Getting the Most from Virtualization

an internet.com Networking eBook
Achieving Agility Through Virtualization
Pam Baker

Get a Handle on Virtual Machine Sprawl
Charlie Schluting

Practical Virtual Machine Architecture: How Do You Scale?
Charlie Schluting

Virtual Machine Architecture: Ensure High Availability
Charlie Schluting

Tackling Virtualized Environment Security
Richard Adhikari

So Much for Simplicity
Drew Robb
Global competition and faltering economies are increasing the pressure on CIOs to do more with less and to do it at lightning speed. The latest tool in the CIO arsenal to step up agility in response to this market squeeze is virtualization, but like most technologies, it’s turning out to be a mixed blessing.

The promise of unprecedented reaction time without additional hardware expense is “undeniably strong,” said Andi Mann, research director at Enterprise Management Associates. “You can re-use what you have with virtualization and can be set up in a manner of minutes to move to a new project.”

Beyond the obvious savings in hardware and time costs, there is the added plus of lower power expenses. “As well as providing increased resource sharing and utilization advantages, virtualization is becoming an attractive option to businesses looking to improve their carbon footprint,” said Jonathan Robinson, chief operating officer at NetBenefit, a provider of Managed Hosting services in the U.K. “By allowing you to stack numerous virtual servers, this technology reduces the demand for power; sometimes by the equivalent of 10 machines down to just one, which is great news for the environment.”

These alone are heady reasons for jumping on the bandwagon but the excitement can lead to disaster. “People dive in with irrational exuberance. Be a little more circumspect,” advises Mann. “The cost and effort can outweigh advantages if it gets out of control.”

Even vendors agree the buzz around virtualization can lead to trouble if excitement overwhelms basic management controls. “Virtualization changes a lot of things but it doesn’t change the need for basic datacenter management,” said Frances Guida, manager for Virtualization Programs at HP.

As a matter of fact, extra control is advised to avoid a virtual pile-up. “Put a control step in place so you don’t create too many virtual centers and remember to retire them when you’re done,” advises Guida.

The latest tool in the CIO arsenal to step up agility in response to this market squeeze is virtualization, but like most technologies, it’s turning out to be a mixed blessing.
Providing virtualization is implemented and managed properly, the benefits are immediate and measurable.

"The one area that really delivers the most consistent and immediate benefits in terms of both costs and agility is server virtualization in the datacenter," said Frank Gillett, vice president and principal analyst at Forrester Research. He explains that server virtualization "revolves around Windows-based x86 servers and other common low-end servers with Intel and AMD processors, not the high-end servers."

Beyond initial implementation, Gillett advises the network be setup to automatically reconfigure to handle the addition of virtual servers and the changing flow of work. "You’ll need to upgrade systems management to take advantage of virtualization."

Mann agrees that server virtualization offers CIOs the most agility. Virtual servers quickly "setup to handle short-term projects and then just as quickly scale down again."

Guida agrees that virtualization is such a boon to IT agility it is causing the "whole transformation of next-gen datacenters." She says the virtualized adaptive infrastructure is a certainty in the near future.

HP is putting money on that bet. In March, HP launched two new products -- HP ProLiant iVirtualization, which integrates virtualization capabilities into the HP ProLiant server platform, allowing customers to rapidly and simply virtualize their server environments.

In May, HP unveiled what it calls "the industry’s first software" to analyze and optimize physical and virtual resources in the same way. HP estimates that the software, called HP Insight Dynamics - VSE, will reduce the cost of common datacenter tasks by as much as 40 percent.

Even so, not everyone understands how virtualization works and how to glean its benefits. A poll conducted by NetBenefit among delegates at the recent Internet World show in London revealed that 30 percent of respondents are already incorporating virtualization as part of their IT strategy and one in three regards virtualization as a key part of their organization’s future, but 67 percent are still not really familiar with the concept.

"This poll has demonstrated the rising interest in virtualization, which we believe people will be hearing a lot more about over the next 12 to 18 months," said Robinson. "However, there is clearly still a really big requirement for education on the subject of virtualization and its benefits."

