Thoughts on Intel® Xeon® E5-2600 v2 Product Family Performance Optimisation – component selection guidelines

August 2014

Marcin Kaczmarski

Romley EP/EN Platforms

Intel® Xeon® Processor E5-2600 v2/2400 v2 Product Families

Ivy Bridge CPUs

Socket R: Up to 12 cores / socket

Socket B2: Up to 10 cores / socket

QPI

Socket R: 2 QPI links

Socket B2: 1 QPI link

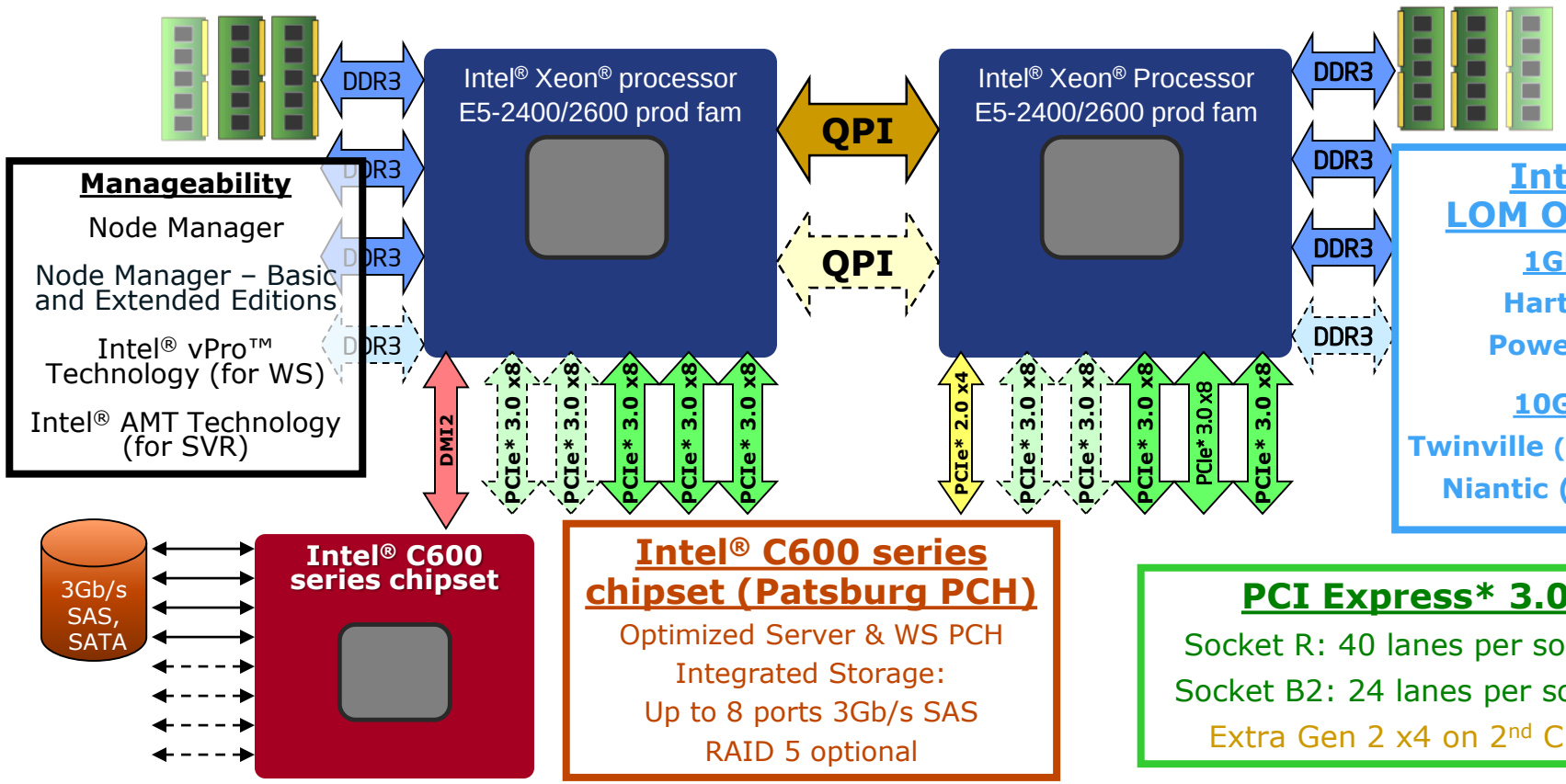
Memory

DDR3 & DDR3L

RDIMMs & UDIMMs, LR DIMMs

Socket R: 4 channels per socket, up to 3 DPC; speeds up to DDR3 1866

Socket B2: 3 channels per socket, up to 2 DPC; speeds up to DDR3 1600



Manageability

Node Manager

Node Manager – Basic and Extended Editions

Intel® vPro™ Technology (for WS)

