

Lenovo

Balanced Memory Configurations for 2-Socket Servers with 3rd-Gen Intel Xeon Scalable Processors

Demonstrates balanced memory guidelines for third-generation Intel Xeon SP “Ice Lake” processors

Compares the performance of balanced and unbalanced memory configurations

Explains memory interleaving and its importance

Provides tips on how to balance memory and maximize performance

Nathan Pham
Joseph Jakubowski
Larry Cook



Abstract

Configuring a server with balanced memory is important for maximizing its memory bandwidth and overall performance. Lenovo® ThinkSystem 2-socket servers running Intel 3rd Generation Xeon Scalable processors (formerly codenamed “Ice Lake”) have eight memory channels per processor and up to two DIMM slots per channel, so it is important to understand what is considered a balanced configuration and what is not.

This paper defines three balanced memory guidelines that will guide you to select a balanced memory configuration. Balanced and unbalanced memory configurations are presented along with their relative measured memory bandwidths to show the effect of unbalanced memory. Suggestions are also provided on how to produce balanced memory configurations.

This paper is for customers and for business partners and sellers wishing to understand how to maximize the performance of Lenovo ThinkSystem 2-socket servers with Intel 3rd Generation Xeon Scalable processors.

At Lenovo Press, we bring together experts to produce technical publications around topics of importance to you, providing information and best practices for using Lenovo products and solutions to solve IT challenges.

See a list of our most recent publications at the Lenovo Press web site:

<http://lenovopress.com>

Do you have the latest version? We update our papers from time to time, so check whether you have the latest version of this document by clicking the **Check for Updates** button on the front page of the PDF. Pressing this button will take you to a web page that will tell you if you are reading the latest version of the document and give you a link to the latest if needed. While you’re there, you can also sign up to get notified via email whenever we make an update.

Contents

Introduction	3
Memory interleaving	4
Balanced memory configurations	4
About the tests	5
Applying the balanced memory configuration guidelines	5
Summary of the performance results	10
Maximizing memory bandwidth	10
Summary	11
Authors	12
Notices	13
Trademarks	14

Introduction

The memory subsystem is a key component of Intel 3rd Generation Xeon Scalable server architecture which can greatly affect overall server performance. When properly configured, the memory subsystem can deliver extremely high memory bandwidth and low memory access latency. When the memory subsystem is incorrectly configured, memory bandwidth available to the server can become limited and overall server performance can be reduced.

This paper explains the concept of balanced memory configurations that yield the highest possible memory bandwidth from the Intel 3rd Generation Xeon Scalable architecture. Memory configuration and performance for all supported memory configurations are shown and discussed to illustrate their effect on memory subsystem performance.

This paper specifically covers the Intel 3rd Generation Xeon Scalable processor family, code named “Ice Lake”.

Cooper Lake processors: The 4-socket-capable Intel 3rd Generation Xeon Scalable processors, formerly codenamed “Cooper Lake”, follow the same rules/performance guidance as the 2nd Generation 2-socket processors (“Cascade Lake”).

For a discussion about balanced memory configurations for those processors, as used in the ThinkSystem SR860 V2 and SR850 V2, see the paper *Balanced Memory Configurations with Second-Generation Intel Xeon Scalable Processors*, available from <https://lenovopress.com/1p1089>.

Figure 1 illustrates how Intel 3rd Generation Xeon Scalable processor’s memory controllers are connected to memory DIMM slots.

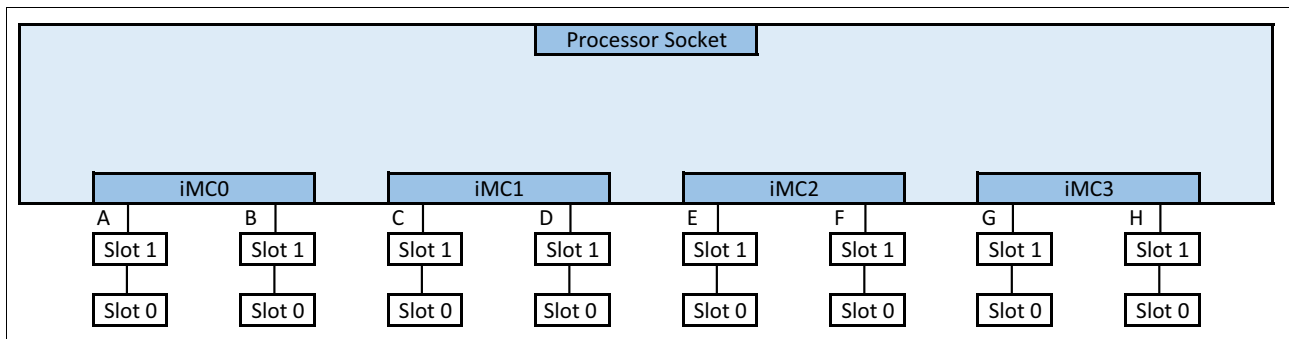


Figure 1 Intel 3rd Generation Xeon Scalable processor with four integrated Memory Controllers (iMC), eight memory channels, and sixteen memory DIMM slots.

