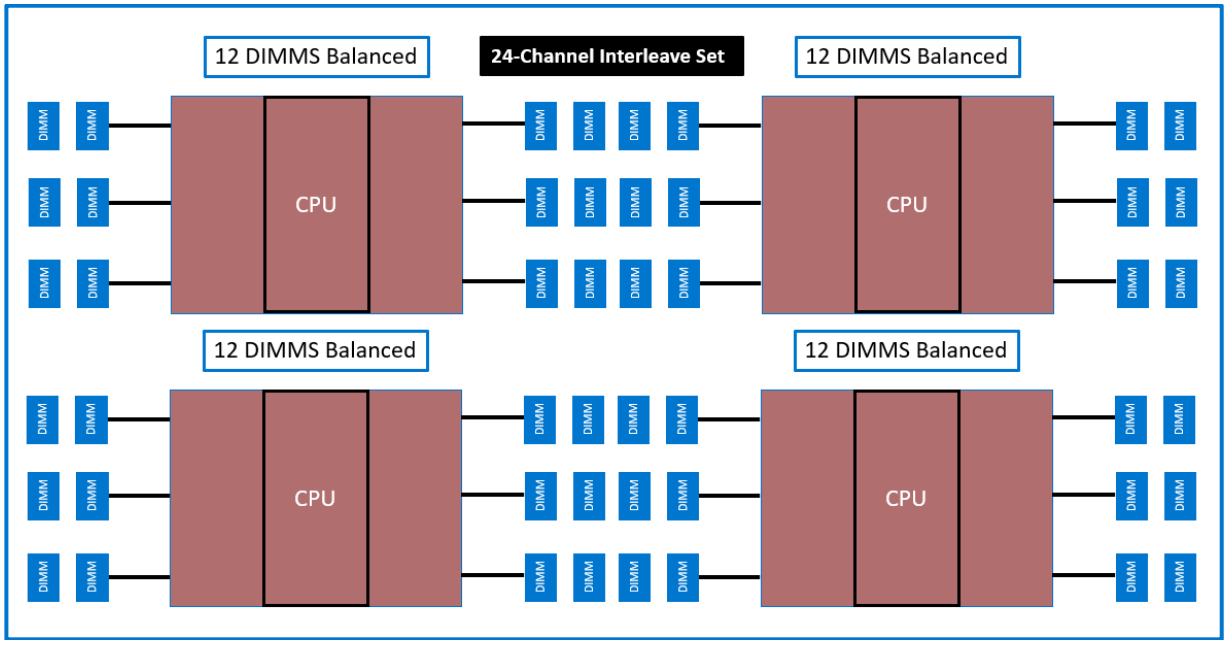

Figure 5: One interleave set is being leveraged across twenty-four channels (four sockets) because they are all identical, balanced configurations

5. Balanced Configurations (Recommended)

5.1 Traditional DIMMs

Memory controller logic was designed around having all memory slots populated to return the highest memory bandwidth, so it should come as no surprise that the top recommendation is a balanced configuration populated with twelve DIMMs. If a balanced and fully populated configuration with twelve DIMMs cannot be executed, then a balanced configuration with six DIMMs is the alternative recommendation, as memory bandwidth is reduced by only 3% when compared to twelve DIMMs.

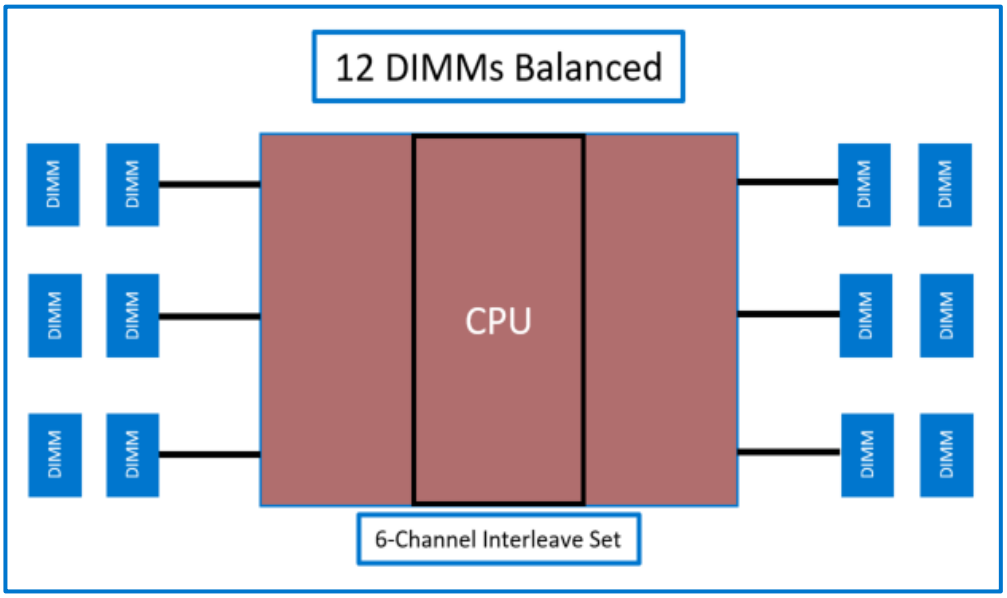

Figure 6: Populating all twelve slots with identical capacity DIMMs for a balanced configuration is recommended for highest memory bandwidth and lowest memory access latency