Of those familiar with virtualization, most are deploying it in servers, but a few are venturing into other areas as well.

"Virtualization tends to be deployed for multiple reasons and tends to deliver multiple outcomes," said Mann. The primary areas where virtualization is being used beyond servers include disaster recovery and storage. Other areas are being targeted as well.

Gillett said that client virtualization is "more hodgepodge, more complicated, less mature and a separate issue entirely." Virtual storage, on the other hand, he says is an obvious next step in virtualization efforts.

Guida said test and development and infrastructure are common areas CIOs move to first. The key to success, however, is implementing one area at a time to ensure you have it under control before you move on.

"Don’t get overwrought around the buzz," concludes Gillett. "The way virtualization is applied in different realms varies greatly and you need to approach it in a highly controlled manner or you’ll see more problems rather than a better solution."

Agility, after all, means the ability to move quickly in a new direction. Just be aware that the new direction may not be where you intended to go. Consider resurrecting the adage “look before you leap” but this time such wisdom applies more to where you go with virtualization rather than whether you should deploy it at all.
Virtualization is great; that much we can all agree on. Virtual machines (VMs) can tend to grow out of control, however, now that it’s so easy to create them. This should not be all that surprising, but apparently many small to medium businesses are also dabbling in VMs, and they are suddenly overwhelmed by the VM growth.

Each VM is another server that an administrator must manage. Security updates must be applied and global configuration changes now need to be propagated to all these new machines. While it’s easy to create three or four (or more) servers on one physical piece of hardware, you’ll certainly struggle if you aren’t already set up to scale.

Unfettered Growth

The number of physical machines in a small company may drop dramatically, perhaps as much as 40 percent, when virtualization is implemented. Unfortunately, the number of operating system (OS) instances will generally increase by two-fold or more at the same time. The power and cooling savings are realized, as was promised by virtualization, but taking 20 servers to 12 servers, for example, will mean you may soon have 40 OS instances to manage.

The reasons for VM proliferation depend on your culture, but the most common reason is that delegating control of an entire OS is easier than managing an application for customers. IT customers, be they engineers, application developers, or smaller IT units within an organization, frequently need more access than central IT is willing to give. The easy solution: Give them a server of their own. Test environments, too, are well served by virtual machines.

To keep hardware (and power and cooling) costs down, many companies introduce policies about the implementation of new services. New applications and servers need to be run on VMs first, unless they really require their own server. Policies such as these are good because they limit wastefulness, but they do tend to exacerbate VM sprawl.

Sprawl aside, it’s worth noting that higher utilization levels on your servers do not mean that they’ll use an appreciably larger amount of power. In fact, the power savings claims are really true, and can be even

While it’s easy to create three or four (or more) servers on one physical piece of hardware, you’ll certainly struggle if you aren’t already set up to scale.
greater if your utilization is low and you use VirtualCenter’s power management features. VMWare can migrate VMs to fewer servers if utilization isn’t high enough, and actually power off unnecessary servers. This works best with Dell hardware, but other large vendors are supported as well. Imagine all your virtual machines migrating to a few blades in a blade server during the nighttime, and then as utilization increases during the day, blades quickly boot up and take the load as needed. Granted, I don’t personally know any enterprise environments that are brave enough to try it yet, but in theory the concept is wonderful.

Dealing

Something magical happens when a company grows to around 50 operating system instances. That’s too many to manage by simply logging in and running commands, so people start to write scripts. In Windows land, if it hasn’t already happened, you must implement Active Directory. For the Unix/Linux servers, configuration management becomes even more important. Writing a script that SSHs to each server and runs a command doesn’t scale, no matter how hard people want it to. You need a real configuration management system (such as puppet or cfengine) to ensure that servers are configured exactly how you want.

If you already operate in a large environment with good automated installations and configuration management systems, chances are scaling to 100-fold won’t be a problem. Barring scaling issues with the management software itself, that is. A good network-booting deployment system is only half the battle, because every server isn’t going to be configured identically. If you’re “doing it right,” you should be able to arbitrarily reinstall any server, walk away, and know that it’ll come back up patched and running all the services it’s supposed to. Servers, or rather the OS that runs on them, should be truly disposable.

