Best Practice Guidelines for ProLiant Servers with the Intel Xeon 5500 processor series
Engineering Whitepaper, 1st Edition

Table of Contents

Executive summary

This white paper gives rules, best practice recommendations, and performance data on DDR3 memory in certain ProLiant servers. Because of significant differences in the system architecture of sixth generation (G6) Intel Xeon-based ProLiant servers over prior generations of Intel Xeon-based servers, more rules and trade-offs must be evaluated in order to optimize your choice of memory type and amount.

Target audience

This white paper is intended for HP partners and customers configuring HP ProLiant servers.

Architecture

Sixth generation (G6) two-processor ProLiant servers that use the Intel Xeon 5500 processor series use DDR3 memory technology. New features of DDR3 memory, the Intel Xeon 5500 processor series, and the design of Intel Xeon-based G6 ProLiant servers all contribute to changes in how memory subsystem performance should be optimized.

DDR3 Memory Technology

DDR-3, the third-generation of DDR SDRAM technology, makes improvements in bandwidth and power consumption over DDR-2. Improvements in DDR-3 can yield up to 40% power savings over DDR-2 at the same speed, and 66% higher bandwidth over DDR-2.

DDR-3 DIMMs use the same 240-pin connector as DDR2 DIMMs, but the notch key is in a different position.

DDR-3 DIMMs can run at different speeds, often called frequencies. DDR-3 operates at data rates from 800 Mbits/s to 1600 Mbits/s. DDR3-1600 DIMMs will be available in 2010. See **Table 1**.

Table 1 : DDR3 Speeds

JEDEC Name	Common Name	Bit Data Rate*	Maximum DIMM Throughput
PC3-12800	DDR3-1600	1600 MT/s	12.8 GB/s
PC3-10600	DDR3-1333	1333 MT/s	10.6 GB/s
PC3-8500	DDR3-1066	1066 MT/s	8.5 GB/s
PC3-6400	DDR3-800	800 MT/s	6.4 GB/s

** The value of the Bit Data Rate is commonly referred to as the DIMM "frequency", quoted in Megahertz (MHz). The actual clock frequency on a DIMM is one half of the bit data rate value. For clarity, this document will adhere to the common definition, using "MHz" when actually referring to the bit data rate.*

The common name, derived from the maximum throughput (in megabits), can be translated into the number portion of the JEDEC name by multiplying by 8. Therefore, a DDR3-1067 DIMM is also known as a PC3-8500 1066 x 8 Bytes ~= 8500).

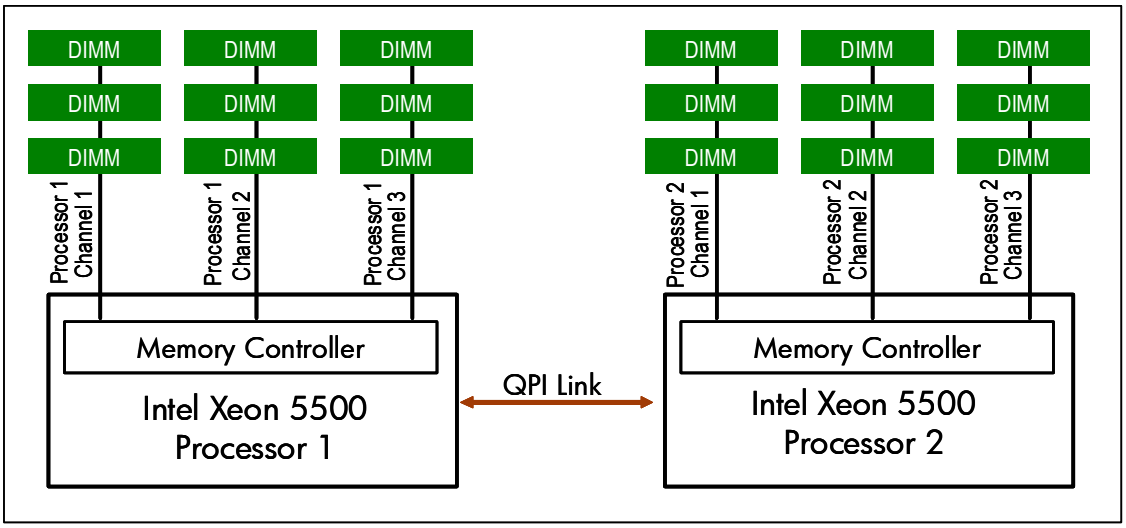
To increase performance and reduce power consumption of DDR-3, several key enhancements have been made:

- DDR-3 DIMMs operate at 1.5V, compared to 1.8V for DDR-2 DIMMs. The DDR-3 specification anticipates future availability of Low Voltage DIMMs, which will operate at even lower voltages.
- An 8-bit prefetch buffer stores more data before it is needed than the 4-bit buffer for DDR2 does.
- Fly-by topology (for the commands, addresses, control signals, and clocks) improves signal integrity by reducing the number of stubs and their length. This feature requires the controller to support "write leveling" on DDR-3 DIMMs.
- A thermal sensor integrated on the DIMM module signals the chipset to throttle memory traffic to the DIMM if DIMM temperature exceeds a programmable critical trip point.

Intel Xeon 5500 series processor

Each Intel Xeon 5500 series processor includes three integrated memory controllers. Integrated memory controllers yield more bandwidth than prior Intel Xeon processor architectures, which had a separate memory controller outside of the processor.

Figure 1 : Intel Xeon 5500 Memory Architecture



Each integrated memory controller has an independent path, called "channel". Each channel can connect to up to 3 DIMM slots. This means each processor can have up to 9 directly connected DIMMs, or 18 total DIMMs for a 2-processor system (see **Figure 1**).

Because the memory controller is integrated into the processor, DIMM slots can only be used if a processor is directly connected to them. That means that in a two-processor server with 18 DIMM slots, such as the ProLiant DL380 G6, both processors must be installed in order to use all 18 DIMM slots. If only one processor is installed, then a maximum of 9 DIMM slots can be used.

The integrated memory controller means an Intel Xeon 5500 based server has a NUMA (Non Uniform Memory Access) architecture. In NUMA architectures, each processor has direct access to a different region of memory. In order to access all memory, processors must sometimes pass memory requests to other processors, instead of the directly connected DIMM slots. One implication to this design is that the time it takes to access

memory can vary considerably, depending on whether the memory being accessed is local to the processor, or requires another processor to handle the request.

The Intel Xeon 5500 series processor supports interleaving. Interleaving is a method of distributing memory across all the three channels of a CPU. This can greatly increase overall memory bandwidth, but puts some limitations on how memory must be installed in the server.

There are several models of the Intel Xeon 5500 series processor family. Models differ in their number of cores, maximum processor frequency, amount of cache memory, and features supported (such as Intel Hyper-Threading Technology). In addition, different models support different maximum memory speeds; which in turn has an impact on performance and power.

Table 2 : Intel Xeon 5500 Series Processor Models

Processor Model Number	CPU Frequency	Level 3 Cache Size	Maximum Memory Speed	Maximum Memory Throughput (per channel)
X5570	2.93 GHz	8 MB	1333 MHz	10.6 GB/s
X5560	2.80 GHz	8 MB	1333 MHz	10.6 GB/s
X5550	2.66 GHz	8 MB	1333 MHz	10.6 GB/s
E5540	2.53 GHz	8 MB	1066 MHz	8.5 GB/s
E5530	2.40 GHz	8 MB	1066 MHz	8.5 GB/s
E5520	2.26 GHz	8 MB	1066 MHz	8.5 GB/s
E5506	2.13 GHz	4 MB	800 MHz	6.4 GB/s
E5504	2.00 GHz	4 MB	800 MHz	6.4 GB/s
E5502*	1.86 GHz	4 MB	800 MHz	6.4 GB/s
L5520	2.26 GHz	8 MB	1066 MHz	8.5 GB/s
L5506	2.13 GHz	4 MB	800 MHz	6.4 GB/s

** The E5502 model is a dual core processor. All other models shown are quad core processors.*

ProLiant G6 Servers

As shown in **Table 3** , there are several models of ProLiant G6 servers that use the Intel Xeon 5500 series processor and DDR-3 memory. These ProLiant servers have 9, 12, or 18 DIMM slots, depending on model.

Table 3 : ProLiant servers using DDR-3 Memory

HP ProLiant server model	Max CPUs	DIMM slots	Max Memory (GB)	Supports Mirrored Memory?	Supports Lock-Step Mode?
ML370 G6	2	18	144	yes	yes
ML350 G6	2	18	144	yes	yes
ML330 G6	2	18	144	yes	yes
ML150 G6	2	12	96	no	no

DL380 G6	2	18	144	yes	yes
DL360 G6	2	18	144	yes	yes
DL320 G6	1	9	72	no	no
DL180 G6	2	12	96	no	no
DL160 G6	2	18	144	no	no
BL460c G6	2	12	96	yes	yes
BL490c G6	2	18	144	yes	yes
BL280c G6	2	12	96	yes	yes

DIMM Types

Unbuffered Memory (UDIMMs) and Registered Memory (RDIMMs)

Manufacturers of DDR-3 SDRAM DIMMs produce two types of DIMMs: Unbuffered DIMMs (UDIMM) and Registered DIMMs (RDIMM). UDIMMs represent the most basic type of memory module and offer lower latency and (relatively) low power consumption but are limited in capacity. Unbuffered DIMMs with ECC are identified with an E suffix in the manufacturer's module name (example PC3-8500E). UDIMMs are applicable for systems with low DIMM counts and where low power is required and large memory capacities are not required. RDIMMs offer larger capacities than UDIMMs and include address parity protection. Registered DIMMs are identified with an R suffix in the module manufacturer's name (example PC3-8500R).