Each integrated Memory Controller (iMC) supports two memory channels as below:

- ▶ iMC0 supports channels A and B
- ▶ iMC1 supports channels C and D
- ▶ iMC2 supports channels E and F
- ▶ iMC3 supports channels G and H

To illustrate various memory topologies for a processor, different memory configurations will be designated as A:B:C:D:E:F:G:H where each letter indicates the number of memory DIMMs populated on each memory channel. As an example, a 2:2:2:2:1:1:1:1 memory configuration has 2 memory DIMMs populated on channels A, B, C, D and 1 memory DIMM populated on channels E, F, G, H.

Memory interleaving

The Intel 3rd Generation Xeon Scalable processor family optimizes memory accesses by creating interleave sets across the memory controllers and memory channels. For example, if two memory channels have the same total memory capacity, a 2-channel interleave set is created across the memory channels.

Interleaving enables higher memory bandwidth by spreading contiguous memory accesses across more memory channels rather than sending all memory accesses to one memory channel. In order to form an interleave set between two channels, the two channels are required to have the same channel memory capacity.

If one interleave set cannot be formed for a particular memory configuration, it is possible to have multiple interleave sets. When this happens, performance of a specific memory access depends on which memory region is being accessed and how many memory DIMMs comprise the interleave set. For this reason, memory bandwidth performance on memory configurations with multiple interleave sets can be inconsistent. Contiguous memory accesses to a memory region with fewer channels in the interleave set will have lower performance compared to accesses to a memory region with more channels in the interleave set.

Figure 2 illustrates a 4-channel interleave set which results from populating identical memory DIMMs on channels A, C, E, G. This 4-channel interleave set interleaves across memory controllers and between memory channels. Consecutive addresses alternate between memory controllers with every fourth address going to each memory channel.

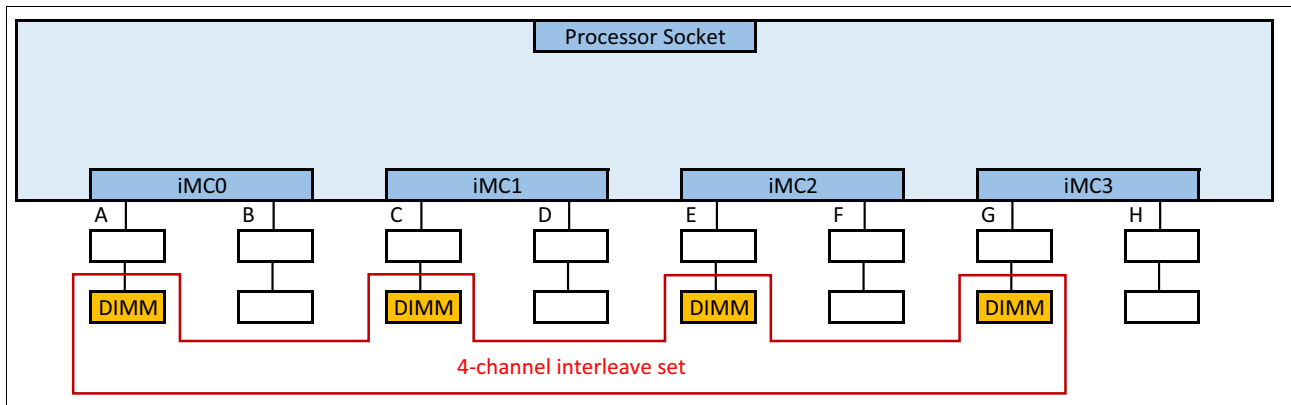


Figure 2 4-channel interleave set across memory controllers and between memory channels

Balanced memory configurations

Balanced memory configurations enable optimal interleaving which maximizes memory bandwidth. Per Intel memory population rules, channels A, E, C, G must be populated with the same total capacity per channel if populated, and channels B, D, F, H must be populated with the same total capacity per channel if populated.

