WHITE PAPER
Intel® Xeon® Processor
Virtualization

How Processor Core Count Impacts Virtualization Performance and Scalability

Using benchmarks to determine the optimal configuration for application performance

Marco Righini
Intel Corporation
marco.righini@intel.com

EXECUTIVE SUMMARY

Driven by Moore's Law, CPU architectures have advanced rapidly over the last decade. We have moved from discussing server performance purely in terms of GHz to a discussion where parallelism and the number of cores and software threads have become more important to achieving optimal application performance.

As virtualization has become more pervasive, two frequent questions are:

1. What is the best way to use all available CPU resources?
2. How can we use benchmarks to determine the optimal configuration instead of having the simplistic view of using the amount of GHz and number of cores, particularly when comparing different architecture generations?

This paper looks into these challenges and also addresses the migration path of the virtual machine (VM) footprint re-evaluation during migration.

OVERALL PERFORMANCE CHANGES

Over the last three to four years, CPU and server platform performance have increased dramatically. The Intel server platform has gone through a complete evolution, which has brought about significant changes and redesign of CPUs, memory subsystems, I/O subsystem, and graphics.

We have increased the CPU core count from two to 10 cores and reintroduced Intel® Hyper-Threading Technology (Intel® HT), which doubles the number of software threads per CPU core to a maximum of 20 in today's highest-performing CPUs. The CPU execution pipeline has changed, instructions issued per clock cycle have increased, and new features such as Turbo Mode have been introduced.

The memory controller has been integrated into the CPU and the memory structure has changed, moving from Uniform Memory Access (UMA) architecture to Non-Uniform Memory Access (NUMA) architecture.

Therefore, comparing older-generation CPUs and platform architectures is not an easy task. The difficulty increases with virtualization, where oversubscription of resources is now possible, heavily modifying non-virtualized behavior and increasing the difficulty of making a comparison.

If a VM migration path includes either cold or live migration without some architectural assessment, VMs with the same load run the risk of oversubscribing resources and, therefore, have the potential to use those resources inefficiently.

Performance Comparison across Generations of CPUs

The idea for this white paper came after a visit to one of the largest Telcos in Europe, where some very interesting questions came up:

- Should I still create or keep the number of vCPUs on the VM once I migrate it from an old system to a new one?
- Should I keep Intel HT turned on or off?
- What is the impact of performance on my applications if I combine my new host and the old host in an Intel® Virtualization Technology FlexMigration (Intel® VT FlexMigration) pool? Would that impact the previous question?
- Is there any way to forecast performance?
- Should we go with a CPU with higher core counts? What is the impact of core count versus frequency?

All of these questions deserve an answer; however, there is not a single answer that is right for each user. With this white paper, we will try to establish a process for some basic analysis using official benchmarks.

The throughput of each tile is measured, and then more tiles get added until the throughput no longer increases and/or the CPU utilization reaches its maximum. Figure 1 shows a SpecVirt tile.

How Does SpecVirt Work?

A good place to start is with SpecVirt* (**www.spec.org**), which is becoming the *de facto* standard benchmark for virtualization, since it is open rather than aligned with an individual software vendor and allows us to compare various types of hypervisors, which are the fundamentals of an infrastructure as a service (IaaS) cloud.

The idea of SpecVirt is to run a set of VMs, each with a well-defined workload. This set of VMs is called a "tile" (Figure 1). In real deployments, virtualized environments run many VMs with different workloads at the same time. This benchmark model represents the closest way to replicate what happens in real life.