Table 4 : RDIMM versus UDIMM

	RDIMMs	UDIMMs
DIMM Sizes Available	2GB, 4GB, 8GB	1GB, 2GB
Low power version of DIMMs available	√	
ECC Support	√	√
Advanced ECC support	√	√
Address parity	√	
Memory Mirroring and Lock-Step Mode support	√	
Relative Cost	Higher	Lower
Maximum capacity on a server with 9 DIMM slots	72GB	12GB
Maximum capacity on a server with 12 DIMM slots	96 GB	24 GB
Maximum capacity on a server with 18 DIMM slots	144 GB	24 GB

DRAM technology

DIMMs are made up of DRAM chips that are ganged together. Each DRAM chip contains arrays of individual bit storage locations. A DRAM chip with one billion storage locations is called 1 Gigabit (1Gb) technology. Note the lower case b in Gb. Eight 1Gb chips ganged together will provide 1 GigaByte (1GB) of memory. Note the upper case B in GB.

DDR3 DIMMs are currently made up of 1Gb and 2Gb DRAM chips. It is not possible to mix DRAM technologies on the same DIMM. DDR3 does not support DIMMs made up of 512Mb DRAM chips.

A DRAM chip may have 4 data I/O signals or 8 data I/O signals. These are called x4 or x8, pronounced "by four" or "by eight" respectively.

Ranks

A Rank is a group of DRAM chips that are ganged together to provide 64 bits (8 Bytes) of data on the memory bus. All chips in a rank are controlled simultaneously by the same Chip Select, Address and Command signals. DDR-3 DIMMs are available in single-, dual- and quad-ranks (1, 2, and 4 ranks respectively.) Eight x8 DRAM chips or 16 x4 chips form a rank. DIMMs with ECC use nine x8 chips and 18 x4 chips for each rank.

Speed

Speed refers to the frequency of the memory clock. The memory subsystem uses a different clock than the processor cores, and this clock is used to coordinate data transfers between the memory controller and the DIMMs. The actual speed at which this clock operates in a particular server depends on five factors:

- **Rated memory speed of the processor**. Each Intel Xeon 5500 series processor model support different maximum memory speeds. See **Error! Reference source not found.**.
- **Rated memory speed of the DIMM**. DDR-3 DIMMs can run at different speeds, often called frequencies. See **Error! Reference source not found.**. HP offers two speeds of DDR3 memory: DDR3-1333 and DDR3-1066
- **Number of ranks on the DIMM**. Each rank on a memory channel adds one electrical load. As the electrical loads increase, the signal integrity degrades. To maintain the signal integrity the memory channel may be forced to run at a lower speed.
- **Number of DIMMs populated**. The number of DIMMs attached to a memory controller impacts the loading and signal integrity of the controller's circuits. In order to maintain signal integrity, the memory controller may operate DIMMs at lower than rated speed. Generally, the more DIMMs that are populated, the lower the speed that DIMMs can be run.
- **BIOS settings**. Enabling certain BIOS features can impact memory speed. For example, the ROM Based Setup Utility (RBSU) in HP ProLiant servers includes a user-selectable setting to force memory to run at slower than rated speed. See the chapter entitled BIOS Settings for details.

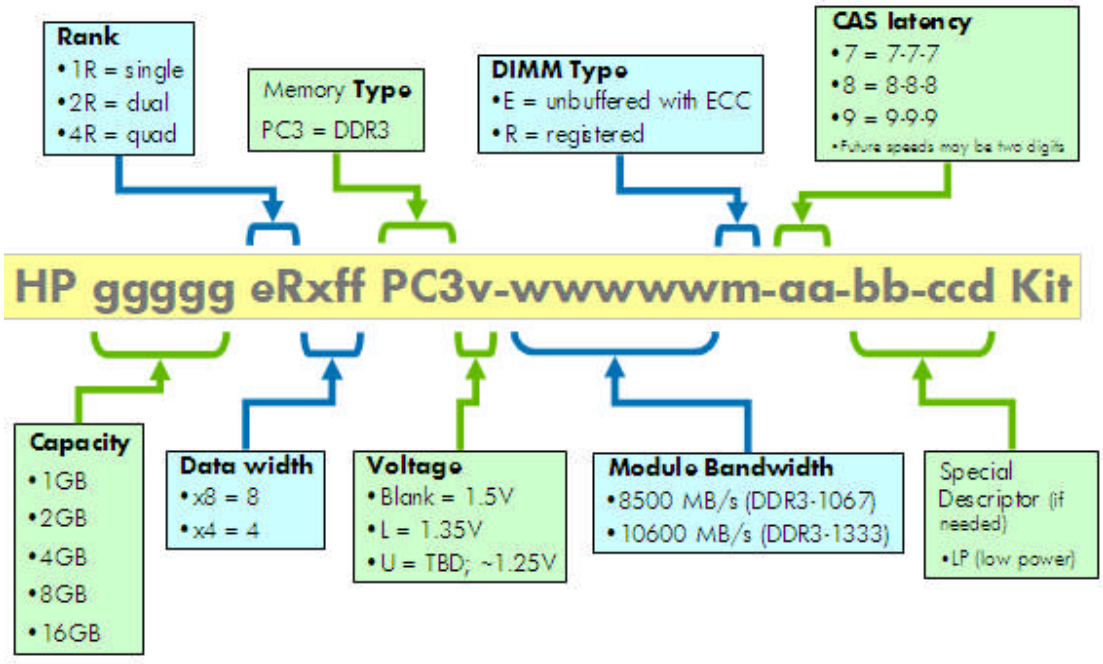
DDR3 Memory for ProLiant

Table 5 shows the DDR3 DIMMs that are qualified in Intel Xeon 5500-based HP ProLiant servers. HP Part descriptions use codes from the JEDEC standard for specifying DIMM type and speed; see Figure 2 for an explanation of part descriptions.

Table 5 : HP DDR3 DIMMs

Registered DIMMs (RDIMMs)	HP Part Number
HP 2GB 2Rx8 PC3-10600R-9	500656-B21
HP 4GB 2Rx4 PC3-10600R-9	500658-B21
HP 4GB 4Rx8 PC3-8500R-7 LP	500660-B21
HP 8GB 2Rx4 PC3-10600R-9	500662-B21
HP 8GB 2Rx4 PC3-8500R-7	516423-B21
Unbuffered DIMMs (UDIMMs)	
HP 1GB 1Rx8 PC3-10600E-9	500668-B21
HP 2GB 2Rx8 PC3-10600E-9	500670-B21

Figure 2: HP DDR3 Memory Part Number Decoder



Memory Protection Features

Data stored on DDR3 memory in HP ProLiant servers is protected from certain DIMM failures by ECC (Error Correcting Code) technology. In addition, certain advanced memory protection techniques are also available.

ECC

All HP qualified DDR3 DIMMs provide ECC data protection. Normal ECC can correct single-bit errors and detect double-bit errors.

Advanced ECC

HP's term "Advanced ECC" means that the server corrects multi-bit errors, how many bits depends on the implementation. Typically, this is up to a 4-bit error within a nibble. Multi-bit errors that go across nibble boundaries, but remain within two nibbles will all be detected. Advanced ECC is not the exact same thing as ChipKill™ 1 , which is an IBM term. In the Xeon 5500 processor family, Advanced ECC is equivalent to ChipKill and Lock-Step mode is equivalent to x8 or 8-bit ChipKill.

Address Parity

With RDIMMs the memory controller drives the address signals, command signals and a parity bit to a common register on the DIMM. The register re-drives the address and command signals to each DRAM chip on the DIMM. The register calculates the parity of the address and command signals and compares it to the parity bit it received. If there is a difference, it will signal the memory controller that there was an address parity error. The system will then shutdown to prevent silent data corruption. This feature is only available with RDIMMs.

Memory Mirroring

Mirrored memory mode is a fault-tolerant memory option that provides a higher level of availability than Advanced ECC protection. Mirrored Memory mode provides full protection against any single-bit and multi-bit errors. With Mirrored Memory mode enabled, identical data is written to two channels simultaneously. If a memory read from one channel returns incorrect data due to an uncorrectable memory error, the system automatically retrieves the data from the other channel. Mirroring is not lost due to a transient or soft error in one channel, and operation continues until the highly unlikely case of a simultaneous error in exactly the same location on a DIMM and its mirrored DIMM. Mirrored Memory mode reduces the amount of memory available to the operating system by 50 percent since only one of the two populated channels provides data.