VMWare promises management of a “golden image,” probably because ITIL mentions it, but it doesn’t really help in practice. You have to create your images (somehow). There’s no mechanism to update a golden image with security patches and apply them to existing systems; you’ll generally have to reinstall the OS instances. That’s what you should do periodically, but without some kind of configuration management system, you’ll also be manually installing and configuring the services that the VMs used to provide in order to restore service functionality.

VM growth, therefore, is no different from server growth. It may be easier and cheaper, but from the OS management viewpoint, you’re doing the same thing. Likewise, the availability of your services is also in danger. Running five VMs on a single piece of hardware means that a hardware failure takes out five servers instead of one. VMWare and Xen can both be clustered and run from shared storage, such that a hardware failure will result in the VMs immediately (instantly, even) being migrated to other servers. The problem is that VMotion requires the most expensive VMWare license, and a VirtualCenter server. Shared storage isn’t as big of an issue these days with iSCSI, but it’s still another aspect that must be configured.

The point is: dealing with VM sprawl is no different than dealing with scaling up to support more physical servers. Use whatever mechanisms are available on your given platforms, and do it right. A VM is, and always will be, just another server.
Virtualization allows you to better utilize your servers, and effectively increasing that utilization can be a big cost-saver. Now that it is so easy to duplicate server instances using virtual machines, it makes sense to start talking about the best way to take advantage of this new ability. Improved application availability and performance can now be realized without many of the complications that used to plague us. How you go about it, however, may vary depending on the application in question.

Out vs. Up

I’d like to first spend some time talking about the philosophical difference between scaling “out” and scaling “up.” Some applications can be scaled up, that is they can be run on faster hardware to support more transactions. We also call this “scaling vertically.” Say we have a database server that can handle one million requests per second but, due to new demands, we need the database to handle at least two million per second. Databases are well suited for scaling up, since bottlenecks are frequently RAM and CPU. If the bottleneck happens to be disk IO, then you can also scale up the storage system the DB is using.

Scaling out (horizontally) means to add more servers and spread the load across multiple machines. In the database example, this may be extremely difficult, since all database servers will need to use the same data and it will have to be replicated. Scaling out application servers, however, is a common practice.

Before deciding whether to scale up or out, you must realize that scaling out presents its own problems. Web applications require session data so that a load-balanced cluster of servers will have the same state. A common example is authentication: if a user is authenticated with one server, and the load balancer decides to serve that user’s next request (page click) via another server, it could fail to recognize the user is logged in.

Here are a few questions to ask before deciding that scaling out is the right solution:

- Does the application operate properly in a load-
balanced environment?
• Will the application scale up to serve enough users without load balancing?
• Can I run many instances of the same application in an automatable and manageable way?

If you find yourself in the situation of having to scale up because of application limitations, you probably shouldn’t be using virtualization at all. An application that requires its own server is not a candidate for virtualization. The overhead of virtualization, as small as it may be these days, will contribute to limiting your performance. Furthermore, you will gain none of the benefits of virtualization, such as consolidation and migration between physical servers, because the application must run on its own dedicated server anyway. The migration argument, in case of a hardware failure, is a weak reason to use virtualization since failover setups can easily be configured between two physical servers.

That said, if you can scale horizontally (out), you’d probably benefit greatly from virtualization. It’s easier to manage virtual machines than physical hardware, you can take down hardware with zero downtime, and your utilization can be maximized.

Load Balancing or Separation of Duties

Assume we have a Web infrastructure that hosts a few thousand Web sites with the Apache Web server. Anyone with that many sites has probably already scaled out to a certain extent, by hosting maybe 500 sites on four servers. There are a few problems with that:
• Any machine failure means that 500 Web sites are down.
• Apache is horrible at handling that many sites and it may take a very long time to restart a Web server.
• Utilization on each server is likely very high.

