

The True Cost of Virtual Server Solutions

March 2009



In March of 2009, VMware contacted Taneja Group to conduct an assessment of how VMware ESXi 3.5 and Microsoft Hyper-V compare when it comes to hosting multiple virtual machines – what we refer to in this Technology Validation Report as “*VM Density*.” At VMware’s request, Taneja Group executed a set of tests to evaluate hypervisor performance under different types of workloads.

Our hands-on validation has revealed that VMware ESXi 3.5 can achieve a significantly higher VM density than Microsoft Hyper-V, allowing users to consolidate more servers on a VMware platform. While VMware’s advantage may not be surprising, given its 10-year history of leadership in this space, the magnitude of the difference is striking: VMware ESXi 3.5 demonstrated at least a 50% (1.5 to 1) density advantage over Hyper-V across a variety of application workloads, and in some cases achieved a 100% (2 to 1) advantage. Thus, for a given application workload, if users can safely consolidate 8 servers on Hyper-V, they will comfortably be able to consolidate 12 to 16 servers on VMware ESXi.

Why is this important? Because it has a considerable impact on virtual infrastructure acquisition costs. A higher VM density translates to a lower number of physical servers required to support the virtual infrastructure. A higher VM density also leads to reduced guest OS licensing costs, which tend to be dependent on the number of hypervisor instances in an installation. Finally, a higher VM density can also reduce the cost of management servers and licensing, which with some solutions can be complex and expensive, especially when they are licensed per hypervisor host. Together, these costs outweigh any individual component or hypervisor costs, and paint a picture of total cost of acquisition.

Our Findings: Based upon our testing of VM density, and contrary to competitive claims and some prevailing end user perceptions, the total solution cost of a VMware Infrastructure 3 environment is not significantly higher than that of a Microsoft virtualization environment, which consists of Hyper-V and Microsoft System Center. Given the VM density advantages of ESXi 3.5 demonstrated in this study, we found the cost per application to be between 5% and 29% *less expensive* in a VMware Infrastructure 3 environment than in a Microsoft environment consisting of Hyper-V and System Center.

In this study, we’ll first assess the relative VM density of VMware and Microsoft hypervisors. Then we will discuss a methodology for assessing acquisition cost or *cost per application*. In our view, end users should look to cost per application to streamline and simplify their assessment of the total acquisition costs of a server virtualization environment.

Charge of the Hypervisors

While server virtualization has come to embody many of the most innovative capabilities in today's IT infrastructures, the technologies are still evolving. Over the past few years, there has been a seemingly never-ending charge of new hypervisors into the market. Each newcomer comes with its own architectural approach and set of performance characteristics. Today, most of the hypervisors on the market are available in a fully functional free version, but free does not mean they are the same.

We've long held to our position that there are marked differences between hypervisor architectures. Some of these differences are based on maturity levels and others on fundamental choices in hypervisor architecture. In the end, the architectural differences between hypervisors can have a major impact on total solution cost. Hypervisor architecture will determine efficiency in processing workloads, and architecture plays an important role in determining whether an infrastructure is flexible and adaptable.

Yet architecture isn't the entire story. Just as there are differences in hypervisor efficiency, there are also significant differences in management infrastructure efficiency.

Taken together, total differences in efficiency – including how well a hypervisor solution makes use of supporting hardware and licensing – will drive the overall cost of a server virtualization solution. What many users do not realize until time of implementation is that their choice of

hypervisor may lead to sprawl in virtual hosts along with a complex management infrastructure when compared to another choice. A big picture view of the total solution – hypervisor architecture and management infrastructure – often yields a vastly different picture of cost than that bandied about in public by emerging vendors.

The Customer's Dilemma – Understanding Total Cost of Acquisition

Server virtualization customers are faced with a complex set of tradeoffs in evaluating virtual server infrastructures, along with a landslide of marketing noise that makes assessing true cost seem like a wild goose chase on a foggy day.