Intel® AMT Technology (for SVR)

Intel® LOM Options

1GbE:

Hartwell

Powerville

10GbE:

Twinville (10GBase-T)

Niantic (SfI/KR)

Intel® C600 series chipset (Patsburg PCH)

Optimized Server & WS PCH

Integrated Storage:

Up to 8 ports 3Gb/s SAS

RAID 5 optional

PCI Express* 3.0

Socket R: 40 lanes per socket

Socket B2: 24 lanes per socket

Extra Gen 2 x4 on 2nd CPU

Intel® Xeon® Processor E5-2600 v2 Product Family

LGA2011 (E5)

Advanced

- 8C-20MB, 10C-25MB cache
- 8.0 GT/s QPI
- Intel® HT Technology
- DDR3-1866 (skt R)
- Intel® Turbo Boost Technology

10C 3.0GHz	130W	E5-2690 v2
10C 2.8GHz	115W	E5-2680 v2
10C 2.5GHz	115W	E5-2670 v2
10C 2.2GHz	95W	E5-2660 v2
8C 2.6GHz	95W	E5-2650 v2

Standard

- 6C-15MB, 8C-20MB cache
- 7.2 GT/s QPI
- Intel® HT Technology
- DDR3 1600
- Intel® Turbo Boost Technology

8C 2.0GHz	95W	E5-2640 v2
6C 2.6GHz	80W	E5-2630 v2
6C 2.1GHz	80W	E5-2620 v2

Basic

- 4C-10MB, 2C-5MB cache
- 6.4 GT/s QPI
- DDR3 1333

4C 2.5GHz	80W	E5-2609 v2
4C 1.8GHz	80W	E5-2603 v2

T Tray only SKUs

LGA2011 (E5)

Segment Optimized

12C 2.7GHz 30M	130W	E5-2697 v2
12C 2.4GHz 30M	115W	E5-2695 v2

- 2.5MB/c cache

More cache as noted

- 8.0 GT/s QPI
- Intel® HT Technology
- DDR3-1866
- Intel® Turbo Boost Technology

8C 3.4GHz	150W	E5-2687W v2
8C 3.3GHz 25M	130W	E5-2667 v2 T
6C 3.5GHz 25M	130W	E5-2643 v2 T
4C 3.5GHz 15M	130W	E5-2637 v2

Workstation Only SKU

intel inside XEON

- 2.5MB/c cache
- 10C 8.0 GT/s QPI
- 6C 7.2 GT/s QPI
- DDR3-1600
- Intel® HT Technology
- Intel® Turbo Boost Technology

Low Power

10C 1.7GHz	70W	E5-2650L v2 T
6C 2.4GHz	60W	E5-2630L v2 T

Q1: What are the key performance differences on 12C & 10C E5-2600 v2 SKUs?

There are 3 key differences that affect performance:

1. Number of Memory Controllers
   - 12C SKUs have 2 memory controllers, all other SKUs have 1 memory controller
   - The 2nd memory controller gives a 3-4% improvement in memory bandwidth[1]
   - More details provided in: Intel Xeon Processor E5 Family Romley Memory Config Guide
2. Turbo Frequencies
   - Top-bin 10C part (E5-2690 v2) has a higher max frequency than the 12C SKUs
   - Frequency remains higher for top-bin 10C vs. 12C SKUs at low core counts

Processor SKU	TDP (W)	Base Freq. (GHz)	# of Cores	Maximum Frequency (+ x00 MHz over base frequency)											
				Number of Active Cores											
				1	2	3	4	5	6	7	8	9	10	11	12
E5-2697 v2	130W	2.7	12	+8	+7	+6	+5	+4	+3	+3	+3	+3	+3	+3	+3
E5-2695 v2	115W	2.4	12	+8	+7	+6	+5	+4	+4	+4	+4	+4	+4	+4	+4
E5-2690 v2	130W	3.0	10	+6	+5	+4	+3	+3	+3	+3	+3	+3	+3	N/A	N/A