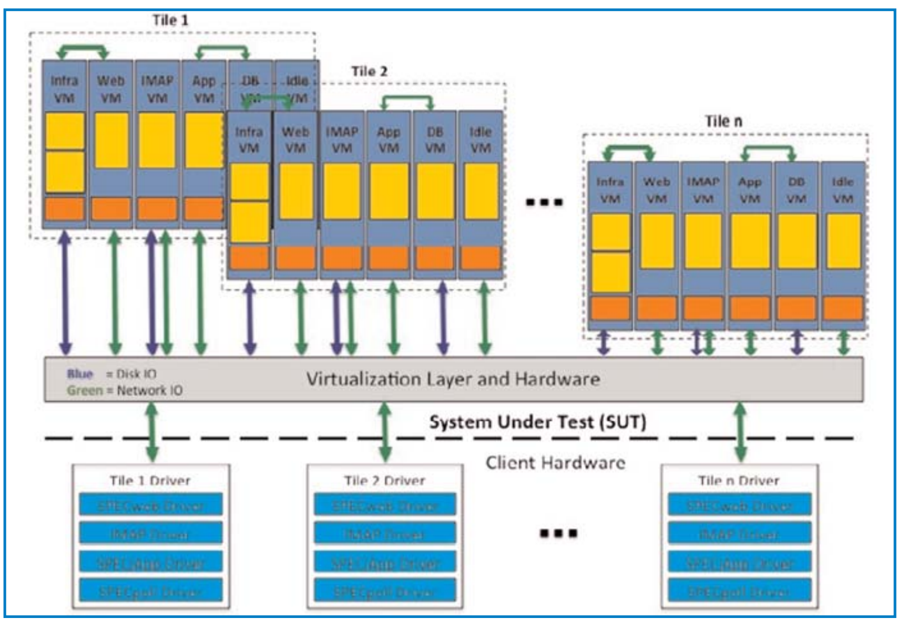


Figure 1. SpecVirt Tile

Unfortunately—since SpecVirt is wide open, has many different hardware and software configurations, and is new to the marketplace—there is a lack of data and published results for worthwhile analysis.

A better option for our purposes is VMmark* 1.x, which has been retired in favor of VMmark 2.0 (which does not measure single-server performance) but has a lot of good data for comparison and calculation.

VMmark 1.x methodology is nearly the same as SpecVirt. The main differences are:

- **Profiles of VMs.** While SpecVirt is based on one vCPU VM, VMmark 1.x has several workloads running on vSMP VMs.
- **Type** of application.
- **VMmark 1.x** supports only the VMware ESX* hypervisor.

VMmark uses the same approach as SpecVirt in measuring each tile's throughput and stops loading tiles when the physical server is saturated. Therefore, the throughput of the tiles either stays constant or starts dropping.

The outcome of the VMmark benchmark is a score with a certain number of tiles. The higher the score, the better the system has performed.

Our goal is to expand those two numbers (VMmark score and number of tiles) and see what other, more tangible numbers are sitting behind the overall score.

We know that a higher VMmark number means better performance and probably a higher number of tiles. Since we know that each tile has a certain number of VMs with a fixed number of vCPUs, we can calculate some interesting things:

- **How many vCPUs** a certain type of core can manage (core/vCPU ratio)
- **How efficient** a certain core is (core/VMmark)
- **How many vCPUs** a system can manage
- **What our migration path** should be from a certain CPU to a newer CPU
- **How many clock cycles** our computations require (the lower the better)
- **Frequency**/VMmark

Some might argue these numbers are not comparable, since the hypervisor may have also changed. However, this is part of the evolution of the industry. The idea is to ensure we migrate efficiently and understand the best path for a specific usage.

Another challenge is that the memory configuration may differ. Here, things become a bit more complex. We can assume the memory increase is also part of evolution, and that memory differences in the same time frame mean that increasing memory would have given either little or no benefit or increased the overall cost of the system.

Results and Calculations

Table 1 contains all public available VMmark 1.x results, plus some data from calculations. Columns include:

- **Date:** The date of the publication.
- **Vendor:** Vendor and model of the server being measured.
- **CPU:** Type of processor.
- **Memory:** Memory of the system.
- **GHz:** Frequency of the specific processor.
- **Cores:** Number of cores in that type of CPU.
- **Socket No.:** Number of sockets in the system.
- **VMmark:** VMmark score (higher is better).
- **No. tiles:** The number of tiles with which the server reached the VMmark score.

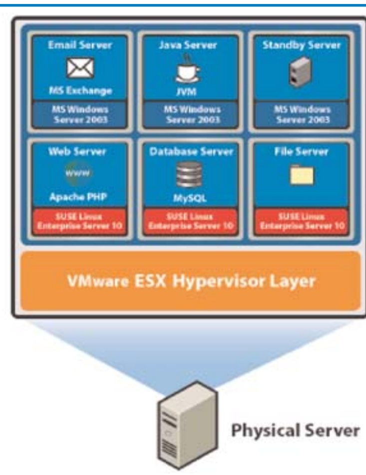
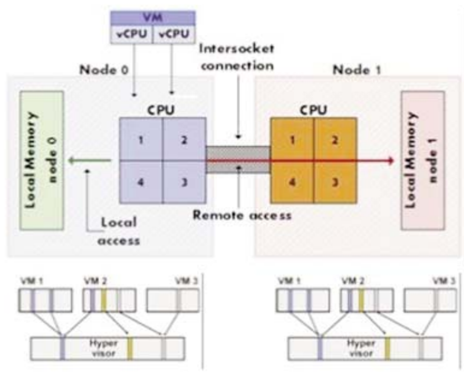