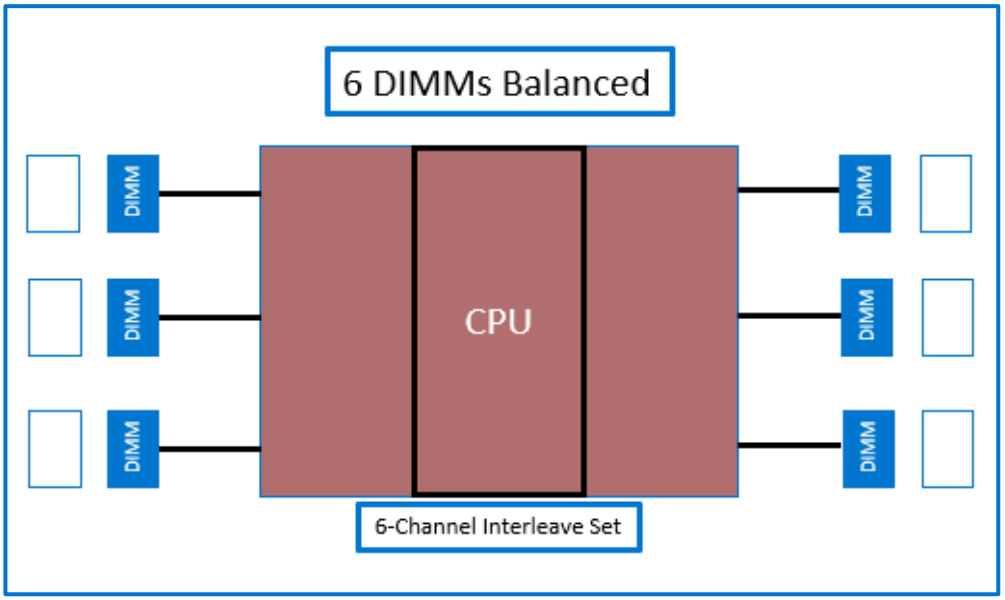

Figure 7: Populating the first six slots with identical capacity DIMMs is recommended for the second highest memory bandwidth and lowest memory access latency

5.2 Traditional DIMMs with DCPMMs

For mixed memory configurations containing both traditional DIMMs and persistent memory modules, populating six traditional DIMMs into the first channels slots and six DCPMMs into the second channels slots will reap the highest memory bandwidth and lowest memory access latency. The benefits gained from DCPMMs persistence and increased capacity will offset the memory bandwidth degradation introduced from having two interleave sets.

For more information about Intel Optane DC Persistent Memory, please see the following link.

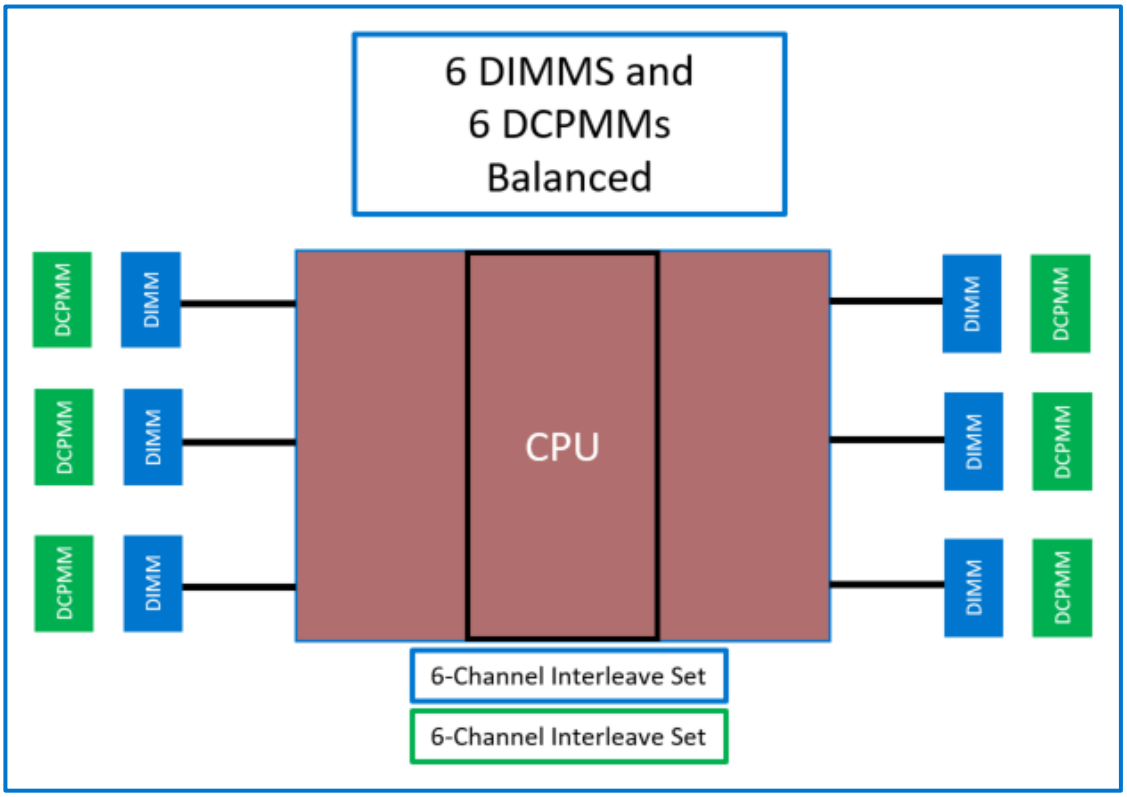

Figure 8: When using both DIMMs and DCPMMs, populating the first six slots with identical traditional DIMMs and the last six slots with identical DCPMMs is recommended. Two interleave sets are created because DIMMs and DCPMMs are unique drive types. Despite the memory bandwidth degradation that occurs from having two interleave sets, the persistence and increased capacity gained from DCPMMs will still enable various potential benefits

