

Fujitsu Server PRIMERGY

Memory performance of EPYC 9004 / 9005 Series Processor (Genoa/Turin) based Systems

This white paper explains the essential features of the memory architecture and the latest improvements in the EPYC 9004 Series Processor (Genoa) and EPYC 9005 Series Processor (Turin) based Fujitsu Server PRIMERGY and quantifies their effect on the performance of commercial applications.

Version

1.1

2025-04-08



Contents

- Introduction 3**
- Memory architecture 5**
 - DIMM slots, memory controller, and NUMA configuration5
 - Available DIMM types8
 - Memory transfer rate 10
 - BIOS parameters..... 12
 - Performant memory configurations..... 14
- Quantitative effects on memory performance 17**
 - The measuring tools..... 18
 - Memory interleaving settings..... 19
 - Memory transfer rate 21
 - Influence of the DIMM types 22
 - NUMA settings..... 24
 - Access to remote memory 25
 - Power-related BIOS setting 26
 - Impact of memory encryption setting 27
- Literature 28**

Introduction

EPYC processors have achieved significant scalability improvements with each generation by integrating multiple semiconductor dies into a single package. The 4th generation EPYC processor (Genoa) and the 5th generation EPYC processor (Turin) inherit the features of the 3rd generation EPYC processor (Milan) and includes the following improvements:

- Microarchitecture improvements: The 4th generation EPYC processor features improved branch prediction, support for AVX-512 instructions, and an increased cache size. In addition, the 5th generation EPYC processor includes additional instruction execution units and an increased L1 cache size and bandwidth.
- Increased number of cores through advanced manufacturing processes: The 4th generation EPYC processor offers models with up to 128 cores, and the 5th generation EPYC processor offers models with up to 192 cores. This is a significant increase compared to the 3rd generation EPYC processors, which offered up to 64 cores per processor.

In terms of the memory architecture, the following improvements have been made over the previous generation of EPYC processors:

- With the latest DDR5 memory, the maximum memory transfer rate is significantly improved from the previous DDR4 (1.5 times for the 4th generation EPYC, 1.9 times for the 5th generation EPYC). The 5th generation EPYC processor supports memory transfer rates of up to 6,000 MT/s.
- The number of memory channels increased to twelve, 1.5 times more than the previous generation.

4th and 5th generation EPYC processor based PRIMERGY servers now have 6 TB of memory per system with support for large 256 GB 3DS RDIMMs.

In addition, the Infinity Fabric interface, which is responsible for connecting dies, is now up to twice as fast. PCI Express for I/O connectivity has also evolved from Gen4 to Gen5, with twice the data transfer rate.

In this document, we will look at the new memory system functions of the latest server generation. On the other hand, as in the earlier issues of this white paper, this document also provides basic knowledge about the NUMA-based memory architecture which is essential when configuring powerful systems. We are dealing with the following points here:

- Due to the NUMA architecture, the memory of each processor should be configured as identical as possible. The aim of this is, in principle, to enable each processor to perform processing on its local memory.
- In order to parallelize memory access and further speed it up, the adjacent area of the physical address space is distributed to several components of the memory system. In technical terms, this is called interleaving. Interleaving is done in two dimensions. First, there are twelve memory channels per processor. Optimal interleaving over memory channels is achieved by setting the number of DIMMs installed in each processor to a multiple of twelve. In addition, interleaving among individual memory channels is realized. The definitive memory resource for this is the so-called number of ranks. The number of ranks is a DIMM sub-structure, and a group of DRAM (Dynamic Random Access Memory) chips are integrated here. Individual memory access always refers to such groups.
- Memory transfer rate affects performance. Depending on DIMM type, the number of DIMM per memory channel, and BIOS settings, they can be either 6,000, 5,600, 5,200, 4,800, 4,400, 4,000, 3,600 or 3,200 MT/s.

In this white paper, factors that affect memory performance are taken up and quantified. For quantification, we use the STREAM and SPECrate2017 Integer benchmarks. STREAM measures the memory bandwidth. SPECrate2017 Integer is used as a model for the performance of commercial

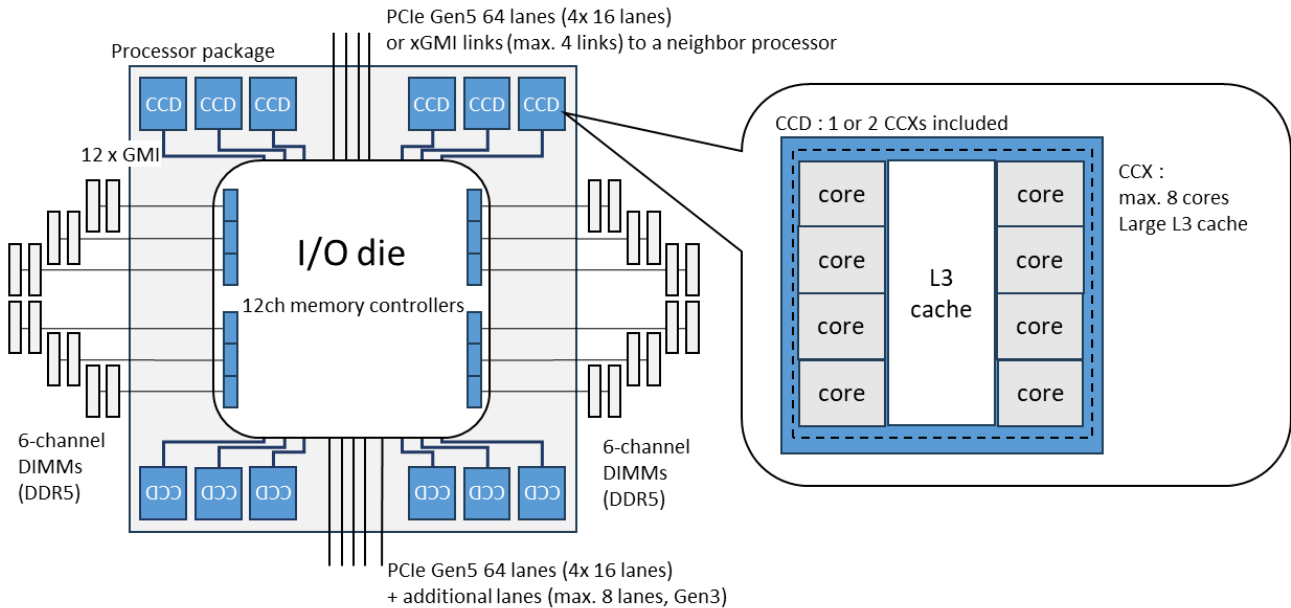
applications. Memory performance under considering BIOS settings related to power and encryption feature is also summarized at the end of this document.

Memory architecture

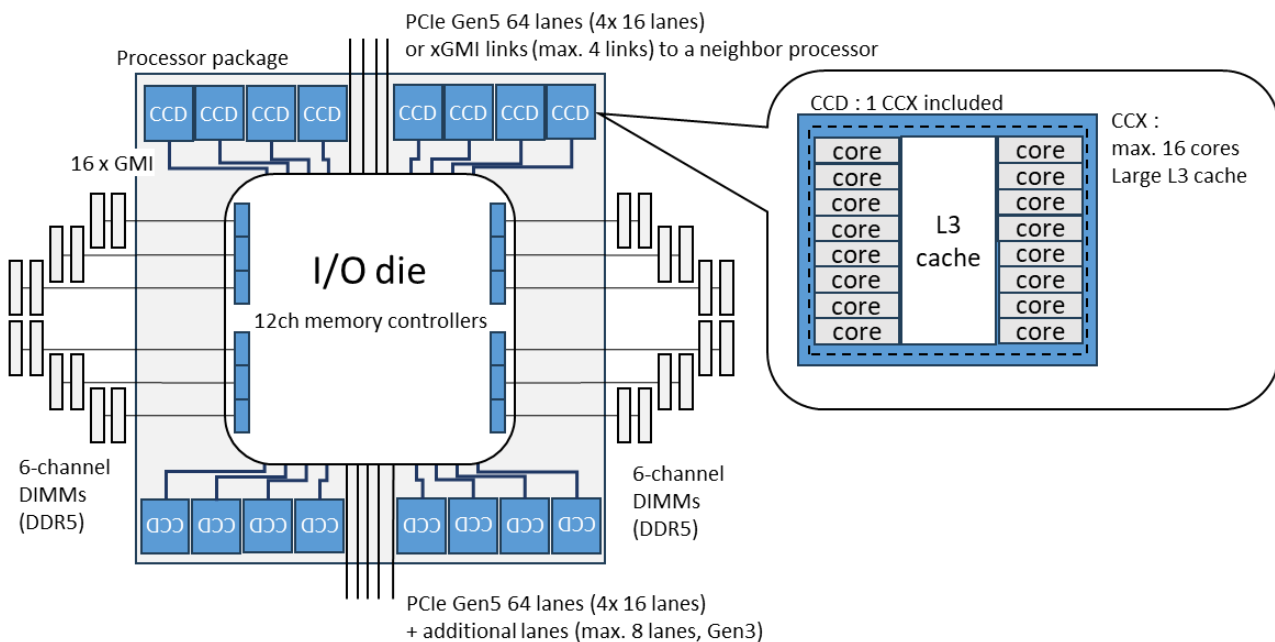
This section explains the outline of the memory system with five parts. First, we will explain the components related to the memory system in the block diagram. The second section shows the available DIMM types. The following third section describes the effect on the effective memory transfer rate. The fourth section describes the BIOS parameters that affect the memory system. The last section lists examples of memory performance optimized DIMM configuration.