On a server based on the Intel Xeon 5500 series processor that support Memory Mirroring, only two channels may be populated when enabling Memory Mirroring. For mirroring, channel 3 remains unpopulated. Channels 1 and 2 are populated identically. See Table 3 for a list of ProLiant G6 servers that support Memory Mirroring.

Lock Step Mode

Lockstep memory mode uses two memory channels at a time and provides an even higher level of protection. In lockstep mode, two channels operate as a single channel—each write and read operation moves a data word two channels wide. The cache line is split across both channels to provide 2x 8-bit error detection and 8-bit error correction within a single DRAM. In three-channel memory systems including servers based on the Intel Xeon 5500 series processors, the third channel is unused and left unpopulated.

---

1 Chipkill is a trademark of International Business Machines Corporation.

See Table 3 for a list of ProLiant G6 servers that support Lock Step Mode.

The Lockstep Memory mode is the most reliable, but it limits the available system memory to a third of the maximum capacity.

Positives of Lock Step Mode:

- Achieves the same level of protection as ChipKill, so there are some additional scenarios in which the system can correct memory errors. Note that Advanced ECC is equivalent to 4-bit ChipKill. Lockstep gets us to 8-bit ChipKill. ChipKill just indicates that an entire DRAM chip can die and the server will keep running.

Negatives of Lock Step Mode:

- You have to leave one of the three memory channels on each processor un-populated, so you cut your available number of DIMM slots by 1/3.
- Performance is measurably slower than normal Advanced ECC mode.
- You can only isolate uncorrectable memory errors to a pair of DIMMs (instead of down to a single DIMM).

Enabling Advanced Memory Protection Features

Advanced ECC memory protection is default. The server's ROM Based Setup Utility (RBSU) can be used to enable Memory Mirroring and Memory Lock-Step Mode.

Population Rules

Certain rules must be followed when populating DDR-3 DIMMs, in order to guarantee that the server will function correctly. In most cases, if any of the following rules are violated, the server will halt during boot and display an error message.

- For servers with eighteen (18) memory slots:
  - There are three (3) channels per processor; six (6) channels per server
  - There are three (3) DIMM slots for each memory channel; eighteen (18) total slots
  - Memory channel 1 consists of the three (3) DIMMs that are closest to the processor
  - Memory channel 3 consists of the three (3) DIMMs that are furthest from the processor
- For servers with twelve (12) memory slots,
  - There are three (3) channels per processor; six (6) channels per server
  - There are two (2) DIMM slots for each memory channel; twelve (12) total slots.
  - Memory channel 1 consists of the two (2) DIMMs that are closest to the processor
  - Memory channel 3 consists of the two (2) DIMMs that are furthest from the processor
- Unbuffered DIMMs (UDIMMs) and Registered DIMMs (RDIMMs) cannot be used together.
  - A server must be either all UDIMMs or all RDIMMs.
- DIMM slots can only be used if the corresponding processor is installed.
  - Do not install DIMMs if the corresponding processor is not installed
  - If only one processor is installed in a 2 processor capable server, only half of the DIMM slots are available
- Up to eight (8) ranks may be installed per channel

    - Quad-rank DIMMs are limited to two per channel.
    - If a quad-rank DIMM is used, it must be installed as the first DIMM on that memory channel. To help identify slots, in a ProLiant G6 server the first DIMM slot in each memory channel is colored white.
    - If a single quad-rank DIMM is installed anywhere in the system, then no channel is allowed to have 3 DIMMs.
- Up to two UDIMMs are allowed per channel.
- Populating Rules for Memory Mirroring and Lock Step Mode:
    - For memory mirroring mode, channel 3 must be unpopulated. Channels 1 and 2 are populated identically
    - For lock-step mode, channel 3 must be unpopulated. DIMMs in channels 1 and 2 will be installed in pairs. The paired slots will be 1,4; 2,5; 3;6 on an 18-slot system or 1,4; 2,5; on a 12-slot system
- At least one DIMM must be installed on a server. It is recommended but not required that at least one DIMM be installed on each processor.
- DIMMs must be populated in a specific order; see below.

Population Order

- Populate DIMMs from heaviest load (quad-rank) to lightest load (single-rank) within a channel
- Heaviest load (DIMM with most ranks) within a channel goes furthest from the processor

You can ensure adherence to these population rules by using the following charts. Populate DIMMs starting with "A" first, "B" second, "C" third, etc. For all servers slots A, B and C are the white slots and should be populated first.

Table 6 : Required DIMM population order for 18 DIMM Server

	CPU1		CPU2	
	slot #	population order	slot #	population order
Channel 1	1	G	1	G
	2	D	2	D
	3	A	3	A
Channel 2	4	H	4	H
	5	E	5	E
	6	B	6	B
Channel 3	7	I	7	I
	8	F	8	F
	9	C	9	C

Table 7 : Required DIMM population order for 12 DIMM Server

	CPU1		CPU2	
	slot #	population order	slot #	population order

Channel 1	1	D	1	D
	2	A	2	A
Channel 2	3	E	3	E
	4	B	4	B
Channel 3	5	F	5	F
	6	C	6	C

DIMM slots in ProLiant G6 servers supporting DDR-3 memory are color coded to ease population. DIMM slots that are colored white represent channel 1, and should be populated first.

DIMM Capacity and Speed

Table 8 shows how fast ProLiant G6 servers will operate DDR3 DIMMs. (This chart does not include impacts of BIOS settings, or the rated memory speed of the processor.)

Table 8 : Populated Memory Speed

DIMM Type	Registered DIMMS (RDIMMs)					Unbuffered DIMMs (UDIMMs)	
DIMM Rank	Dual Rank (2R)				Quad Rank (4R)	Single Rank (1R)	Dual Rank (2R)
DIMM Capacity	2GB	4GB	8GB	8GB	4GB	1GB	2GB
DIMM Native Speed (MHz)	1333	1333	1066	1333	1066	1333	1333
SLOTS THAT CAN BE POPULATED							
9 slot servers	9	9	9	9	6	6	6
12 slot servers	12	12	12	12	12	12	12
18 slot servers	18	18	18	18	12	12	12
MAXIMUM CAPACITY (GB)							
9 slot servers	18	36	72	72	24	6	12
12 slot servers	24	48	96	96	48	12	24
18 slot servers	36	72	144	144	48	12	24
POPULATED DIMM SPEED (MHz)							
1 DIMM Per Channel	1333	1333	1066	1333	1066	1333	1333
2 DIMM Per Channel	1066*	1066*	1066	1066*	800	1066**	1066**
3 DIMM Per Channel	800	800	800	800	not supported	not supported	not supported

** Supported at 1333 via setting in ROM based setup utility (RBSU).*
*** Supported at 1333 via setting in RBSU for 12 DIMM servers only.*

Additional notes on populated memory speed:

- Mixing DIMM speeds is allowed, but the system processor speed rules always override the DIMM capabilities
- If you do mix DIMM speeds, the memory bus will default to the minimum clock rate of all DIMMs in the server.
- Both processors will operate at the same memory clock rate.
- If you install 1x 1066MHz DIMM in channel 1 and 1x 1333MHz DIMM in channel 2, you still run at 1066MHz
- If you install 1x 1066MHz DIMM in channel 1 and 5x 1333MHz DIMMs with 1 DIMM Per Channel (DPC) in each of the other channels, you run at 1066MHz
- If you install 3DPC in one channel (if applicable) and 1DPC in all other channels, you run at 800MHz

General Population Recommendations

Overall, a good balance between performance, power usage, and cost can be achieved by following these general guidelines:

- Populate all 3 channels of each processor. (This means for a 2-processor configuration, populate in groups of 6 identical DIMMs.)
- Use DIMMs with the lowest number of ranks.
- For 24GB or less use UDIMMs.

Optimizing for Capacity

Maximum capacity can be achieved with RDIMMs.

Table 9 : Maximum Memory Capacity

Number of DIMM slots	DIMM Type	Maximum Capacity	Configuration
18	UDIMM	24 GB	12 x 2GB 2R
18	RDIMM	144 GB	18 x 8GB 2R
12	UDIMM	24 GB	12 x 2GB 2R
12	RDIMM	96 GB	12 x 8GB 2R
9	UDIMM	12 GB	6 x 2GB 2R
9	RDIMM	72 GB	9 x 8GB 2R

Optimizing for Performance

The two primary measurements of memory subsystem performance are latency and throughput.

Latency

Memory latency is a measure of how long it takes from the time the processor's Arithmetic Logic Unit (ALU) in the core requests data until it receives the data. Unloaded latency is measured when the system is idle. Loaded latency is measured when the memory sub-section is saturated.

Factors impacting latency include DIMM type, speed, ranks, and CAS timing.