Figure 2. VMmark 1.x Tile

Table 1. Results and Calculations

Date	Vendor	CPU	Mem.	GHz	Cores	Sockets	VMmark	Tiles	Threads	vCPU	vCPU/ Core	Core Eff.	Freq./ vCPU
11/17/09	NEC Express5800/A1160*	X7460	256	2,667	4	16	48,23	30	64	300	4,6	0,7	568,9
07/28/09	NEC Express5800/A1160	X7460	128	2,667	6	8	34,05	24	48	240	5,0	0,7	533,4
06/16/09	IBM System x3950 M2*	X7460	256	2,667	6	8	33,85	24	48	240	5,0	0,7	533,4
10/19/10	Fujitsu RX600 S5*	X7560	1024	2,260	8	4	77,29	51	64	510	15,9	2,4	141,8
09/07/10	Cisco UCS* C460 M1	X7560	512	2,260	8	4	76,1	51	64	510	15,9	2,4	141,8
08/24/10	HP ProLiant* DL580 G7	X7560	1024	2,260	8	4	75,01	50	64	500	15,6	2,3	144,6
07/27/10	Dell PowerEdge* R910	X7560	512	2,260	8	4	74,34	50	64	500	15,6	2,3	144,6
07/28/10	Lenovo WQ* R680 G7	X7560	384	2,260	8	4	73,2	50	64	500	15,6	2,3	144,6
04/20/10	IBM System x3850 X5	X7560	384	2,260	8	4	71,85	49	64	490	15,3	2,2	147,5
08/10/10	Dell PowerEdge R810	L7555	256	1,870	8	4	59,21	41	64	410	12,8	1,8	145,9
08/24/10	Dell PowerEdge M910	L7555	256	1,870	8	4	58,37	41	64	410	12,8	1,8	145,9
10/02/08	IBM System x3950 M2	X7350	128	2,930	4	8	24,62	18	32	180	5,6	0,7	520,8
10/19/10	Fujitsu RX600 S5	E7540	1,024	2,000	6	4	55,88	39	48	390	16,2	2,3	123,0
03/24/09	IBM System x3850 M2	X7460	128	2,660	6	4	20,5	14	24	140	5,8	0,8	456,0
02/10/09	Dell PowerEdge R900	X7460	96	2,660	6	4	19,99	14	24	140	5,8	0,8	456,0
10/19/10	Fujitsu RX600 S5	X7560	512	2,260	8	2	40,52	28	32	280	17,5	2,5	129,1
10/19/10	Cisco UCS B230 M1	X7560	256	2,260	8	2	39,19	27	32	270	16,8	2,4	133,9
10/19/10	HP ProLiant BL620c G7	X7560	256	2,260	8	2	37,92	28	32	280	17,5	2,3	129,1
07/27/10	Dell PowerEdge R810	X7560	256	2,260	8	2	37,28	26	32	260	16,2	2,3	139,0
06/29/10	Fujitsu BX960 S1	E7520	256	1,860	4	4	32,82	22	32	220	13,7	2,0	135,2
09/05/08	HP ProLiant DL580 G5	X7350	64	2,930	4	4	14,14	10	16	100	6,2	0,8	468,8
07/08/08	Dell PowerEdge R900	X7350	64	2,930	4	4	14,05	10	16	100	6,2	0,8	468,8
03/26/08	IBM System x3850 M2 HP	X7350	64	2,930	4	4	13,16	9	16	90	5,6	0,8	520,8
08/31/07	ProLiant DL580 G5 HP ProLiant	X7350	64	2,930	4	4	11,54	8	16	80	5,0	0,7	586,0
