Managing virtualization at a large scale is fraught with hidden challenges.

BY EVANGELOS KOTSOVINOS

Virtualization: Blessing or Curse?

VIRTUALIZATION IS OFTEN touted as the solution to many challenging problems, from resource underutilization to data-center optimization and carbon emission reduction. However, the hidden costs of virtualization, largely stemming from the complex and difficult system administration challenges it

poses, are often overlooked. Reaping the fruits of virtualization requires the enterprise to navigate scalability limitations, revamp traditional operational practices, manage performance, and achieve unprecedented cross-silo collaboration. Virtualization is not a curse: it can bring material benefits, but only to the prepared.

Al Goodman once said, "The perfect computer has been invented. You just feed in your problems and they never come out again." This is how virtualization has come to be perceived in recent years: as a panacea for a host of IT problems. Bringing virtualization into the enterprise is often about *reducing costs* without compromising quality of service. Running the same workloads as virtual machines (VMs) on fewer servers can improve server utilization and, perhaps more importantly, allow the

deferral of data-center build-outs—the same data-center space can now last longer.

Virtualization is also meant to enhance the *manageability* of the enterprise infrastructure. As virtual servers and desktops can be live-migrated with no downtime, coordinating hardware upgrades with users or negotiating work windows is no longer necessary—upgrades can happen at any time with no user impact. In addition, high availability and dynamic load-balancing solutions provided by virtualization product families can monitor and optimize the virtualized environment with little manual involvement. Supporting the same capabilities in a nonvirtualized world would require a large amount of operational effort.

Furthermore, enterprises use virtualization to provide IaaS (Infrastruc-

ture as a Service) cloud offerings that give users access to computing resources on demand in the form of VMs. This can improve developer productivity and *reduce time to market*, which is key in today's fast-moving business environment. Since rolling out an application sooner can provide first-mover advantage, virtualization can help boost the business.

The Practice

Although virtualization is a 50-year-old technology,[3] it reached broad popularity only as it became available for the x86 platform from 2001 onward—and most large enterprises have been using the technology for fewer than five years.[1,4] As such, it is a relatively new technology, which, unsurprisingly, carries a number of less-well-understood system administration challenges.

Old Assumptions. It is not, strictly speaking, virtualization's fault, but many systems in an enterprise infrastructure are built on the assumption of running on real, physical hardware. The design of operating systems is often based on the principle that the hard disk is local, and therefore reading from and writing to it is fast and low cost. Thus, they use the disk generously in a number of ways, such as caching, buffering, and logging. This, of course, is perfectly fair in a nonvirtualized world.

With virtualization added to the mix, many such assumptions are turned on their heads. VMs often use shared storage, instead of local disks, to take advantage of high availability and load-balancing solutions—a VM with its data on the local disk is a lot more difficult to migrate, and doomed if the local disk fails. With virtualization, each read and write operation travels to shared storage over the network or Fiber Channel, adding load to the network interface controllers (NICs), switches, and shared storage systems. In addition, as a result of consolidation, the network and storage infrastructure has to cope with a potentially much higher number of systems, compounding this effect. It will take years for the entire ecosystem to adapt fully to virtualization.

System Sprawl. Conventional wisdom has it that the operational workload of managing a virtualized server

running multiple VMs is similar to that of managing a physical, nonvirtualized server. Therefore, as dozens of VMs can run on one virtualized server, consolidation can reduce operational workload. Not so: the workload of managing a physical, nonvirtualized server is comparable to that of managing a VM, not the underlying virtualized server. The fruits of common, standardized management—such as centrally held configuration and image-based provisioning—have already been reaped by enterprises, as this is how they manage their physical environments. Therefore, managing 20 VMs that share a virtualized server requires the same amount of work as managing 20 physical servers. Add to that the overhead of managing the hypervisor and associated services, and it is easy to see that operational workload will be higher.

More importantly, there is evidence that virtualization leads to an increase in the number of systems—now running in VMs—instead of simply consolidating existing workloads.[2,5] Making it easy to get access to computing capacity in the form of a VM, as IaaS clouds do, has the side effect of leading to a proliferation of barely used VMs, since developers forget to return the VMs they do not use to the pool after the end of a project. As the number of VMs increases, so does the load placed on administrators and on shared infrastructure such as storage, Dynamic Host Configuration Protocol (DHCP), and boot servers.