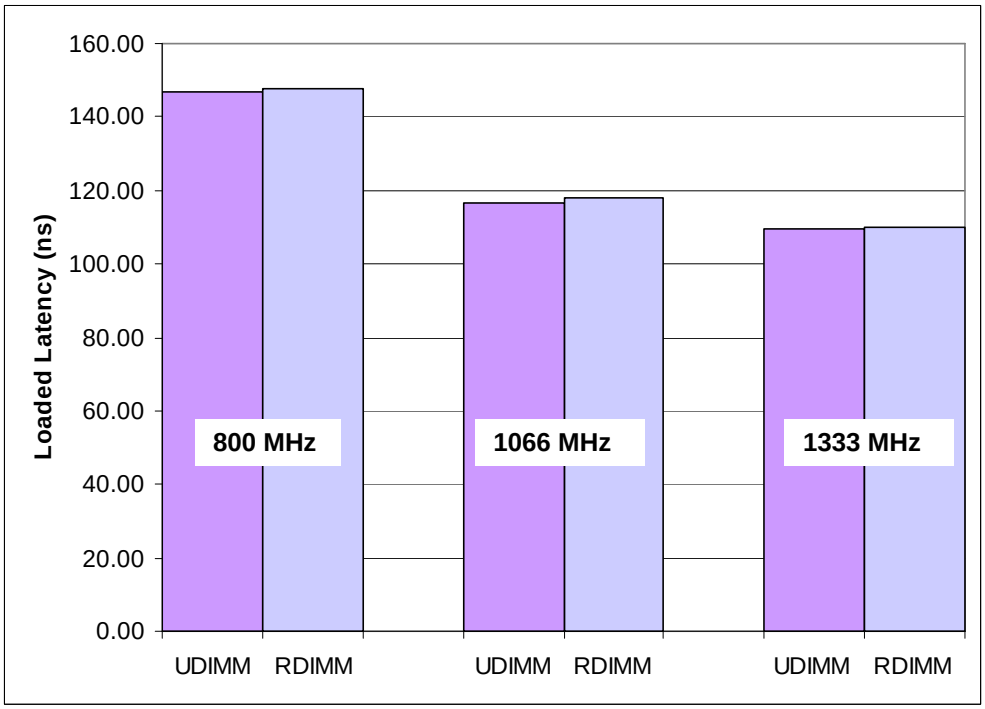
Rated DIMM Speed. To minimize latency, use the fastest DIMMs available. 1333MHz DIMMs, populated with no more than 2 DIMMs per channel (i.e. maximum of 12 DIMMs), achieve

the lowest latency. Memory running at 1333MHz will have a 35% lower latency than memory running at 800MHz.

Note that memory speed is also processor dependent; for lowest latency, use a processor that supports 1333MHz DIMM speed.

Latency is nearly identical between UDIMM and RDIMMs.

Figure 3 : Loaded Latency by Speed and DIMM type



Latency measurements with 6 x 2GB, dual-rank DIMMs.

– **CAS latency** - CAS (Column Address Strobe) latency refers to the DRAM read response time in number of bus clocks from the Column Address to the DRAM providing data on the memory bus. CAS latency is specified in a number, such as "6" or "7". A lower latency means a faster response time. The latency in seconds can be calculated with the CAS latency number and the clock speed.

Table 10 : CAS Latency Timings

DIMM Type	DIMM Speed (MHz)	CAS Latency (CL), in clocks	Single DIMM Latency (in ns)
UDIMM	800	6	15.00
UDIMM	1066	7	13.13
UDIMM	1066	8	15.00
UDIMM	1333	9	13.50
RDIMM	800	6	15.00

RDIMM	1066	7	13.13
RDIMM	1066	8	15.00
RDIMM	1333	9	13.5

All HP single rank and dual rank DIMMs are rated at DDR3-10600R/E -9. When running at DDR3-1333 the CL is 9 clocks or 13.5 ns. At DDR3-1066 the CL will be adjusted to 7 or 13.13 ns. At DDR3-800 the CL will be adjusted to 6 or 15 ns.

All HP quad-rank DIMMs are rated at DDR3-8500R -7. When running at DDR3-1066 the CL is 7 or 13.13 ns. At DDR3-800 the CL will be adjusted to 6 or 15 ns.

- **Ranks**: The number of ranks on a DIMM have a small impact on latency. For the same speed, more ranks will result in lower latency. The analogy is similar to the number of cashiers at the supermarket. Few cachiers result in long queues and long latency. More cashiers result in shorter queues and latency. More ranks on a channel give the memory controller more places it may go address to get the desired memory. Fewer ranks will result in more requests queuing up.

Table 11 : Latency and Throughput by DIMM Rank

	Unloaded Latency (ns)	Loaded Latency (ns)	Throughput (GB/s)
DDR3-800 Dual Rank	82.50	148.80	30.70
DDR3-800 Quad Rank	87.21	151.19	29.89
DDR3-1066 Dual Rank	87.36	118.46	37.10
DDR3-1066 Quad Rank	85.40	117.44	37.35
DDR3-1333 Dual Rank	86.54	110.45	38.52

- **Processor affinity**. In the NUMA architecture, accesses from one processor to a memory region attached to a different processor will have much longer latency (slower response) than accesses to a memory region attached to the local processor.

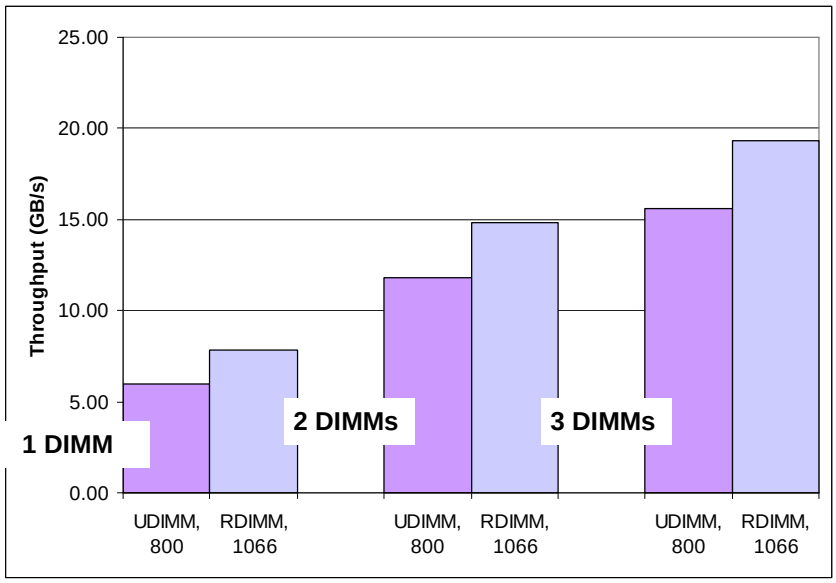
Throughput

Factors impacting memory throughput include the number of memory channels populated, and the speed at which the memory runs.

Memory Channels. The biggest impact on throughput is the number of memory channels populated. By interleaving memory access across multiple memory channels, the integrated memory controller of the Intel Xeon 5500 series processor is able to greatly increase memory throughput. HP recommendations that you always populate all three channels on each installed processor. For ProLiant G6 servers using the Intel Xeon 5500 series processor, this means populating DIMMs in sets of threes for each processor.

As shown in Figure 4, adding a second DIMM (and thus populating the second memory channel) increases bandwidth by 85% to 95%. Adding a third DIMM further increases bandwidth by 30% to 35%.

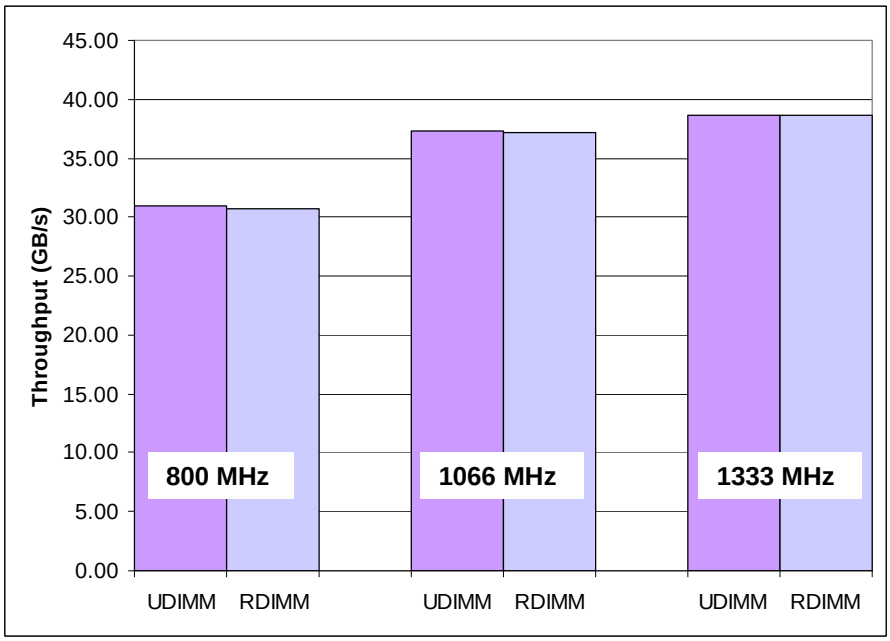
Figure 4: Throughput with 1, 2, and 3 Channels



Memory Throughput with 1 processor, using 1 GB x 1 Rank UDIMM at 800MHz, and 8GB x 4 Rank RDIMM at 1066MHz.