08/31/07	BL680 G5	E7340	64	2,930	4	4	10,17	7	16	70	4,3	0,6	669,7
10/19/10	Fujitsu BX924 S2	X5680	192	3,300	6	2	40,86	30	24	300	25,0	3,4	132,0
10/19/10	HP ProLiant DL380 G7	X5680	192	3,300	6	2	38,97	28	24	280	23,3	3,2	141,4
10/05/10	Dell PowerEdge R710II	X5680	192	3,300	6	2	38,39	27	24	270	22,5	3,2	146,6
10/19/10	PowerEdge M610x	X5680	192	3,300	6	2	38,38	27	24	270	22,5	3,1	146,6
10/19/10	Cisco UCS B250 M2	X5680	192	3,300	6	2	38,04	27	24	270	22,5	3,1	146,6
10/19/10	Fujitsu BX924 S2	X5677	144	3,460	4	2	30,05	20	16	200	25,0	3,7	138,4
10/19/10	HP ProLiant DL380 G7	X5677	192	3,460	4	2	29,46	20	16	200	25,0	3,6	138,4
09/21/10	SGI* C2104-TY3	W5590	96	3,300	4	2	25,67	18	16	180	22,5	3,2	146,6
04/06/10	HP ProLiant ML370 G6	W5580	96	3,200	4	2	25,29	18	16	180	22,5	3,1	142,2
04/20/10	HP ProLiant BL490c G6	X5570	96	2,930	4	2	25,27	17	16	170	21,2	3,1	137,8
11/09/09	Fujitsu RX300 S5	W5590	96	3,300	4	2	25,16	17	16	170	21,2	3,1	155,2
09/22/09	HP ProLiant BL490c G6	X5570	96	2,930	4	2	24,54	17	16	170	21,2	3,0	137,8
09/08/09	Dell PowerEdge R710Supermicro*	X5570	96	2,930	4	2	24,27	17	16	170	21,2	3,0	137,8
03/30/09	6026-NTR+	X5570	72	2,930	4	2	14,22	10	16	100	12,5	1,7	234,4
10/09/08	HP ProLiant ML370 G5	X5470	48	3,300	4	2	9,15	7	8	70	8,7	1,1	377,1
09/24/08	Dell PowerEdge M600	X5470	32	3,300	4	2	8,97	6	8	60	7,5	1,1	440,0
08/31/07	Dell PowerEdge 2950Dell	X5365	32	3,000	4	2	7,03	5	8	50	6,2	0,8	480,0
08/31/07	PowerEdge 2950	X5160	32	3,000	2	2	3,89	3	4	30	7,5	0,9	400,0
05/01/09	Intel EPSD	E5540	96	2,530	4	2	21,64	16	16	160	20,0	2,7	126,5
04/26/11	Intel EPSD	X5690	144	3,460	6	2	34,19	28	24	280	23,3	2,8	148,2
04/26/11	Intel EPSD	X5675	144	3,060	6	2	33,5	27	24	270	22,5	2,7	136,0
04/26/11	Intel EPSD	E5645	144	2,400	6	2	29,64	24	24	240	20,0	2,4	120,0
04/26/11	Intel EPSD	E5649	144	2,530	6	2	30,24	25	24	250	20,8	2,5	121,4
04/26/11	Intel EPSD	W5647	144	2,930	4	2	25,48	23	16	230	28,75	3,1	101,9
04/26/11	Intel EPSD	E5607	144	2,260	4	2	13,74	14	16	140	17,5	1,7	129,1
04/26/11	Intel EPSD	E5606	144	2,130	4	2	13,34	13	16	130	16,25	1,6	131,0