In recent years, virtualization vendors have confused the marketplace by introducing “free” hypervisors, surrounded by claims that they are cost leaders, while focusing little attention on performance and management. Other vendors have claimed that the choice of hypervisor is irrelevant. All of this has muddied the waters for IT buyers.

Marketing hype has convoluted not only the differences between hypervisor architectures, but between overall virtual infrastructure management capabilities as well. As a result, users often have a difficult time understanding what it takes to properly set up a functioning virtual server infrastructure in a real world data center.

Though the variation between hypervisors and virtual server infrastructures is seldom

TECHNOLOGY VALIDATION

apparent in surface-level comparisons, it has an enormous impact upon both Total Cost of Acquisition (TCA) as well as Total Cost of Ownership (TCO). TCO is often fuzzy and complex territory, but when it comes down to the reality of today's hard dollars involved in Total Cost of Acquisition, hidden costs simply must be understood.

To that end, we have carried out this Technology Validation exercise. We will first examine the relative performance of two hypervisors – VMware ESXi and Microsoft Hyper-V – with a series of performance tests. Using our assessment of performance, we will then examine how relative performance contributes to total virtual infrastructure costs, and conduct a comparative assessment of TCA. But first, we'll discuss how the industry is in need of a lens of simplification with which to view hypervisor selection, and discuss what we consider the best framework for capturing total cost of acquisition – *virtual server infrastructure cost per application*.

Cost Per App: Why VM Density Matters

Administrators have long been used to evaluating new IT initiatives on a cost per application basis, but this practice has seldom carried over to an evaluation between virtual server infrastructures. Yet this same rigorous approach can easily be applied to compare virtual infrastructures, once the

VM density: The number of guest VMs that can be effectively consolidated on a hypervisor without disruptive performance impact.

total components that make up a solution are understood.

With this in mind, we've distilled two core elements that drive virtual infrastructure solution cost. These are:

1. **Hypervisor efficiency**, which determines how many hypervisors will be needed to support a given number of VMs running a particular application workload.
2. **Virtual Infrastructure Management Efficiency**, which determines how much management software is needed to manage the hypervisor environment.

These two elements drive total solution cost because they determine what an environment will require in the form of:

1. **Licensing.** Hypervisor and management infrastructure efficiency determine how much licensing will be required to support hypervisors, operating systems, and management components.
2. **Physical Infrastructure.** Hypervisor efficiency determines how much physical infrastructure will be required to run the virtual infrastructure, and in turn this also determines how much starting up that virtual infrastructure will cost, in the form of power, cooling, and floor space.

The biggest influencing factor behind total solution cost is hypervisor efficiency, or what we call *VM density* in this Technology Validation report. VM density will determine the total physical equipment, management infrastructure, and software licensing required for a virtual server infrastructure.

The reason VM density matters is simple. In an infrastructure with as few as 30 workloads, a hypervisor that can host even 50% more workloads may have a considerable cost advantage. A 50% advantage in efficiency may allow a customer to reduce their number of hypervisor instances by one-third or more. Since management infrastructures are often licensed according to number of hypervisor instances, and operating systems are licensed per hypervisor instance, the number of hypervisor instances will have a dramatic influence on the total solution cost.

Testing Density: Not All Hypervisors – or Virtual Infrastructures – are the Same

With this background in mind, we tested two leading hypervisors to assess their efficiency in hosting virtual machines. While architectural differences between VMware ESXi and Microsoft Hyper-V hypervisors are now generally well understood, the impact of these differences on achievable VM density is not. Our goal is to determine the VM density supported by these two hypervisors based on a set of typical application workloads.

Once we have determined the differences in VM density at the hypervisor level, we will then use these results to evaluate cost per application differences between VMware Virtual Infrastructure 3 and Microsoft Hyper-V plus System Center.