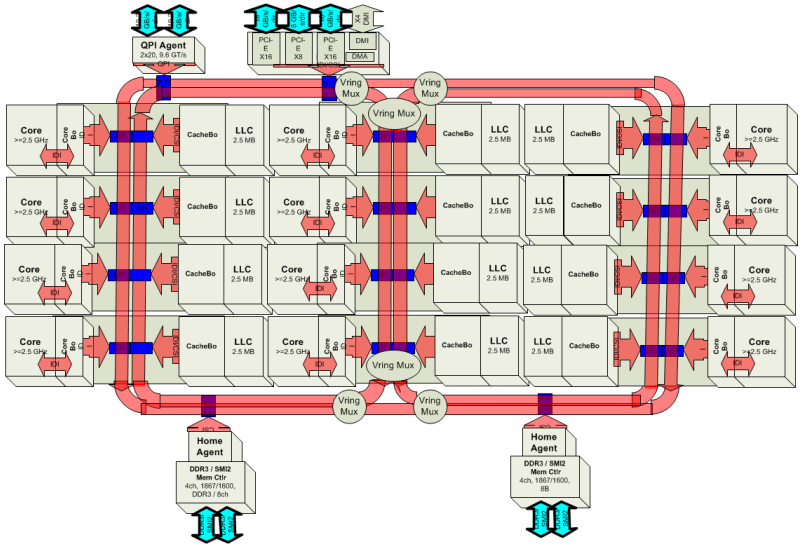
[1]Source: Intel internal measurements as of Sept 2013 on STREAM Triad v5.4. E5-2697 v2, 8x16GB DR-DDR3-1866 DIMMS vs. E5-2690 v2, 8x16GB DR-DDR3-1866 DIMMS. Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. **For more information go to** http://www.intel.com/performance

Q1: What are the key performance differences on 12C & 10C E5-2600 v2 SKUs?

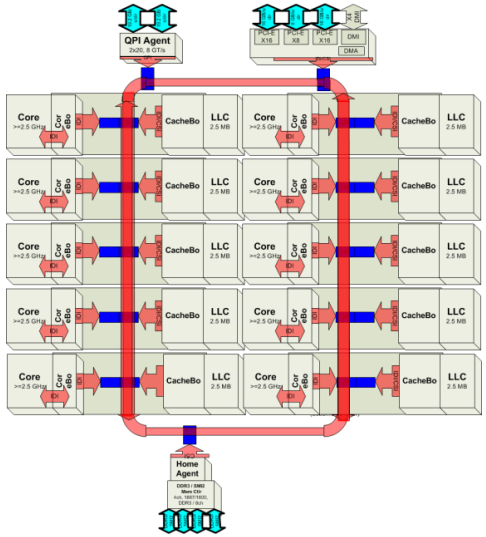
Latency for sequential reads (ns)	LLC	Local Memory	Remote Socket Memory
E5-2690 v2 (10C, 3.0GHz, DDR3-1866)	13 ns	58 ns	113 ns
E5-2697 v2 (12C, 3.0GHz, DDR3-1866)	16 ns	61 ns	116 ns

3. Last Level Cache (LLC) Latency
   - At the same frequency, 12C SKUs have 20% higher LLC latency than 10C SKUs
   - Differences in LLC latency are due to the die configurations
     - 12C die has 3 columns of cache & core, with the 3rd column adding to LLC latency
     - 10C die has 2 columns of cache & core, similar to E5-2600 (SNB-EP)

12C Die Configuration



10C Die Configuration



Q2: At low thread counts, how does performance differ on 12C & 10C E5-2600 v2 SKUs?