Note: Recent hypervisor exposure of NUMA topology to VM makes this less impactful.

Figure 3. NUMA Architecture: Impact on VMs

- **No. Threads:** Number of overall threads of the system. This includes symmetric multi-threading technology.
- **Total vCPU:** This is the total number of vCPUs the system was able to sustain to get to the specific VMmark score.
- **vCPU/core:** This is the number of vCPUs that a single physical core is able to sustain.
- **Core Efficiency:** This is a ratio between VMmark score and number of cores (higher is better).
- **Freq. /vCPU:** This ratio tries to show how computing is performed in terms of CPU cycles (lower is better).

We can try to make some comparisons and informed decisions based on performance data that are important factors when deciding which VMs to migrate from older systems to newer ones.

For comparison, Table 1 uses servers based on the year they were on the market, the segment, and the technology. Results are calculated based on VMmark 1.x results.

Please note that Table 1 does not show all results, but only the top result per segment and vendor. You can find a complete list of results on **www.vmware.com**.

Total vCPUs per System

The total vCPU value per system is calculated based on the profile of the VMs running VMmark 1.x and the number of tiles. This value is a great way to determine how many VMs and vCPUs a system is able to handle (the higher the better). However, it is only a rule of thumb, since you need to consider that if the number of vCPUs is greater, then the actual number of cores per socket and NUMA topology may heavily affect performance.

NUMA topology becomes a potential impacting factor due to the memory allocation of a process, which may execute on a remote core, therefore adding latency.

Figure 3 shows a typical NUMA system with a VM running on top of it.

vCPU per Core Ratio

The number of vCPUs per system is a key value but does not provide a full picture. vCPU per core shows:

- **How many vCPUs** can be executed on that specific core.
- **The differences** in cache, instruction set, VM entry, exit latencies, and number of instructions issued per clock cycles.
- **The number of cores** and how they interconnect, which may make core performance more or less efficient.

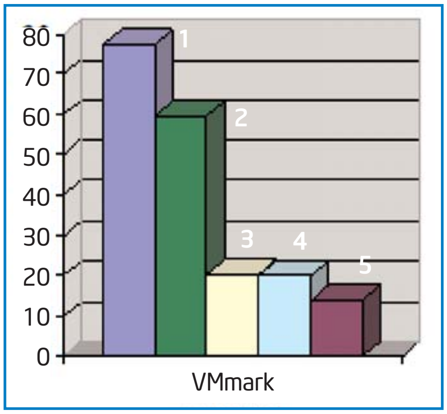


Figure 4. VMmark Results

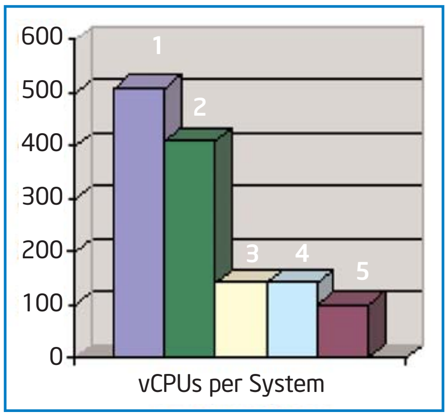


Figure 5. vCPUs per System

Figures 4 through 7:
1 = Intel® Xeon® processor X7560
2 = AMD 6174* processor
3 = Intel Xeon processor X7460
4 = Intel Xeon processor X7460

Core Efficiency

Core efficiency, or the system efficiency ratio, is actually calculated based on the sum of the number of cores and the VMmark score result. It provides a ratio of how efficient that core is.

Frequency per vCPU

The number of cores and frequency are not good enough metrics on their own. Besides the vCPU/core ratio, it is interesting to consider a vCPU/GHz ratio. This will help us to see how VMs and hardware-assisted features and core efficiency affect application performance using VMmark 1.x results.

The frequency per vCPU essentially shows how many CPU cycles are used to execute the workload on a per-vCPU basis. To get this number, multiply the nominal frequency of the core by the number of cores in the system and then divide it by the number of vCPUs the system is able to execute.

A lower number means fewer clock cycles are needed to accomplish the benchmark and, therefore, the system is more efficient.

Unfortunately, the results are application-dependent and, therefore, almost impossible to predict. (For examples, read the Intel white paper **"Enabling Intel Virtualization Technology Features and Benefits."**)

Table 2. 1 Four-Socket Segment Results

Vendor	CPU	VMmark	vCPU per System	vCPU per core	Core Efficiency Ratio	Freq./vCPU	Tot. GHz
Fujitsu* RX600 S5	Xeon X7560	77,29	510	15,93	2,415	141,80	72.320
Dell PowerEdge* R815	AMD* 6174	59,74	410	8,54	1,244	257,56	105.600
IBM* x3850 M2	Xeon X7460	20,5	140	5,83	0,854	456	63.840
Dell R900	Xeon X7460	19,99	140	5,83	0,832	456	63.840
HP* DL580 G5	Xeon X7350	14,14	100	6,25	0,883	468,8	46.880