Our Methodology

To evaluate VM density between Hyper-V and ESXi, we turned to two tests,

DBHammer and SPECjbb. The first test – DBHammer – simulated SQL Server workloads within multiple VMs in order to shed light on how total hypervisor differences may show up when virtualizing common enterprise workloads. The second test – SPECjbb – used an atypical, extremely aggressive workload to level the playing field around memory management differences. By doing so, SPECjbb allowed us to peer into vCPU and driver efficiency to assess how effectively Hyper-V and ESXi can deal with worst-case workloads.

We selected these two tests to validate how Hyper-V and ESXi performed under a variety of conditions. We intended to make sure no single set of optimizations or proprietary technology gave either hypervisor an unfair advantage with a particular workload. We'll turn next to a discussion of the individual results from these tests, and then with this data in hand, we will discuss how VM density differs between Hyper-V and ESXi.

Our Observations

Arriving at these tests, and rolling up and comparing results between the selected hypervisors was in fact no easy task.

The two hypervisors under consideration, Microsoft's Hyper-V and VMware's ESXi, are remarkably different in capabilities and features. Since Hyper-V is a relatively young hypervisor, we routinely ran into Hyper-V architectural limitations (e.g. limited SMP vCPU support on older Windows OSs and Linux, and no memory oversubscription). Consequently, each of our tests was designed

T E C H N O L O G Y V A L I D A T I O N

specifically with the optimal configuration and performance of Hyper-V in mind.

Moreover, even when ESXi was faced with artificially limited test configurations, it still delivered markedly different levels of performance and scalability that made distilling results a challenge. How do you fairly compare results when Hyper-V is limited to less than half the number of VMs that ESXi can host due to the way it assigns memory? Based on a broad collection of data, here are the salient points relevant to our VM density testing.

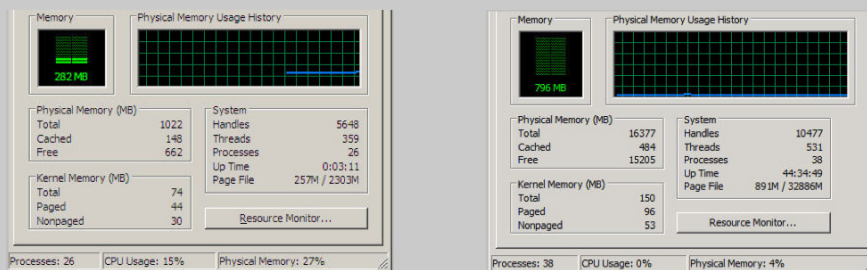
DBHammer Testing: *Under typical workloads, ESXi can run twice as many VMs as Hyper-V.*

As illustrated in Figure 1 on page 6, under typical workloads, VMware ESXi scaled well past twice the number of virtual machines that Hyper-V can host using the same amount of physical memory.

As we tested Hyper-V we encountered a hard limit to the number of VMs we could start up. Our test system was configured with 16 physical gigabytes of memory, and once

A note on Server Core

Microsoft advertises Server Core and Hyper-V Server as the ideal platforms for Hyper-V, and both have smaller memory footprints than Windows Server 2008, which is the product we used in our testing. For the purposes of this exercise both Server Core and Hyper-V Server suffered from a number of limitations that led us to stick with Windows Server 2008. Server Core's lack of configuration tools was a major obstacle to carrying out testing, but it shouldn't be overlooked that Hyper-V Server also has several built-in restrictions that led us to conclude it is an unequal competitor. Specifically, Hyper-V Server is more limited in memory and number of processors, and lacks common features such as Microsoft Cluster Services. Even so, we wanted to fairly consider any potential advantages that Hyper-V Server may have had. Consequently, we assessed the impact of the Hyper-V Server / Server Core memory footprint separately and we saw a difference of about 500MB.



Server Core with Hyper-V (left) and Windows Server 2008 with Hyper-V (right)

When we ran into Hyper-V's VM limits (discussed in the sidebar: "The Usefulness of Memory Oversubscription"), we had 1.3GB of free memory available. Nonetheless, this was not enough to start another 1GB VM. But since an extra 500MB may have created enough headroom to host one more virtual server, we've given Hyper-V unvalidated credit for being able to host 12 VMs instead of 11. Moreover, we've assumed that all 12 guests would deliver the same performance, since we did not see significant performance drop-offs in ESXi at this number of guests. Your mileage may vary.