DIMM slots, memory controllers, and NUMA configuration

The following figure shows the memory system architecture of the 4th generation EPYC processor (Genoa) and the 5th generation EPYC processor (Turin) based systems.



The memory system architecture of the 4th generation EPYC processor



The memory system architecture of the 5th generation EPYC processor

4th and 5th generation EPYC processors consists of a CCD (Core Complex Die) that contains the processor core and an IOD (I/O Die) that facilitates communication between CCDs and processors and provides access to memory and I/O. In the 4th generation EPYC processor, a CCD consists of one or two Core Complexes (CCX), each with up to 8 processor cores and up to 96 MB of L3 cache. In contrast, in the 5th generation EPYC processor, a CCD has one CCX with up to 16 processor cores and 32 MB L3 cache.

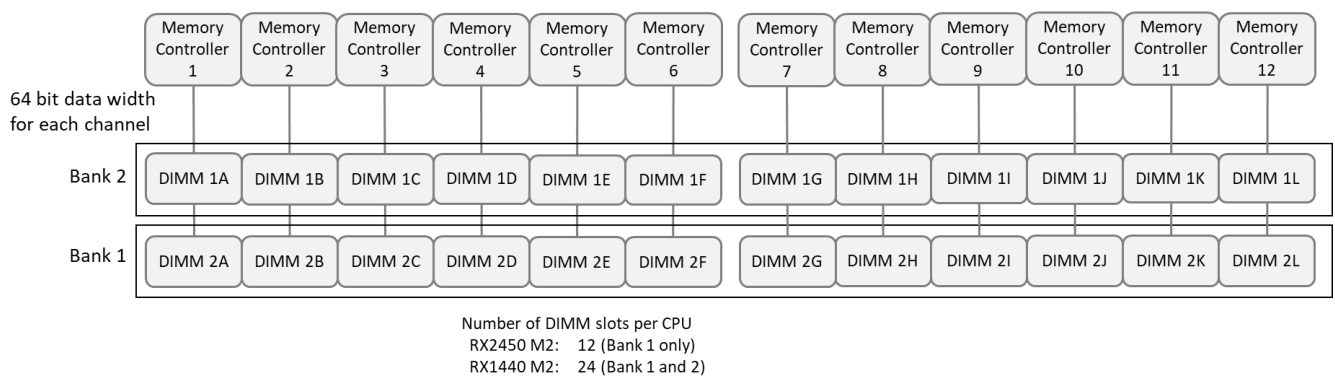
The I/O die is connected by Global Memory Interconnect (GMI) to up to twelve CCDs in the 4th generation EPYC and up to sixteen CCDs in the 5th generation EPYC. It features twelve memory controllers supporting DDR5 memory and up to 128 lanes of the latest PCI Express Gen 5 for I/O connectivity^{1,2}. There are up to four Infinity Fabric links (xGMI, External Global Memory Interconnect) for inter-processor communication. These connections form the data fabric.

DDR5 memory was first supported on the 4th generation EPYC processor. The current DDR5 specification has a maximum transfer rate of 8,800 MT/s, which is up to 2.75 times faster than the previous DDR4 memory. Although the data path width of the DDR5 DIMM is 64 bits, same as DDR4, it operates as two independent 32 bit subchannels. This greatly improves the parallel access performance over DDR4.

In 4th and 5th generation EPYC processors, each memory controller has one memory channel, for a total of twelve memory channels. Up to two DIMMs can be installed per channel for up to 24 DIMMs per processor. The PRIMERGY RX1440 M2, one processor system, supports up to 24 DIMMs, while the PRIMERGY RX2450 M2, two processors system, supports up to twelve DIMMs per processor, for a total system of up to 24 DIMMs.

On EPYC processors, changing the value of DPC (This term is used hereafter), the number of DIMMs per channel, causes a change in the memory transfer rate and affects memory performance. This is important to note. On the 5th generation EPYC processor, the number of DIMM slots per memory channel also affects the memory transfer rate.

We also use the term “memory bank” in the following. In the figure below, a group of twelve DIMMs distributed across multiple channels forms one bank. When distributing DIMMs via available slots per processor, allocating them sequentially from Bank 1 provides optimal interleaving across the entire channel. Interleaving is the main factor affecting memory performance.



The relation between memory banks and memory channels

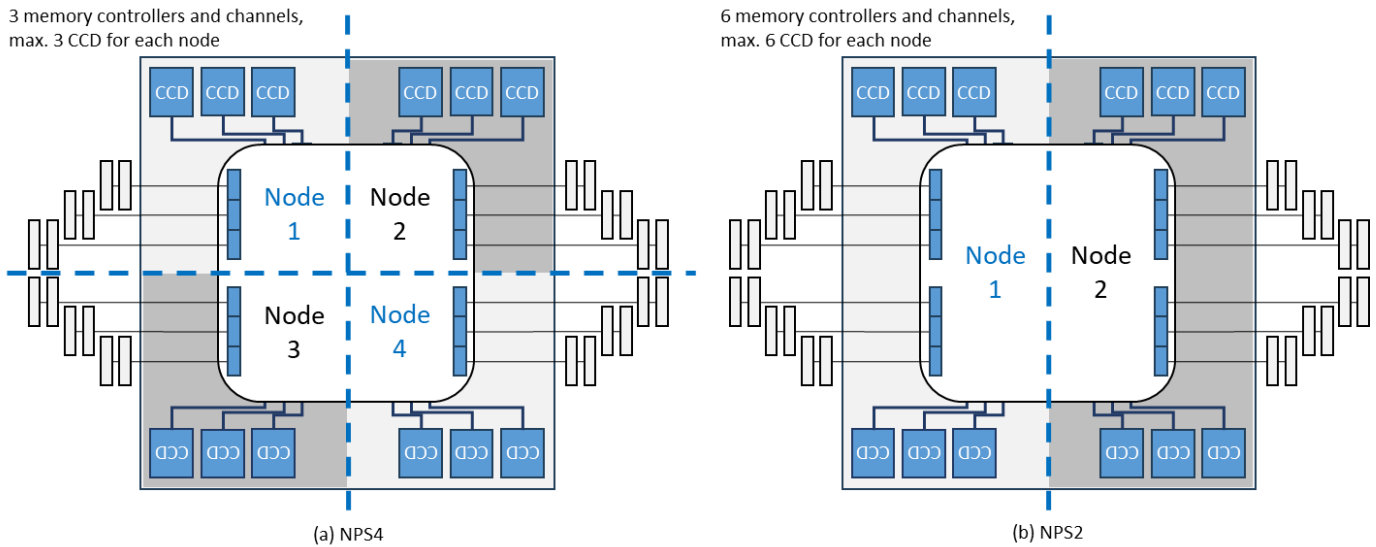
¹ The actual number of lanes available is limited by the number of PCI Express slots in the system and the number of lanes of the slots. Refer to the specification of each system.