6. Near Balanced Configurations

Near balanced configurations also only have one interleave set. Populating DIMMs one, two and three on the same channel column will naturally only create one interleave set, but once four or more DIMMs are introduced, guideline 4.2 must be satisfied to maintain a near balanced composure. Mirrored columns must be identically populated with the same DIMMs. The four and eight DIMM illustrations below demonstrate what this looks like. Regardless, near balanced configurations are not fully populating memory channels with one or two DIMMs, and therefore **can reduce memory bandwidth by up to 33% from its maximum potential.** Dell EMC does not recommend populating memory in a near balanced configuration.

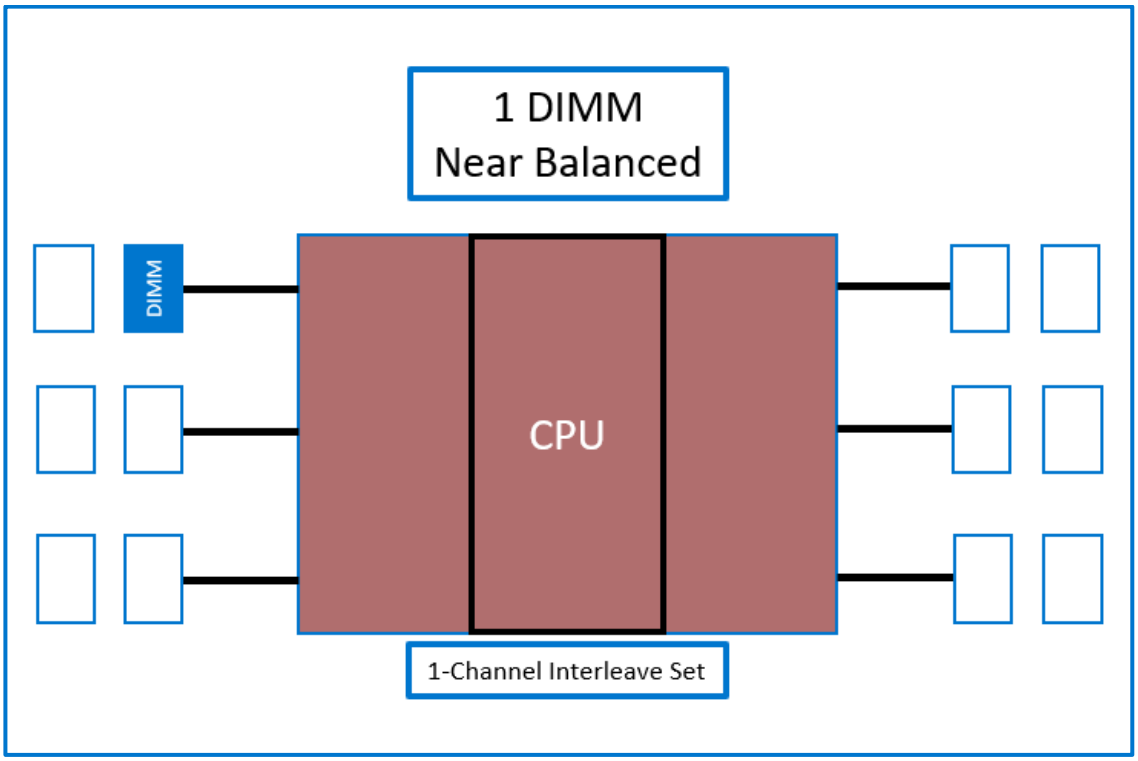

Figure 9: This configuration has one interleave set because one memory module only requires one interleave set. The absence of all channels being populated makes this configuration near balanced.