T E C H N O L O G Y V A L I D A T I O N

16GB was used by the host OS or *assigned* to a running VM no more VMs could be started. In our test case, this limited Hyper-V to running 11 1GB-memory VMs at once. But *assigning* memory to a VM does not mean it is being fully utilized. This in turn limited VMs and created underutilization of memory, CPU, and IO. (See sidebar: The Usefulness of Memory Oversubscription)

This hard limit was a surprise to us, as we've long become accustomed to assigning more memory to VMs than is physically available to the hypervisor. As an example, using VMware ESXi we easily assigned over 24GB of virtual server memory (24 servers with 1 GB allocations each) using only 16 physical gigabytes of memory. Real utilization of

memory barely approached 16GB, and we could have started more VMs until at some point we decided we were witnessing significant performance impact.

But the story doesn't end there. While Hyper-V was limited by a hard cut-off in the number of VMs it can host, ESXi also demonstrated how it leverages VMware's advanced memory management technologies to further out-scale Hyper-V. When we virtualized workloads, we wanted to assess the ability of a hypervisor to host guests without performance impact. What we found was that ESXi easily scaled to twice as many guests as Hyper-V with no performance impact. But further analysis beyond 2 times as many guests (22) led to an intriguing conclusion:

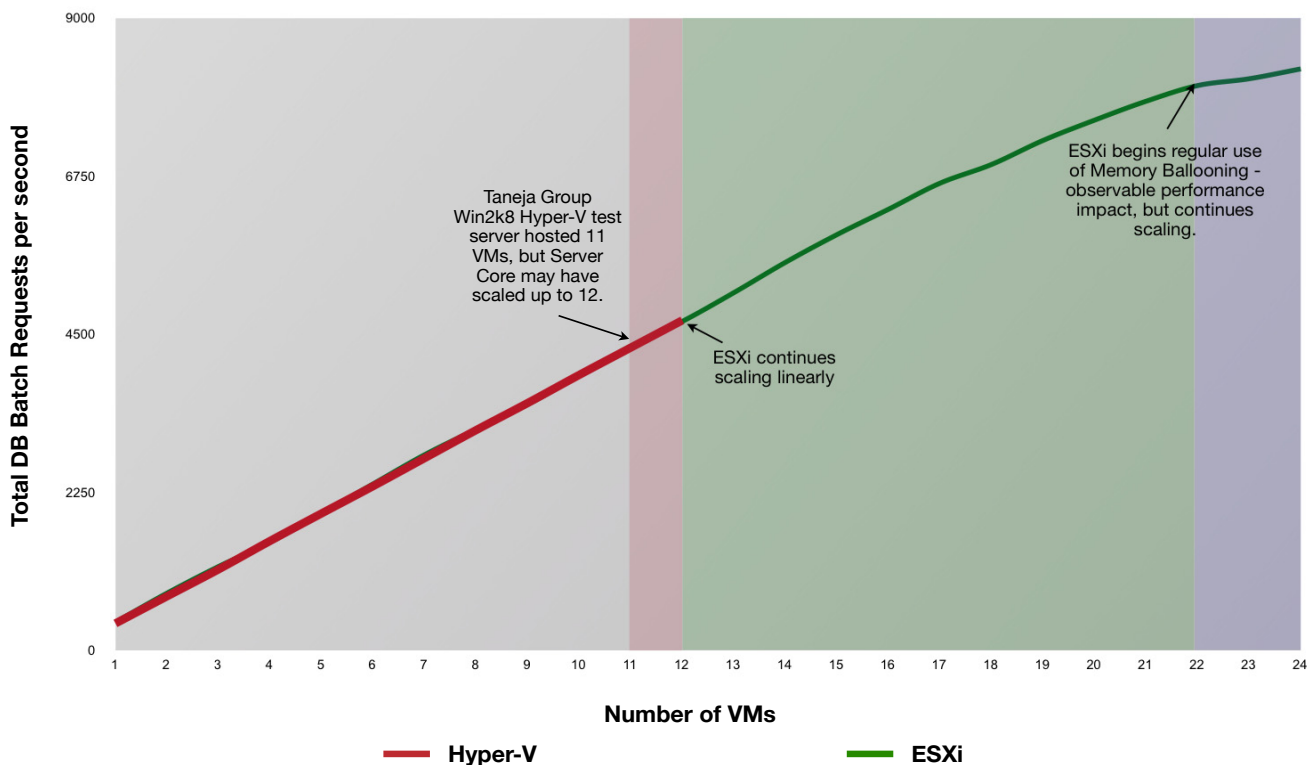


Figure 1 : How VMs running DBHammer scaled in performance on Hyper-V and ESXi

T E C H N O L O G Y V A L I D A T I O N

Key features in ESXi may enable much greater than 2:1 scalability with only an inconsequential drop-off in per-guest performance.

As we scaled performance past 2 times as many virtual guests, or 22 VMs, ESXi regularly employed a feature called memory ballooning. Memory ballooning works to keep only active VM memory pages in memory, and returns inactive pages to a free pool that can be shared among other VMs. This feature helps increase VM density, and makes a VMware infrastructure better able to virtualize rapidly changing workloads. Our test results demonstrate that even with static performance loads, memory ballooning can police, clear out, and share enough memory space to make a substantial difference in the VM density of a hypervisor.