The basic guidelines for a balanced memory subsystem are as follows:

1. All populated memory channels should have the same total memory capacity and the same number of ranks per channel.
2. All memory controllers on a processor socket should have the same configuration of memory DIMMs.

3. All processor sockets on the same physical server should have the same configuration of memory DIMMs.

Tip: We will refer to the above guidelines as *Balanced Memory Guidelines 1, 2 and 3* throughout this brief.

About the tests

STREAM Triad is a simple, synthetic benchmark designed to measure sustainable memory bandwidth. Its intent is to measure the highest memory bandwidth available. STREAM Triad will be used to measure the sustained memory bandwidth of various memory configurations supported by Intel 3rd Generation Xeon Scalable processors. Unless otherwise stated, all test configurations were done at the same memory speed, 3200MHz.

For more information about STREAM Triad, see the following web page:

<http://www.cs.virginia.edu/stream/>

Applying the balanced memory configuration guidelines

We will start with the assumption that balanced memory guideline 3 (described in “Balanced memory configurations” on page 4) is followed: all processor sockets on the same physical server have the same configuration of memory DIMMs. Therefore, we only have to look at one processor socket to describe each memory configuration.

When installing memory DIMMs into your server, follow the DIMM installation sequence for your specific server model. The following rules must be followed when populating memory DIMMs on Intel 3rd Generation Xeon Scalable processor platform:

- ▶ Max two different DIMM capacities can be supported per system.
- ▶ Channel A, C, E, G must be populated with the same total capacity per channel if populated.
- ▶ Channel B, D, F, H must be populated with the same total capacity per channel if populated.
- ▶ Populate higher electrical loading (higher rank, higher capacity) in slot 0 (further away from CPU).

In our lab measurements, all memory DIMMs used were 32GB dual-rank (2R) RDIMMs. The examples in this brief follows the recommended memory population sequence as shown in Table 1 below.

Table 1 Memory population sequence for Intel 3rd Generation Xeon Scalable Processors

iMC		iMC0				iMC1				iMC2				iMC3			
Channel		A		B		C		D		E		F		G		H	
DIMM slot		0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
Number of DIMMs	1	X															
	2	X				X											
	4	X				X				X				X			
	6	X		X		X				X		X		X			
	8	X		X		X		X		X		X		X		X	
	12	X	X	X	X	X	X			X	X	X	X	X	X		
	16	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

For a more complete memory rules and population guide, please refer to product guide for each specific ThinkSystem V2 server that uses the 3rd Gen processors:

<https://lenovopress.com/search#term=icelake&rt=product-guide>

Tip: Some ThinkSystem V2 servers only implement 1 DIMM per channel. Take that into consideration when reviewing the memory recommendations.

Configuration with 1 DIMM – unbalanced

We will start with one memory DIMM which yields the 1:0:0:0:0:0:0:0 memory configuration shown in Figure 3.

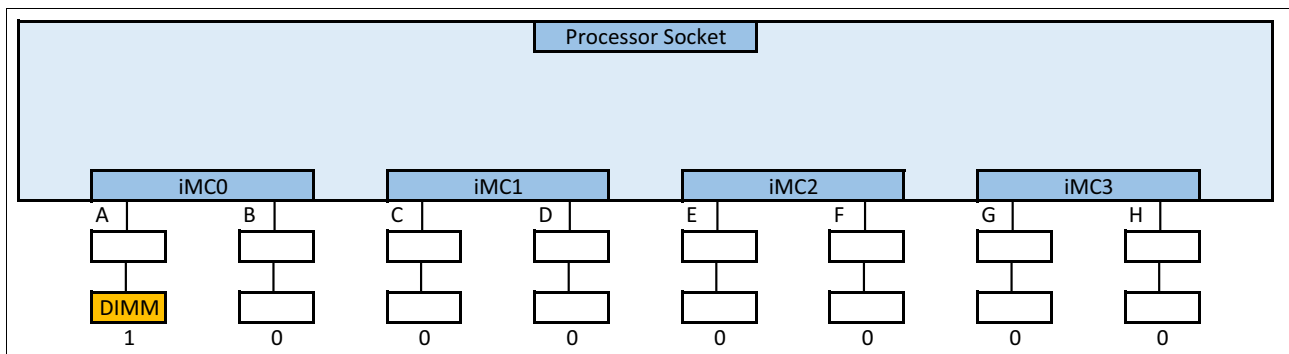


Figure 3 1:0:0:0:0:0:0:0 memory configuration, relative memory bandwidth: 13%

Balanced memory guideline 2 is not followed with only one iMC populated with memory DIMM. This is an unbalanced memory configuration.