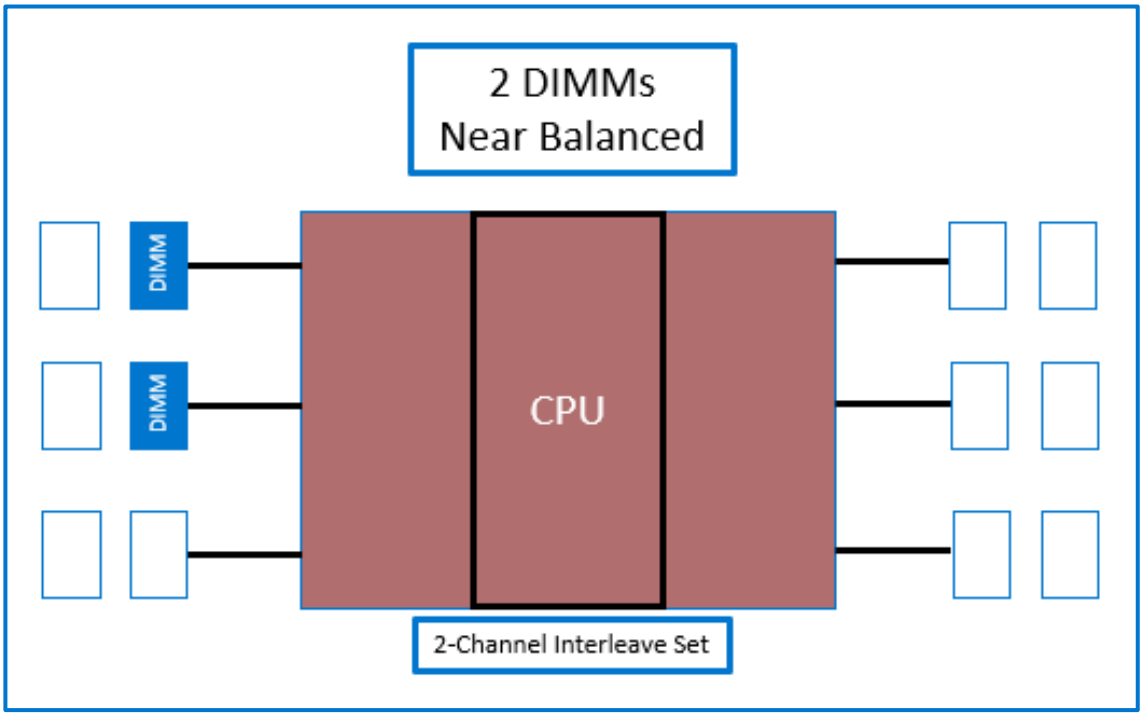

Figure 10: This configuration has one interleave set because identical memory modules are populated in the same channel column. The absence of all channels being populated makes this configuration near balanced.

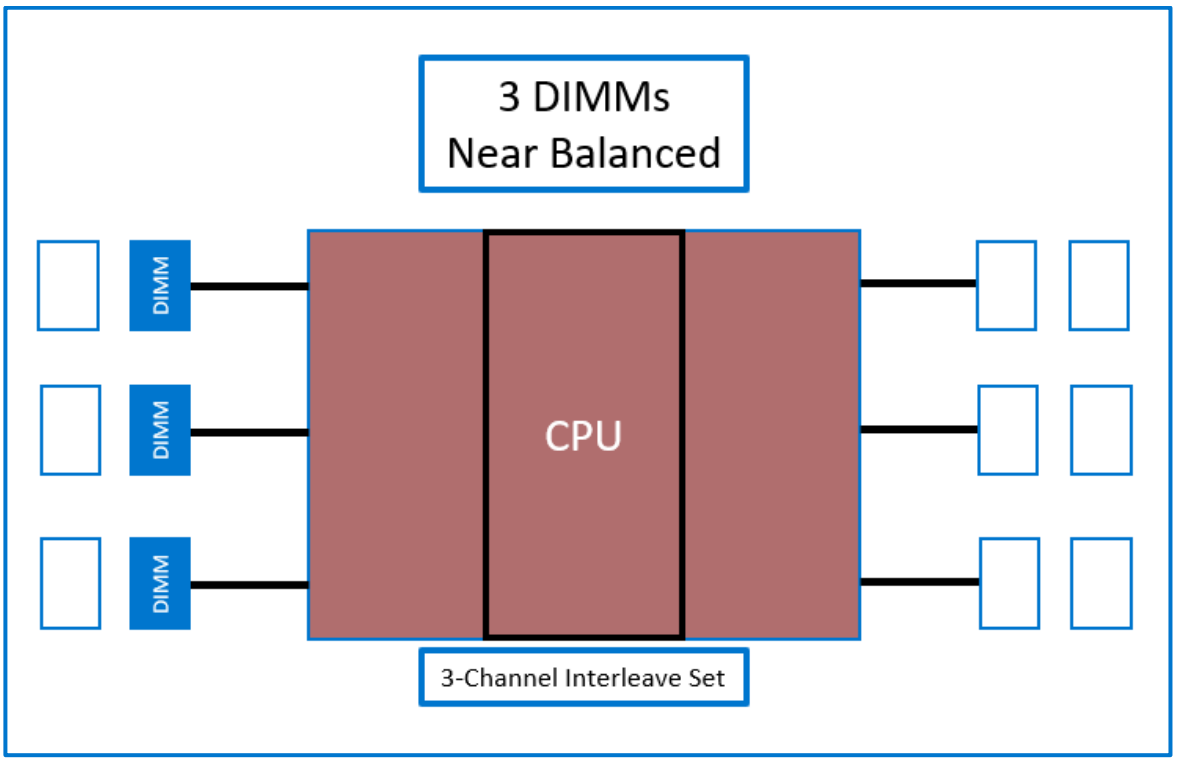

Figure 11: This configuration has one interleave set because identical memory modules are populated in the same channel column. The absence of all channels being populated makes this configuration near balanced.