Once heavily used, memory ballooning processes are not without overhead, and evidently created an 11% performance impact on per VM performance. But more significantly, the 11% performance impact happened as soon as memory ballooning started, and did not change significantly with the addition of more VMs. Memory ballooning allowed us to easily increase our total VMs up to 24. Moreover, we did not stop at 24 VMs due to any detectable server workload impact. At 24 VMs we noticed slightly increased vCenter GUI latency, but we ran out of provisioned servers as well as the time and patience to provision more. We suspect we could have added several more VMs to this host without additional performance impact. But eventually more significant performance degradation will set in when memory oversubscription drives the

hypervisor memory system to turn to disk swapping to meet demands. Since our environment consisted of unusually consistent workloads, we see the performance impact associated with memory ballooning as inconsequential. Dense virtual infrastructures have widely varying workloads that will only occasionally peak at the same time, so reducing total system performance in favor of hosting more workloads that aren't saturating system resources is a reasonable tradeoff.

Our conclusion: Based on these typical workloads, in some situations ESXi can virtualize twice as many VMs as Hyper-V.

SPECjbb Testing: *ESXi delivers 24% more performance from the same hardware.*

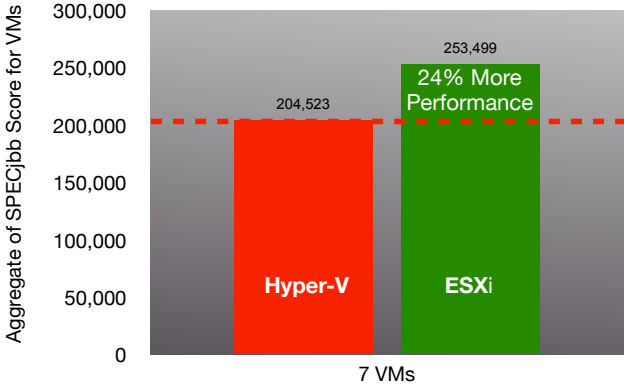


Figure 2 : Difference in the aggregate of SPECjbb performance score running in 7 VMs on a single host

When we turned to SPECjbb testing, we confirmed that VMware does indeed deliver more performance through a more efficient hypervisor architecture. Using 7 VMs configured with 2 vCPUs and 2GB of memory each on an AMD Opteron 2384 8 core

T E C H N O L O G Y V A L I D A T I O N

system, our test results demonstrate that ESXi was 24% more efficient than Microsoft Hyper-V. But 24% is once again only part of the story. Looking back, the data gathered from the SPECjbb testing demonstrates interesting patterns.