Most enterprise users of virtualization implement their own VM reclamation systems. Some solutions are straightforward and borderline simplistic: if nobody has logged on for more than three months, then notify and subsequently reclaim if nobody objects. Some solutions are elaborate and carry the distinctive odor of over-engineering: analyze resource utilization over a period of time based on heuristics; determine level of usage; and act accordingly. Surprising as it may be there is a lack of generic and broadly applicable VM reclamation solutions to address sprawl challenges. In addition, services that are common to all VMs sharing a host—such as virus scanning, firewalls, and backups—should become part of the virtualization layer itself. This has already

started happening with such services entering the hypervisor, and it has the potential to reduce operational workload substantially.

Scale. Enterprises have spent years improving and streamlining their management tools and processes to handle scale. They have invested in a backbone of configuration management and provisioning systems, operational tools, and monitoring solutions that can handle building and managing tens or even hundreds of thousands of systems. Thanks to this—largely home-

grown—tooling, massively parallel operational tasks, such as the build-out of thousands of servers, daily operating system checkouts, and planned data-center power-downs, are routine and straightforward for operational teams.

Enter virtualization: most vendor solutions are not built for the large enterprise when it comes to scale, particularly with respect to their management frameworks. Their scale limitations are orders of magnitude below those of enterprise systems, often because of fundamental design flaws—such as overreliance on central components or data sources. In addition, they often do not scale out; running more instances of the vendor solution will not fully address the scaling issue, as the instances will not talk to each other. This chal-

lenge is not unique to virtualization. An enterprise faces similar issues when it introduces a new operating system to its environment. Scaling difficulties, however, are particularly important when it comes to virtualization for two reasons: first, virtualization increases the number of systems that must be managed, as discussed in the section on system sprawl; second, one of the main benefits of virtualization is central management of the infrastructure, which cannot be achieved without a suitably scalable management framework.

tion and size of operational teams.

Interoperability. Many enterprises have achieved a good level of integration between their backbone systems. The addition of a server in the configuration-management system allows it to get an IP address and host name. The tool that executes a power-down draws its data about what to power off seamlessly from the configuration-management system. A change in a server's configuration will automatically change the checkout logic applied to it. This uniformity and tight integra-

physical infrastructure.

To be sure, some enterprises are fortunate enough to have a homogeneous environment, managed by a product suite for which solid virtualization extensions already exist. In a heterogeneous infrastructure, however, with more than one virtualization platform, with virtualized and nonvirtualized parts, and with a multitude of tightly integrated homegrown systems, the introduction of virtualization leads to *administration islands*—parts of the infrastructure that are managed differ-

As a result, enterprises are left with a choice: either they live with a multitude of frameworks with which to manage the infrastructure, which increases operational complexity; or they must engineer their own solutions that work around those limitations—for example, the now open source Aquilon framework extending the Quattor toolkit (http://www.quattor.org). Another option is for enterprises to wait until the vendor ecosystem catches up with enterprise-scale requirements before they virtualize. The right answer depends on a number of factors, including the enterprise's size, business requirements, existing backbone of systems and tools, size of virtualized and virtualizable infrastructure, engineering capabilities, and sophistica-

tion massively simplifies operational and administrative work.

Virtualization often seems like an awkward guest in this tightly integrated enterprise environment. Each virtualization platform comes with its own APIs, ways of configuring, describing, and provisioning VMs, as well as its own management tooling. The vendor ecosystem is gradually catching up, providing increased integration between backbone services and virtualization management. Solutions are lacking, however, that fulfill all three of the following conditions:
▸ They can be relatively easily integrated with homegrown systems.
▸ They can handle multiple virtualization platforms.
▸ They can manage virtual as well as

ently from everything else. This breaks the integration and uniformity of the enterprise environment, and increases operational complexity.

Many enterprises will feel like they have been here before—for example, when they engineered their systems to be able to provision and manage multiple operating systems using the same frameworks. Once again, customers face the "build versus suffer" choice. Should they live with the added operational complexity of administration islands until standardization and convergence emerge in the marketplace, or should they invest in substantial engineering and integration work to ensure hypervisor agnosticism and integration with the existing backbone?