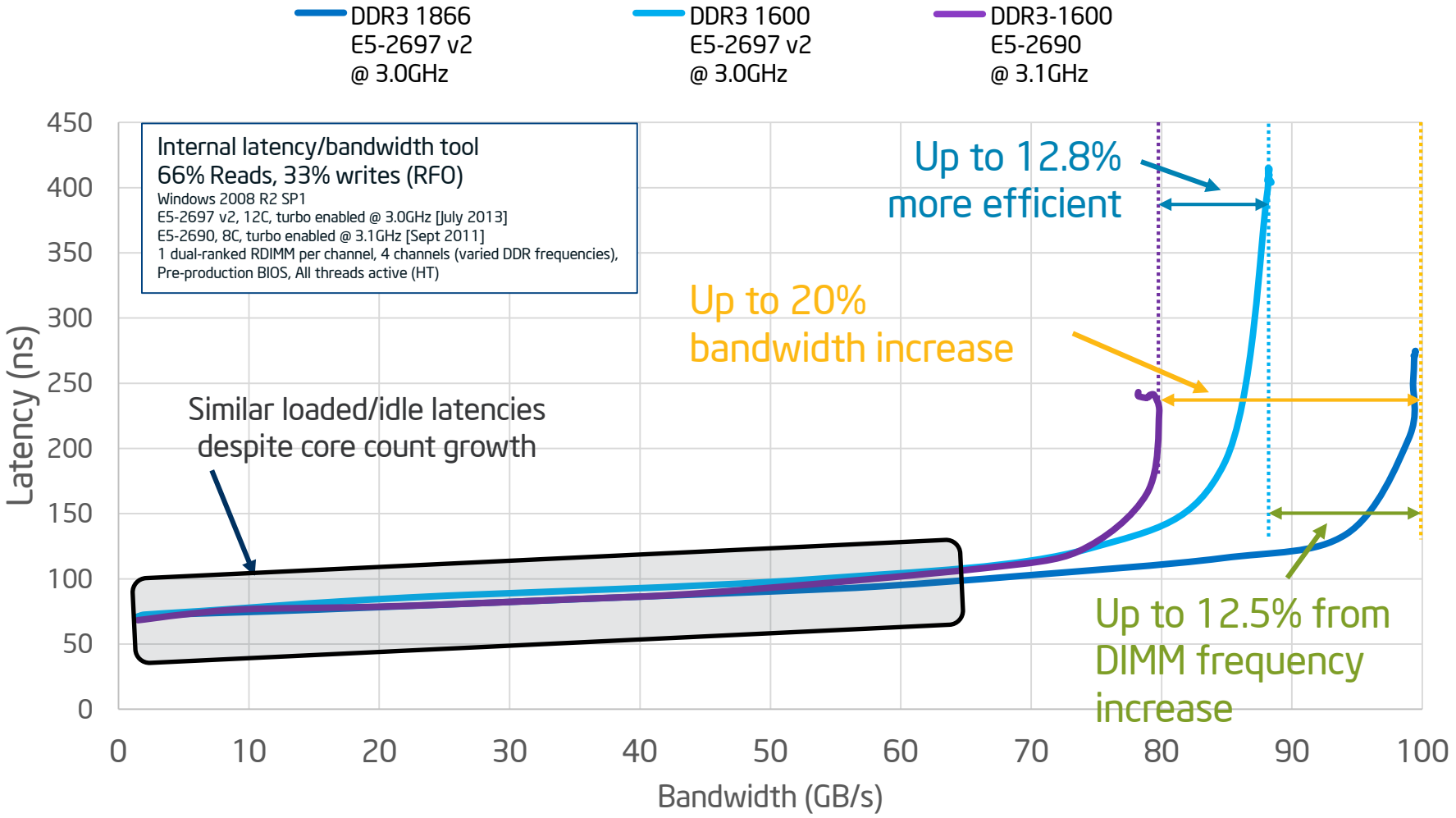
- When software doesn't have enough threads to take advantage of extra cores, performance is more sensitive to memory latency & frequency.
- At low thread counts, 10C SKUs may have the performance advantage over 12C SKUs due to lower LLC latencies & higher turbo frequencies.

Q3: What is difference in the memory bandwidth of DDR3 DIMMs running at 1866 MHz compared to 1600 MHz?

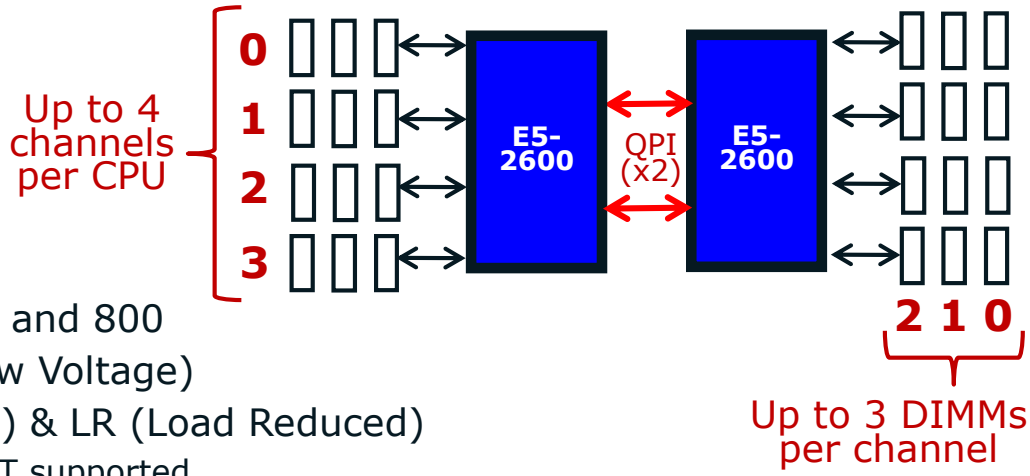
- With all cores enabled, there is up to12.5% increase in memory bandwidth on E5-2697 v2 using DDR3-1866 MHz vs. DDR3-1600 MHz DIMMs.
  - More details provided in: Intel Xeon Processor E5-2600 v2 Product Family Performance
- When not all cores are active, the full 12.5% increase in memory bandwidth from 1866 MHz DIMMs may not be provided.
  - There is a limitation on the number of memory transactions that can be issued by each core. When only a few cores are active, there are not enough memory transactions to saturate the memory controller and take advantage of the higher frequency.