We’ve made some assumptions, but let’s suppose the above problems are all true in our fictional environment. There are two solutions to the problem. Load balancing isn’t such a great solution in this case. As previously mentioned, load-balanced setups require that the application keep state in a central, shared location. It isn’t likely that all 2,000 sites have

---

Enterprise Spending on Virtualization to Rise

By Richard Adhikari

Recession? Virtualization remains a hot seller, according to a survey by ChangeWave Research. In fact, the faltering economy may have a lot to do with virtualization’s continued strong sales. Virtualization is designed to help companies reduce the number of physical servers they need, which means less spending on maintenance of servers, power, and cooling. Fewer servers also reduce space requirements. It makes sense that companies would continue spending on technology that saves them money, Paul Carton, director of research at ChangeWave, told InternetNews.com.

Of the 1,956 respondents to the second-quarter survey, 18 percent said they would increase purchases of virtualization software over the next 90 days, up five points from the 13 percent recorded in January’s survey.

The planned investments in virtualization come at a time when overall software spending has tanked: 25 percent of the respondents planned to spend less on software over the second quarter of 2008, and only 12 percent planned to spend more, ChangeWave found.

That’s a drop of three points and four points, respectively, from January.

These surveys are conducted among the ChangeWave Alliance’s members, who total 15,000 senior technology and business executives in leading companies in select industries.

Carton said that with a survey base of about 2,000 software buyers, changes of one or two percent are “significant.”

Thirteen percent of respondents said a general slowdown in business conditions and capital budgets was driving their companies’ purchasing decisions—four points more than in January and double the percentage of six months ago.
the ability to do this. A good load balancer can track incoming requests and always send them to the same back-end server, which helps a bit, but if a server goes offline, those clients must reauthenticate. Load balancing, therefore, is best suited for scaling specific applications that support horizontal scaling.

The separation of duties approach is required in the Apache-hosting-too-many-sites scenario. Ideally, we’d like to see no more than 100 Web sites per Apache instance. You won’t have Apache fail because it has too many open configuration files (one for each vhost, most likely), and you won’t have to deal with load balancer issues. It depends on the hardware, but it’s quite possible that you can run four to five VMs on each server, which gets you 500 sites per server again.

Wait a minute, that’s the same utilization level we had before, but with virtualization overhead and the extra CPU and RAM requirements for five virtual OS instances. Indeed, but each set of 100 Web sites won’t be utilizing resources the same.

When we scale out, we can realize better utilization by taking advantage of auto-migration functionality in VMWare based on server load. The more you scale out, the more opportunities there are to optimize. If you’re using Xen with some sort of cluster management software, similar rules can be constructed.

(Disclaimer: in reality, you will probably want to throw another server in the mix to handle future growth.)

In short, scaling horizontally using virtualization is much easier than doing it with physical servers, with the added benefit of shuffling around the load at-will. Scaled-out services that can be separated into many smaller portions will offer the most flexibility and benefit for a multi-VM scaling strategy. However, if a single, mission-critical application properly supports a load-balanced configuration, load balancing is probably your best bet.
Now that we understand how to handle managing VM sprawl, and we know we need to scale an application horizontally, let’s take the next step and build a VM infrastructure that’s capable of fault-tolerant operation.

Fault tolerance means that a physical machine failure (fault) will not interrupt the operation of any running services. This isn’t practical if the hardware actually fails, since VMs running on it will immediately crash. But if you want to perform maintenance or if you sense a non-instant failure looming, you can migrate VMs to other servers. A cluster can, however, automatically restart crashed VMs.

Before we get too far ahead of ourselves, let’s first explore the concept behind VM migration. Manually or automatically migrating VMs requires some special configuration.

There are essentially two requirements, or layers, involved in setting up high-availability VM environments. First, you must have shared storage and a clustered file system. VMFS in VMware or GFS in Red Hat Linux will do the trick. Second, you need something to manage the cluster, such as VirtualCenter from VMware, or the Red Hat Cluster Suite, which will keep track of where VMs are running and perform migrations. Sounds easy enough, right?