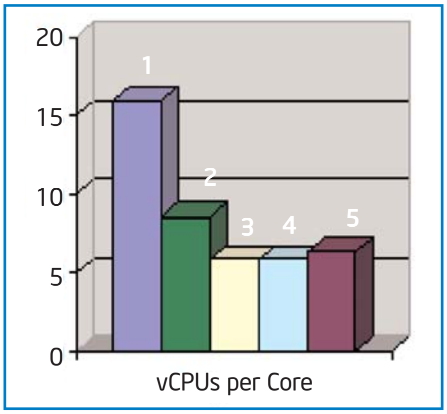


Figure 6. vCPUs per Core

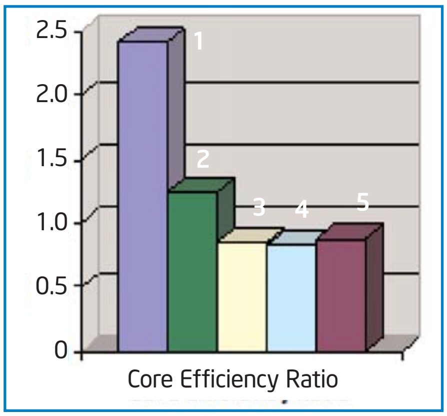


Figure 7. Core Efficiency Ratio

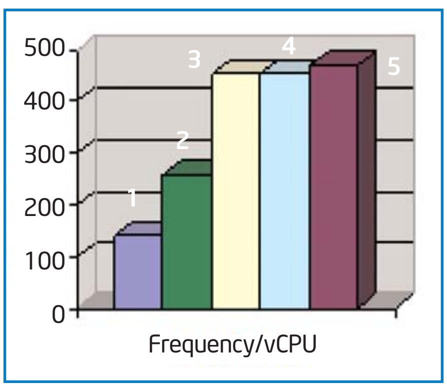


Figure 8. Frequency/vCPU

Table 2 and figures 4 through 8 show four-socket performance results. Table 3 and figures 9 through 11 show two-socket segment results.

ANSWERING THE QUESTIONS

After analyzing all of the graphs, we can now answer some of the questions we had at the beginning of this paper:

- **Should I still create or keep the number of vCPUs on the VM once I migrate it from an old system to a new one?** This really depends on the application, but there is no doubt that re-evaluating the number of vCPUs capable on previous-generation platforms should be a best practice, especially if you look at the performance increase, the efficiency of the cores, and at the overall architectural performance boost. Having 20 percent more performance probably means dropping the cost even more than 20 percent if we take into account connectivity, management, power, and all aspects of the cost structure of a cloud environment.
- **Should I keep Intel HT turned on or off?** You should consider Intel HT as an advantage if the hypervisor knows how to make good use of it. In a migration path from a non-Intel HT-enabled system to one where HT is enabled, the only consideration should be to make sure the hypervisor is HT-aware and consider overall system performance without trying to understand the impact compared to real cores. Intel HT will boost performance between 10 and 20 percent versus the same server without Intel HT.

Table 3. Two-Socket Segment Results

Vendor	CPU	VMmark	vCPU per System	vCPU per core	Core Efficiency Ratio	Freq./vCPU	Tot. GHz
Fujitsu* BX924 S2	Xeon X5680	40,86	300	25	3,405	132	39.600
Dell PowerEdge* R715	AMD* 6176 SE	32,44	220	9,1	1,351	250,90	55.200
SGI* C2104-TY3	Xeon W5590	25,67	180	22,5	3,208	146,66	26.400
HP* ML370 G6	Xeon W5580	25,29	180	22,5	3,161	142,22	25.600
HP BL490c G6	Xeon X5570	25,27	170	21,25	3,158	137,88	23.440
HP DL385 G5	AMD 2384	11,28	80	10	1,41	270	21.600
HP ML370 G5	Xeon X5470	9,15	70	8,75	1,143	377,14	26.400
Dell PowerEdge R805	AMD 2360 SE	7,96	60	7,5	0,995	333,33	20.000
Dell PowerEdge 2950	Xeon X5365	7,03	50	6,25	0,878	480	24.000
Dell PowerEdge 2950	Xeon X5160	3,89	30	7,5	0,972	400	12.000