Memory Latency & Bandwidth

Faster Memory, Better Efficiency, Similar Latency

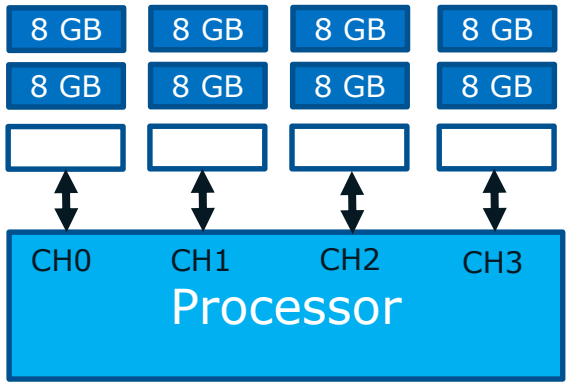


Platform Memory Overview
(Intel® Xeon® processor E5-2600 Product Family)

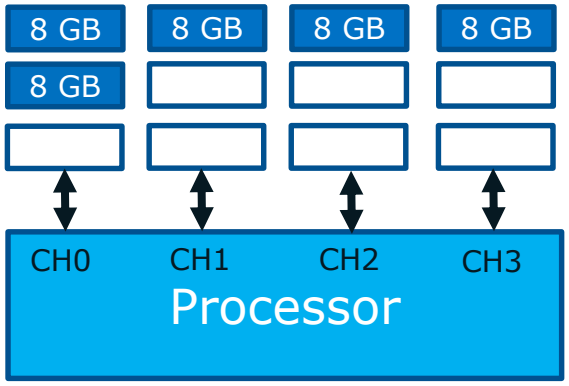


- Platform capability (up to 24 DIMMs):
  - Up to 4 memory channels per CPU
  - Up to 3 DIMMS per channel
- Memory types supported:
  - DDR3 1866 (v2 only), 1600, 1333, 1066, and 800
  - 1.5V and 1.35V DIMMs (1.35V = LV = Low Voltage)
  - Registered (RDIMM), Unbuffered (UDIMM) & LR (Load Reduced)
    - Mixing of UDIMM, RDIMM or LR-DIMMS is NOT supported
  - 1Gb, 2Gb & 4Gb memory technologies
  - Single-rank (SR), dual-rank (DR), quad-rank (QR)
    - QR DIMMs limited to 2 DPC and only 800 or 1066MHz
- System memory speed (i.e. the speed at which the memory is *actually* running) is set by BIOS depending on:
  - CPU capability
  - DIMM type(s) used (memory speed, U/RDIMM, SR/DR/QR)
  - DIMMs populated per channel
  - See subsequent foils for exact configurations supported (Speed, # of DIMMS etc.)
- All memory channels in a system will run at the slowest *common* frequency

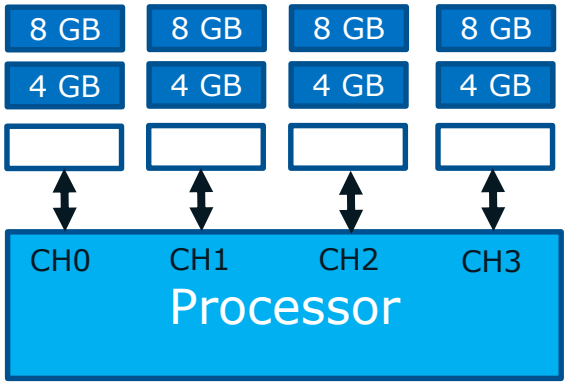
Memory Configuration Definitions



Balanced



Un-balanced



Near Balanced

Memory Configuration	All DIMMs Identical	All Memory Channels Populated	All Channels Identically Populated
Balanced	Yes	Yes	Yes
Un-Balanced	Maybe	Maybe	No
Near-Balanced	No	Yes	Yes