Well, it’s not easy with Red Hat, which requires the sysadmin to become extremely familiar with some quirks. Things like documented GFS commands returning “not yet implemented,” or ‘clustat -v’ returning “clustat version DEVEL,” or Conga generating a cluster.conf that the other tools claim is invalid, all in a stable release of RHEL 5.2, can result in blank stares at the terminal. That said, you could get it working. VMware, on the other hand, tightly controls its product, and it generally works as advertised.

Just what do we mean by “migrations” anyway? Migrating a VM from one server to another can be done one of two ways: live or not. Non-live migration of a VM involves the copying of a virtual machine’s disk image to another server, and then starting it up on the new hardware. This can be automated, but it results in the VM having to shut down
and then boot up again.

To perform a live migration with zero downtime, you must be on a shared file system so that disk image copying isn’t necessary. Then the program managing the migration must copy the entire memory space of the VM over to the new server and quickly "start" the virtual machine. This happens without interruptions to the OS or the applications running, and without the end-users ever knowing.

**Clustered File Systems**

You can’t just use NFS to share out the disk images to all your servers. It's too slow, and locking issues would be unbearable during frequent migrations. We need a clustered file system.

Conceptually, a clustered file system is a file system that supports multiple operating systems mounting and writing to it at the same time. It’s a very tricky business, and getting it wrong will instantly corrupt the file system your VMs are stored on.

Isolating or "fencing" a cluster node is required to ensure that your file system does not corrupt. Fencing is abrupt; most methods of fencing will talk to the hardware management interface of a server (via IPMI most likely) and immediately remove power from the server. A cluster will generally decide that a node needs to be fenced when it stops responding to heartbeat messages. Cluster nodes can also fence themselves if they discover an inconsistency in the file system they weren’t expecting. The details of setting up fencing can be tricky, so be sure to understand all your options for your given platform.

You have two options these days for creating a clustered volume: iSCSI or an FC SAN. The iSCSI route is certainly cheaper, and should perform sufficiently as long as your VMs aren’t doing tons of I/O. A busy database server is not a candidate for virtualization.

After configuring the LUNs to be accessible on multiple hosts, you then create the clustered file system and make it available on all nodes. The steps vary based on some decisions you make. You can choose to use CLVM or not, and you also get to decide if GNBD is right for you. Spend some serious time reading all of the Red Hat documents before even setting up a test environment. It’s frequently the case that you find out after the fact you’ve implemented something incompatible with another critical technology you needed to use.

**Clustering**

Red Hat Cluster Suite is stable and mostly works. It can manage services, which for the sake of this article are virtual machines. You can define which services run primarily on which physical servers, and the cluster will automatically restart services if they crash or disappear. This ability to ensure that VMs can survive hardware failure, or at least be automatically restarted to minimize downtime, is the minimum required functionality. Manually migrating a virtual machine with zero downtime is also extremely useful in a production environment. With careful configuration, RHCS does this well.

VMware's VMotion is much more evolved. It can do all of that, and with VirtualCenter’s help, you can configure load shuffling rules. Too much CPU load or RAM utilization on server A? Move a VM or two to server B. The opposite is possible as well: It can power off unneeded servers.

Both technologies strive to achieve the same thing, which in the end amounts to many extremely intelligent service-specific VMs that maintain overwhelming uptime levels. This is only the beginning of autonomic computing and self-healing infrastructures.

We’d recommend VMware for mission-critical applications. The Red Hat Cluster Suite is getting there. If you are budget constrained, use it. Just make sure you understand the limitations, and test every failure scenario you can think of before putting your Red Hat cluster into production.
As enterprises gallop ahead toward virtualizing their IT infrastructures, security and compliance issues are going to slow them to a crawl.

That’s because virtual environment security is nothing like security in the physical environment.

Security measures in the physical environment are based on the servers being fixed, having a constant identity, and being easy to check on, but the virtual environment is always fluid, always changing and difficult to get a handle on.

Worse still, the tools and processes that ensure security in the physical environment just don’t work in a virtualized one.