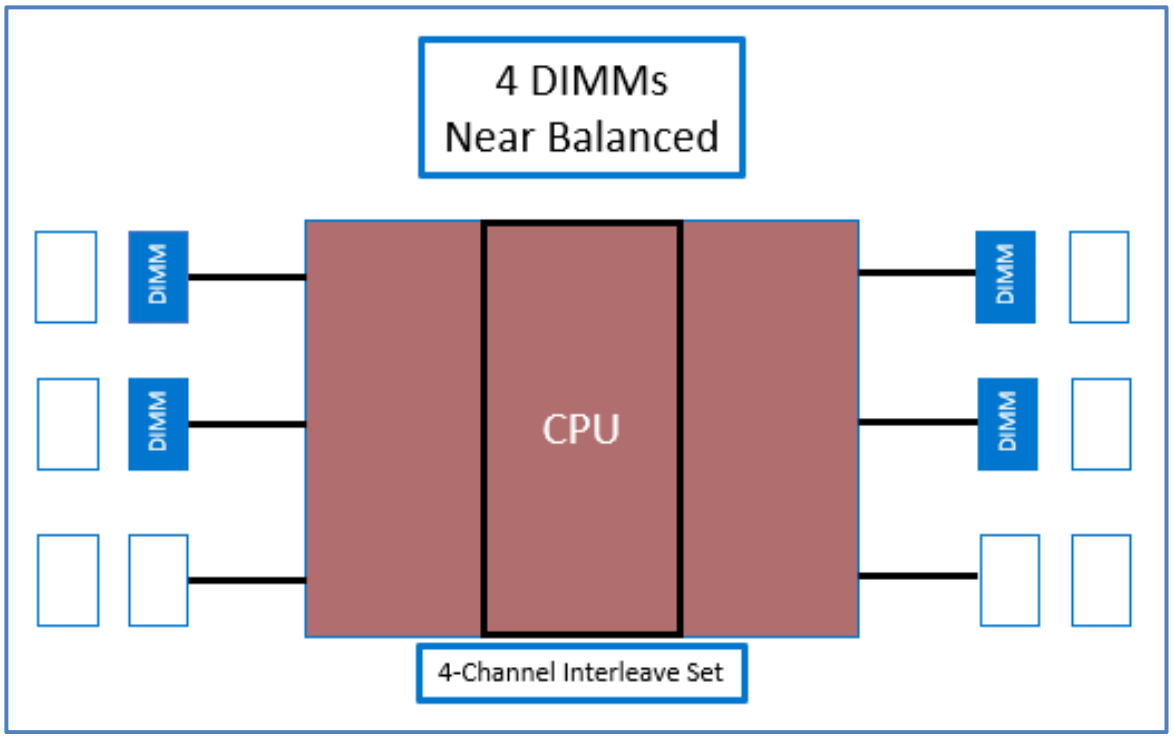

Figure 12: This configuration has one interleave set because identical memory modules were distributed evenly across the mirrored channels. The absence of all channels being populated makes this configuration near balanced.

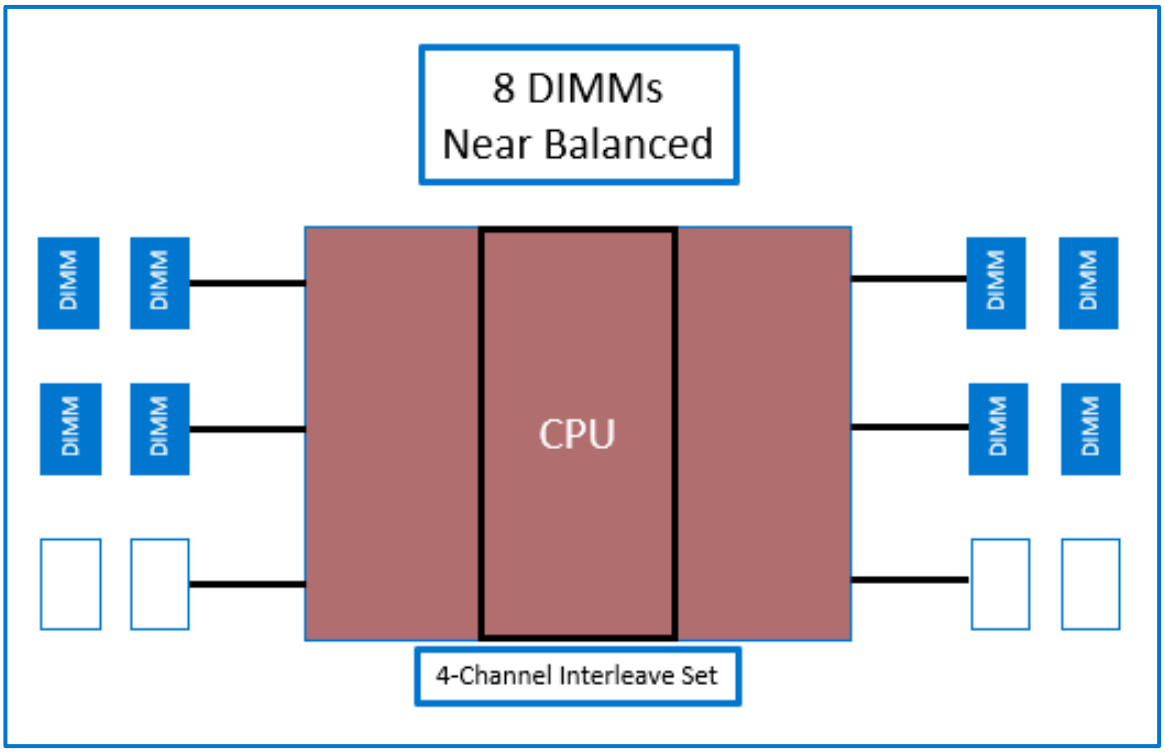

*Figure 13: This configuration has one interleave set because identical memory modules were distributed evenly across both mirrored channels **and** columns. The absence of all channels and columns being populated makes this configuration near balanced.*

7. Unbalanced Memory Topology Illustrations

Unbalanced configurations create more than one interleave set by unequally distributing DIMMs across mirrored memory channel columns. These configurations **can reduce memory bandwidth by up to 33% from its maximum potential.** Dell EMC does not recommend populating memory in an unbalanced configuration.