As illustrated in Figure 3, optimal performance was achieved at 4 VMs with either hypervisor. With 4 VMs running SPECjbb Hyper-V closes the performance gap to only 20%. Our conclusion is that 4 2vCPU processors running on an 8-core system was the easiest workload to distribute and balance across all processors.

When looking at workloads that are harder to balance across all processors, VMware ESXi widens the performance gap. Moreover, this doesn't only occur with high numbers of VMs. With only 1 VM ESXi turned in performance that was 28% better.

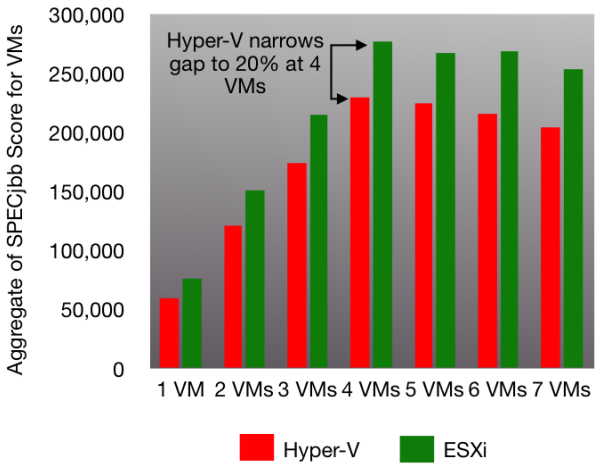


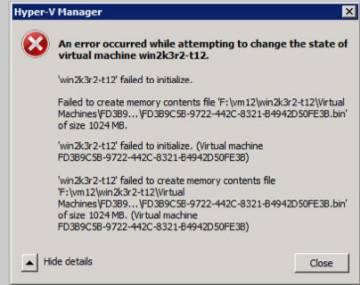
Figure 3 : Difference in aggregate of SPECjbb running in VMs, at different number of simultaneously running VMs

What this means to us, is that less optimally load balanced infrastructures are more inefficient on Hyper-V. We guess that *if*

The Usefulness of Memory Oversubscription

Many users take memory oversubscription for granted while using it extensively. Based on our discussions with end users, as many as 70% of VMware customers make use of memory oversubscription. Seasoned users of VMware have come to expect that they won't be restricted by physical memory when starting VMs. Many users plan around their ability to temporarily impact performance by moving active machines when they have already assigned 100% or more of a hypervisor's available memory - this is in fact the premise of many virtual server availability and disaster recovery plans today.

At the time we encountered the error below, we seemed to have sufficient memory available to launch another VM. Task manager identified 1.3GB free memory and our guest required 1GB.



Since memory is one of the most pressured subsystems in a virtual host, elegantly managing memory use is an important aspect of maximizing system performance. Memory may be occupied but not productively used, with large amounts of performance remaining on the system. As a testament to this, we saw overall virtual host performance increase on ESXi by another 92% when additional machines were fired up beyond the 11 limit imposed by Microsoft's hard physical memory restriction. By this assessment, Hyper-V was utilizing only 52% of total available system performance.

TECHNOLOGY VALIDATION

Hyper-V did not run into hard memory limits and could run more VMs, we would see a widening gap in performance that would be much more than 24% in ESX's favor.

In real world virtual infrastructures experiencing constant peaks and troughs in demands, this difference between hypervisors will have significant impact. Hyper-V inefficiency means that administrators must more carefully allocate resources and not place too many demands on Hyper-V. In turn, this limits VM density with Hyper-V.

Our conclusion: ESXi's performance gives it a distinct edge in hosting more VMs with dynamic workloads.

What's the real VM density factor?