A single 1-channel interleave set is formed. Having only one memory channel populated with memory greatly reduces the memory bandwidth of this configuration which was measured at 13% or about one eighth of the full potential memory bandwidth.

The best way to increase the memory bandwidth of this configuration is by using more memory DIMMs. Two 16GB RDIMMs populated on two channels A and C would provide the same memory capacity while nearly doubling the memory bandwidth.

Configuration with 2 DIMMs – unbalanced

The recommended memory configuration with 2 memory DIMMs is the 1:0:1:0:0:0:0:0 configuration as shown in Figure 4.

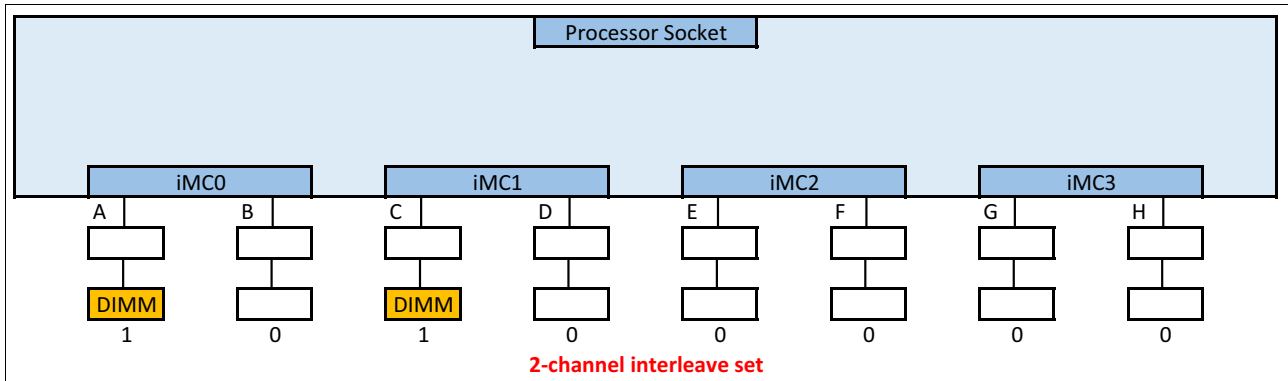


Figure 4 1:0:1:0:0:0:0:0 memory configuration, relative memory bandwidth: 25%

This memory configuration follows guideline 1, but not guideline 2 since not all iMCs were populated with identical memory configuration. This is an unbalanced memory configuration.

A single 2-channel interleave set is formed. Only two memory channels are populated with memory which greatly reduces the memory bandwidth of this memory configuration to about 25% or one fourth of the full potential memory bandwidth.

Configuration with 4 DIMMs – balanced

The recommended configuration with 4 memory DIMMs is 1:0:1:0:1:0:1:0 memory config as shown in Figure 5.

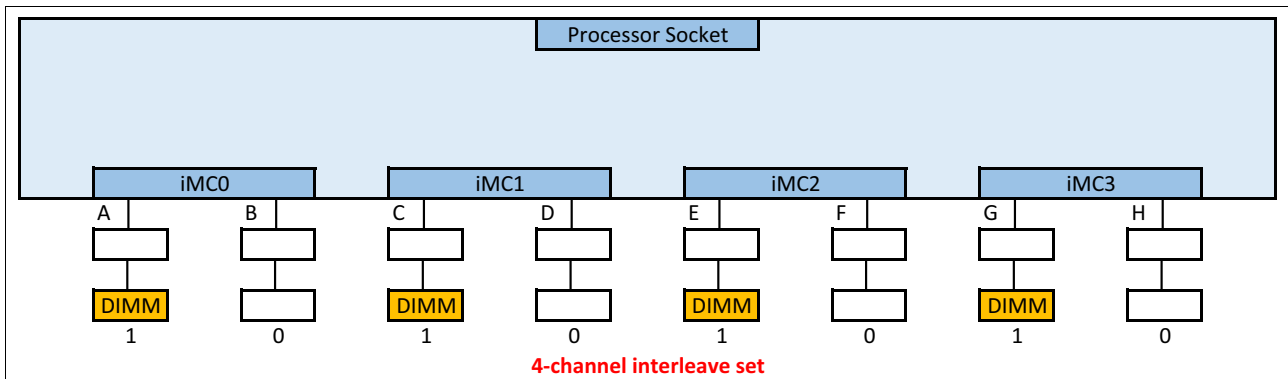


Figure 5 1:0:1:0:1:0:1:0 memory configuration, relative memory bandwidth: 50%

This memory configuration follows both memory population guidelines 1 and 2. All populated channels have the same channel capacity, and memory configurations are identical with all iMCs. This is a balanced memory configuration.