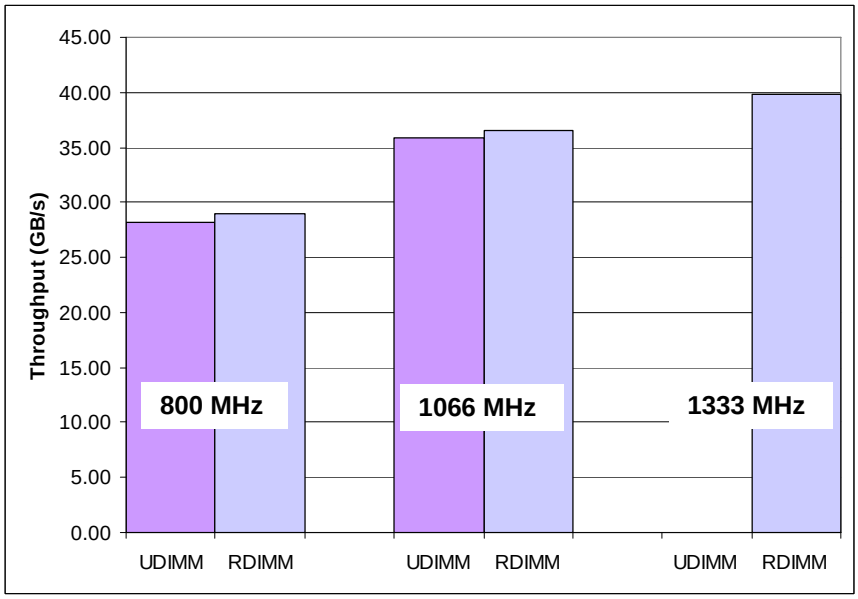
– **Memory Speed**. Higher memory speeds increase throughput. **Figure 5**, **Figure 6**, and **Figure 7** show that throughput at 1066MHz memory speed is 20% to 25% higher than at 800MHz. Throughput at 1333MHz memory speed is 3.5% to 9% higher than 1066MHz.

Figure 5: Throughput by DIMM Speed, 6x2GB DIMMs



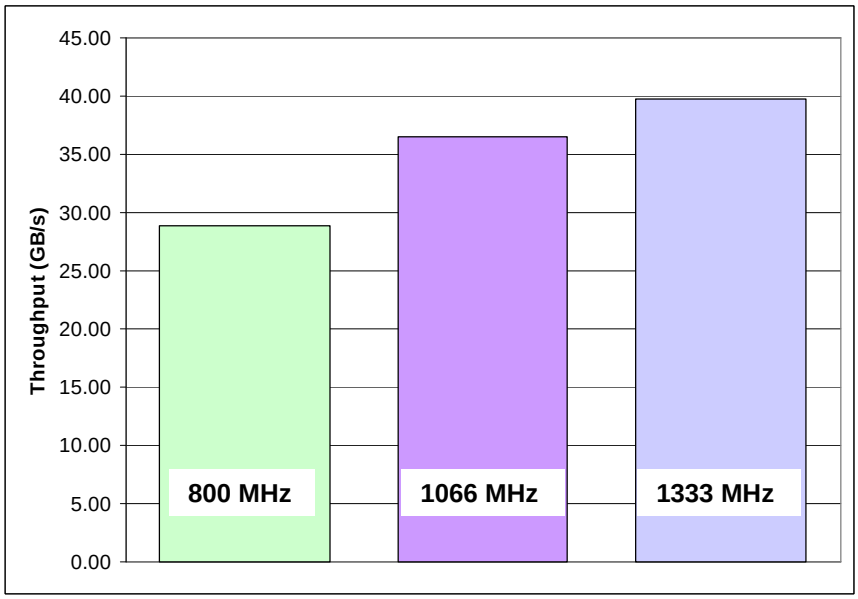
Throughput measurements on a 2-processor configuration using 6 x 2GB dual-rank DIMMs.

Figure 6: Throughput by DIMM Speed, 12x2GB DIMMs



Throughput measurements on a 2-processor configuration using 6 x 2GB dual-rank DIMMs.

Figure 7: Throughput by DIMM Speed, 12x4GB DIMMs



Throughput measurements on a 2-processor configuration using 12 x 4GB dual-rank DIMMs.

– **Two DIMMs per Channel at 1333:** Although Intel limited its design to only support 1 DIMM per channel when running at 1333MHz, HP has engineered its ProLiant G6 servers to reliability operate at 1333MHz speeds with 2 DIMMs per channel.

  Enabling 2 DIMMs per channel at 1333MHz increases throughput by about 10% and decreases latency by 15% over 2 DIMMs per channel operating at 1066MHz. It consumes about 5% more power.

Table 12 : Impact of Two DIMMs Per Channel at 1333MHz

	2DPC @1066	2DPC @1333	Delta
Idle Latency (ns)	86.5	87.4	1%
Loaded Latency (ns)	125.3	109.3	-15%
Throughput (GB/s)	36.5	39.7	8%
Idle Power (W)	37.7	42.3	11%
Loaded Power (W)	120.0	132.8	10%

Mixing DIMM sizes

There are no performance implications for mixing sets of different capacity DIMMs at the same operating speed. For example, latency and throughput will not be negatively impacted by installing 6 x 4GB Dual-Rank DDR3-1333 DIMMs (one per channel), plus 6 x 2GB Dual-Rank DDR3-1333 DIMMs (one per channel)

Optimizing for Power

The power used by a DIMM is impacted by the DIMM technology, capacity, number of ranks, and speed. To a lesser extent, power consumed is also impacted by the efficiency of the voltage regulators supplying power to the DIMMs and to the DDR3 interface of the memory controller.

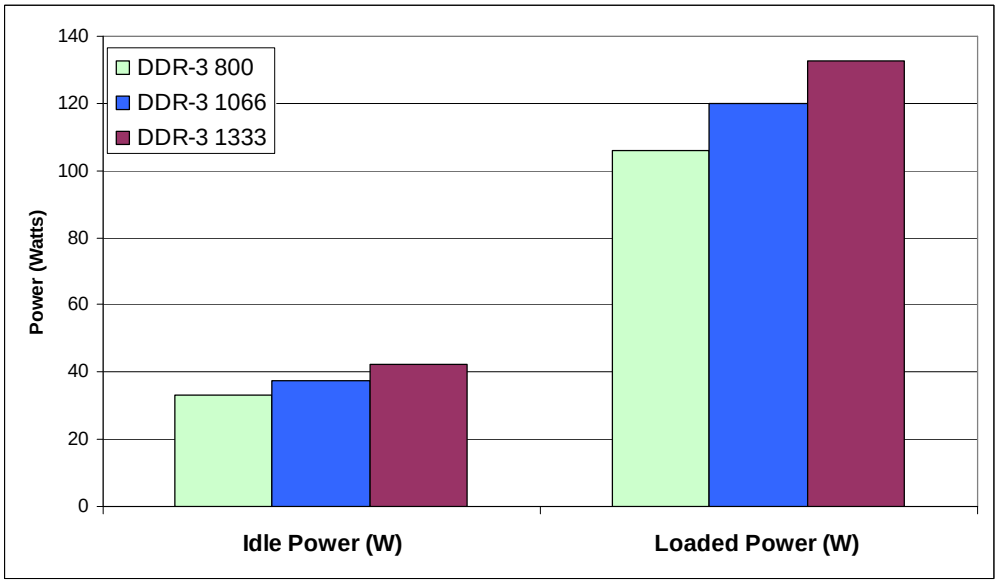
Generally, the following rules should be used to optimize for power:

- Use UDIMMs instead of RDIMMs.
- Use the smallest number of DIMMs possible, by using the highest capacity DIMM.
- Choose quad-rank DIMMs over dual-rank DIMMs.
- Run memory as the slowest speed possible.

Factors impacting power usage:

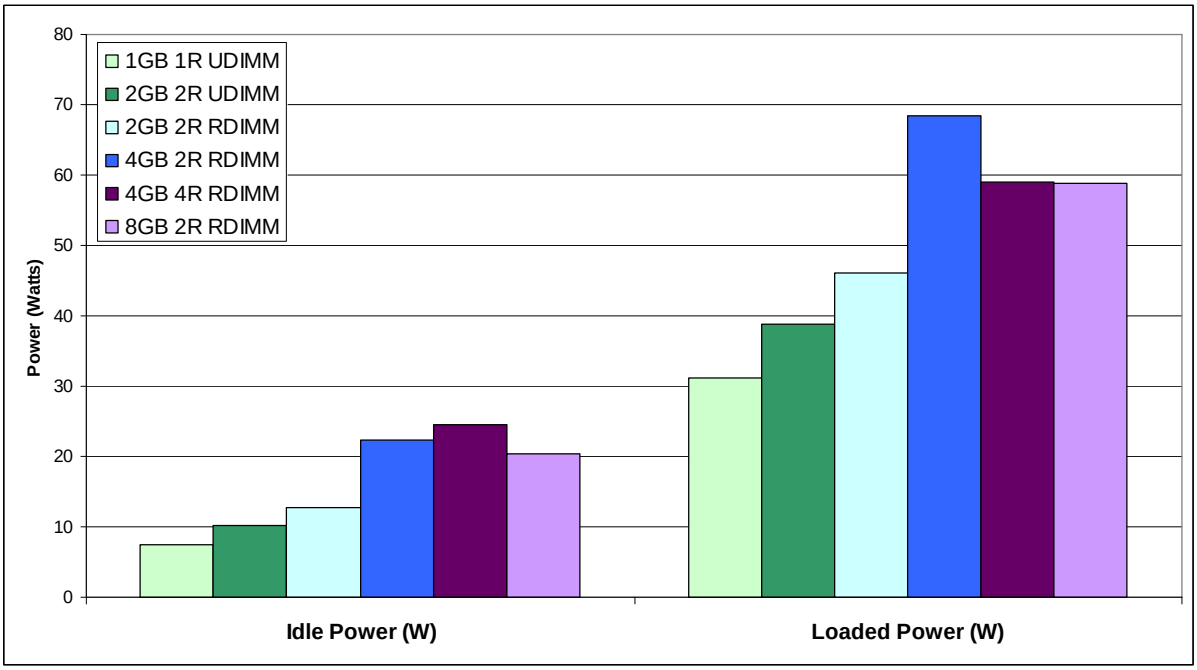
– **Memory Speed**: Memory operating at 1333MHz speed consumes about 25% more power under loaded conditions than memory operating at 800MHz speed.