We tested ESXi and Hyper-V to comprehensively evaluate their architectures and determine how many VMs each could effectively host with the same hardware. One of these tests yielded a VM density that was 2:1 in ESXi's favor. The other test emphasized a more performant hypervisor architecture and driver stack when memory management technologies were factored out of the picture. While no infrastructure continuously runs at 100% utilization, the point of this test was that VMware's architecture makes more efficient use of hardware resources, and performs better under duress.

Based on our observations during testing as well as our familiarity with a broad range of virtual server infrastructures, we know there

are many realistic workloads under which ESXi will demonstrate a 2:1 VM density advantage when compared to Hyper-V. Even so, we also realize that exceptional workloads, similar to our SPECjbb testing, will occur in the real world. Consequently, we've tempered our expectations for 2:1 VM density with the 1.25:1 processing advantage seen under the aggressive SPECjbb testing, and have selected a comfortable mid-point. Our conclusion is that a 1.5 VM density advantage – or 1.5 times as many VMs as Hyper-V – is simply a balanced point at which ESXi can easily consolidate servers under a wide range of workloads.

Our Findings:

VMware ESXi has at least a 1.5:1 VM density advantage over Hyper-V.

The Next Step: Total Solution Cost – a walkthrough of assessing Cost Per App

With our assessment of VM density in mind, we return to the goal of understanding how hypervisor density affects total virtualization solution cost, including physical equipment, licensing, and first year power/cooling/floorspace costs. We examined the Microsoft Hyper-V plus System Center and VMware Infrastructure 3 solutions with the help of a VMware calculator that fully captures a variable and complex Microsoft management infrastructure. This tool is applicable to environments of up to 1000 VMs, and is also available on the VMware website. Let's look at what the total costs are for each solution.

T E C H N O L O G Y V A L I D A T I O N

With some time in the market, and a specialized focus on virtual infrastructure management, VMware supports ESXi with 3 versions of their VMware Infrastructure 3 platform. Those solutions today include Foundation, Standard, and Enterprise editions, and each are licensed by the number of ESX or ESXi hypervisor CPUs (on a 2CPU basis). While these solutions vary in capabilities and cost, they have in common an architecture that requires only two management servers – a vCenter Server and a VMware Update Manager database server.

Microsoft’s management infrastructure is more complex in comparison. Microsoft System Center requires several elements to

obtain similar functionality. Those elements include System Center Operations Manager, System Center Configuration Manager, and System Center Virtual Machine Manager. While relatively recent MS licensing changes have simplified client licenses – virtual hosts can now get client licenses for all of these management components with a single Server Management Suite Enterprise license – the management servers themselves, and associated database servers, each need their own client licenses for the various management tools. In addition, the number of management servers and database servers will change with the size of the hypervisor environment.

	VMware Infrastructure 3	Microsoft Hyper-V plus System Center
Total number of VMs, including Mgmt Servers	51	53
VMs per host	18	12
Number of 2 CPU hosts	3	5
Total OS licensing (Datacenter for VM Hosts)	\$26,991	\$44,985
Total Cost of Management Infrastructure	\$22,517	\$13,109
Total Cost of Servers, Network, and Storage	\$90,750	\$111,250
Data center space, power and cooling	\$3,736	\$5,070
Total	\$143,994	\$174,413
Cost Per Application	\$2,880	\$3,488
Savings %	17%	0%

Figure 4 : Breakdown of Microsoft Hyper-V plus System Center and VMware Infrastructure 3 solution costs for a 50 VM environment, using 18 VMs per server for VI3, and 12 VMs per server for Microsoft.

TECHNOLOGY VALIDATION

Using cost-optimized, best practice management infrastructures for each of the respective vendors, and our validated VM density that in turn determined the number of hosts, amount of hardware, and the size of the management infrastructure, we assessed the costs of a variety of different sized environments.

