

Strategic Advisory

An Executive Perspective on Server Virtualization Benefits, Risks, Financial Issues

Glenn Miller, VP of Technology
Macquarie Equipment Finance

REVISED JULY 2010

Server virtualization has been commercially available since 1972, starting out on mainframes (with VM/370), moving quickly to the mini-computers of the late 70's (e.g. DEC's VMS), and into the Intel/AMD x86 worlds in the mid-late 90's. The technology marketplace for virtualization products for the x86 architecture world is very active, with competitive offerings appearing almost weekly, and management eco-systems being built out rapidly.

Some of the technology advances in the Intel/AMD world actually go beyond the more mature capabilities of earlier generations of products. A good example is the ability to take a disk-file 'picture' of an entire physical server (business software, operating system, system hardware and peripherals) and run that 'picture' on another physical system. The disk-file image is called a 'guest' and the physical system that it runs on is called the 'host'.

As with any powerful technology, there are tradeoffs and issues which must be considered. In a period of rapid technology change and progress, attention must be paid to the underlying technology architectures used in each of the many competing approaches. However, many of the benefits, risks, and financial issues will be same for most of the competing architectures.

Expected Benefits

There are many benefits to virtualization, some of which have been widely realized and some of which are still only in the 'early-adopter' phase. From a management perspective, each of these benefits should be examined for measurability and recovery. In other words, each benefit should leave 'tracks across the financials' somewhere, and should be 'recovered' for re-investment in the business. Management needs to identify where these 'tracks' are, and then how to harvest (where possible) these benefits for business advantage.

Bring legacy operating systems and applications into modern, more manageable infrastructure

Virtualization of an OS (operating system) or application typically makes it 'portable'. This allows the software to be hosted on newer systems, with better management features. Modern backup devices and software, newer HA (High Availability) solutions, performance monitors, and workload management utilities can all be brought to bear on the newly migrated OS/applications. This retains the original functionality, but surrounds it with a better management infrastructure. For example, critical applications running on an older, initial release version of Windows Server 2000 could be migrated onto a newer server with encryption, better security, better backup processes, and failover.

Bring legacy operating systems and applications into modern, higher performance infrastructure

Similar to the first point, a virtualized OS/application can be moved from a slower performing system to a much faster one, allowing significant performance gains to be realized. Even with the overhead of virtualization, the net effect can be a significant performance gain for

the application. A three-year gap in server hardware, for example, typically represents a price-performance improvement of over 40%, while virtualization overhead is typically less than 10%. In addition, the other system elements, such as NIC (network interface card) speed, disk I/O speed, and backup software speed are also considerably faster on the newer system, adding to the gains. This actually allows IT to increase realized performance, while consolidating footprint.

Augment legacy applications by leveraging new features of newer (impersonating) physical systems

In some cases, peripheral systems attached to the newer system can increase the value or decrease the operational cost of the legacy application. The newer device 'impersonates' the older device (which was attached to the original server), yet adds value by features unknown to the virtualized OS/application. Examples of this would be where a disk-hungry application is moved to a server which has a thin-provisioning SAN (storage area network) attached, or where an OS with a small-memory footprint (therefore with heavy paging usage) is run on a server with solid-state disk.

Facilitate consolidation/centralization for those standard benefits

A very typical usage scenario is where many individual, distributed servers are each virtualized and consolidated onto a single server (or server complex). Apart from physically moving the original servers into a centralized facility, this is essentially the only way to achieve a practical level of centralization benefits. The original servers are decommissioned (and returned to the lessor, disposed of, or recommissioned for other usage), but users continue to access

the server/application images from the new, consolidated server. This allows the generic/standard benefits of centralization/consolidation to be reaped: leverage of labor/skills, leverage of peripherals, leverage of physical and network security, economies of scale, negotiation power with vendors, easier DR (disaster recovery) planning, etc.

Increase asset utilization

It is well known that our distributed servers (in 'server sprawl' cases) were woefully underutilized. Traditional wisdom has it that classic x86 servers ran at less than 20% utilization. This represents, of course, significant asset waste. By consolidating 2-4 physical servers onto only one of those physical units, utilization increases dramatically. Of course, overall asset utilization doesn't change (i.e., you still have a couple of 100% idle decommissioned servers), but once you dispose of the idle systems, then the utilization story is fully accurate. [The same applies for all the peripherals attached to the consolidated unit—they are leveraged more highly as well.]