Optimizing for Memory Performance:
General Guidelines and Priorities

1. Use identical DIMM types throughout the platform:
   - Same size, speed, and number of ranks
2. Maximize number of channels populated for highest bandwidth
3. Use a "balanced" platform configuration:
   - All available memory channels populated equally
   - Identical DIMMs in all locations (size/speed/rank)
4. Use a "near-balanced" platform configuration:
   - All available memory channels and sockets populated equally
   - Identical DIMMs in each "row", but different sized DIMMs in row #1 vs. row #2
5. For performance…use Dual Rank DIMMs (DR) then SR or LR-DIMMs
   - Use QR DIMMs as a last option (the memory bus will run slower with QR DIMMs)

Guidance to using Memory Performance Data

- Not all applications are equally sensitive to memory performance (in particular memory bandwidth)
  - Some applications are highly sensitive (typically HPC, Financial)
  - Some are moderately sensitive (Virtualization, Web)
  - And some are relatively insensitive (Database)
- The information provide here is favors applications that are moderately to highly sensitive to memory bandwidth
  - As such, the info here may not be applicable to all situations
    - E.g. database applications are sensitive to memory capacity and latency, and not so much to memory bandwidth. As such, database apps will typically get more performance out of larger capacity memory configurations, even if such a configuration is identified as lower performance here
- In addition, the performance difference between one configuration and another may be relatively small, but the cost of such a configuration is relatively large.

Use the data here as a guide; apply appropriately to your situation

Q4: What type of DIMMs should be used in E5-2600 v2 systems for the best memory bandwidth?

For best performance, use pTRR compliant DDR3 DIMMs and enable the pTRR feature.

- When non-pTRR compliant DIMMs are used, the E5-2600 v2 system defaults into double refresh mode, which has longer memory latency/DIMM access latency and can lower memory bandwidth by up to 2-4%.
- As a debug option only, disable the penalty of double refresh mode for non-pTRR DIMMs by setting:
  - Advanced->Memory Configuration->Memory RAS Configuration
  - DRAM Maintenance -> Manual
  - Note: Intel does not recommend this setting in production systems since this exposes the DDR3 pass gate issue that can lead to memory errors

Background:

- DDR3 DIMMs are affected by a pass gate charge migration issue (also known as Row Hammer) that may result in a memory error.
- The Pseudo Target Row Refresh (pTRR) feature introduced on Ivy Bridge processor families (2S/4S E5 v2, E7 v2) helps mitigate the DDR3 pass gate issue by automatically refreshing victim rows.
- More details provided in Doc #522542 – Intro to DDR3 Memory Refresh Management

Q5: What is the recommended setting for A7 Address Mode?

- A7 Address Mode is a new BIOS option introduced on E5-2600 v2 systems.
- Enabled by default, A7 Address Mode changes the interleave from the default cache line interleave between channels to a two cache line interleave between channels.
- For applications with a cache line stride, the second access is mapped to the same channel as the first access which maps to the same page in memory, resulting in a page hit. A page hit saves multiple cycles of latency, as the page does not need to be opened before it is read.
- For example: If a system has 4 memory channels populated with proper channel interleave in A7 Address Mode, a 512-byte stride will send all requests to the same channel.

Q5: What is the recommended setting for A7 Address Mode?

A7 Address Mode should be disabled when three of the memory channels are populated on E5-2400 v2 SKUs or E5-2600 v2 SKUs with less than 12 cores to prevent memory bandwidth degradation in these configurations

- More details provided in Doc #503836 - Ivy Bridge BIOS Writer's Guide

A7 Address Mode Recommendations		# of memory channels populated per socket			
		1 channel	2 channels	3 channels	4 channels
E5-2600 v2 Product Family (IVB-EP)	E5-2697 v2 or E5-2695 v2 (12 cores)	Default (enable A7)	Default (enable A7)	N/A – Does not support 3 channel interleave	Default (enable A7)
	All other E5-2600 v2 SKUs (4,6,8,10 cores)	Default (enable A7)	Default (enable A7)	Disable A7	Default (enable A7)
E5-2400 v2 Product Family (IVB-EN)	All E5-2400 v2 SKUs	Default (enable A7)	Default (enable A7)	Disable A7	N/A

Key: Intel recommended configuration for best performance