Dragon Slayer Consulting

Marc Staimer

W H I T E P A P E R

Overcoming Virtualization's Unintended Consequences

An integrated Virtualization Approach for the Mass Market • June 2010

Overcoming Virtualization's Unintended Consequences

An Integrated Virtualization Approach for the Mass Market

Marc Staimer, President & CDS of Dragon Slayer Consulting
marcstaimer@comcast.net 503-579-3763

Introduction

Virtualization is an incredible tool. It breaks the permanent association between an asset's image and the actual physical asset. This can solve a lot of problems and issues for servers, storage, and even storage networks. The two key benefits are a dramatic increase in effective asset utilization and a corresponding decrease in disruptive scheduled application downtime. Increased effective asset utilization is most noticeable in server virtualization. Per ESG and Gartner, average non-virtualized server utilization is between 7 and 15%. When virtualized, utilization increases to between 90 and 100% or roughly ten-fold. By utilizing many of server virtualization's key features such as online live migration of applications while operating, boot from disk, golden images, snapshots, etc., application scheduled downtime can be drastically cut.

Whereas storage and SAN virtualization do not provide the same dramatic improvements as server virtualization, there are substantial increases in asset utilization and significant decreases in application downtime. This is a good thing.

Conventional wisdom would have everyone believing that the primary reason for IT virtualization adoption is consolidation. Conventional wisdom is wrong. User surveys by Dragon Slayer Consulting show that the primary rationale for virtualization deployment is increased application availability (more than 72%). Asset consolidation is the principal *financial* rationale.

So why then is the adoption in small to medium businesses and enterprises occurring at such a glacial pace? It has everything to do with the law of unintended consequences. Unintended consequences are the proposition that unforeseen outcomes can often overshadow the primary desired result. These unforeseen outcomes can be and often are undesirable and generally accompany every activity, however well intentioned. In layperson's terms this is known as Murphy's Law: "Whatever can go wrong, will." Virtualization has added new meaning to Murphy's Law in that nothing in virtualization is as simple as it first seems.

Word is getting out. Smaller IT organizations have neither the expertise nor bandwidth to resolve frustrating virtualization issues. This paper will briefly examine those unintended problems, current industry workarounds, as well as industry solutions and why they're raising market anxiety. It will then examine a more effective and efficient way to provide the benefits of virtualization without the hazardous side effects.

Virtualization Integration Problems: Connecting the Dots

Virtualization benefits are well documented whereas its problems are not. Some of the problems that often occur include:

- Too much oversubscription resulting in inconsistent VM performance.
- Difficult multi-vendor troubleshooting. Especially, troubleshooting of intermittent problems such as variable performance.
- Pathing errors between VMs and storage LUNs. More urgently problematic when discovered on failover.
- Cabling mismanagement (incorrectly marked or not properly identified)

Many of these problems can be directly correlated to awkward, unwieldy, or poorly thought out interactions between virtual servers, virtual storage, virtual SANs, and infrastructure. Tuning one area to its optimum performance and value creates problems in the other areas. Not that these problems do not occur in the physical world. They do. However, with the rush to virtualize, these types of problems are more frequent and more complicated to resolve. A closer look at these specific virtualization problems makes clear as to why.

- **Too much oversubscription resulting in inconsistent VM performance**
  Oversubscription is the assignment of more potential utilization demand than a resource can possibly handle if all that potential utilization demand were to take place simultaneously. The underlying assumption being that this possibility is exceedingly remote. When this assumption is true, oversubscription makes good technical and financial sense because it demonstrably increases resource utilization. Increased resource utilization means significantly lower capital (less hardware and software), and operating expenditures (less maintenance, real estate, power, cooling, and management).

  Historically, bad things happen when there is too much oversubscription. Oversubscription means the possibility of utilization demand being greater than the available resources becomes probable. When that happens, the resource cannot respond quickly enough or at all to utilization requests. Most people experience too much oversubscription when they dial a telephone number and received a “fast busy” signal. This is an example where the network oversubscription level is greater than the current demand resulting in an inability to handle the request.