² In a two-socket configuration, the number of lanes available is reduced because the Infinity Fabric link for the socket-to-socket connection uses PCI Express links.

For a 64-bit bandwidth of the data, the individual DRAM chips on the DIMM are responsible for 4 bits or 8 bits each (see code x4 or x8 for type name). Such a chip group is called a rank. There are DIMM types of one, two, four, or eight ranks.

The corresponding processor must be available in order to use the DIMM slots. If CPU installation does not have the maximum configuration, slots assigned to empty CPU sockets cannot be used.

In NUMA-based systems, a single processor typically forms a single NUMA node, but with 4th and 5th generation EPYC processors, you can split a processor into multiple NUMA nodes just as you would with the 3rd generation processor.



The example of splitted NUMA nodes (the 4th generation EPYC processor)

Divided NUMA nodes are associated with CCDs (processor cores and L3 cache), memory controllers, and memory channels to reduce access latency when accessing memory in the node. On the other hand, access to memory outside the node is slightly slower.

There are two ways to split NUMA nodes: four (NPS4, 4 NUMA nodes Per Socket) and two (NPS2). For NPS4, each node has three memory controllers, memory channels associated with it, and up to three CCDs. For NPS2, six memory controllers, memory channels, and up to six CCDs are associated. Interleaving among memory channels is performed between the memory channels belonging to each node. For example, in the case of NPS4, each node has a three-channel interleaving.

When a processor is divided into multiple NUMA nodes, you need to pay attention to the design of the system to ensure that the resources of all NUMA nodes are utilized effectively to achieve the performance potential of the processor.

Available DIMM types

A new DDR5 SDRAM memory module was introduced in the 4th generation EPYC processor based PRIMERGY servers. This brings the following improvement to 4th and 5th generation EPYC processor based systems.

- The EPYC processor based PRIMERGY servers now supports memory transfer rate up to 6,000 MT/s with DDR5 SDRAM. The 3rd generation EPYC processor based systems supported up to 3,200 MHz using DDR4 SDRAM. It improves the maximum memory transfer rate by the factor of 1.5 for the 4th generation EPYC processor based PRIMERGY servers, and by 1.9 times for the 5th generation EPYC processor based PRIMERGY servers.
- 4th and 5th generation EPYC processor based systems can be equipped with up to 6 TB of DRAM per system with 256 GB 3DS RDIMMs.

The following table shows the DIMMs supported by 4th and 5th generation EPYC processor based PRIMERGY servers. In DIMMs, there are Registered DIMM (RDIMM) and 3DS Registered DIMM (3DS RDIMM) types. RDIMM x4, RDIMM x8, and 3DS RDIMM cannot be mixed.

DIMM type	Control	Max. memory Transfer rate (MT/s)	Volt (V)	# of Ranks	Capacity
16GB (1x16GB) 1Rx8 DDR5-4800 R ECC	Registered	4,800	1.1	1	16 GB
32GB (1x32GB) 2Rx8 DDR5-4800 R ECC	Registered	4,800	1.1	2	32 GB
32GB (1x32GB) 1Rx4 DDR5-4800 R ECC	Registered	4,800	1.1	1	32 GB
64GB (1x64GB) 2Rx4 DDR5-4800 R ECC	Registered	4,800	1.1	2	64 GB
128GB (1x128GB) 4Rx4 DDR5-4800 3DS R ECC	3DS Registered	4,800	1.1	4	128 GB
256GB (1x256GB) 8Rx4 DDR5-4800 3DS R ECC	3DS Registered	4,800	1.1	8	256 GB
16GB (1x16GB) 1Rx8 DDR5-5600 R ECC	Registered	5,600	1.1	1	16 GB
32GB (1x32GB) 2Rx8 DDR5-5600 R ECC	Registered	5,600	1.1	2	32 GB
32GB (1x32GB) 1Rx4 DDR5-5600 R ECC	Registered	5,600	1.1	1	32 GB
64GB (1x64GB) 2Rx4 DDR5-5600 R ECC	Registered	5,600	1.1	2	64 GB
96GB (1x96GB) 2Rx4 DDR5-5600 R ECC	Registered	5,600	1.1	2	96 GB
256GB (1x256GB) 8Rx4 DDR5-5600 3DS R ECC	3DS Registered	5,600	1.1	8	128 GB
16GB (1x16GB) 1Rx8 DDR5-6400 R ECC	Registered	6,400	1.1	1	16 GB
32GB (1x32GB) 2Rx8 DDR5-6400 R ECC	Registered	6,400	1.1	2	32 GB
32GB (1x32GB) 1Rx4 DDR5-6400 R ECC	Registered	6,400	1.1	1	32 GB
64GB (1x64GB) 2Rx4 DDR5-6400 R ECC	Registered	6,400	1.1	2	64 GB
96GB (1x96GB) 2Rx4 DDR5-6400 R ECC	Registered	6,400	1.1	2	96 GB
128GB (1x128GB) 2Rx4 DDR5-6400 R ECC	Registered	6,400	1.1	2	128 GB

That being said, the essential features of the two DIMM types are as follows:

- RDIMM: The control commands of the memory controller are buffered in the register (that gave the name), which is in its own component on the DIMM. This relief for the memory channel enables configurations with up to 2DPC (DIMMs per channel).
- 3DS RDIMM: This is a RDIMM with multiple silicon dies laminated by Through Silicon Via technology based on the Three-Dimensional Stack (3DS) standard. Only one die called a master exchanges signals with the outside, and the other dies adopt an architecture that exchanges signals only with the master as a slave, enabling higher capacity and higher speed.

Which type of RDIMM or 3DS RDIMM is desirable is usually determined by the memory capacity required and the performance requirements. For example, 64GB RDIMMs with 2DPC configuration are usually cheaper than a 128GB RDIMM with 1DPC configuration, but the latter has a better performance advantage, as discussed later.

Note that some DIMM types may not be available depending on system model, processor model, and sales region.

Memory transfer rate

There are five types of memory transfer rate on the 4th generation EPYC processor based PRIMERGY servers: 4,800, 4,400, 4,000, 3,600, and 3,200 MT/s. The 5th generation EPYC processor based PRIMERGY has seven types: 6,000, 5,600, 5,200, 4,800, 4,400, 4,000, and 3,600 MT/s. The transfer rate is defined by the BIOS when the system is switched on and applies per system, not per processor.

In general, the memory transfer rate is affected by the maximum memory transfer rate of the processor model, memory specifications, the DPC configuration, and the BIOS setting. For the 5th generation EPYC processor, the memory transfer rate is also affected by the number of DIMM slots per channel. The maximum memory transfer rate of 4th and 5th generation EPYC processors is 4,800 MT/s and 6,000 MT/s, respectively, regardless of the processor model. However, the PRIMERGY RX1440 M2 server is limited to a maximum memory transfer rate of 4,800 MT/s, even with the 5th generation EPYC processor. This limitation is due to the configuration with 24 DIMM slots (2 DIMM slots per channel).

In a 2DPC configuration, the memory transfer rate is significantly reduced and varies depending on the memory type. The following table summarizes the maximum memory transfer rates for processors with different memory types and DPC configurations.