A single 4-channel interleave set is formed. Although it is a balanced memory configuration, only four memory channels were populated with memory DIMMs. Memory bandwidth is measured at 50% of one half of the full potential memory bandwidth.

Configuration with 6 DIMMs – unbalanced

The recommended configuration with 6 memory DIMMs is 1:1:1:0:1:1:1:0 memory configuration as shown in Figure 6.

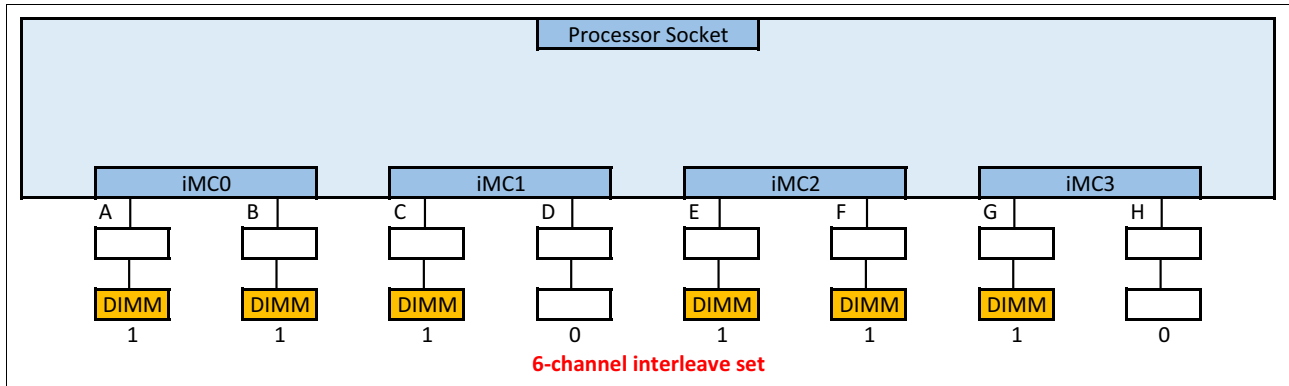


Figure 6 1:1:1:0:1:1:1:0 memory configuration, relative memory bandwidth: 75%

This memory configuration follows memory population guideline 1, but not guideline 2. Memory configurations are not identical between the iMCs. This is an unbalanced memory configuration.

A single 6-channel interleave is formed, and memory bandwidth measured at 75% of the full potential memory bandwidth. Even though a 6 DIMM population is unbalanced, the 3rd Generation Intel Xeon Scalable Processor design includes the technology to create a single 6-channel interleave which results in relatively good performance. Without this technology, multiple interleave sets would have been formed and performance would be degraded.

Configuration with 8 DIMMs – balanced

The recommended configuration with 8 DIMMs is 1:1:1:1:1:1:1:1 memory configuration as shown in Figure 7.

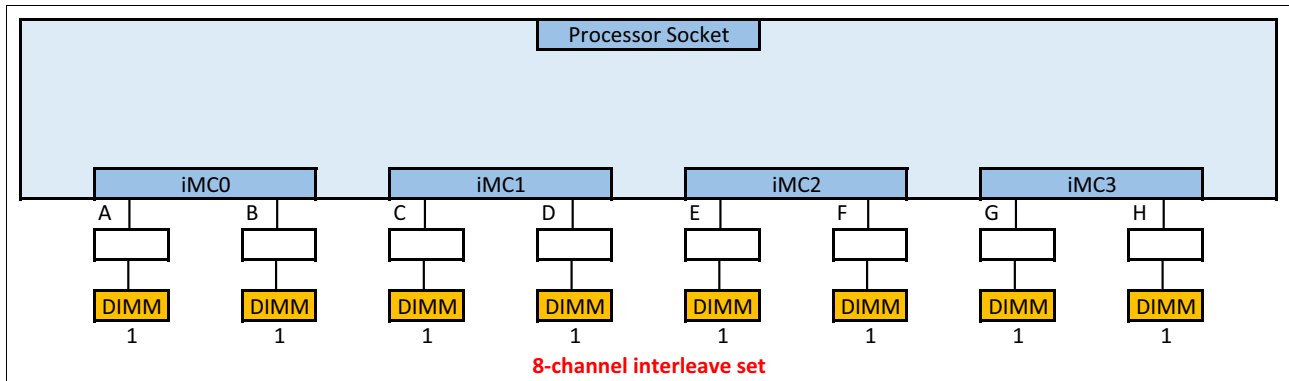


Figure 7 1:1:1:1:1:1:1:1 memory configuration, relative memory bandwidth: 100%

This memory configuration follows both guidelines 1 and 2. All channels were populated with the same memory capacity, and memory configurations were identical between all iMCs. This is a balanced memory configuration.