  Oversubscription is the principle behind server virtualization’s much increased server hardware utilization and its value. The performance problem occurs when the oversubscribed virtualized servers are overlaid on top of oversubscribed networked storage. Instead of being additive, the oversubscription is multiplicative. A simple example illustrates this concept clearly. Common best SAN practices have the target SAN storage array ports oversubscribed on average at 8:1. This means there are 8 physical server initiator ports connecting to every SAN storage array target port over the SAN fabric. For hypothetical purposes, let’s make those physical servers virtual servers now with no more than 5 virtual guest servers each. If the SAN storage array target port oversubscription doesn’t change, the actual oversubscription ration has just increased by 5x making it 40:1. The odds of a given target port being busy increased by 5x. Those VMs are not going to get the storage performance they were expecting. And if the target ports get too busy, the SCSI protocol is not known for being forgiving and will time out. The virtual application is then told its storage is not there and will crash. This is a great way to get the IT administrator’s phones to light up like a Christmas tree.

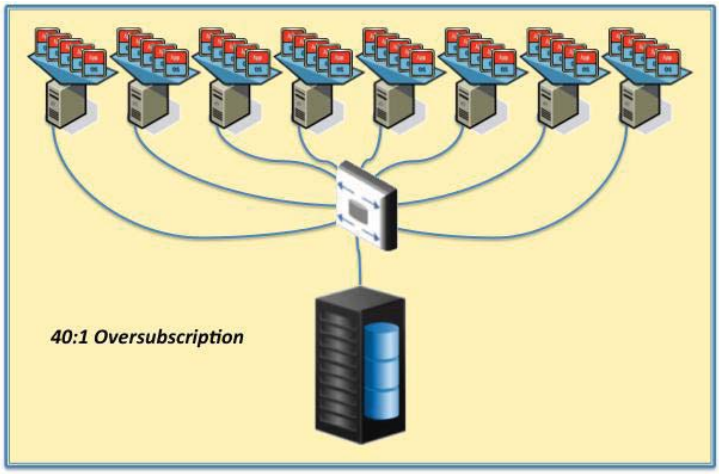

Figure 1: 40:1 Virtualization infrastructure Oversubscription Example

Another aspect of virtualized server oversubscription that gets overlooked (while occurring quite frequently) is LUN or volume oversubscription. The server virtualization hypervisor has a storage virtualization capability. This capability allows the hypervisor to take the physical volumes assigned by the storage admin, and carve them up into virtual volumes for the virtual machine guests. The problem occurs when multiple virtual applications start pounding on those virtual volumes simultaneously. Each virtual application is requesting IO from what it believes is its own unique disks when in reality all the virtual applications are being serviced by the same shared disks. This can get very ugly very fast. IO requests are queued by both the storage system and the HDDs. As the queues or buffers fill (which happens a lot faster with SATA than SAS or Fibre Channel drives because they have much smaller queue depth), response times degrade. When the queues become full, the virtual application IO request does not get serviced at all (fast busy). Once again the SCSI protocol will very likely time out and crash the virtual application.

- **Difficult multi-vendor troubleshooting (even difficult with single-vendor)**
  Just as virtualization masks an asset from its image, it also makes troubleshooting much more difficult. The virtualization layer creates a firewall that causes enormous challenges in trapping, capturing, and assessing the root cause to problems. Take for example a "too much" oversubscription problem. Too much oversubscription isn't just a storage phenomenon, it can and will occur in the TCP/IP network as well leading to severe congestion events decimating user and application performance. So when an application begins to noticeably fall below SLA performance requirements, how will the admin know where to look first? Is the "too-much" oversubscription in the network? Is it in the physical server? Is it in the SAN? Is it in the virtualized storage? Is it in the volume? Is it in the storage system? Is it in all the above? This is a troubling problem that can lead to vendor finger-pointing when discrete products are put together in a non-integrated approach.

Figure 2: Troubleshooting

The integration problems get much worse and even more difficult to diagnose when scaling virtualization. There gets to a point when there are just not enough IT professionals or hours in the day to manage the ongoing integration and ongoing task issues. It is like squeezing a balloon. Fix or squeeze something here and it bulges out there.

