insideHPC

InsideHPC Guide to

Virtualization, the Cloud and HPC

POWERING RESEARCH AGILITY

---

by John Kirkley,
Features Editor, insideHPC
August 2014

BROUGHT TO YOU BY vmware

Introduction

Over the past several years, virtualization has made major inroads into enterprise IT infrastructures. And now it is moving into the realm of high performance computing (HPC), especially for such compute intensive applications as electronic design automation (EDA), life sciences, financial services and digital media entertainment.

It took a while. Virtualization dates back to the 1960s when some of the earliest mainframes claimed center stage in the "glass house" data centers of the day. Back then the term meant dividing up mainframe's resources to handle various applications. However, it wasn't until the late 1990s when a group of Stanford University researchers figured out how to virtualize the x86 instruction set that virtualization was able to move from the mainframe to the mainstream. Those same researchers went on to form VMware and release the first virtualization product for the x86 ISA in 1999.

What is Virtualization?

Hardware virtualization refers to the creation of a number of self-contained virtual servers that are resident on the physical server, or host machine. Virtualization allows multiple applications to be run on the same machine while providing security and fault isolation. Typically an administrator decides how much of each resource — CPU, memory, networking — to allocate to the virtual machine (VM), while assigning priorities to different classes of users. The virtual infrastructure dynamically enforces these policies to ensure that each VM gets its fair share of resources.

Contents

Hardware abstraction provides the ease of use and agility that physical (bare metal) clusters lack.

Another key attribute of x86 virtualization important to HPC is hardware abstraction. With virtualization, there is no need to retool the IT infrastructure's OS and drivers every time new hardware is added to a cluster. The OS always sees the same virtual hardware regardless of the underlying physical hardware type. Hardware abstraction provides the ease of use and agility that physical (bare metal) clusters lack.

As virtualization became more prevalent, both Intel and AMD have continued to add more hardware support for the technology, including CPU, memory and I/O virtualization. These hardware advancements, coupled with increasingly sophisticated virtualization software, have resulted in improved performance for a wide variety of workloads, including HPC-based applications.

Server virtualization used in the enterprise data center for IT infrastructure workloads has yielded a number of benefits.

Some typical benefits include:

- Significantly increases server utilization
- Improves business agility by providing the additional computational resources that enable companies to bring products to market faster and be more competitive in global markets
- Provides significant reductions in CapEx and OpEx for each server virtualized
- Highly resilient in dealing with power outages, equipment failure or other disruptions
- Enhances security

Virtualizing Networking Services

With the development of software-defined networks (SDN) over the past decade, the principles of virtualization have been applied to network resources — abstracting, pooling and automating them to overcome the limits of rigid physical architecture.

Network services are assigned to each application and elastically adapt to meet the application's changing requirements. The choreography between the compute and network virtualization provides an unparalleled degree of automation, security, flexibility and agility.

VMware's approach uses "bare metal" virtualization — the hypervisor interfaces directly with the computer hardware without the need for a host operating system — making the VMware platform one of the most secure in the industry.

Security in a Virtualized Environment

Security has turned out to be one of the major benefits associated with virtualization. It is delivered as a software-defined service that is decoupled from physical devices, aggregated, and targeted to wherever it is needed without requiring a hardware upgrade. This allows workloads to be scaled and moved without security constraints or the need for specialized appliances.

VMware's approach uses "bare metal" virtualization — the hypervisor interfaces directly with the computer hardware without the need for a host operating system. This safeguards the virtualization software from OS-related vulnerabilities, making the VMware platform one of the most secure in the industry.

Integrated firewalling and optimized gateway services protect the network edge. And VMware's open architecture allows data center managers to integrate existing security investments and the latest third party innovations.

Challenges on the Road to HPC Virtualization

Although both the enterprise and HPC can benefit from virtualization, the two have had dissimilar requirements. Until recently, typical enterprise applications were not resource intensive; it was the HPC crowd running complex modeling and simulation applications that required ever larger supercomputers and more powerful clusters.

For HPC environments — unlike the enterprise — massive consolidation of resources is not an option. In fact, the opposite is true: HPC users are always looking to add more hardware so they can solve bigger problems more quickly as they push the envelope of scientific and engineering knowledge.

Traditional HPC clusters run a single, standard OS (often a flavor of Linux) and software stack across all nodes. By operating in a uniform, homogenous cluster environment, the job scheduler is free to place jobs anywhere on the cluster as long as the target node is not overloaded with other jobs. This uniformity allows data center managers to maintain one image and distribute it to all the nodes at system set up and then occasionally for maintenance and updates. However, this approach limits the flexibility of their computational resources, especially when trying to accommodate multiple user populations.