“‘The existing tools for remediation, discovery, and so on aren’t for the virtual world,' said Chris Farrow, director of product strategy at virtualization policy management vendor Fortisphere. ‘They don’t understand the virtual architecture is dynamic, virtual machines can be turned on or off, and typical scanning and provisioning tools don’t understand the concept of machines being able to migrate on the fly, an entire machine that you can capture on a thumb drive. They expect a box that’s on 24x7, is always sitting on a rack somewhere and not dynamically changing its identity and nature or being moved easily from one host to another.’"

The procedures for regularly assessing the IT environment, finding out which boxes are running what software, for patch management and for provisioning are "great for the physical world, but not for the virtual world," Farrow said.

“You can have a physical box with 20 virtual machines on it talking to each other all day long and there’s no way to get inside the network and find out what’s going on, so all the tools people have bought over the last 10 years or so have to be re-instrumented.”

There are three facets to the problem, said David Lynch, vice president of marketing at Embotics. These are the loss of identity; mobility; and the loss of control by the IT security team.

Security measures in the physical environment are based on the servers being fixed, having a constant identity, and being easy to check on, but the virtual environment is always fluid, always changing and difficult to get a handle on.
In the physical world, a server is identified in the environment by its physicality -- the rack or row number, or something associated with the physical machine -- and, when it's virtualized, "you, in essence, remove its identity," he said.

To make things worse, cloning a virtual machine results in several identical copies, and that creates system management, maintenance, and updating problems because it’s difficult to identify and differentiate the various clones of a VM from one another.

**Adhering to compliance**

Ensuring virtual machines are adhering to compliance and separation rules is also difficult because VMs are highly mobile, and can be migrated automatically to a different physical server if the resources of the one they're on are inadequate.

For example, an enterprise's human resources systems or credit card systems could end up running on a server where they could be potentially accessed by a Web server application when the VM they are running on is kicked over automatically to a new physical server.

Consolidation, which is the main reason corporations opt for virtualization, can also lead to this problem because "you might have had separate VLANs (virtual local-area networks) and segments for different kinds of data -- customer data, credit card data, and so on -- but when you consolidate 20 physical servers into a single ESX host, all that data is on the same virtual switch so, more often than not, your data and network segmentation are lost," said Michael Berman, Catbird's chief technology officer.
Virtualization is exploding in popularity. Virtual machine deployments are expected to grow from 540,000 in 2006 to more than 4 million in 2009 according to research house IDC. While the benefits are widely advertised, the complexities have not been comprehensively discussed. As a result, there appears to be a prevailing attitude of complacency in the VM arena. Because you can potentially do so much and the gains are often so spectacular, server administrators might not be taking the same precautions on performance management as before.

“There is a perception that VMware VirtualCenter and basic resource throttling is enough, but this doesn’t give the full picture,” said Andi Mann, an analyst at Enterprise Management Associates. “The bundled tools for managing VMs are not enough to guarantee SLAs (service level agreements) based on business performance objective.”

Virtualization, after all, adds another layer into what is already a complex environment. You start with an OS, applications, Web servers, middleware, databases, interfaces, etc., and you add to that a hypervisor layer, which is largely deficient in fundamental management tools and capabilities like performance/capacity management. Since it takes only seconds to add a new VM, they tend to proliferate if left unchecked and this creates VM sprawl — an uncontrolled proliferation of virtual machines. The result is you have multiplied the volume of systems you need to manage, increased the depth of management required, and yet has insufficient tools to do so.

According to Mann, the management tools in VirtualCenter and other virtualization platforms do not provide a broad view of performance across multiple hosts and subnets. Nor do they help administrators to understand physical performance issues. They are not really aware enough of applications; let alone the interactions of multiple components in a composite application (with a separate app server, database server and Web server, for example). VM tools also tend to miss the boat with regard to business services and priorities.

Since it takes only seconds to add a new VM, they tend to proliferate if left unchecked and this creates VM sprawl — an uncontrolled proliferation of virtual machines.
"If you do not properly manage performance, you can end up with a single VM overusing or saturating resources in a host," said Mann. "An overactive application can saturate the channels to the database, using 95 percent of the network interface, which slows down I/O for all other VMs on the same host."