For the 4th generation EPYC processor			
DIMM type	Config uration	# or ranks per channel	Max. memory transfer rate (MT/s)
16GB (1x16GB) 1Rx8 DDR5-4800 R ECC	1DPC	1	4,800
	2DPC	2	4,000
32GB (1x32GB) 2Rx8 DDR5-4800 R ECC	1DPC	2	4,800
	2DPC	4	3,600
32GB (1x32GB) 1Rx4 DDR5-4800 R ECC	1DPC	1	4,800
	2DPC	2	4,000
64GB (1x64GB) 2Rx4 DDR5-4800 R ECC	1DPC	2	4,800
	2DPC	4	3,600
128GB (1x128GB) 4Rx4 DDR5-4800 3DS R ECC	1DPC	4	4,800
	2DPC	8	3,600
256GB (1x256GB) 8Rx4 DDR5-4800 3DS R ECC	1DPC	8	4,800
	2DPC	16	3,600

For the 5th generation EPYC processor			
DIMM type	Config uration	# or ranks per channel	Max. memory transfer rate (MT/s)
16GB (1x16GB) 1Rx8 DDR5-5600 R ECC	1DPC	1	5,600
	2DPC	2	4,000
32GB (1x32GB) 2Rx8 DDR5-5600 R ECC	1DPC	2	5,600
	2DPC	4	3,600
32GB (1x32GB) 1Rx4 DDR5-5600 R ECC	1DPC	1	5,600
	2DPC	2	4,000
64GB (1x64GB) 2Rx4 DDR5-5600 R ECC	1DPC	2	5,600
	2DPC	4	3,600
96GB (1x96GB) 2Rx4 DDR5-5600 R ECC	1DPC	2	5,600
	2DPC	4	3,600
256GB (1x256GB) 8Rx4 DDR5-5600 3DS R ECC	1DPC	8	5,600
	2DPC	16	3,600

For the 5th generation EPYC processor			
DIMM type	Config uration	# or ranks per channel	Max. memory transfer rate (MT/s)
16GB (1x16GB) 1Rx8 DDR5-6400 R ECC	1DPC	1	6,000
	2DPC	2	- ³
32GB (1x32GB) 2Rx8 DDR5-6400 R ECC	1DPC	2	6,000
	2DPC	4	- ³
32GB (1x32GB) 1Rx4 DDR5-6400 R ECC	1DPC	1	6,000
	2DPC	2	- ³
64GB (1x64GB) 2Rx4 DDR5-6400 R ECC	1DPC	2	6,000
	2DPC	4	- ³
96GB (1x96GB) 2Rx4 DDR5-6400 R ECC	1DPC	2	6,000
	2DPC	4	- ³
128GB (1x128GB) 2Rx4 DDR5-6400 R ECC	1DPC	2	6,000
	2DPC	4	- ³

The BIOS parameter “Memory Clock” allows you to choose whether to prioritize performance or power consumption to a limited extent. The available options are “DDR6000”, “DDR5600”, “DDR5200”, “DDR4800”, “DDR4400”, “DDR4000”, “DDR3600”, “DDR3200”, (corresponding to 6,000 MT/s, 5,600 MT/s, 5,200 MT/s, 4,800 MT/s, 4,400 MT/s, 4,000 MT/s, 3,600 MT/s, 3,200 MT/s respectively), and “Auto”. The default option is “Auto”, which sets the maximum possible memory transfer rate for the system configuration. Note that if you set a speed that exceeds the maximum memory transfer rate that can be configured, “Auto” is assumed. Since reducing the memory transfer rate will also reduce system performance (as described in the second part of this documented), it is recommended that you test the impact of making this setting before applying it to production.

³ On PRIMERGY RX2450 M2, only 1DPC configuration is available. DDR5-6400 DIMMs are not supported on PRIMERGY RX1440 M2.

BIOS parameters

Having looked at the BIOS parameter “Memory Clock” in the previous section, we now turn to the other BIOS options that affect the memory system. These parameters are in the submenu, Memory Configuration, under Advanced menu.

Memory parameters under Memory Configuration

The following 9 parameters are explained. The default is underlined each time.

- Memory Clock: Auto / DDR3200 / DDR3600 / DDR4000 / DDR4400 / DDR4800 / DDR5200 / DDR5600 / DDR6000
- Memory interleaving: Auto / Disabled / Enabled
- Chipselect Interleaving: Auto / Disabled / Enabled
- DRAM Scrub Time: 1 hour / 4 hours / 6 hours / 8 hours / 12 hours / 16 hours / 24 hours / 48 hours / Disabled
- TSME: Disabled / Enabled
- Power Down Enable: Disabled / Enabled
- Power Profile Selection: Efficiency Mode / High Performance Mode / Maximum IO Performance Mode / Balanced Memory Performance Mode
- NUMA nodes per socket: NPS0 / NPS1 / NPS2 / NPS4
- ACPI SRAT L3 Cache As NUMA Domain: Disabled / Enabled

The first parameter “Memory Clock” concerns the memory transfer rate and was dealt with in the last section in detail.

The next two parameters, “Memory interleaving” and “Chipselect Interleaving”, are the settings for the memory interleaving.

“Memory interleaving” configures the interleaving between the memory channels of the processor. “Enabled” enables the channel interleaving between memory channels in the NUMA node of the processor. Typically “Auto” or “Enabled” is set.

“Chipselect Interleaving” configures interleaving between DIMM ranks. For configurations with multiple ranks in a memory channel, the interleaving between ranks is enabled by setting it to “Enabled”. In some configurations, the “Enabled” setting may reduce performance. Basically, it is recommended that you leave it “Auto”, the default setting.

The fourth parameter “DRAM Scrub Time” periodically searches main memory for correctable errors and makes corrections as needed. This prevents the accumulation of memory errors that will make automatic correction impossible. If you have sensitive performance indicators, you may be affected by this feature. However, it may be difficult to demonstrate the effect on performance.

The fifth parameter “TSME” (Transparent Secure Memory Encryption) is a security feature. Encrypting the data to be stored in DIMMs transparently from the OS prevents data snooping through physical access to DIMMs.

The sixth parameter “Power Down Enable” sets the Power Down Mode for DDR5 memory. Enabling it reduces power consumption during inactive state.

The seventh parameter “Power Profile Selection” sets the power management policy of the processor. The setting “Efficiency Mode” is a power-efficient setting. It improves performance per

power by adjusting the operating frequency of the processor core and the data fabric. The original performance of the processor may not be achieved due to reduced operating frequency. "High Performance Mode" is a performance-oriented policy. It controls the processor core frequency to maintain it to a high level. "Max IO Performance Mode" is another performance-oriented policy and it strives to keep the operating frequency of the data fabric at a high level for high-volume I/O operations. As a result, the operating frequency of the processor core may be reduced, resulting in some performance degradation. "Balanced Memory Performance Mode" adjusts the performance of memory and data fabric to match bandwidth and latency required by workload.

The last two parameters are the settings related to the NUMA configuration in the processor.

"NUMA nodes per socket" (NPS) is a parameter that specifies how many NUMA nodes the processor described in the DIMM slots, memory controllers, and NUMA configuration section is divided into. Four options are available⁴: "NPS1", "NPS2", "NPS4", and "NPS0". "NPS1" is set by default.

When set to "NPS4", the processor is split into four NUMA nodes, and when set to "NPS2", it is split into two. For "NPS1", the processor is treated as a single NUMA node that is Uniform Memory Access (UMA). By splitting NUMA nodes, access to L3 cache and memory from cores in a NUMA node improves its latency. Therefore, this setting is recommended for NUMA-optimized applications.

"NPS0" is only available on 2-socket servers. This differs from the other options in that it treats the two processors as a single NUMA node. Memory is interleaved across the memory channels of the two processors. Therefore, the two processors must have identical memory configurations. Configuring "NPS0" is generally not recommended unless you have special requirements, such as a non-NUMA-aware OS.

If you use a processor with more than 64 logical CPUs in Windows, you must change the NPS. The processor group that Windows uses to manage logical CPUs is capped at 64 logical CPUs, so any logical CPUs that exceed that limit are managed as a separate processor group. This results in uneven processor group size, which is a performance disadvantage. You can divide NUMA nodes into equally sized processor groups by the setting.

The "ACPI SRAT L3 Cache As NUMA Domain" (L3AsNUMA) parameter subdivides the processor into multiple NUMA nodes per CCX. This results in a configuration of up to 16 NUMA nodes per processor. If you run workloads that fit within the L3 cache and the cores that share it, you may see performance improvements.

⁴ Some options are not available depending on the processor model.

Performant memory configurations

The memory transfer rate and the number of memory channels used greatly affect memory performance. We dealt with the memory transfer rate above. 4th and 5th generation EPYC processors has twelve memory channels for each processor and in order to realize high memory performance, it is necessary to place DIMMs in as many memory channels as possible.

Furthermore, there are several configuration features that affect memory performance. The interleaving, NUMA configuration setting, etc. In the Part 2 of this document, we will report the test results of these topics.

Performance Mode configurations

Another factor which should always be observed is the influence of the DIMM placement. There are a range of memory configurations between the minimum configuration (one 16 GB DIMM per configured processor) and the maximum configuration (full configuration with 256 GB DIMMs) which are ideal regarding memory performance. The following table lists the particularly interesting configurations of this type (it is not necessarily complete).