Reduce power, heat, and cooling requirements footprint

When you move from many small servers to fewer large ones, there is often a reduction in aggregate energy requirements for power, heating and cooling. However, this is a difficult reduction to measure, and even harder to capitalize on, due to the fact that the energy used by the distributed servers was not 'consolidated' in one place—the energy costs for the individual servers were 'invisible' since they were not broken out as individual line-items on the financials. But in the case of a centralized IT facility, the energy costs will appear to have grown substantially because now all the energy demands

are concentrated in one, centralized budget. But, at a theoretical—although typically nonmeasurable—level, overall energy usage can be reduced.

Reduce actual floor space required for servers

This is largely a function of centralization (but requires removal of decommissioned servers to free up their real estate), but can be much greater with virtualization technologies since the new systems require zero additional floor space (after the first one).

Reduce physical network load

In many cases, interdependent servers can be virtualized and hosted on the same physical machine (like is done on mainframes in logical partitions—one partition will host the database, one the web access layer, and another the business logic layer). In such cases, the network connections between the VMs (virtual machines) run 'inside' the physical machine (often in internal memory) in virtual LANs. These 'software networks' can be very, very fast—speeding up overall performance—but they also shift some network traffic from the physical net to the internal virtual net. This can free up network bandwidth capacity, in specific subnets. [There is a downside to this, though, in that requirements of the multiple VMs for external data will now be consolidated onto a single NIC and perhaps overload the shared NIC.]

Implement policy-based resource management

In many cases, the virtualization technology used is supported by an eco-system of management products and tools from multiple vendors. Some of the tools are focused on workload management, implementing various types of policy-based logic. [Many of these tools migrated down from the mainframe and high-end UNIX worlds.]

Typical examples would be tools that move a VM from one host to another (for reasons of proximity to data) and tools that allocate more processors to a given VM during anticipated peak-load periods. When servers are virtualized, IT can take advantage of workload and provisioning tools to simply the IT management workload, and increase overall performance/reliability.

Facilitate rapid reconfigurations

One of the major advantages of software-based virtualization is the speed at which it can occur. The ability to move VMs from one physical system to another ('on the fly') means that rapid reconfiguration of the server complex is feasible. This can come in handy in situations of unforeseen downtime, peak load adjustments, and tech refresh.

Consolidate/centralize system integrity management functions

(e.g., patch management, security). In addition to the standard centralization benefits available in consolidation scenarios, virtual infrastructures can also be protected more efficiently, through centralized software tools. Tools exist today to allow network-wide application of updates, patches, software agents, security scans, problem containment, failover, etc.

Expedite fault isolation measures

When a VM is compromised by malware and/or becomes a 'carrier' of malware to other systems—creating risks to other systems on the network—it is very easy to shut it down and restart a fresh, uncorrupted copy of the VM. This has been done with physical servers for years and years, but it is much simpler and faster in a virtualized server environment.

Implement accountability/chargeback options

The current generation of virtualization technologies are starting to produce tools to allow measurement and reporting of resource usage. This will allow IT departments to implement cost-awareness programs, and eventually (if desired), cost-chargeback programs for system resource usage.

Reduce cost of test and development equipment

One of the earliest uses of server virtualization was in the Test/Dev departments. IT personnel would have virtual copies of all the server OS/application stacks, running on only one physical server. When a needed upgrade, patch, hotfix, or new application became available, these VM copies of the production servers were where these changes were tested first. This reduced quality costs, of course, since the patches weren't pushed out to servers which would have problems with the upgrade. Additionally, it was easy to create a new VM for application development work. This strategy was popular because it reduced the hard invoice costs to buy extra physical systems for Test/Dev groups.

Simplify and improve disaster recovery processes

Since the state of a physical server can be captured in a 'snapshot' to a disk file using virtualization tools, it is a fairly simple process to create full-system 'restore points' for entire server rooms. And, as disk files, these can be replicated over WAN connections to DR failover sites. Since we often allow reduced performance—but not reduced functionality—during a catastrophic outage, the physical server equipment in the DR site can be less 'dollar intensive' than the live, production server site. This helps contain some DR costs. Furthermore,

since the snapshot process can be done often, the ‘currency’ of the systems in the DR site can be very fresh.

Reduce cost for ‘spot’ appliances

IT often needs ‘spot’ solutions to problems not addressed by the big vendors yet, and these solutions have classically showed up in the marketplace as ‘appliances’. These are dedicated physical servers running some point-product software to meet a specific need. Early examples were firewalls, TCP/IP accelerators, and XML accelerators. In the virtualization marketplace, many of these solutions can now be downloaded as fully-contained VMs, ready to be provisioned on a generic host server. These software-only appliances contain their own specifically configured OS—like the physical versions of the appliance—but are available at a fraction of the cost of the physical unit. This represents both a less-expensive and a faster way to solve spot problems in an IT infrastructure.