Figure 8: Power by Memory Speed (12x4GB Registered 2-Rank DIMMs)



DIMM type and capacity: Generally, large capacity DIMMs consume more power. However, on a per-gigabyte basis, large capacity DIMMs are more efficient. DIMM made with 2Gb DRAM, such as the 8GB 2R DIMM are the most efficient. At equal sizes and ranks, UDIMMs consume less power than RDIMMs. At DDDR3-1066 speed under loaded conditions, each 2GB UDIMM uses about 1.2 watts less than a comparable 2GB RDIMM.

Figure 9 : Power by DIMM Capacity (Six DDR3-1066 DIMMs)



DIMM ranks: Quad-ranked DIMMs consume about 15% less power than dual-rank DIMMs.

Figure 10: Impact of Ranks on Power Used (DDR3-1066)

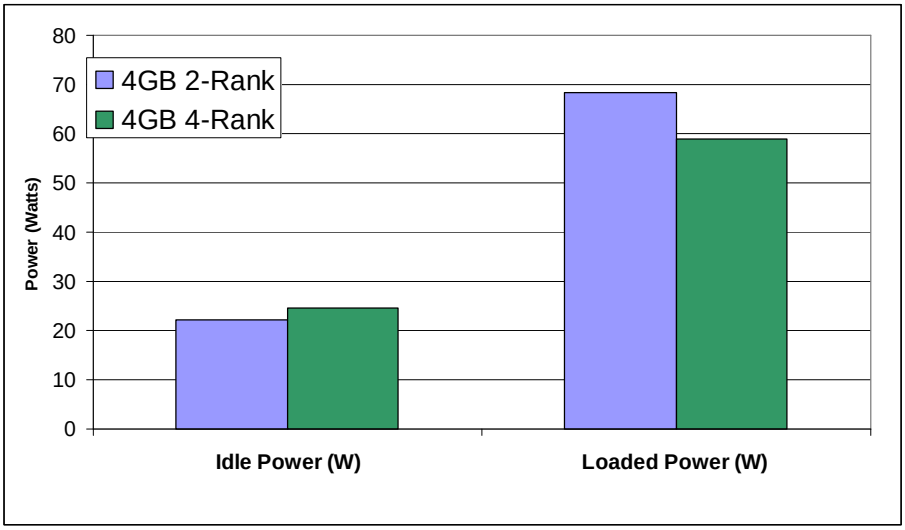


Table 13 and Table 14 compare the loaded power used by the different configurations that achieve 24GB and 48GB total memory capacity.

Table 13 : Power of various 24GB Configurations

Total Memory (GB)	Memory Config	Number of DIMMs	DIMM Size	DIMM Rank	UDIMM or RDIMM	DIMM Speed	DDR3 Loaded Power (W)
24	6x4G4R_800_R	6	4GB	4	RDIMM	800	51.82
24	6x4G4R_1067_R	6	4GB	4	RDIMM	1066	58.96
24	6x4G2R_800_R	6	4GB	2	RDIMM	800	59.34
24	12x2G2R_800_U	12	2GB	2	UDIMM	800	59.88
24	12x2G2R_1067_U	12	2GB	2	UDIMM	1066	67.33
24	6x4G2R_1067_R	6	4GB	2	RDIMM	1066	68.45
24	12x2G2R_800_R	12	2GB	2	RDIMM	800	72.99
24	6x4G2R_1333_R	6	4GB	2	RDIMM	1333	75.24
24	12x2G2R_1067_R	12	2GB	2	RDIMM	1066	80.29
24	12x2G2R_1333_R	12	2GB	2	RDIMM	1333	87.04

Table 14 : Power of various 48GB Configurations

Total Memory (GB)	Memory Config	Number of DIMMs	DIMM Size	DIMM Rank	UDIMM or RDIMM	DIMM Speed	DDR3 Loaded Power (W)
48	6x8G2R_800_R	6	8GB	2	RDIMM	800	52.79
48	6x8G2R_1067_R	6	8GB	2	RDIMM	1066	58.92
48	6x8G2R_1333_R	6	8GB	2	RDIMM	1333	62.96
48	12x4G4R_800_R	12	4GB	4	RDIMM	800	90.23
48	12x4G2R_800_R	12	4GB	2	RDIMM	800	106.08
48	12x4G2R_1067_R	12	4GB	2	RDIMM	1066	119.97

48	12x4G2R_1333_R	12	4GB	2	RDIMM	1333	132.80

Low Power DIMMs: HP offers a performance 4GB 2R PC3-10600 and a low power 4G 4R PC3-8500 RDIMM. The dual rank DIMM is capable of running up to DDR3-1333 date rate or 10.6GB/s. The quad-rank DIMM is capable of running up to DDR3-1066 or 8.5GB/s. In the dual-rank DIMM half of the DRAM chips are active at a time. In the quad-rank DIMM, a quarter of the DRAM chips are active at a time resulting in 15% less power consumption.

Optimizing for Cost

Generally, for a given memory capacity, cost is minimized by using the smallest DIMM size possible to achieve that capacity. However, memory prices often change, so it's important to calculate the price of different memory configurations using the latest pricing.
General rules to achieve the lowest cost on memory:

- For configurations of 24GB or below, use UDIMMs.
- Choose many 4GB DIMMs instead of higher capacity DIMMs. (Generally, 1GB, 2GB, and 4GB DIMMs cost about the same per gigabyte. DIMMs higher than 8GB have a cost-per-GB premium.)
- Use standard-power memory instead of low-power memory.

Performance Comparisons

These charts summarize latency, bandwidth, and power of three different configurations.

For clarity, these charts abbreviate DIMM type and configuration. DIMM count and capacity, number of ranks, memory speed, and DIMM type are abbreviated. For example, "6x2G2R_800_U" refers to six 2GB DIMMs, each of which is a 2-rank (dual-rank), DDR3-800 UDIMM.