With these configurations, all twelve memory channels per processor are the same. In each bank configuration, the same type of twelve DIMMs set is used. This ensures that memory accesses are evenly distributed among these memory system resources. Technically speaking, the optimum 12-way interleaving is realized via the memory channel. In this document, this is called Performance Mode configuration.

The 4th generation EPYC processor based PRIMERGY server Performance Mode configuration						
1 CPU system ⁵	2 CPU system	DIMM type	DIMM size (GB) Bank 1	DIMM size (GB) Bank 2 ⁶	Max. memory transfer rate MT/s	Comment
192 GB	384 GB	DDR5-4800 R	16	-	4,800	
384 GB	-	DDR5-4800 R	16	16	4,000	
384 GB	768 GB	DDR5-4800 R	32	-	4,800	Best configuration for benchmark
576 GB	-	DDR5-4800 R	32	16	3,600	Mixed configuration
768 GB	-	DDR5-4800 R (1Rx4)	32	32	4,000	
768 GB	-	DDR5-4800 R (2Rx8)	32	32	3,600	
768 GB	1,536 GB	DDR5-4800 R	64	-	4,800	
1,152 GB	-	DDR5-4800 R	64	32	3,600	Mixed configuration
1,536 GB	3,072 GB	DDR5-4800 R	64	-	4,800	
3,072 GB	-	DDR5-4800 3DS R	128	128	3,600	
3,072 GB	6,144 GB	DDR5-4800 3DS R	128	-	4,800	Max. configuration of RX2450 M2 server Max. configuration at memory transfer rate of 4,800 MT/s
6,144 GB	-	DDR5-4800 3DS R	256	256	3,600	Max. configuration of RX1440 M2 server

⁵ For PRIMERGY RX1440 M2 server.

⁶ The Bank 2 DIMM slots are not available for PRIMERGY RX2450 M2.

The 5th generation EPYC processor based PRIMERGY RX1440 M2 server Performance Mode configuration					
1 CPU system	DIMM type	DIMM size (GB) Bank 1	DIMM size (GB) Bank 2	Max. memory transfer rate MT/s	Comment
192 GB	DDR5-5600 R	16	-	4,800	
384 GB	DDR5-5600 R	16	16	4,000	
384 GB	DDR5-5600 R	32	-	4,800	Best configuration for benchmark
576 GB	DDR5-5600 R	32	16	3,600	Mixed configuration
768 GB	DDR5-5600 R (1Rx4)	32	32	4,000	
768 GB	DDR5-5600 R (2Rx8)	32	32	3,600	
768 GB	DDR5-5600 R	64	-	4,800	
1,152 GB	DDR5-5600 R	64	32	3,600	Mixed configuration
1,536 GB	DDR5-5600 R	96	-	4,800	
3,072 GB	DDR5-5600 3DS R	96	96	3,600	
3,072 GB	DDR5-5600 3DS R	256	-	4,800	Max. configuration at memory transfer rate of 4,800 MT/s
6,144 GB	DDR5-5600 3DS R	256	256	3,600	Max. configuration

The 5th generation EPYC processor based PRIMERGY RX2450 M2 server Performance Mode configuration					
1 CPU system	2 CPU system	DIMM type	DIMM size (GB) Bank 1	Max. memory transfer rate MT/s	Comment
192 GB	384 GB	DDR5-5600 R	16	5,600	
		DDR5-6400 R		6,000	
384 GB	768 GB	DDR5-5600 R	32	5,600	
		DDR5-6400 R		6,000	Best configuration for benchmark
768 GB	1,536 GB	DDR5-5600 R	64	5,600	
		DDR5-6400 R		6,000	
1,152 GB	2,304 GB	DDR5-5600 R	96	5,600	
		DDR5-6400 R		6,000	
1,536 GB	3,072 GB	DDR5-6400 R	128	6,000	Max. configuration at memory transfer rate of 6,000 MT/s
3,072 GB	6,144 GB	DDR5-5600 3DS R	256	5,600	Max. configuration

The table is organized according to the total memory capacity of the left end. The total capacity is defined in one or two processor configurations. It is assumed that the memory configuration is the

same for all the processors. The next column is the DIMM type used. RDIMM, or 3DS RDIMM technology is the determinant. The next two columns show the DIMM size by bank. This is because it is using the Performance Mode configuration and therefore groups the DIMMs into sets of twelve per bank.

The smallest configuration in the table has 192 GB for one processor because the twelve 16 GB DIMMs (i.e., 192 GB) must be counted for each processor. The Performance Mode configuration requires an identical DIMM group of twelve per bank, but it does not forbid different DIMM sizes in different banks if the following restrictions are observed:

- RDIMMs and 3DS RDIMMs must not be mixed.
- RDIMMs of type x4 and x8 must not be mixed.

The second column from the right of the table shows the maximum memory frequency that can be achieved with each configuration. The maximum memory transfer rate depends on the value of DPC, as described above.

Independent Mode configurations

This covers all the configurations that are not in Performance Mode. There are no restrictions other than the followings but please refer to the respective configurator for details.

- RDIMMs and 3DS RDIMMs must not be mixed.
- RDIMMs of type x4 and x8 must not be mixed.
- The number of the DIMM on a processor is limited to one, two, four, six, eight, ten, twelve, sixteen, twenty, or twenty four.⁷

You also need to pay attention to configurations where the number of DIMMs per processor does not become a multiple of twelve, that is, less than the minimum number required for the Performance Mode configuration. This configuration may be done for reasons such as power saving and a low memory capacity. Cost savings may be realized by minimizing the number of DIMMs. From the quantitative evaluation showing the influence of the interleave configuration to the memory channel on the system performance introduced below, operation with one or two DIMMs configuration is not recommended.

Symmetric memory configurations

Finally, a separate section is to once again highlight that all configured processors are to be equally configured with memory if possible and the default setting of the BIOS is not to be changed without a convincing reason.

It goes without saying that preinstallation at the factory takes this circumstance into account. The ordered memory modules are distributed as equally as possible across the processors.

These measures and the related operating system support create the prerequisite to run applications as far as possible with a local, high-performance memory. The memory accesses of the processor cores are usually made to DIMM modules, which are directly allocated to the respective processor.

In order to estimate the performance merit of this, although the memory of the 2-way server is configured symmetrically, the measurement results when the BIOS parameter "NUMA nodes per socket" is set to "NPS0" are shown below. Statistically, one out of every two memory accesses is done to the remote memory. In the case where the application is executed on 100 % remote memory, it is necessary to estimate double the performance loss when executed on local memory and remote memory at a ratio of 50 %/50 %.

⁷ For PRIMERGY RX2450 M2, the maximum number of DIMMs populated per processor is twelve.

Quantitative effects on memory performance

After the functional description of the memory system with qualitative information, we now have specific statements about with which gain or loss in performance differences are connected in the memory configuration. As a means of preparation, the first section deals with the two benchmarks that were used to characterize memory performance.

This is followed - in order of their impact - by the already mentioned features interleaving of the memory channels, memory frequency, influence of the DIMM types and cache coherence protocol. At the end we then have measurements for the case of memory performance under redundancy.

The measurements were made on a PRIMERGY RX1440 M2 or a PRIMERGY RX2450 M2 under the Linux operating system. The following table shows the details of the configuration used for quantitative testing.