A single 8-channel interleave is formed, and memory bandwidth measured at 100% of the full potential memory bandwidth.

Configuration with 12 DIMMs – unbalanced

The recommended configuration with 12 DIMMs is 2:2:2:0:2:2:2:0 memory configuration as shown in Figure 8.

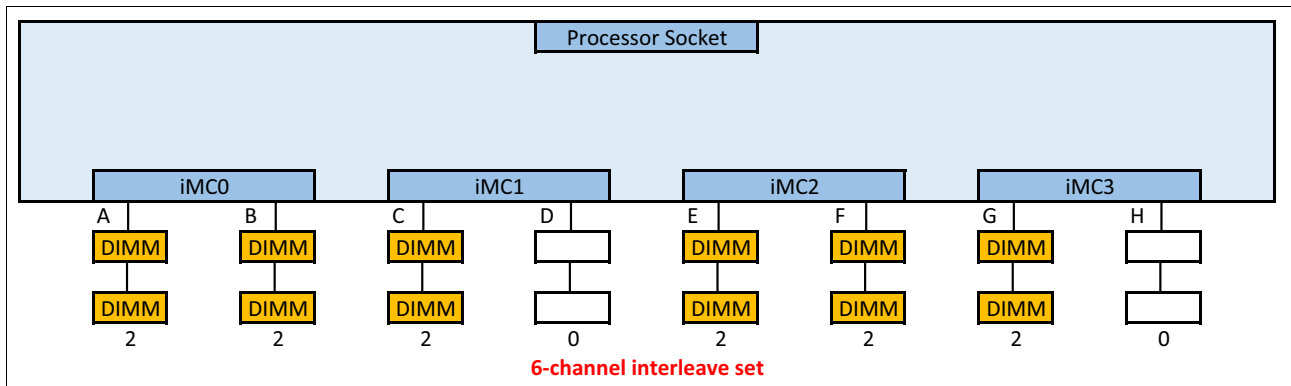


Figure 8 2:2:2:0:2:2:2:0 memory configuration, relative memory bandwidth: 72%

This memory configuration follows guideline 1, but not guideline 2. This is an unbalanced memory configuration.

A single 6-channel interleave is formed, and memory bandwidth measured at 72% of the full potential memory bandwidth. Similar to the unbalanced 6 DIMM population, the 3rd Generation Intel Xeon Scalable Processor design includes the technology to create a single 6-channel interleave which results in relatively good performance. Without this technology, multiple interleave sets would have been formed and performance would be degraded.

Configuration with 16 DIMMs – balanced

This is a fully populated configuration 2:2:2:2:2:2:2:2 as shown in Figure 9.

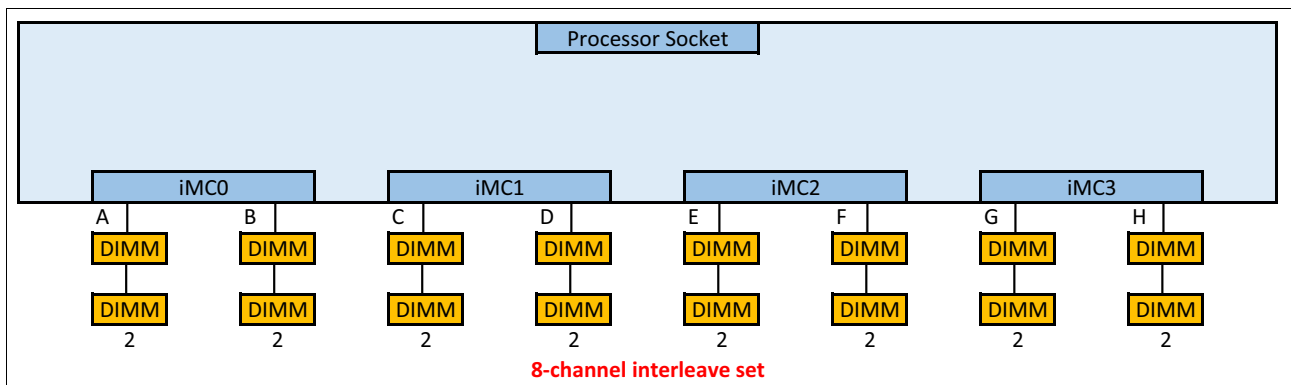


Figure 9 2:2:2:2:2:2:2:2 memory configuration, relative memory bandwidth: 98%

This is a fully populated memory configuration. It follows both guideline 1 and 2. This is a balanced memory configuration.