Figure 11: Performance of 12GB Configurations

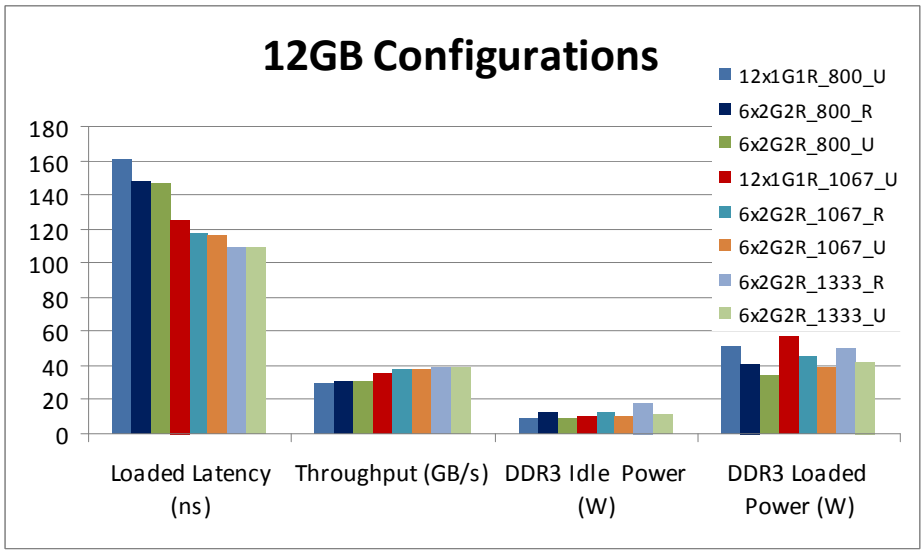


Figure 12: Performance of 24GB Configuration

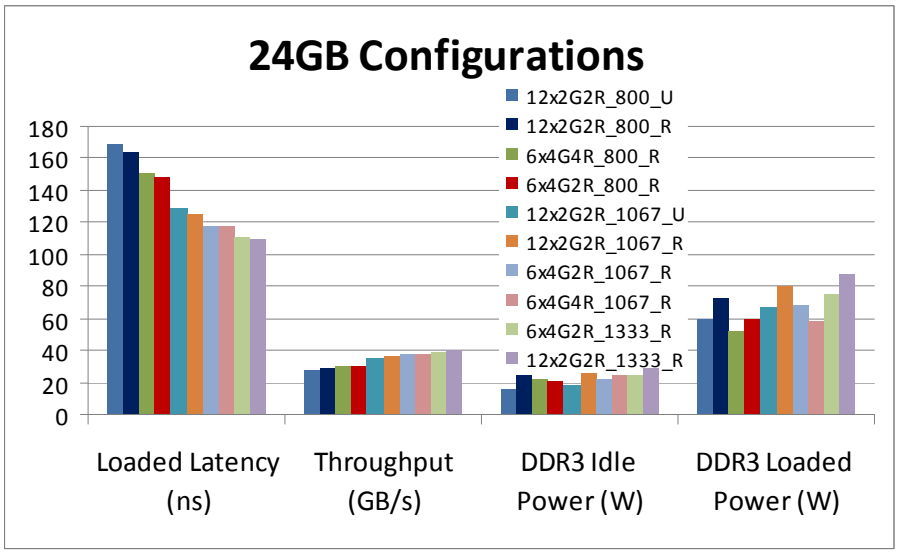
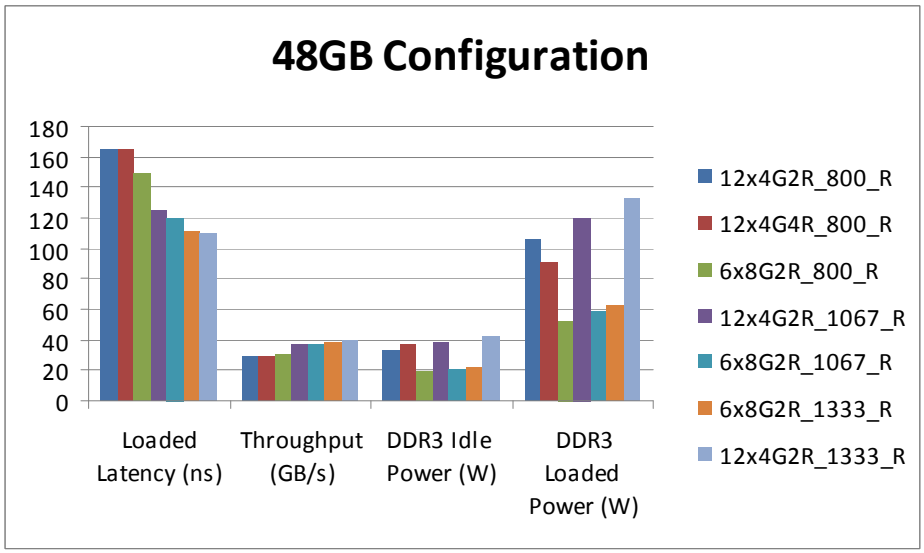


Figure 13: Performance of 48GB Configuration



Unbalanced Configurations

There are two categories of unbalanced configurations.

- *Unbalanced across channels*: A different amount of memory on each of the 3 channels of each processor. For example, putting 4x4GB DIMMs on each processor yields an unbalanced configuration across channels.
- *Unbalanced across processors*: A different amount of memory on each processor. For example, putting 3x4GB on one processor and 6x4GB on the second processor.

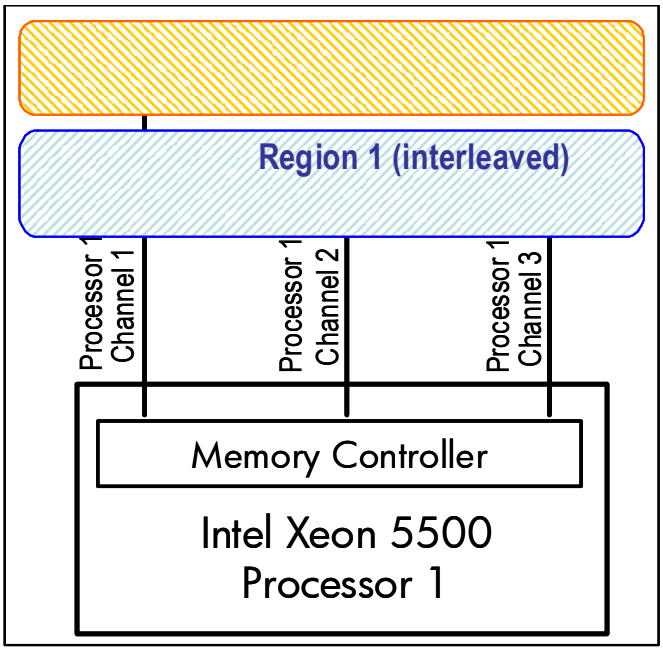
In unbalanced configuration across channels, the memory controller will split memory up into regions. Each region of memory will have different performance characteristics. The memory controller groups memory across channels as much as possible to create the regions. It will create as many regions as possible with DIMMs that span all three memory channels. Next, it will move to create regions that span two memory channels and then to just one.

Unblanaced Across Channels

The overall system bandwidth in each type of region will be different. In regions that span all three channels, the memory performance will be best. When moving to a region that spans two channels, the performance will drop by about 1/3 (due to dropping 1 of 3 channels). Moving to the next type of region that spans only one channel, the performance drops again by a similar amount.

For example, take a 2-processor, 64GB configuration using 8x8GB DIMMs. Four DIMMs will be put onto each processor:

Figure 14 : Memory Unbalanced Across Channels



1. Performance will be highest in Region 1, where the memory spans 3 memory channels.
2. Performance will be worse in Region 2, where the memory is only on the first channel; that is, on the memory in the 4th DIMM of each processor. In addition, this memory channel will also have increased latency.

Comparing the performance of this example (8x8GB) versus a balanced configuration (6x8GB plus 6x4GB):

In both the 8x 8GB and the 6x 8GB + 6x 4GB solutions, the memory bus will run at the same speed - 1067. This equates to max theoretical bandwidth of 8.5GB/s per channel.

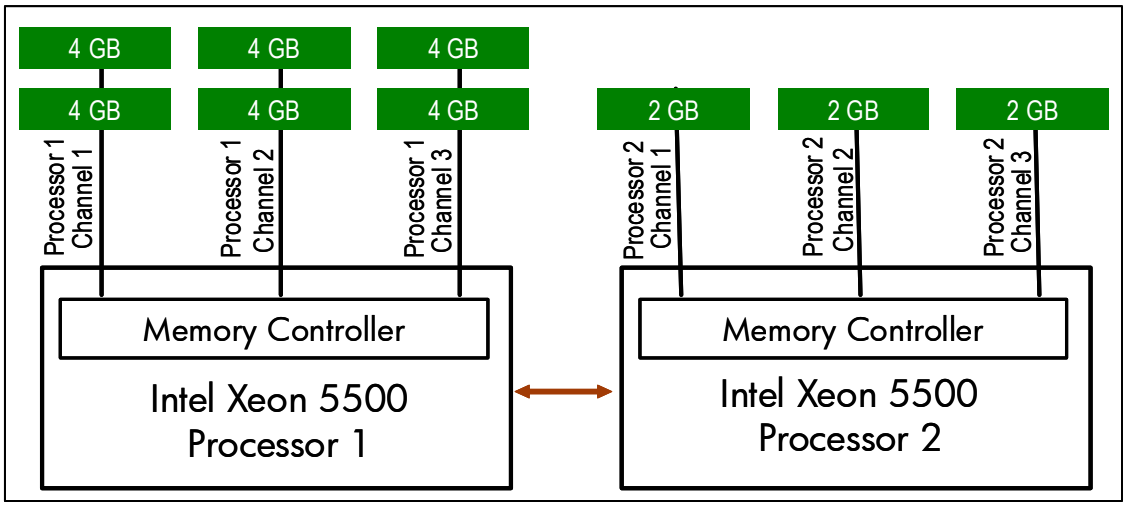
For the 6x8GB + 6x4GB solution, and in that region of memory of the 8x8GB solution that spans 3 channels, the region maximum bandwidth will be 25.5GB/s per processor. For that region of the 8x8GB solution that only spans one channel, the region maximum bandwidth will be 8.5GB/s per processor. This means peak bandwidth would be 66% lower on the unbalanced configuration.

Total bandwidth for the balanced 6x8GB + 6x4GB solution would be 51GB/s. If you assume an application spreads memory access equally across all memory in the unbalanced 8x8GB configuration, with both processors accessing all memory regions, average system throughput in the unbalanced configuration would be about 42GB/s, or about 16% lower bandwidth on average.

Unbalanced Across Processors

In unbalanced configuration across processors, you will have different performance characteristics for threads running on each processor.

Figure 15 : Memory Unbalanced Across Processors



Application specific considerations

For virtualization, use as few high-capacity DIMMs as possible to meet the memory requirement.