System Under Test (SUT)	
Hardware	
Model	PRIMERGY RX1440 M2 PRIMERGY RX2450 M2
Processor	1x EPYC 9554P (4th gen EPYC, 64 cores, 3.1 GHz, Max. memory transfer rate 4,800 MT/s) 2x EPYC 9534 (4th gen EPYC, 64 cores, 2.45 GHz, Max. memory transfer rate 4,800 MT/s) 1x EPYC 9555 (5th gen EPYC, 64 cores, 3.2 GHz, Max. memory transfer rate 6,000 MT/s)
DIMM types	16GB (1x16GB) 1Rx8 DDR5-4800 R ECC 32GB (1x32GB) 2Rx8 DDR5-4800 R ECC 32GB (1x32GB) 1Rx4 DDR5-4800 R ECC 64GB (1x64GB) 2Rx4 DDR5-4800 R ECC 128GB (1x128GB) 4Rx4 DDR5-4800 3DS R ECC 256GB (1x256GB) 8Rx4 DDR5-4800 3DS R ECC 32GB (1x32GB) 2Rx8 DDR5-5600 R ECC 64GB (1x64GB) 2Rx4 DDR5-5600 R ECC 256GB (1x256GB) 8Rx4 DDR5-5600 3DS R ECC 32GB (1x32GB) 2Rx8 DDR5-6400 R ECC
Disk subsystem	1x SAS 12G SSD 1.6 TB (via SAS RAID controller)
Software	
BIOS	R1.1.0 (for 4th gen EPYC) R2.1.0 (for 5th gen EPYC)
Operating system	SUSE Linux Enterprise Server 15 SP4 (for 4th gen EPYC) SUSE Linux Enterprise Server 15 SP6 (for 5th gen EPYC)

A PRIMERGY RX1440 M2 with an EPYC 9554P processor and 32 GB 2Rx8 RDIMMs was primarily used for the test set described below. A PRIMERGY RX2450 M2 with an EPYC 9555 processor was used to evaluate the impact of memory transfer rate and DIMM type using DDR-5600 or DDR-6400 DIMMs. All the other DIMMs listed in the table were used only in the test set for the impact of the DIMM type except for the evaluation of the impact of interleaving across the memory channels and that of remote memory access. A PRIMERGY RX2450 M2 with EPYC 9534 processors were used in the test set for the impact of remote memory access.

The following table shows relative performance. The absolute measurement values for the STREAM and SPECrate2017 Integer benchmarks under ideal memory conditions, which are usually equivalent to the 1.0 measurement of the tables, are included in the Performance Reports of each EPYC processor based PRIMERGY server.

The measuring tools

Measurements were made using the benchmarks STREAM and SPECrate2017 Integer.

STREAM Benchmark

The STREAM benchmark (Developer: Mr. John McCalpin) is a tool to measure memory throughput. This benchmark implements copying and arithmetic operations on a large array of double type data, and provides four types of access results: Copy, Scale, Add and Triad. For access types other than Copy, arithmetic operations are included. Results are always indicated with throughput in GB/s. In general, the value of Triad is best quoted. Afterwards, the measured value of STREAM's benchmark is the Triad access value, and the unit is GB/s.

STREAM is the industry standard for measuring the memory bandwidth of a server and can apply a large load to the memory system using a simple method. In particular, this benchmark is suitable for investigating the effect on memory performance in complex configurations. STREAM shows the effect of the configuration on memory and the resulting performance (degradation or improvement) caused by it. The value related to the STREAM benchmark described below shows the degree of influence on performance.

The memory impact on application performance is distinguished by the latency of each access and the bandwidth required by the application. Since the latency increases as the memory bandwidth increases, both are related. The degree to which the latency is canceled by parallel memory access also depends on the application and the quality of the machine code created by the compiler. For this reason, it is very difficult to make a general forecast for all application scenarios.

SPECrate2017 Integer Benchmark

The SPECrate2017 Integer benchmark has been added as a model for commercial application performance. This is part of the Standard Performance Evaluation Corporation (SPEC) SPEC CPU2017. SPEC CPU2017 is the industry standard for evaluating system processors, memory and compilers. It is the most important benchmark in the server field because a large number of measurement results are released and used for sales projects and technical investigation.

SPEC CPU2017 consists of two independent test sets that use a lot of integer operations and floating-point operations. The integer operation portion is equivalent to a commercial application and consists of 10 types of benchmarks. The floating-point operation portion is equivalent to a scientific application and consists of 10 or 13 types of benchmarks. In either case, the benchmark execution result is the geometric mean of the individual results.

A distinction is also made in the suites between the speed run with only one process and the rate run with a configurable number of processes working in parallel. The second version is evidently more interesting for servers with their large number of processor cores and hardware threads.

In addition, depending on the type of measurement, the optimization allowed for the compiler differs. For the peak result the individual benchmarks may be optimized independently of each other, but for the more conservative base result the compiler flags must be identical for all benchmarks, and certain optimizations are not permitted.

This is the summary of SPEC CPU2017. The SPECrate2017 Integer suite was selected, because commercial applications predominate in the use of PRIMERGY servers.

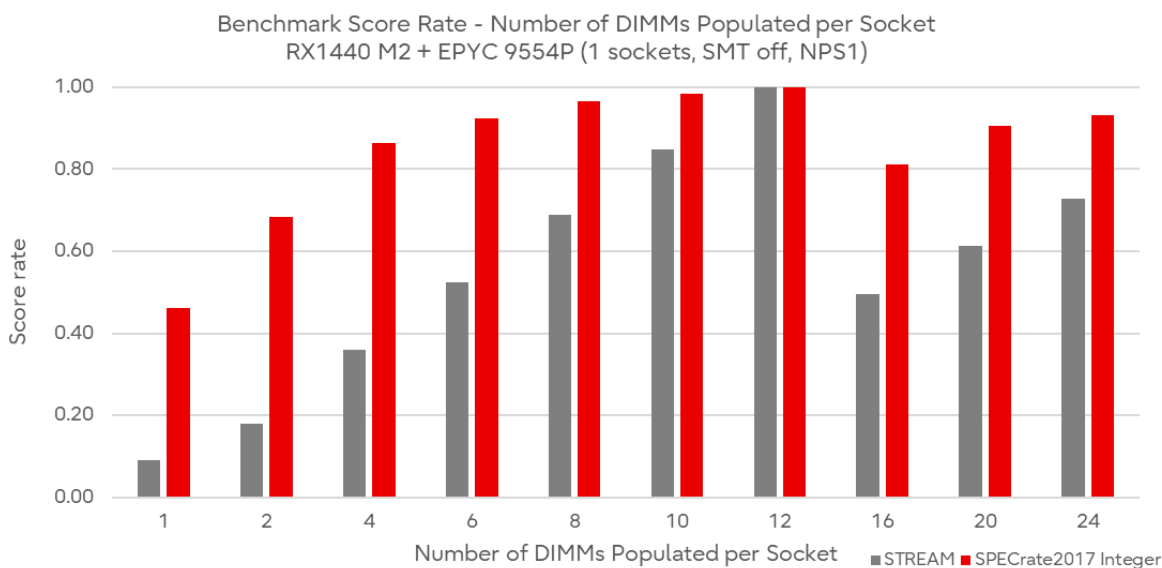
Memory interleaving settings

Number of DIMMs and channel interleaving

Interleaving among memory channels is a method of setting a physical address space so that up to twelve memory channels are sequentially used for each processor, such that the first block is on the first channel, the second block is on the second channel, and so on. Memory access is mainly done in the adjacent memory area according to the locality principle. Access to this contiguous range of physical addresses is distributed across all channels by the above interleaving. This results in the performance improvement.

The following figure shows the ratio of the performance, when DIMMs are not mounted in a set of twelve pieces per processor and the ideal 12-way interleave is not performed; the value is considered as 1 when the number of DIMMs is twelve. The number of DIMMs populated per one processor is limited to 1, 2, 4, 6, 8, 10, 12, 16, 20, and 24 for 4th and 5th generation EPYC processor based PRIMERGY servers^{8,9}. The results of the measurement with the setting “NUMA nodes per socket = NPS1”, which is available for any number of DIMMS above, are shown here.

The DIMM type used for this test was 128 GB DDR5-4800 4Rx4 3DS RDIMM. In addition, SMT (Simultaneous MultiThreading) was disabled. These were chosen to ensure that there was enough memory to satisfy the working set of benchmark tests.



In particular, marked declines are seen in the STREAM index that measures memory throughput. When the number of DIMMs is equal to or less than twelve, the performance is improved according to the increase in the number of DIMMs. With sixteen DIMMs or more, the difference in the number of memory channels used and the reduced memory transfer rate caused by 2DPC configuration results in about 30% to 50% drop in the performance of STREAM.

Evaluation on SPECrate2017 Integer concerns the performance of commercial applications. The relationships of the memory bandwidth as expressed by STREAM should be understood as extreme cases, which cannot be ruled out in certain application areas, especially in the HPC (High-Performance Computing) environment. However, such behavior is improbable for most commercial

⁸ For PRIMERGY RX2450 M2, the maximum number of DIMMs populated per processor is twelve.