What becomes readily apparent is the operational budget was grossly underestimated. Operating expenditures in time, maintenance, and human assets quickly exceed all expectations. It's further exacerbated by power and cooling requirements of discrete systems not designed to cooperate in optimizing energy consumption.

- **Pathing errors between VMs and storage LUNs**
  Storage network pathing errors are extraordinarily trying at best and maddeningly annoying at worse. A simple Google search reveals thousands to millions of user posted pathing problems. They can be as simple as an incorrect FC/FCoE zone or as difficult as a down level driver/microcode. Pathing problems are serious. These errors will prevent a VM from booting,

stop a VM dead in its tracks during a primary path fault, or reduce performance because it can't load balance.

Regardless of the root cause of a pathing problem, finding that root cause is an exercise in manually intensive, time consuming, detail oriented focus. An exercise that is likely to be performed on weekends, late nights, and holidays.

- **Cabling mismanagement**
  Cable mismanagement problems occur when cables are poorly labeled, incorrectly labeled, or not labeled at all. They crop up during server, SAN, or storage hardware installs, changes, and additions. Few administrators appreciate cable management until there is a problem. Tracing cables is not a job for the faint of heart or those with ADD or ADHD. It is tedious, time consuming and enormously manually intensive. It's extraordinarily difficult to even determine if a bad cable, cable connector, or transceiver is the root cause of a pathing or performance problem.

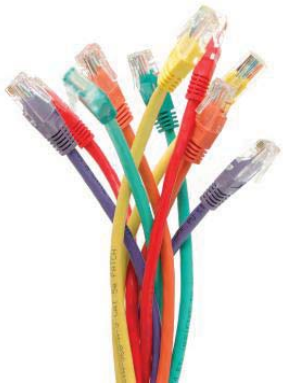

Figure 3: Cable Management

Common Virtualization Problem Fixes or Workarounds

Most IT organizations consider three options to solving the unintended consequences of virtualization

1. Hire Dedicated Resources to Fill a Custom Need: Larger organizations with mission-critical applications and uncapped budgets tend to throw people at the problem. This is the domain of large consultant groups and specialized trained professionals (employees or contractors) that have the expertise and experience to complete the extensive manually intensive tasks both during the day and off hours. While throwing people at the problem is the best known and most comfortable approach, it carries a steep price in time, resources, and capital. It's a point-in-time solution that doesn't help when problems reoccur (which they do). In a nutshell, it's expensive, inefficient in resolving these problems, and offers no guarantee of success.

2. Use off-the-shelf tools to simplify integration: When less customization is required, an internal staff or a systems integrator can often deploy off the shelf management tools and custom scripts. The positive aspects of upfront integration are that it eliminates implementation headaches as well as some of the ongoing management and growth issues. The negative aspects come from lock-ins to the integrator or homegrown customization usually in the form of scripts or custom software. The scripts/software are typically unsupported (e.g. no patches, hot fixes, upgrades, ongoing testing and QA, etc.), and undocumented. Script/software changes are difficult if not impossible to make. It tends to be inflexible.

3. Multi-vendor proprietary integrated solutions: More recently a third option emerged. It's a factory integrated and supported virtualized servers, SAN, and storage packaged as an all-in-one system. The first came from Cisco with its Unified Computer Systems (UCS). HP and IBM thought UCS was such a good idea that they offered variations of their blade server systems packaged similarly.

   These factory-packaged systems are essentially x86 blade servers with extended memory, VMware vSphere4 pre-installed, a SAN switch (Fibre Channel or FCoE), and preconfigured external storage systems (EMC or NetApp for Cisco, EVA or LeftHand P4000 for HP, or LSI for IBM). There are differences in the complete end-to-end management software. Cisco includes a virtual switch for IP networks; HP and IBM utilize the one packaged with vSphere4.

This option has a lot to like. It's integrated at the factory. It's supported by the biggest names in the industry. Since their reputations are on the line, they are determined to make it work. Cisco and EMC have even put together a joint venture support organization. These packages also eliminate the majority of implementation heartburn, ongoing management issues, and vendor finger pointing.