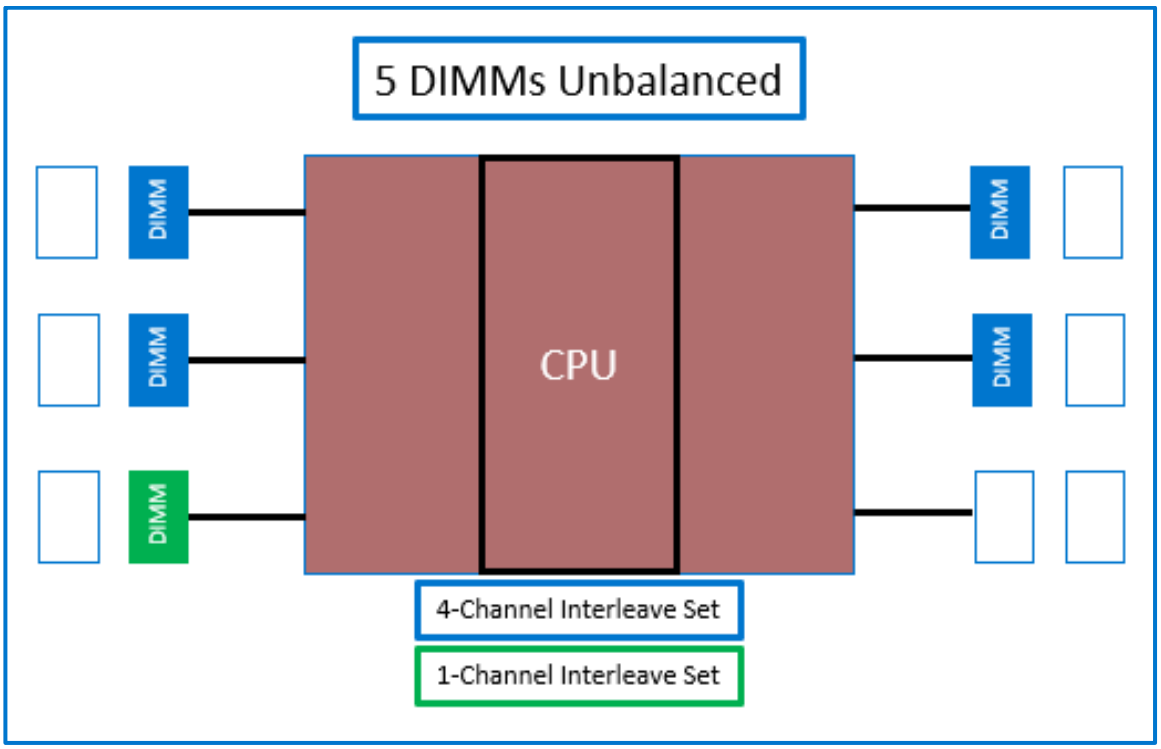

Figure 14: This configuration is unbalanced because the bottom left channel is populated while the bottom right channel is not. An additional, undesired interleave set has been introduced to accommodate the isolated memory module in the bottom left.

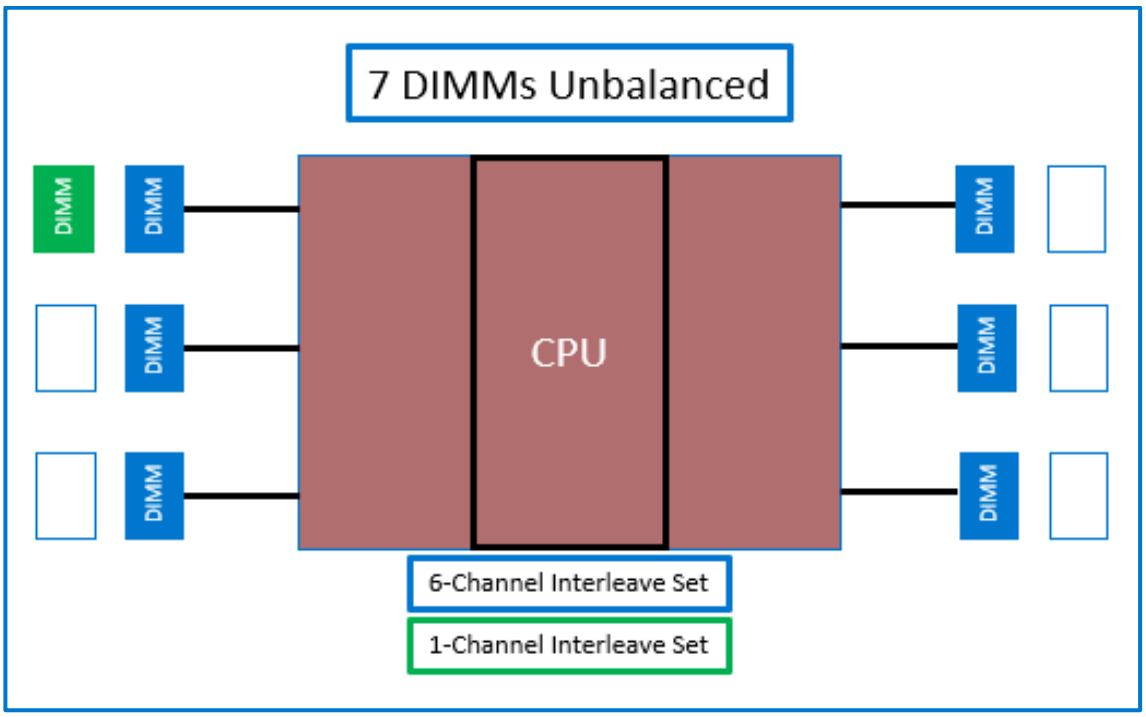

Figure 15: This configuration is unbalanced because the top left channel has both slots populated while the top right channel does not. An additional, undesired interleave set has been introduced to accommodate the isolated memory module in the top left.