⁹ Refer to Upgrade and Maintenance Manual of the respective servers for the DIMM location.

loads. This assessment of the interpretation quality of STREAM and SPECrate2017 Integer not only applies for the performance aspect dealt with in this section, but also for all following sections.

There may be good reasons for choosing a 8-way or 10-way interleave, where the performance degradation of SPECrate2017 Integer is gentle. In other words, the required memory capacity is small or the number of DIMMs is kept to a minimum because of low power consumption. 1-way interleaving is not recommended. (Strictly speaking this is not interleaving, it is only called as such in the classification.) In this case, the performance of the processor and the memory system are not well balanced.

Although the channel interleaving can be disabled by “Memory interleaving” parameter in the BIOS settings, it is not recommended. Without the channel interleaving, effective use of the twelve memory channels depends on which portion of system memory is used. If only memory on a particular channel is used, this may cause a significant performance degradation, as the result of STREAM below. This is because STREAM uses less memory, about 20% of total system memory, so memory on certain memory channels is disproportionately used. The SPECrate2017 Integer, which uses more memory, has a smaller impact because the used memory is distributed across more memory channels.

Benchmark	Memory interleaving = Enabled (default)	Memory interleaving = Disabled
STREAM	1.00	0.34
SPECrate2017 Integer	1.00	0.96

Chipselect Interleaving

The “Chipselect Interleaving” parameter configures the interleaving between DIMM ranks. For the configurations with multiple ranks in a memory channel, setting it to “Enabled” enables the interleaving between ranks. In general, we recommend that you leave the default setting of “Auto”.

Benchmark	Chipselect Interleaving = Auto (default)	Chipselect Interleaving = Enabled	Chipselect Interleaving = Disabled
STREAM	1.00	1.00	0.95
SPECrate2017 Integer	1.00	1.00	1.00

Note that in a 2-DPC configuration using two, four, and eight rank DIMMs, setting this to “Disabled” may slightly improve performance compared to “Enabled”.

Memory transfer rate

For the 4th generation EPYC processor based PRIMERGY servers, the DPC configuration and BIOS parameters affect the memory transfer rate. Furthermore, used DIMM types and server models also affect it on the 5th generation EPYC processor based PRIMERGY servers. You can reduce some power consumption by lowering the memory transfer rate.

The following is a comparison of the performance impact of changing the effective memory transfer rate using “Memory Clock” parameter in the BIOS settings. The values in the table assume 1.0 for the best case, i.e., the performance at the maximum memory transfer rate.

The number in parentheses in the table means how much power consumption is reduced compared with the best case. These figures are provided as reference data. These results are not always expected because the power consumption is affected by various factors, such as workloads, system utilization, system configuration, and ambient temperature.

For the 4th generation EPYC processor							
Benchmark	Processor type	Maximum memory transfer rate	3,200 MT/s	3,600 MT/s	4,000 MT/s	4,400 MT/s	4,800 MT/s
STREAM	EPYC 9554P	4,800 MT/s	0.68 (-41 W)	0.77 (-26 W)	0.85 (-20 W)	0.92 (-17 W)	1.00 (0 W)
SPECrate2017 Integer	EPYC 9554P	4,800 MT/s	0.92 (-55 W)	0.95 (-28 W)	0.97 (-25 W)	0.99 (-17 W)	1.00 (0 W)

For the 5th generation EPYC processor									
Benchmark	Processor type	Maximum memory transfer rate	3,600 MT/s	4,000 MT/s	4,400 MT/s	4,800 MT/s	5,200 MT/s	5,600 MT/s	6,000 MT/s
STREAM	EPYC 9555	6,000 MT/s	0.64 (-25 W)	0.71 (-25 W)	0.77 (-22 W)	0.83 (-16 W)	0.88 (-15 W)	0.95 (-10 W)	1.00 (0 W)
SPECrate2017 Integer	EPYC 9555	6,000 MT/s	0.93 (-35 W)	0.95 (-31 W)	0.96 (-25 W)	0.97 (-18 W)	0.99 (-15 W)	0.99 (-9 W)	1.00 (0 W)

These measurements used a PRIMERGY RX1440 M2 with an EPYC 9554P and 32 GB 2Rx8 DDR-4800 RDIMMs and a PRIMERGY RX2450 M2 with an EPYC 9555 and 32 GB 2Rx8 DDR-6400 RDIMMs.

Influence of the DIMM types

PRIMERGY servers with 4th and 5th generation EPYC processors support a total of 18 DIMM types. Refer to the system configuration guide for each server model to determine the available DIMM configurations.

The following table shows the differences in performance between these DIMM types under otherwise identical conditions:

- Measurements were taken using a single-processor configuration with either a PRIMERGY RX1440 M2 (EPYC 9554P) or a PRIMERGY RX2450 M2 (EPYC 9555). As noted above, the PRIMERGY RX1440 M2 with DDR5-5600 DIMMs is limited to a maximum memory transfer rate of 4,800 MT/s. Therefore, its performance is generally comparable to that of DDR5-4800 DIMMs.
- For these measurements, all memory channels were configured identically in Performance Mode. Twelve DIMMs were installed for 1DPC measurements, and twenty-four DIMMs were installed for 2DPC measurements.
- All the measurements were performed at the maximum memory transfer rate supported by each processor. Specifically, DIMMs operated at 4,800 MT/s, 5,600 MT/s, or 6,000 MT/s in 1DPC configurations, and at 3,600 MT/s or 4,000 MT/s in 2DPC configurations.
- The table is normalized to the 1DPC configuration using the 32 GB 2Rx8 RDIMMs (highlighted in bold), which are expected to provide the best memory performance. This DIMM type is recommended for benchmarking when its capacity meets the application's memory requirements.

For 4th generation EPYC processors					
DIMM type	Config uration	# of ranks per channel	Max. memory transfer rate (MT/s)	STREAM	SPECrate 2017 Integer
16GB (1x16GB) 1Rx8 DDR5-4800 R ECC	1DPC	1	4,800	0.95	0.98
	2DPC	2	4,000	0.85	0.97
32GB (1x32GB) 2Rx8 DDR5-4800 R ECC	1DPC	2	4,800	1.00	1.00
	2DPC	4	3,600	0.71	0.92
32GB (1x32GB) 1Rx4 DDR5-4800 R ECC	1DPC	1	4,800	0.91	0.98
	2DPC	2	4,000	0.83	0.97
64GB (1x64GB) 2Rx4 DDR5-4800 R ECC	1DPC	2	4,800	0.97	1.00
	2DPC	4	3,600	0.70	0.92
128GB (1x128GB) 4Rx4 DDR5-4800 3DS R ECC	1DPC	4	4,800	0.99	0.99
	2DPC	8	3,600	0.72	0.90
256GB (1x256GB) 8Rx4 DDR5-4800 3DS R ECC	1DPC	8	4,800	1.00	0.97
	2DPC	16	3,600	0.69	0.87

For 5th generation EPYC processors (RX2450 M2) ¹⁰					
DIMM type	Config uration	# of ranks per channel	Max. memory transfer rate (MT/s)	STREAM	SPECrate 2017 Integer
16GB (1x16GB) 1Rx8 DDR5-5600 R ECC	1DPC	1	5,600	0.90*	0.97*
32GB (1x32GB) 2Rx8 DDR5-5600 R ECC	1DPC	2	5,600	0.95	0.99
32GB (1x32GB) 1Rx4 DDR5-5600 R ECC	1DPC	1	5,600	0.86*	0.97*
64GB (1x64GB) 2Rx4 DDR5-5600 R ECC	1DPC	2	5,600	0.90	0.99
256GB (1x256GB) 8Rx4 DDR5-5600 3DS R ECC	1DPC	8	5,600	0.93	0.97
16GB (1x16GB) 1Rx8 DDR5-6400 R ECC	1DPC	1	6,000	0.95*	0.98*
32GB (1x32GB) 2Rx8 DDR5-6400 R ECC	1DPC	2	6,000	1.00	1.00
32GB (1x32GB) 1Rx4 DDR5-6400 R ECC	1DPC	1	6,000	0.91*	0.98*
64GB (1x64GB) 2Rx4 DDR5-6400 R ECC	1DPC	2	6,000	0.96*	0.99*

(*estimated)

In a 1DPC configuration, there are performance differences depending on the DIMM type in spite of the same maximum memory transfer rate as 4,800 MT/s. The performance differences shown here are mainly due to the different number of rank interleaves. The number of rank interleaves is equal to the number of ranks per memory channel and depends on the DIMM type and DPC value. For example, a 1-DPC configuration with dual-rank DIMMs shown in the table allows 2-way rank interleaving, and a 2-DPC configuration allows 4-way interleaving.