Each channel is populated with two dual-rank (2R) DIMMs, so the total number of ranks per channel is 4R. Memory bandwidth measured at 98% of the full potential memory bandwidth.

Both 8-DIMM and 16-DIMM memory configurations have all channels populated. The 8-DIMM configuration has one DIMM populated per channel, or total 2R per channel while the 16-DIMM configuration has two DIMMs populated per channel, or total 4R per channel. In previous Xeon processor generations, having 4R per channel resulted in slightly better memory bandwidth performance compared to having 2R per channel. This behavior is reversed with Intel 3rd Generation Xeon Scalable processors. In this set of measurements, having 4R per channel with the 16-DIMM configuration resulted in slightly lower memory bandwidth compared to having 2R per channel with the 8-DIMM configuration. This behavior has also been confirmed by Intel.

Summary of the performance results

Table 2 shows a summary of the relative memory bandwidth of all the memory configurations that were evaluated. It also shows the number of interleave sets formed for each and whether it is a balanced or unbalanced memory configuration.

Table 2 Summary of all memory configurations with their relative memory bandwidth performance

Number of DIMMs Populated	Configuration	Number of Interleave Set(s)	Relative Performance	Balanced or Unbalanced
1	1:0:0:0:0:0:0	1	13%	Unbalanced
2	1:0:1:0:0:0:0	1	25%	Unbalanced
4	1:0:1:0:1:0:1:0	1	50%	Balanced
6	1:1:1:0:1:1:1:0	1	75%	Unbalanced
8	1:1:1:1:1:1:1:1	1	100%	Balanced
12	2:2:2:0:2:2:2:0	1	72%	Unbalanced
16	2:2:2:2:2:2:2:2	1	98%	Balanced

When using the same memory DIMM, only memory configurations with 8 or 16 memory DIMMs provide the full potential memory bandwidth. These are the best memory configurations for performance. Balanced memory configuration can also be achieved with four memory DIMMs, but this configuration does not populate all the memory channels which reduces its memory bandwidth and performance.

Balanced memory configurations are the only memory configurations that should be used if memory bandwidth and performance are important.

Maximizing memory bandwidth

To maximize the memory bandwidth of a server, the following rules should be followed:

1. Balance the memory across the processor sockets – all processor sockets on the same physical server should have the same configuration of memory DIMMs.
2. Balance the memory across the memory controllers – all memory controllers on a processor socket should have the same configuration of memory DIMMs.
3. Balance the memory across the populated memory channels – all populated memory channels should have the same total memory capacity and the same total number of ranks.

Peak memory performance is achieved with 8 or 16 DIMMs per socket. Given a memory capacity requirement per server, follow these steps to get an optimal memory bandwidth configuration for your requirement:

1. Determine your needed memory capacity per socket.
2. Divide this memory capacity by eight to determine the minimum memory capacity needed per DDR channel.
3. Round this calculated memory capacity per channel up to the closest capacity available with 1 DIMM Per Channel (DPC) or 2 DPC combination, and
4. Populate your server with eight identical DIMM combination per channel derived from step 3.
5. When two DIMMs are populated on a memory channel, the total number of ranks per channel should be an even number to achieve the best memory performance.

For example, two 2-rank DIMMs per memory channel will have higher performance when compared to one 2-rank DIMM plus one 1-rank DIMM per memory channel.

Examples:

- ▶ If 512GB of total memory capacity is needed per socket, you can populate each socket with 16x 32GB DIMMs.
- ▶ If 768GB of total memory capacity is needed per socket, each DDR channel needs to be populated with 96GB (768/8) of memory, which can be achieved using one 64GB DIMM + one 32GB DIMM. You can populate each socket with 8x 64GB DIMMs and 8x 32GB DIMMs. Using all 2-rank DIMMs will result in the best performance.