Avoid or defer future server costs

Another early use of server virtualization was to reduce the requests for additional future servers by the business units. Once virtualization was in place, each new request for a small x86 physical server would first be evaluated for possible implementation as a virtual server (involving only software licensing costs). In the cases where a VM was an acceptable alternative to the physical asset, hard invoice costs were avoided (or at least deferred until the workload grew to require dedicated machines).

Facilitate implementation of new/next systems

Migration to new operating systems and infrastructure elements often is hampered by the need to support

legacy applications. In some cases in the past, servers which hosted multiple applications could not be upgraded, due to the limitations of only one of the applications. All the other applications had to remain on the ‘old OS,’ because the legacy app could not be migrated. Virtualization of that system would allow the legacy app to run on the ‘old OS’—now running as a guest VM on a newer OS—while the other applications could be moved to running natively on the new OS. This would allow progressive modernization of most applications, without jeopardizing the business in the process.

Facilitate the future move to wider auto-provisioning

Eventually we will get to a more automatic provisioning model in the IT shop (sometimes called ‘in-house utility computing’). This is an extension of the point above about ‘policy-based resource management’, but introduces the notion of such virtualization of all server-related resources too: SAN, NAS, networks, routers, etc. These virtualization technologies are already in development (and in some cases, in deployment) in their respective technology areas, and server virtualization will be a necessary part in the truly autonomic/adaptive data center.

Anticipated Risks

Just as benefits need to be analyzed for measurability and exploitation, risks need to be evaluated for both probability of occurrence and business impact. The level of exposure in these risks will vary by industry, and therefore mitigation tactics will need to be developed for at least some of these.

Standard centralization/consolidation risks

These are fairly well-known to IT practitioners, and include all the ‘eggs

in one basket’ risks, concentrated physical plant risk, span-of-control risks, etc. These are not unique to virtualization.

Virtual “server sprawl”

The fact that it only takes a second or two to create a full VM means that this process needs to be controlled. Virtualization—as a consolidation play—was very successful in helping IT contain (and even reverse in many cases) physical ‘server sprawl’, but some shops are seeing that ‘virtual server sprawl’ is starting to happen as well. This has created reduced manageability and lower leverage of physical assets. If not managed, virtual server farms may become less manageable than were the physical farms – since VMs cannot be physically seen! This issue extends to making sure patches are applied, software agents installed, etc—all the things which were required for physical machines.

Performance bottlenecks

These include the generic risks of consolidation of workloads, but in virtualization there are some additional aspects. Since all the virtual components eventually map down to shared physical components, the physical components can easily get overloaded, if there are too many VMs. For example, a NIC that handles network traffic well at 20% utilization can become a bottleneck when trying to run six VMs (each at 20% utilization) on it. This concentration-of-workload problem shows up in external network access, SAN/NAS I/Os, and tape/backup systems. This is manageable of course—we are supposed to be doing ‘capacity planning’ even for VMs—but it is something to watch carefully. The technology marketplace is beginning to see offerings and developments to help alleviate these potential problems.



Higher equipment costs

To consolidate 20 physical systems onto 2-3 larger systems requires capital outlay for the new systems—often without ‘getting any money back’ for the decommissioned systems. These new larger systems always require more/faster disk (to support the consolidated workload), faster backup devices, more redundancy, considerably more memory, etc.

Extra planning workload

IT shops must give forethought to their virtual architecture design, to avoid major management headaches downstream. For example, naming conventions should be designed early and creation/alteration/retirement procedures and controls need to be defined and implemented. Since many IT shops implemented VMs early without such procedures and planning, additional effort will be required at converting pre-standards VMs (e.g. “Sleepy”, “Sneezy”, and “Dopey”) into standardized naming schemes (e.g., “CustMas1”, “CreditValD8”, and “VendorDB”).

Regulatory and privacy issues

In the physical system world, organizations sometimes have to isolate systems from one another—to avoid ‘data closeness’. The regulatory environment might require sensitive data from one server to be ‘at arms length’ from sensitive data on a different server. If, in a virtualization initiative, such organizations host both of those servers as VMs on the same physical server, then they might face a compliance risk issue. Technologically speaking, VMs are isolated ‘sandboxes’, but this may be difficult to defend to a regulatory body. Additionally, many administrative roles for these servers are sometimes consolidated into fewer hands, when the legal environment might require stronger separation of roles.

Skillset scarcity

As dependence on virtualization grows, so too will dependence on skilled personnel (or outside firms) to manage and exploit our infrastructure. Due to the newness of the technology eco-system (for enterprise-scale environments), the personnel with deep experience in the field are fairly scarce. This can create a hiring, cost, and retention challenge for IT, in the short run.