Troubleshooting. Contrary to con-

ventional wisdom, virtualized environments do not really consolidate three physical machines into one physical machine; they consolidate three physical machines onto several physical subsystems, including the shared server, the storage system, and the network.

Finding the cause of slowness in a physical computer is often a case of glancing at a few log files on the local disk and potentially investigating local hardware issues. The amount of data that needs to be looked at is relatively small, contained, and easily found. Monitoring performance and diagnosing a problem of a virtual desktop, on the other hand, requires trawling through logs and data from a number of sources including the desktop operating system, the hypervisor, the storage system, and the network.

In addition, this large volume of disparate data must be *aggregated or linked*; the administrator should be able to obtain information easily from all relevant systems for a given time period, or to trace the progress of a specific packet through the storage and network stack. Because of this multisource and multilayer obfuscation, resolution will be significantly slower if administrators have to look at several screens and manually identify bits of data and log files that are related, in terms of either time or causality. New paradigms are needed for storing, retrieving, and linking logs and performance data from multiple sources. Experience from fields such as Web search can be vital in this endeavor.

Silos? What Silos? In a nonvirtualized enterprise environment, responsibilities for running different parts of the infrastructure are neatly divided among operational teams, such as Unix, Windows, network, and storage operations. Each team has a clear scope of responsibility, communication among teams is limited, and apportioning credit, responsibility, and accountability for infrastructure issues is straightforward.

Virtualization bulldozes these silo walls. Operational issues that involve more than one operational team—and, in some cases, all—become far more common than issues that can be resolved entirely within a silo. As such, cross-silo collaboration and communication are of paramount importance,

Virtualization holds promise as a solution for many challenging problems. Expectations are running high. Can virtualization deliver?

requiring a true mentality shift in the way enterprise infrastructure organizations operate—as well as, potentially, organizational changes to adapt to this requirement.

Impact of Changes. Enterprises have spent a long time and invested substantial resources into understanding the impact of changes to different parts of the infrastructure. Change-management processes and policies are well oiled and time tested, ensuring that every change to the environment is assessed and its impact documented.

Once again, virtualization brings fundamental change. Sharing the infrastructure comes with centralization and, therefore, with potential bottlenecks that are not as well understood. Rolling out a new service pack that increases disk utilization by 5IOPS (input/output operations per second) on each host will have very little impact in a nonvirtualized environment—each host will be using its disk a little more often. In a virtualized environment, an increase of disk usage by 5IOPS per VM will result in an increase of 10,000IOPS on a storage system shared by 2,000 VMs, with potentially devastating consequences. It will also place increased load on the shared host, as more packets will have to travel through the hypervisor, as well as the network infrastructure. We have seen antivirus updates and operating-system patches resulting in increases in CPU utilization on the order of 40% across the virtualized plant—changes that would have a negligible effect when applied to physical systems.

Similarly, large-scale reboots can impact shared infrastructure components in ways that are radically different from the nonvirtualized past. Testing and change management processes need to change to account for effects that may be much broader than before.

Contention. Virtualization platforms do a decent job of isolating VMs on a shared physical host and managing resources on that host (such as CPU and memory). In a complex enterprise environment, however, this is only part of the picture. A large number of VMs will be sharing a network switch, and an even larger number of VMs will be sharing a storage system. Contention on those parts of the virtualized stack

can have as much impact as contention on a shared host, or more. Consider the case where a rogue VM overloads shared storage: hundreds or thousands of VMs will be slowed down.

Functionality that allows isolating and managing contention when it comes to networking and storage elements is only now reaching maturity and entering the mainstream virtualization scene. Designing a virtualization technology stack that can take advantage of such features requires engineering work and a good amount of networking and storage expertise on behalf of the enterprise customer. Some do that, combining exotic network adapters that provide the right cocktail of I/O virtualization in hardware with custom rack, storage, and network designs. Some opt for the riskier but easier route of doing nothing special, hoping that system administrators will cope with any contention issues as they arise.

GUIs. Graphical user interfaces work well when managing an email inbox, data folder, or even the desktop of a personal computer. In general, it is well understood in the human-computer interaction research community that GUIs work well for handling a relatively small number of elements. If that number gets large, GUIs can overload the user, which often results in poor decision making.[7] Agents and automation have been proposed as solutions to reduce information overload.[6]

Virtualization solutions tend to come with GUI-based management frameworks. That works well for managing 100 VMs, but it breaks down in an enterprise with 100,000 VMs. What is really needed is more intelligence and automation; if the storage of a virtualized server is disconnected, automatically reconnecting it is a lot more effective than displaying a little yellow triangle with an exclamation mark in a GUI that contains thousands of elements. What is also needed is interoperability with enterprise backbones and other systems, as mentioned previously.