Based on our assessment, Microsoft Hyper-V plus System Center only has a cheaper cost per application when a single hypervisor is required. Once more than a single hypervisor is needed, the difference in hypervisor density and licensing immediately moves the balance to VMware Infrastructure 3's favor, by as much as 5% to 29% in our calculations. A variety of factors influence the outcome of these calculations, the most significant of which are:

- Whether management servers are virtualized or physical
- The total size of the environment (this drives up the number of components in a Microsoft infrastructure)
- The cost of data center space and power and cooling, which we factored in for the first year

In Figure 4 on page 10, the advantage of virtualizing management servers reduces the cost of Microsoft's management infrastructure licensing, but even so, the virtual server sprawl created by an inefficient hypervisor drives the cost of just the OS licensing to almost 167% that of the VMware Infrastructure 3 environment.

While our experimental calculations included the first year cost of data center space and power and cooling, even with these figures

factored out, VMware Infrastructure 3 still came out cheaper in every configuration requiring more than one virtual host. In the above example, without power, cooling, and datacenter costs, VMware Infrastructure 3 would still retain a cost advantage of \$582 per application, making it 17% cheaper than Hyper-V plus System Center.

Taneja Group Opinion

As we have found in this Technology Validation exercise, the rush of hypervisors into the marketplace, and the predominance of free hypervisors does not mean the hypervisor is becoming commoditized. Reality is in fact far from this – as we've observed, there are distinct differences between hypervisors. Clearly VMware's time in the market has led to a level of hypervisor maturity that pays off in performance and efficiency with hardware.

Moreover, while in this case testing demonstrated that VM density and performance limitations are related to the architecture and design of Hyper-V, we believe that similar limitations are likely to surface in other hypervisors that use a parent partition-like approach to virtualization.

Meanwhile, observers should realize that VM density is based not just on the hypervisor, but also on the total capabilities of the surrounding virtual infrastructure. While we did not test other virtual infrastructure capabilities nor factor them into our VM density evaluation, as one example, VMware Infrastructure 3 includes Distributed Resource Scheduler (DRS). Across a pool of hypervisors, DRS ensures that each virtual

T E C H N O L O G Y V A L I D A T I O N

server is allocated sufficient resources on an ongoing basis to meet its performance needs. By constantly juggling resources among VMs based on utilization and user defined rules and policies, DRS allows users to safely run more VMs per server without compromising the performance of critical applications, even at high utilization levels. DRS thus enables users to safely increase VM density in VMware Infrastructure 3 environments. Microsoft virtual infrastructures currently do not have an equivalent capability, preventing them from offering this benefit. Due to the limited scope of our testing efforts, we did not model the impact of VMware Infrastructure capabilities such as DRS, but they should be considered as important factors in further increasing VM density.

VM density ultimately determines total solution cost, and today has VMware Infrastructure 3 in the lead. As we've reviewed in this Technology Validation,

hypervisor efficiency can yield total cost differences of up to 29% in VMware Infrastructure 3's favor. Even in small implementations, this is a difference of many thousands of dollars.

In an economic environment that is repeatedly heralding the message "do more with less" the efficiency of hypervisors are an oft-overlooked aspect of virtual infrastructure acquisition that has massive impact on total price. We recommend every organization undertaking a server virtualization initiative sharpen their pencils and review our VM density findings in this report. Then use our findings, along with the associated cost per application assessment approach we've reviewed in this report to identify the true cost of the virtualization solutions you are considering. Viewing the multitude of virtualization solutions on the market through this lens will help you figure out how to do more with less.

***NOTICE:** The information and product recommendations made by the TANEJA GROUP are based upon public information and sources and may also include personal opinions both of the TANEJA GROUP and others, all of which we believe to be accurate and reliable. However, as market conditions change and not within our control, the information and recommendations are made without warranty of any kind. All product names used and mentioned herein are the trademarks of their respective owners. The TANEJA GROUP, Inc. assumes no responsibility or liability for any damages whatsoever (including incidental, consequential or otherwise), caused by your use of, or reliance upon, the information and recommendations presented herein, nor for any inadvertent errors which may appear in this document.*