Baseline issues are shared with physical systems

VMs require at least the same amount of attention and protection as physical machines. This will include OS hardening, administrative role segmentation, VLANs, IDS/IPS security systems (including the ability to see interactions between VMs), firewalls, encryption, physical security of the host, climate control for the host, etc.

Inconsistent application of security policies

IT worked hard to develop security practices for its physical servers, but often those practices are not applied consistently to VMs. This creates security risks similar to that of un-managed physical networks.

Licensing legal risks

There are generally some limitations on the ‘movement’ or redeployment of various software components (Windows being a great example) between physical systems. The market success of virtualization has created economic forces to soften some of these restrictions, but, in general, the more expensive/important the software is, the more restrictions are placed on its mobility. Given the ease of creation and movement of VMs, we will likely see cases where licensing terms are violated.

Financial Issues

Several of the financial issues have already been mentioned above, and can be addressed by pre-planning. However, all of these need to be assessed for importance by the executive team. In some cases, there are technology timing issues in the marketplace which need to be exploited (or avoided!) and some situations are perfect fits for leasing arrangements (for refresh, cost recovery, or basic financing reasons).

The highly visible host system cost

As mentioned above, the host physical system has to be a powerful system, with high performance components and peripherals. And, although price/performance has been increasing dramatically and consistently over the past several years, the overall system costs still end up fairly high (or at least, ‘very noticeable’ on an invoice!). This cost is often increased even further by requiring a lesser-or-equal duplicate of the system to be placed off-site as a disaster recovery capability. Virtualization software itself costs money (at least the fully-supported ‘good stuff’ does), and the software to manage the virtualized infrastructure often costs more than the virtualizing software does. This investment will largely be a cost-avoidance play (i.e., fewer physical servers need to be bought going forward), so ROI calculations will need to know the current physical server procurement run-rate.

Decommissioned servers

These servers are generally idled by virtualization, and as mentioned above, represent investments beyond recovery. They typically are not suitable for trickle-down tactics (e.g., to end-users) since they are typically configured without user-oriented features (e.g., professional-class graphics cards, sound capabilities,



removable media ports), and sometimes are outside their warranty period. Certain tactics can be used, however, to reduce the investment loss. Had a leasing program been in place before the virtualization initiative began, there might be return/refresh options available—which could greatly soften the impact (and even help with financing the newer systems). In situations where the servers have been purchased, some cost recovery might be available through sales to a used equipment dealer (also available through mainstream lessors).

Software licensing

This area is one that is in flux, as software vendors are trying out various pricing models in virtualization contexts. They have mature pricing strategies for monolithic physical servers, but the challenge of multi-core systems and virtual machine systems are still quite significant. There is an opportunity for the end-customer to negotiate 'special deals' and 'interim pricing', given the uncertainty in the software vendors. As mentioned earlier, software licensing costs do not go away in virtualization (like some hardware costs do), but they can be subject of effective negotiations efforts.

Performance gains and market timing

The processor technology 'race' between Intel and AMD in the x86 space has produced an interesting planning challenge for IT shops through the end of 2011. In 2009, both Intel and AMD upgraded their workhorse processors with new designs and new memory (i.e. DDR3). These showed considerable

improvements in price/performance and performance/watt. Since many (if not most) virtualization initiatives in the x86 space are implemented on server blades or on server rack-units, this allows the possibility for blade/rack-unit swap-out. Many firms will be procuring server blades/rack-units today (to start the virtualization initiative) and then swap out some of those blades in late 2012/early 2013 with the newer technology. This will be often be done for reasons of holding the wattage, cooling costs, and floor space constant—while growing the performance of the system to keep up with business demand growth. To change high-performance blades out at the 18-24 month mark can be very expensive financially, unless a suitable fast-refresh leasing structure is in place.

Labor costs

As mentioned above, the labor market for virtualization skills will be highly competitive for hiring companies over the next couple of years. In major urban markets, this could be a significant challenge to find, attract, and retain the necessary personnel. In secondary markets, this may even involve relocation needs. It would be prudent to begin training internal staff in this area (if not already doing so), although this will likely not be adequate for production needs at first. These internal staff would, however, be able to learn critical skills from those hired (with experience), and eventually reduce the dependency on the scarce resource. [Of course, retention strategies would need to be implemented to keep the newly-trained resource from shopping the job market early!]

Conclusion

Although the server virtualization marketplace is churning with changes in technology, pricing, alliances, and channels, the technology has already proven to be of significant business advantage over the past several decades. Potential benefits must be assessed and placed over against potential risk assessments, and the financial investment, funding vehicles, and ROI measurement methods must be developed.