Summary

Overall server performance is affected by the memory subsystem which can provide both high memory bandwidth and low memory access latency when properly configured. Balancing memory across the memory controllers and the memory channels produces memory configurations which can efficiently interleave memory references among its DIMMs producing the highest possible memory bandwidth. An unbalanced memory configuration can reduce the total memory bandwidth to as low as 13% of a balanced memory configuration with 8 or 16 identical DIMMs installed per processor.

Implementing all three of the balanced memory guidelines described in this paper results in balanced memory configuration and produces the best possible memory bandwidth and overall performance.

Lenovo recommends installing balanced memory population with 4, 8, or 16 DIMMs per socket. Peak memory performance is achieved with 8 or 16 DIMMs per processor.

Authors

This paper was produced by the following team of specialists:

Nathan Pham is a Senior Engineer in the Lenovo Infrastructure Solution Group Performance Laboratory in Morrisville, NC. He has spent over 19 years between IBM and Lenovo working on system performance. His focus has been in system performance architecture, modeling, analysis, and optimization. In his current role, he is responsible for all aspects of future Lenovo ISG system architecture and design for optimized performance. Nathan holds a Bachelor of Science degree in Computer Engineering and a Master of Science degree in Computer Engineering, both from North Carolina State University.

Joe Jakubowski is the Principal Engineer and Performance Architect in the Lenovo Infrastructure Solutions Group Performance Laboratory in Morrisville, NC. Previously, he spent 30 years at IBM. He started his career in the IBM Networking Hardware Division test organization and worked on various token-ring adapters, switches and test tool development projects. For more than 25 years, he has worked in the x86 server performance organization focusing primarily on database, virtualization and new technology performance. His current role includes all aspects of Intel and AMD x86 server architecture and performance. Joe holds Bachelor of Science degrees with Distinction in Electrical Engineering and Engineering Operations from North Carolina State University and a Master of Science degree in Telecommunications from Pace University.

Larry Cook is a x86 Performance Engineer in the Lenovo Infrastructure Solution Group Performance Laboratory located in Morrisville, NC. He has 13 years of experience between IBM and Lenovo with enterprise servers. For 8 of those years, his focus has been centered on system performance analysis, development, optimization, and benchmarking. In his current role within the System Performance Verification (SPV) group, he focuses on CPU and memory subsystem readiness on 2-socket servers in preparation for competitive benchmarking. Larry holds a Bachelor of Science degree in Electrical Engineering with a Concentration in Computers from Tennessee State University.

Thanks to the following people for their contributions to this project:

- ▶ David Watts, Lenovo Press

This paper was based on the paper *Balanced Memory Configurations with Second-Generation Intel Xeon Scalable Processors*. Thanks to the authors of that paper:

- ▶ Dan Colglazier
- ▶ Joe Jakubowski
- ▶ Jamal Ayoubi

Notices

Lenovo may not offer the products, services, or features discussed in this document in all countries. Consult your local Lenovo representative for information on the products and services currently available in your area. Any reference to a Lenovo product, program, or service is not intended to state or imply that only that Lenovo product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any Lenovo intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any other product, program, or service.

Lenovo may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:

Lenovo (United States), Inc.
1009 Think Place - Building One
Morrisville, NC 27560
U.S.A.
Attention: Lenovo Director of Licensing

LENOVO PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. Lenovo may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

The products described in this document are not intended for use in implantation or other life support applications where malfunction may result in injury or death to persons. The information contained in this document does not affect or change Lenovo product specifications or warranties. Nothing in this document shall operate as an express or implied license or indemnity under the intellectual property rights of Lenovo or third parties. All information contained in this document was obtained in specific environments and is presented as an illustration. The result obtained in other operating environments may vary.

Lenovo may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Any references in this publication to non-Lenovo Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this Lenovo product, and use of those Web sites is at your own risk.

Any performance data contained herein was determined in a controlled environment. Therefore, the result obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

This document was created or updated on August 23, 2021.

Send us your comments via the **Rate & Provide Feedback** form found at <http://lenovopress.com/lp1517>

Trademarks

Lenovo and the Lenovo logo are trademarks or registered trademarks of Lenovo in the United States, other countries, or both. These and other Lenovo trademarked terms are marked on their first occurrence in this information with the appropriate symbol (® or ™), indicating US registered or common law trademarks owned by Lenovo at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of Lenovo trademarks is available from <https://www.lenovo.com/us/en/legal/copytrade/>.

The following terms are trademarks of Lenovo in the United States, other countries, or both:

Lenovo(logo)®

Lenovo®

The following terms are trademarks of other companies:

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

Other company, product, or service names may be trademarks or service marks of others.