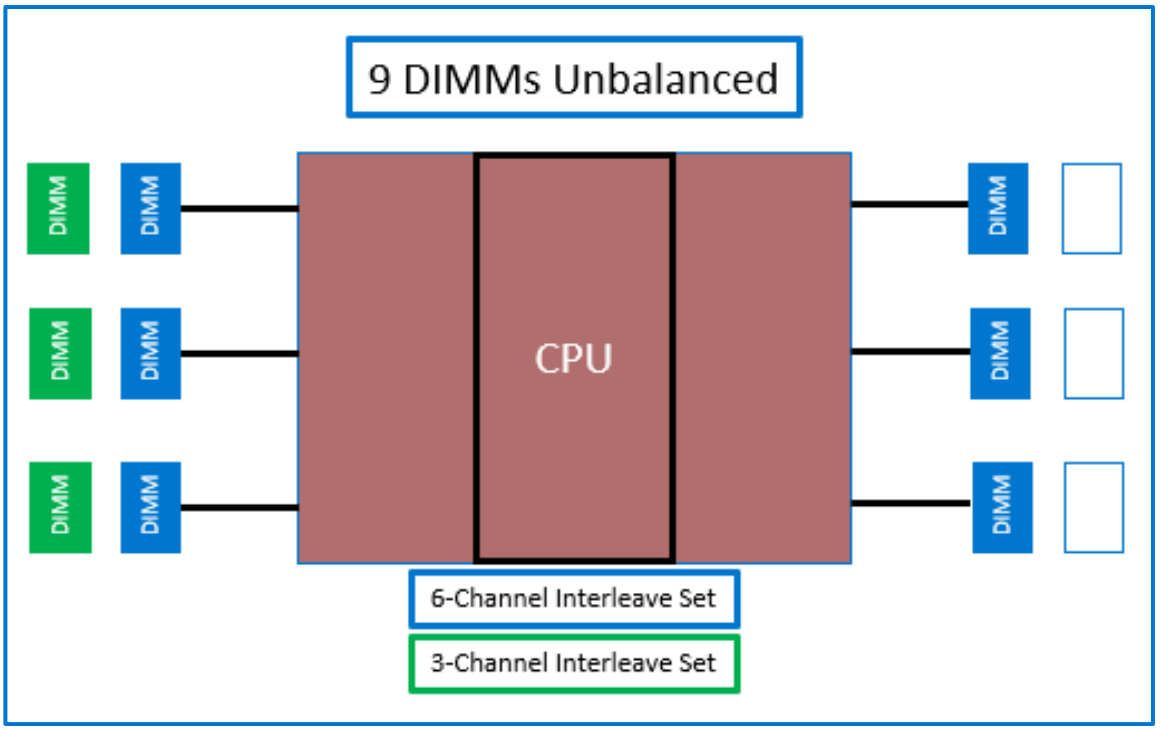

Figure 16: This configuration is unbalanced because both left controller slots are fully populated with memory modules while the right controller slots are not. An additional, undesired interleave set has been introduced to accommodate the isolated memory modules on the far-left column.