VMware's best practice guidelines advise customers that to avoid performance degradation, unbalanced memory configurations should be avoided. *"The host hardware should be configured so that physical host memory is evenly balanced across NUMA nodes."* (http://pubs.vmware.com/vi35/resmgmt/wwhelp/wwhimpl/common/html/wwhelp.htm?context=resmgmt&file=vc_best_practice.13.8.html)

BIOS Settings

These settings in the ROM Based Setup Utility (RBSU) of ProLiant G6 server BIOS provide control over memory configuration:

- Maximum memory speed
  - Used to limit the speed at which memory runs
  - Can be set to Automatic (speed determined according to normal population rules), 1066 MHz, or 800 MHz.
  - Setting this to 1066 MHz or 800MHz can save power, but will incur some performance penalty.
- Memory interleaving
  - Disabling memory interleaving saves some power per DIMM, but will incur some performance penalty.

For more information

Memory technology evolution: an overview of system memory technologies: http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00256987/c00256987.pdf

DDR3 Online Memory Configuration Tool: http://www.hp.com/go/ddr3memory-configurator

HP ProLiant BL460c G6 server blade: http://www.hp.com/servers/proliant/bl460c

HP ProLiant DL380 G6 server: http://www.hp.com/servers/proliant/dl380

Optimizing the Performance of IBM System x and BladeCenter Servers using Intel Xeon 5500 Series Processors: ftp://ftp.software.ibm.com/common/ssi/sa/wh/n/xsw03025usen/XSW03025USEN.PDF

Dell™ PowerEdge™ Server 2009 – Memory : http://www.dell.com/downloads/global/products/pedge/en/server-poweredge-11g-white-paper.pdf

The Intel processor roadmap for industry-standard servers" technology brief http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00164255/c00164255.pdf

Additional HP Technology Papers can be found at http://h18004.www1.hp.com/products/servers/technology/whitepapers/index.html?jumpid=servers/technology

Additional information on HP ProLiant servers can be found at www.hp.com/servers/proliant

Send comments about this whitepaper to Daniel R. Bowers (daniel.r.bowers@hp.com)

Appendix A

Performance Test Results

Measurements on an 18-DIMM slot server with 2 processors installed.

Configuration	Total Memory (GB)	DIMM Size (GB)	Ranks/ DIMM	Data Rate	DIMMs per Channel	Total DIMMs	Unloaded Latency (ns)	Loaded Latency (ns)	Throughput (GB/s)	Idle Power (W)	Loaded Power (W)
6x1G1R_800_U	6	1	1	800	1	6	89.77	145.56	29.49	7.01	28.63
6x1G1R_1067_U	6	1	1	1067	1	6	87.95	123.24	34.00	7.50	31.14
6x1G1R_1333_U	6	1	1	1333	1	6	89.19	120.17	34.95	8.44	34.00
12x1G1R_800_U	12	1	1	800	2	12	82.84	160.57	29.19	9.52	51.14
6x2G2R_800_R	12	2	2	800	1	6	89.16	147.61	30.77	12.72	40.62
6x2G2R_800_U	12	2	2	800	1	6	80.11	146.59	30.98	9.30	33.94
12x1G1R_1067_U	12	1	1	1067	2	12	89.29	125.62	35.75	10.50	56.47
6x2G2R_1067_R	12	2	2	1067	1	6	86.91	117.78	37.19	12.68	46.08
6x2G2R_1067_U	12	2	2	1067	1	6	86.52	116.76	37.38	10.24	38.92
6x2G2R_1333_R	12	2	2	1333	1	6	70.23	109.77	38.60	18.38	50.67
6x2G2R_1333_U	12	2	2	1333	1	6	87.61	109.43	38.69	11.58	42.42
12x2G2R_800_U	24	2	2	800	2	12	83.18	168.58	28.23	16.21	59.88
12x2G2R_800_R	24	2	2	800	2	12	89.70	164.48	28.93	23.76	72.99
6x4G4R_800_R	24	4	4	800	1	6	87.21	151.19	29.89	21.89	51.82
6x4G2R_800_R	24	4	2	800	1	6	82.50	148.80	30.70	20.12	59.34
12x2G2R_1067_U	24	2	2	1067	2	12	89.25	128.52	35.81	18.16	67.33
12x2G2R_1067_R	24	2	2	1067	2	12	86.25	124.77	36.52	26.13	80.29
6x4G2R_1067_R	24	4	2	1067	1	6	87.36	118.46	37.10	22.27	68.45
6x4G4R_1067_R	24	4	4	1067	1	6	85.40	117.44	37.35	24.54	58.96
6x4G2R_1333_R	24	4	2	1333	1	6	86.54	110.45	38.52	24.71	75.24
12x2G2R_1333_R	24	2	2	1333	2	12	86.12	109.60	39.77	28.89	87.04
18X2G2R_800_R	36	2	2	800	3	18	81.13	160.39	29.21	32.50	98.65
12x4G2R_800_R	48	4	2	800	2	12	79.77	165.00	28.88	33.30	106.08
12x4G4R_800_R	48	4	4	800	2	12	90.62	165.00	28.89	36.98	90.23
6x8G2R_800_R	48	8	2	800	1	6	79.77	150.00	30.53	19.18	52.79
12x4G2R_1067_R	48	4	2	1067	2	12	86.50	125.28	36.46	37.70	119.97
6x8G2R_1067_R	48	8	2	1067	1	6	88.19	119.49	36.80	20.40	58.92
6x8G2R_1333_R	48	8	2	1333	1	6	87.20	111.47	38.30	21.78	62.96
12x4G2R_1333_R	48	4	2	1333	2	12	87.41	109.26	39.70	42.25	132.80
18X4G2R_800_R	72	4	2	800	3	18	86.93	161.76	29.20	47.29	145.94
12x8G2R_800_R	96	8	2	800	2	12	80.11	165.00	28.77	30.88	95.33
6x16G4R_800_R	96	16	4	800	1	6	80.11	153.92	29.77	28.92	66.57
12x8G2R_1067_R	96	8	2	1067	2	12	71.93	125.96	36.23	33.30	104.20
6x16G4R_1067_R	96	16	4	1067	1	6	72.61	119.31	37.12	31.30	73.79
12x8G2R_1333_R	96	8	2	1333	2	12	70.91	110.11	39.47	34.51	111.20
18x8G2R_800_R	144	8	2	800	3	18	80.11	161.59	29.04	43.57	131.39
12x16G4R_800_R	192	16	4	800	2	12	80.45	166.53	28.70	49.55	123.32

Appendix B

Recommended Configurations for 2-processor servers

Recommendations for a 2 Processor, 18 DIMM slot configuration optimized for performance:

Capacity Desired	DIMMs to use	
	CPU 1	CPU 2
> 6GB	Not Recommended	
6 GB	3 x 1GB UDIMM	3 x 1GB UDIMM
12 GB	3 x 2GB UDIMM	3 x 2GB UDIMM
18 GB	3 x 1GB UDIMM plus 3 x 2GB UDIMM	3 x 1GB UDIMM plus 3 x 2GB UDIMM
24 GB	3 x 4GB 2R RDIMM	3 x 4GB 2R RDIMM
36 GB	9 x 2GB RDIMM	9 x 2GB RDIMM
48 GB	6 x 4GB 2R RDIMM	6 x 4GB 2R RDIMM
60 GB	6 x 4GB 2R RDIMM plus 3 x 2GB RDIMM	6 x 4GB 2R RDIMM plus 3 x 2GB RDIMM
72 GB	9 x 4GB 2R RDIMM	9 x 4GB 2R RDIMM
96 GB	6 x 8GB RDIMM	6 x 8GB RDIMM
108 GB	6 x 8GB RDIMM plus 3 x 2GB RDIMM	6 x 8GB RDIMM plus 3 x 2GB RDIMM
120 GB	6 x 8GB RDIMM plus 3 x 4GB 2R RDIMM	6 x 8GB RDIMM plus 3 x 4GB 2R RDIMM
144 GB	9 x 8GB 2R RDIMM	9 x 8GB 2R RDIMM

Recommendations for a 2 Processor, 12 DIMM slot configuration optimized for performance:

Capacity Desired	DIMMs to use	
	CPU 1	CPU 2
> 6GB	Not Recommended	
6 GB	3 x 1GB UDIMM	3 x 1GB UDIMM
12 GB	3 x 2GB UDIMM	3 x 2GB UDIMM
18 GB	3 x 1GB UDIMM plus 3 x 2GB UDIMM	3 x 1GB UDIMM plus 3 x 2GB UDIMM
24 GB	3 x 4GB 2R RDIMM	3 x 4GB 2R RDIMM
36 GB	3 x 4GB 2R RDIMM plus 3 x 2GB RDIMM	3 x 4GB 2R RDIMM plus 3 x 2GB RDIMM
48 GB	6 x 4GB 2R RDIMM	6 x 4GB 2R RDIMM
60 GB	3x 8GB RDIMM plus 3 x 2GB RDIMM	3x 8GB RDIMM plus 3 x 2GB RDIMM
72 GB	3 x 8GB RDIMM plus	6 x 8GB RDIMM plus

	3 x 4GB 2R RDIMM	3 x 4GB 2R RDIMM
96 GB	6 x 8GB RDIMM	6 x 8GB RDIMM

May 2009