As shown in the table above, you can see that the performance is better with two ranks per memory channel than with one rank. On the other hand, performance degradation is noticeable in 1DPC configurations with 16 GB 1Rx8 RDIMMs and 32 GB 1Rx4 RDIMMs, i.e. 1-way rank interleaving.

In the 2DPC configuration, you can see a noticeable decrease in performance from the 1DPC configuration. This is because the maximum memory transfer rate in a 2DPC configuration is lower than in a 1DPC configuration. The performance degradation of STREAM is greatly affected by this.

¹⁰ As noted above, the performance of the PRIMERGY RX1440 M2 with DDR5-5600 DIMMs is generally comparable to that of DDR5-4800 DIMMs.

NUMA settings

On EPYC processors, "NUMA nodes per socket" parameter and "ACPI SRAT L3 Cache AsNUMA Domain" (L3AsNUMA) parameter are provided for the processor NUMA settings.

Three options are available for "NUMA nodes per socket": "NPS1", "NPS2", and "NPS4". These divide a processor into one, two, or four NUMA nodes, respectively. For the two-socket server, the another option "NPS0" is also available. Unlike the other options, this treats the two processors as a single node. We'll get to that later.

The "L3asNUMA" parameter is a setting that treats the CCX in the processor as a single NUMA node. Depending on the processor model, there can be up to sixteen nodes per processor. For details, refer to the section on [memory system BIOS options](#).

The following table shows the effect for the two loads (benchmarks) performed in this document. The performance at the default settings is set to 1.00.

The table shows that "NPS2" and "NPS4" have a performance impact in the range of 1% to 9%. When evaluating this table, it should be considered that both benchmarks are extremely NUMA friendly due to careful process binding during test setup. Therefore, a typical commercial application may not have the same effect as these results.

Our results did not show an effect of "L3AsNUMA" parameter, but for workloads that fit within the CCX cores and L3 cache, the performance improvement may be expected.

Benchmark		NPS1 (default)	NPS2	NPS4
STREAM	L3AsNUMA = Disabled (default)	1.00	1.05	1.09
	L3AsNUMA = Enabled	1.00	1.05	1.09
SPECrate2017 Integer	L3AsNUMA = Disabled (default)	1.00	1.01	1.02
	L3AsNUMA = Enabled	1.00	1.01	1.02

Access to remote memory

For the tests using the STREAM and SPECrate2017 Integer benchmarks mentioned above, only the local memory was targeted (the processor accesses the DIMM module of its own memory channel). Modules of adjacent processors are not accessed at all, or only rarely accessed. This situation is representative, insofar as it also exists for the majority of memory accesses of real applications thanks to NUMA support in the operating system and system software.

If you set "NUMA nodes per socket" parameter to "NPS0" on a system with two processors, the two processors are treated as a single node. This makes the application more susceptible to remote memory access.

The following results show the performance impact of remote memory access on the PRIMERGY RX2450 M2 with two EPYC 9534 processors and 64 GB 2Rx4 DIMMs in a 1DPC configuration. The performance at the default settings is set to 1.00.

In a nearly ideal memory configuration of 64 GB 2Rx4 RDIMM 1DPC configuration operating at maximum memory transfer rates, the deterioration in performance occurs because statistically one out of every two memory accesses is to a remote DIMM, i.e., a DIMM allocated to the neighboring processor, and the data must be accessed via the xGMI link.

Benchmark	NPS1 (default)	NPS0
STREAM	1.00	0.70
SPECrate2017 Integer	1.00	0.88

Measurements for "NPS0" indicate that this level of performance degradation can occur on platforms and applications that are not NUMA optimized. These results are useful for estimating the impact if most or all accesses are to remote memory.

Power-related BIOS setting

This section addresses the performance impact of two power-related BIOS settings.

First is the impact of "Power Down Enable". When enabled, the DDR5 memory feature is used to reduce the power consumption at inactive state. The values in the table assume that the performance at the default setting is 1.0. The significant impact in performance is not observed.

The number in parentheses in the table below means how much power consumption was reduced compared with that of the default setting for the reference. In the situation where the memory accesses are low, the power consumption may be slightly reduced. Because the power consumption is affected by various factors such as workload, system utilization, system configuration, and ambient temperature, such results cannot always be expected.

Benchmark	Power Down Enable = Disabled (default)	Power Down Enable = Enabled
STREAM	1.00 (0 W)	1.00 (-8 W)
SPECrate2017 Integer	1.00 (0 W)	1.00 (-15 W)

Next is the impact of "Power Profile Selection". This setting affects the operating frequency of the processor core and the data fabric. "Efficiency Mode" setting optimizes for performance per power and may not be able to deliver the full performance of the processor. In our test set, the setting "High Performance Mode" improved the performance by 9% on STREAM and 14% on SPECrate2017 Integer over the default. "Maximum IO Performance Mode" prioritizes the operating frequency of the data fabric for a large amount of I/O processing, but since I/O processing is small for both STREAM and SPECrate2017 Integer, no difference in performance was observed.

The number in parentheses in the table below means how much power consumption was increased compared with that of the default "Efficiency Mode" setting for the reference. Because the power consumption is affected by various factors such as workload, system utilization, system configuration, and ambient temperature, such results cannot always be expected.

Benchmark	Power Profile Selection			
	Efficiency Mode (default)	High Performance Mode	Maximum IO Performance Mode	Balanced Memory Performance Mode
STREAM	1.00 (0 W)	1.09 (+ 66 W)	1.09 (+ 66 W)	1.09 (+ 50 W)
SPECrate2017 Integer	1.00 (0 W)	1.14 (+ 106 W)	1.14 (+ 97 W)	1.14 (+ 95 W)

Impact of memory encryption setting

4th and 5th generation EPYC processors have a number of features that improve security. This section evaluates the performance impact of enabling "TSME" (Transparent Secure Memory Encryption).

The results show a performance degradation of 1% for STREAM and 3% for SPECrate2017 Integer. By enabling "TSME", data stored in memory are encrypted. This encryption takes time, which increases memory latency. Since memory latency often affects performance in typical applications, the performance of the SPECrate2017 Integer is more affected.

Benchmark	TSME = Disabled (default)	TSME = Enabled
STREAM	1.00	0.99
SPECrate2017 Integer	1.00	0.97

Literature


PRIMERGY Servers

<https://www.fujitsu.com/global/products/computing/servers/primergy/>

Memory performance

This Whitepaper

 <https://docs.ts.fujitsu.com/dl.aspx?id=fd60214d-4b08-4ca1-b789-c5b2e3d42c3f>

 <https://docs.ts.fujitsu.com/dl.aspx?id=fce4cf19-8c42-496d-aa32-1b17511f862c>

Benchmarks

STREAM

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

SPECcpu2017

<https://docs.ts.fujitsu.com/dl.aspx?id=20f1f4e2-5b3c-454a-947f-c169fca51eb1>

BIOS settings

BIOS optimizations for EPYC 9004 and 9005 processors-based systems

<https://docs.ts.fujitsu.com/dl.aspx?id=24cacc7c-b128-4674-91d1-23bbc185bc89>

PRIMERGY Performance

<https://www.fujitsu.com/global/products/computing/servers/primergy/benchmarks/>

Document change history

Version	Date	Description
1.1	2025-04-08	Added the description on the 5th generation EPYC processor (Turin)
1.0	2024-07-02	Initial version

Contact

Fujitsu

Web site: <https://www.fujitsu.com>

PRIMERGY Performance and Benchmarks

<mailto:fj-benchmark@dl.jp.fujitsu.com>

© Fujitsu 2024. All rights reserved. Fujitsu and Fujitsu logo are trademarks of Fujitsu Limited registered in many jurisdictions worldwide. Other product, service and company names mentioned herein may be trademarks of Fujitsu or other companies. This document is current as of the initial date of publication and subject to be changed by Fujitsu without notice. This material is provided for information purposes only and Fujitsu assumes no liability related to its use.