In addition, administrators who are accustomed to the piecemeal systems management of the previrtualization era—managing a server here and a storage element there—will discover they will have to adapt. Virtualiza-

tion brings unprecedented integration and hard dependencies among components—a storage outage could mean that thousands of users cannot use their desktops. Enterprises need to ensure that their operational teams across all silos are comfortable with managing a massively interconnected large-scale system, rather than a collection of individual and independent components, without GUIs.

Conclusion

Virtualization holds promise as a solution for many challenging problems. It can help reduce infrastructure costs, delay data-center build-outs, improve our ability to respond to fast-moving business needs, allow a massive-scale infrastructure to be managed in a more flexible and automated way, and even help reduce carbon emissions. Expectations are running high.

Can virtualization deliver? It absolutely can, but not out of the box. For virtualization to deliver on its promise, both vendors and enterprises need to adapt in a number of ways. Vendors must place strategic emphasis on enterprise requirements for scale, ensuring that their products can gracefully handle managing hundreds of thousands or even millions of VMs. Public cloud service providers do this very successfully. Standardization, automation, and integration are key; eye-pleasing GUIs are less important. Solutions that help manage resource contention end to end, rather than only on the shared hosts themselves, will significantly simplify the adoption of virtualization. In addition, the industry's ecosystem needs to consider the fundamental redesign of components that perform suboptimally with virtualization, and it must provide better ways to collect, aggregate, and interpret logs and performance data from disparate sources.

Enterprises that decide to virtualize strategically and at a large scale need to be prepared for the substantial engineering investment that will be required to achieve the desired levels of scalability, interoperability, and operational uniformity. The alternative is increased operational complexity and cost. In addition, enterprises that are serious about virtualization need a way to break the old dividing lines, fos-

ter cross-silo collaboration, and instill an end-to-end mentality in their staff. Controls to prevent VM sprawl are key, and new processes and policies for change management are needed, as virtualization multiplies the effect of changes that would previously be of minimal impact.

Virtualization can bring significant benefits to the enterprise, but it can also bite the hand that feeds it. It is no curse, but, like luck, it favors the prepared.

Related articles on queue.acm.org

Beyond Server Consolidation
Werner Vogels
http://queue.acm.org/detail.cfm?id=1348590

CTO Roundtable: Virtualization
http://queue.acm.org/detail.cfm?id=1508219

The Cost of Virtualization
Ulrich Drepper
http://queue.acm.org/detail.cfm?id=1348591

References
1. Bailey, M., Eastwood, M., Gillen, A., Gupta, D. Server virtualization market forecast and analysis, 2005–2010. IDC, 2006.
2. Brodkin, J. Virtual server sprawl kills cost savings, experts warn. *NetworkWorld.* Dec. 5, 2008.
3. Goldberg, R.P. Survey of virtual machine research. *IEEE Computer Magazine 7*, 6 (1974), 34–45.
4. Humphreys, J. Worldwide virtual machine software 2005 vendor shares. IDC, 2005.
5. IDC. Virtualization market accelerates out of the recession as users adopt "Virtualize First" mentality; 2010.
6. Maes, P. Agents that reduce work and information overload. *Commun. ACM 37*, 7 (1994), 30–40.
7. Schwartz, B. *The Paradox of Choice.* HarperCollins, NY, 2005.

Evangelos Kotsovinos is a vice president at Morgan Stanley, where he leads virtualization and cloud-computing engineering. His areas of interest include massive-scale provisioning, predictive monitoring, scalable storage for virtualization, and operational tooling for efficiently managing a global cloud. He also serves as the chief strategy officer at Virtual Trip, an ecosystem of dynamic start-up companies, and is on the Board of Directors of NewCred Ltd. Previously, Kotsovinos was a senior research scientist at T-Labs, where he helped develop a cloud-computing R&D project into a VC-funded Internet start-up. A pioneer in the field of cloud computing, he led the XenoServers project, which produced one of the first cloud-computing blueprints.