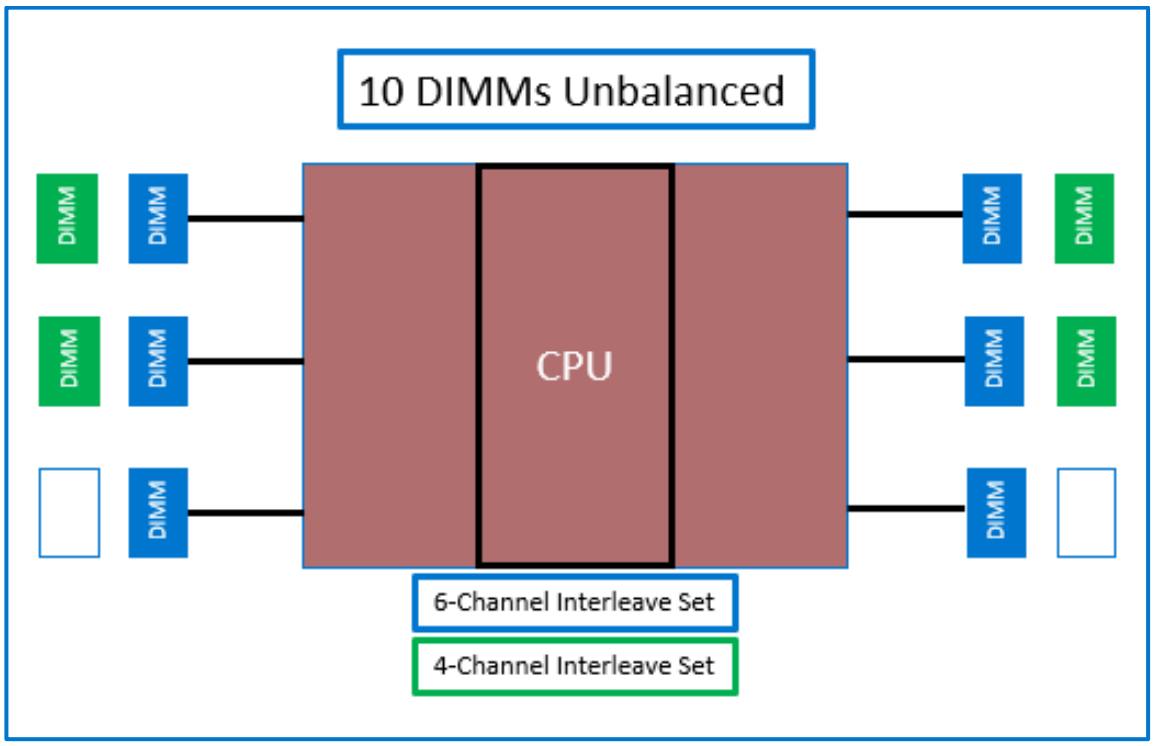

Figure 17: This configuration is unbalanced because the top and middle channels are fully populated, while the bottom channels are only partially populated. An additional, undesired interleave set has been introduced to accommodate the isolated memory modules in the bottom, inner slots.

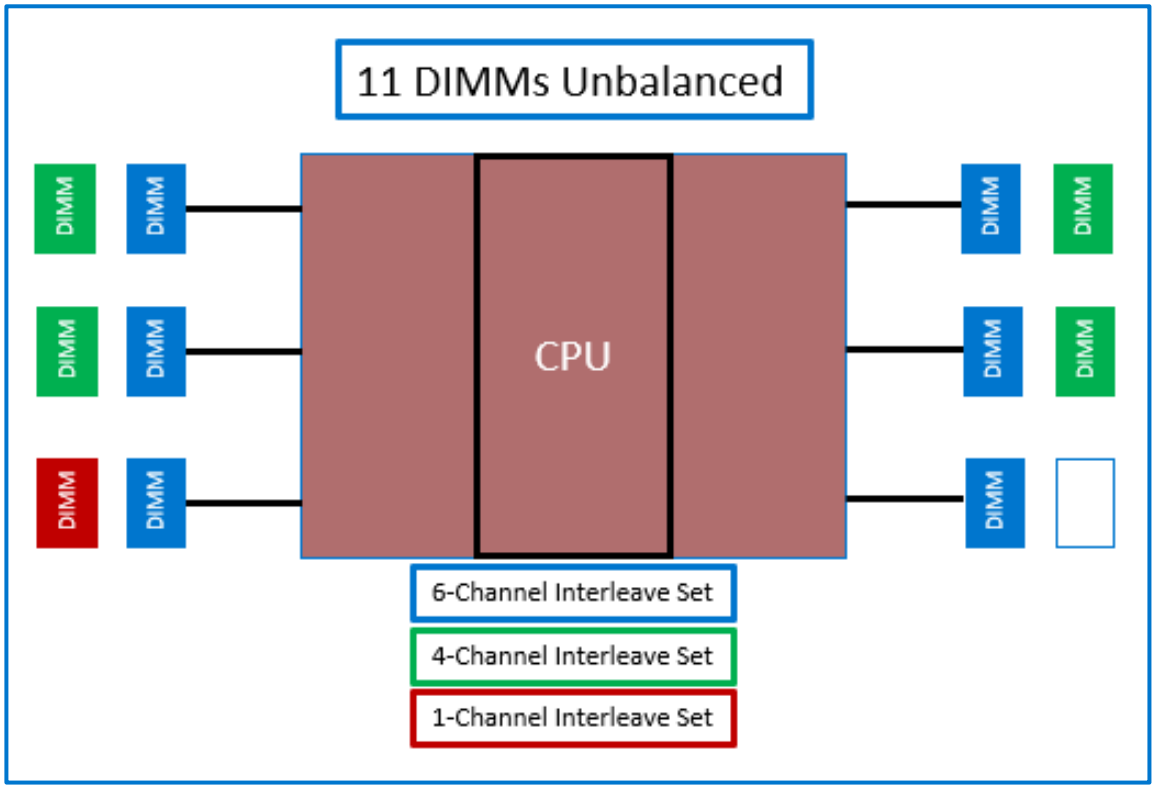

Figure 18: This configuration is unbalanced for two reasons. First, the top and middle channels are fully populated while the bottom channel is not. Second, the bottom left channel is fully populated with two DIMMs while the bottom right channel only has one. This has caused two additional, undesired interleave sets to be introduced to accommodate these isolated memory module groups.

8. Conclusion

Balancing memory with 2nd Generation Intel® Xeon™ scalable processors increases memory bandwidth and reduces memory access latency. If memory is populated into a near balanced or unbalanced configuration, **memory bandwidth can be reduced by up to 33% from its maximum potential.**

Applying the balanced memory guidelines demonstrated in this whitepaper will ensure that both memory bandwidth and memory access latency are optimized, therefore ensuring peak memory performance for Dell EMC PowerEdge servers.

9. References

1 https://www.intel.com/content/www/us/en/products/processors/xeon/scalable.html

2 https://www.streetdirectory.com/travel_guide/124468/hardware/computer_memory_controllers_how_they_work.html

3 https://www.computerhope.com/jargon/d/dual-channel-memory.htm

4 https://www.computerhope.com/jargon/m/memoslot.htm

5 https://www.geeksforgeeks.org/memory-interleaving/