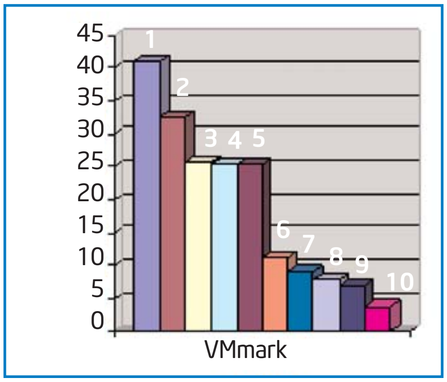


Figure 9. VMmark

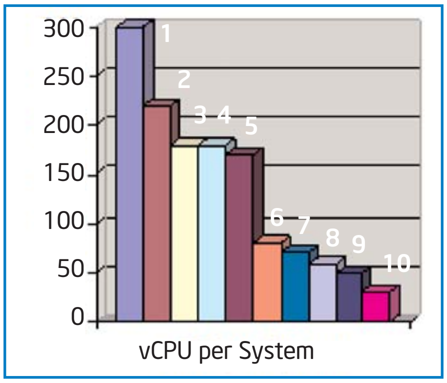


Figure 10. vCPU per System

- **What is the performance impact on my applications if I run my new and old host in an Intel VT FlexMigration pool?** When deciding on a migration strategy between old and new servers in the same resource pool through live migration, you should be aware which new instructions will be hidden and which will be available. (See **"Enabling Intel® Virtualization Technology FlexMigration on VMware ESX*"** for more information.) Even applications that would benefit from new instructions that are hidden due to Intel VT Flex Migration would get a performance boost from the improvement from architectural changes and from non-ring 3 instructions.

Figures 9 through 11:
1 = Intel Xeon processor X5680
2 = AMD 6176 SE* processor
3 = Intel Xeon processor W5590
4 = Intel Xeon processor W5580
5 = Intel Xeon processor X5570
6 = AMD 2384 processor
7 =Intel Xeon processor X5470
8 = AMD 2360 SE processor
9 = Intel Xeon processor 5365
10 = Intel Xeon processor X5160

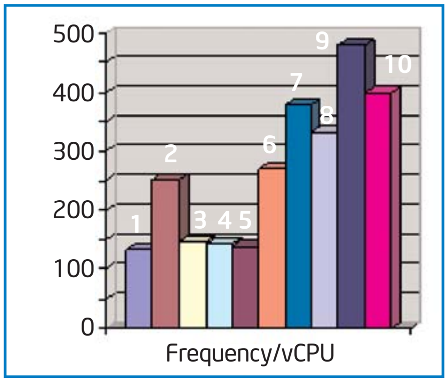


Figure 11. Frequency/vCPU

- **Should we go with a CPU with a higher core count? What is the impact of core count versus frequency?** A system needs to be balanced. Just adding more cores will not help if the entire architectural design is not able to feed those cores. The same applies to frequency. If the architectural design does not meet the frequency requirements, the system ends up throwing CPU cycles away. This is why the core efficiency and the frequency math are interesting. New CPUs can run at higher frequencies than their nominal clock frequency as needed according to power limits and temperatures (Intel® Turbo Boost Technology). This reduces the concerns with frequency versus core count.
- **If we have four vCPU VMs running on an Intel Xeon processor 5160 and we migrate it to the newest server generation, do I still need to assign the same number of vCPUs for the same workload? What is the rationale?** By looking at the numbers in Table 2 and/or using the same methodology, we could use the core efficiency and the frequency/vCPU to make a reasonable forecast. This should provide an idea of what the sizing of the VM should be. Overall, doing those calculations may help you to address the migration path question.

SUMMARY

It is not easy to understand the performance impact of a refresh cycle with certainty. Re-evaluating the VM configuration through the years, and having a performance assessment before migrating, are both essential.

There is no single tool to do that, but looking behind the numbers of traditional official benchmarks may be helpful.

It is key is to look behind the benchmark numbers (e.g., SpecVirt, VMmark) and understand what those numbers really represent.

Find the solution that's right for your organization. Contact your Intel representative, visit Intel's Business Success Stories for IT Managers (**www.intel.com/itcasestudies**), or explore the Intel.com IT Center (**www.intel.com/itcenter**).

About the Author

Marco Righini is a solution architect with Intel Corporation, responsible for enterprise solutions and designs worldwide. With the company since 2006, Marco is primarily responsible for designing and supporting virtualization solutions and cloud computing architectures and enterprise virtual disk image (VDI) architectures in support of Intel's enterprise computing focus areas. He has deep technical background in vSphere*, Linux*, Xen*, unified networking, and security.

0212/SS/MR/PDF