But that’s just one scenario. A highly processor-intensive application can saturate the server, using 95 percent of the CPU. This leaves only five percent for the rest of the applications on the VM. Interestingly, one of the many touted benefits of virtualization, the elimination of under-utilized servers, may actually be one of the consequences of this lack of effective VM management tools.

Under-provisioning, said Mann, tends to happen first, i.e., attempting to squeeze as many workloads as possible onto a single system. "Without accurate performance and capacity tools, under-provisioning is usually the first mistake as administrators and IT managers typically put more VMs on a server than it has resources to deal with," said Mann. "That leads to over-provisioning as they react by making sure they have spare headroom even for exception cases."

Virtualization vs. Capacity Planning

The time-honored practice of capacity planning, then, is essential in any virtualized environment. Unfortunately, many incorrectly assume that as virtualization’s popularity increases, capacity management’s value steadily diminishes. The opposite, however, turns out to be the case.

"Despite propaganda to the contrary, capacity planning is more important than it has ever been," said Jerred Ruble, CEO of TeamQuest Corp. "Technologies such as VMware Distributed Resource Scheduler, utility computing, IBM Workload Manager or grid computing will never eliminate the need for solid capacity planning."

Such tools provide intelligent dynamic resource allocation, continuously balanced computing capacity, real-time server utilization optimization and automated dynamic reconfiguration. They certainly help manage existing environments, add much needed automation and ensure workloads have appropriate resources. They can also be useful in supplying capacity quickly and easily to meet varying usage requirements. But they don’t tell you what you need, don’t relate well to business goals and don’t help you look into the future.

It’s one thing to add more processing power and memory to optimize a specific application or VM. But the last thing you want are such changes being made automatically and uncontrollably as the ROI many not merit the investment.

That’s where capacity and performance management tools come in from vendors like TeamQuest, CA, and BMC Corp. They monitor existing performance levels and enable IT to model different scenarios to determine what changes should be made to better support VMs. Perhaps more importantly, they relate the cost of proposed changes to the performance benefits of implementing them. By viewing this in advance, IT can then find the sweet spot in terms of cost/benefits and implement accordingly.

Take the case of a large insurance firm that demanded a response time of 1.5 seconds for a new application. Using TeamQuest Model, the capacity planner discovered that this solution would cost $15 million. Further modeling revealed that a three-second response time would reduce the budget to $12 million and a five-second response time would cost $10 million. By providing management with this information, they took a second look at their original specifications, concluding that it was better to accept a short delay than to pay an extra $3 million for 1.5 seconds of response.

In addition, capacity planning facilitates the rightsizing of new applications. Ruble tells of an IT service provider implementing a new application where the capacity planner modeled the new app and discovered it would flood the network with log traffic. This data was presented to designers who corrected the bug before it became a problem.

"Capacity planning takes the guesswork out of accommodating future business workloads," said Ruble. "It also ensures that a virtualized infrastructure is configured optimally to meet required service levels."

Virtually Lacking

A recent survey by Netuitive, Inc., supports the lack of virtualization management. VMware customers were polled about their ability to manage VMs. Ninety-four
percent weren’t confident in the tools they currently use to manage their virtual environments. Respondents cited poor visibility into performance, difficulty in isolating root causes, and high administration time as their major complaints.

“A new approach is needed, one that uses sophisticated, real-time analytics to reduce the massive manual effort of managing VM complexity and ultimately creates confidence and restores performance predictability to managing VMs,” said Mann. “That requires collecting metrics across virtualization technologies, vendors, and platforms, and across both guests and hosts, correlating them with each other and with physical metrics, and aligning them with application and business policies.”

Capacity and performance tools fulfill many of these needs. In addition, Mann names Netuitive, Hyperic Inc., InfoVista SA in Paris and eG Innovations Inc., as niche vendors with promising technology in analytics and VM monitoring.

This content was adapted from Internet.com’s Enterprise Networking Planet, CIO Update and InternetNews.com Web sites. Contributors: Richard Adhikari, Drew Robb, Charlie Schluting, and Pam Baker.