Unfortunately, there is also quite a lot that is not so likeable. It's aimed at the very high-end of the market (equals expensive). It requires extensive professional services (even more expensive). It requires the integration of multiple hardware systems more commonly referred to as bolted together integration. Multiple hardware systems do nothing to reduce power and cooling. Worse, bolted together integration makes management considerable more challenging. The management interface is really multiple management systems with a thin GUI layer on top that provides the illusion of an integrated management system. The servers, SAN and storage all have unique management systems with a different look, logic, and feel. Some information is passed between these systems, albeit not much.

The factory-packaged integrated system is conceptually a great idea. Sadly, the current generation integration is only "skin deep". The marketing messaging is far more integrated than the technology. And it is financially out of reach for the vast majority of small to medium businesses and Enterprises.

This has left a pressing market need for an integrated approach to virtualization that brings the benefits of the technology to the mass-market and majority of users without overwhelming their IT with unintended consequences. Pivot3 has developed exactly that with a factory-integrated virtual system including servers, storage and networking that is specifically designed to simplify the deployment of virtualization for the small to medium business enterprise customer. It is called the Pivot3 Unified Virtual System or UVS.

The Pivot3 Unified Virtual System (UVS)

The UVS integrates virtualized server hypervisor (currently packaged with, but not limited to XenServer), Pivot3's market proven and award winning RAIGE® virtualized scale-out SAN storage, virtualized iSCSI SAN, plus virtualized NAS into one unified advanced Intel Nehalem dual quad core server hardware platform.

Figure 4:
Pivot3's All Encompassing UVS vs. Other Less Integrated More Costly Packages + Storage

Each Pivot3 UVS physical node comes with 12GB RAM, up to 16GB flash write cache, and up to 24 TB of capacity. Its scale-out architecture enables it to scale-out to 12 nodes, 144GB of DRAM; 120GB flash cache; up to 288TB of capacity; and 24 Gigabits of aggregated iSCSI bandwidth. A UVS node also supports up to 4 VMs, while providing dual Gbps iSCSI ports and dual Gbps TCP/IP network ports.

Every VM on the Pivot3 UVS system regardless of which node it resides, has the ability to access all UVS storage exactly the same as attaching to an external iSCSI SAN.

Pivot3 UVS has extensive reliability and uptime characteristics without requiring complex custom hardware. The IP SAN storage can tolerate up to 5 concurrent disk failures across 12 nodes, or a simultaneous complete node failure and 2 drive failures without losing any data. Global distributed sparing, predictive sparing, and parallel rebuilds are supported for maximum data protection across physical nodes. Server applications are similarly protected without requiring dedicated cluster hardware or licensing. In the event of a nodal failure the VMs will automatically failover and reboot on another available node in the array. Moreover, the Pivot3 UVS end-to-end management is built-in simplifying implementations, operations, and management.

How then does the Pivot3 UVS solve those pesky unintended virtualized server problems? Too much oversubscription is easily avoided because the storage, SAN, and VMs reside on the same hardware. The hardware is sized and integrated to prevent too much oversubscription. Cable mismanagement issues are avoided because many cables are virtual leaving far fewer cables to implement or manage. VM pathing is a complete non-issue since once again the internal VMs are automatically virtually pathed to their correct storage. Troubleshooting is simplified from both the reduction in actual physical parts and the fact that scaling of VMs, capacity, and performance is provided at an appliance level and is based on Gigabit Ethernet connectivity. The appliance model simplifies configuration and expansion with a simple, intuitive and pay as you grow model that can be dynamically changed without affecting server or storage accessibility.

One important ancillary benefit to the Pivot3 UVS is that to other physical or virtual servers not in the UVS, it looks and feels like an external iSCSI or NAS storage. This allows a Pivot3 UVS to be the networked storage for far more than the 48 VMs it can internalize.

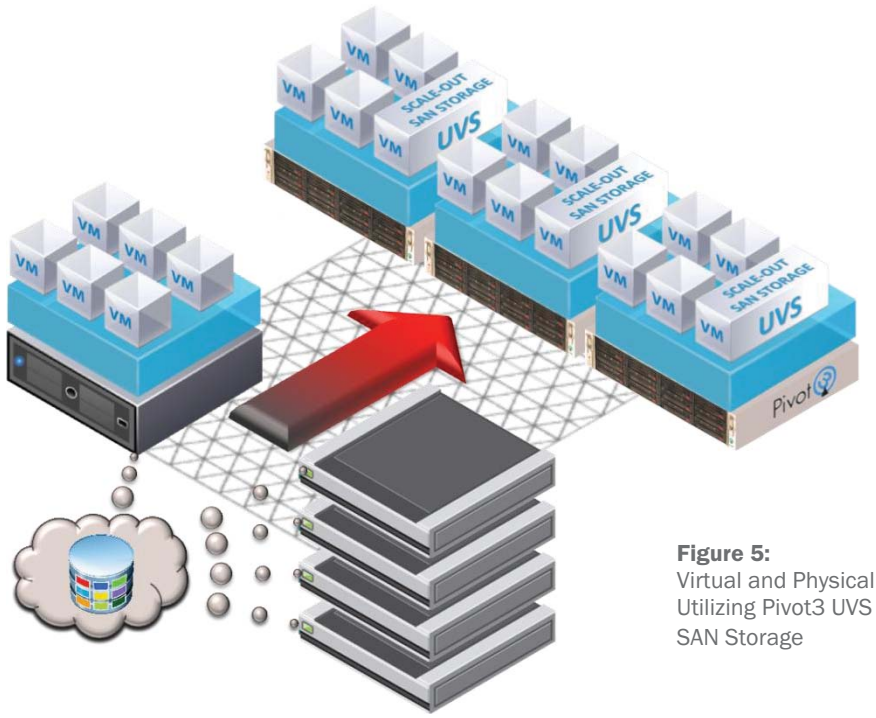

Figure 5:
Virtual and Physical Servers Utilizing Pivot3 UVS Scale-out SAN Storage

Another problem uniquely solved by the Pivot3 UVS vs. other factory packaged integrated systems is far more reduced power and cooling. The combination of Pivot3 UVS storage and server hardware into a single platform eliminates physical powered hardware to reduce power and cooling by as much as 40% over separately powered server and storage systems.

Best of all, the Pivot3 UVS has a much lower initial and total cost of ownership than equivalent discrete commodity systems that aren't integrated and don't solve the unintended problems of server virtualization.

So what's the catch? There has to be a catch. The catch is that Pivot3 UVS is not a brand new product launch to the market. UVS has been quietly proving itself in the extremely competitive surveillance market under the name Pivot3 Cloudbank™. There are now hundreds of installations and it has quietly become the dominant surveillance virtualized storage/server system. This is Pivot3 UVS' coming out party.

Conclusion

There are numerous exasperating unintended consequences of server virtualization. The common approaches of throwing expensive professional people or system integrators at these problems do not effectively solve them in a consistent and ongoing fashion. Whereas the current breed of factory packaged integrated solutions provides some relief, they are only partially effective and financially out of reach for most organizations.

Pivot3 UVS is a new approach to integrated virtualization that solves the most common pain points by collapsing the software stack so that virtual servers and storage are completely integrated. This unique storage-centric approach effectively alleviates all of those pesky aggravating virtualization unintended and troubling consequences. And it does so at a TCO that's lower than non-integrated solutions. Pivot3 UVS is ideal for today's SMB and SME looking to reap the benefits of virtualization and will help drive the use of virtualization beyond the Fortune 100 data center.

For more detailed information, please contact:
info@pivot3.com or go to www.pivot3.com

About the author: Marc Staimer is the founder, senior analyst, and CDS of Dragon Slayer Consulting in Beaverton, OR. The consulting practice of 12 years has focused in the areas of strategic planning, product development, and market development. With over 30 years of marketing, sales and business experience in infrastructure, storage, server, software, and virtualization, he's considered one of the industry's leading experts. Marc can be reached at ***marcstaimer@mac.com***.