For example, individual researchers or engineers may require specific software stacks to run their applications. If this stack is not compatible with the HPC cluster's standard OS, too often the result is an unhappy user and a beleaguered IT organization. One of the side effects of this kind of IT inflexibility is the user creation of separate "islands of compute" scattered across the organization.

This is an inefficient and expensive solution. It also adds to the complexity of cluster management, especially for users new to the world of HPC and those considering migrating from desktop systems to a lower end cluster priced below $250,000. Typically these users do not have in-house HPC experts that they can turn to when problems inevitably arise.

Many Jobs, One OS

In bare metal environments, running multiple user jobs within the same OS can cause more problems over and above data loss or leakage. If a job disrupts the OS by crashing a daemon or other component, saving excessive files to the hard drive, or some other malfunction, other unrelated jobs can be impacted and schedules disrupted.

Because the power of virtualizing the compute, network, and storage infrastructure unlocks the power of automation, there is less likelihood that a manual error or neglect will lead to security breaches or compliance lapses.

Bare metal environments have additional inefficiencies. Consider this scenario relative to the placement of jobs. Several jobs are running on the cluster when a higher priority job is scheduled but no appropriate resources are available. The IT admin can either make the new job wait in queue which, given its status, is not a viable solution, or kill other jobs to run the newcomer — also not a very satisfactory state of affairs. Either solution reduces the cluster's throughput. Also, killing jobs can be quite expensive if they are costly ISV applications — for example, EDA applications licenses can cost hundreds of thousands of dollars.

Life sciences is another sector where a lack of virtualization can cause problems. In fact, bare metal environments can be created that include HIPAA and FISMA (Federal Information Security Management Act) compliance. The virtualized environment provides the requisite compliance along with the benefits of cost savings, agility, self-provisioning, and fault and security separation discussed above. And, because the power of virtualizing the compute, network, and storage infrastructure unlocks the power of automation, there is less likelihood that a manual error or neglect will lead to security breaches or compliance lapses.

Virtualization and the Secure Private Cloud

Virtualization has been proven to be a viable architectural approach that addresses the many challenges mentioned above.

By creating a virtualized infrastructure, the IT organization ensures that:

- Departments, principal investigators and other key stakeholders receive the HPC resources they need when they need them
- Clusters and cluster nodes can be sized to meet specific application requirements
- Different operating systems and software stacks can be hosted simultaneously on the same infrastructure and adjusted dynamically
- IT can make efficient use of the underlying host hardware even though individual user jobs may only require a small number of CPUs
- Hardware can be shared while providing fault and security separation between users
- Policies can be enacted that allow high priority jobs to receive a higher "fairshare" of the underlying resources

For HPC environments, wrapping the virtualization infrastructure in a secure private cloud provides the most value to both the end users and the IT organization.

This approach enables self-provisioning, allowing researchers and engineers to instantiate the resources they need for a particular project without waiting for IT to create the resource for them. To instantiate a virtual HPC cluster, the user applies a previously defined blueprint that specifies the required virtual machine (VM) attributes, the number of machines involved, and the needed software — including the operating system and middleware. Users can fully customize the VMs to meet their specifications. The blueprint also allows the centralized IT organization to enforce corporate IT mandates — for example, security and data protection policies.

This solution assumes that end users will continue to run their familiar HPC batch schedulers within their virtual HPC clusters. At the same time, the VMware Distributed Resource Scheduler (DRS) and other components are able to dynamically manage the placement and priority of virtual machines on the underlying physical resources.

What the end user sees is an HPC cluster that looks just like a standard bare metal cluster running a standard job scheduler — there is no indication that they are interacting with virtual machines. This allows multiple engineering or research clusters to be instantiated on the same physical infrastructure — all available through a private cloud.

Underneath it all, virtualization is handling load balancing, protection, network services and all the other fundamentals that allow for multi-tenancy on the physical hardware while still delivering high performance. Cloud automation provides policy-based governance and logical application modeling to make sure that multi-vendor, multi-cloud services are delivered at the right size and service level for the task that needs to be performed.

Virtualization and cloud automation are fundamental attributes of a software defined data center, which allows IT to create private clouds that deliver agility and economies of scale while maintaining data sovereignty and governance.

Protecting Applications

Virtualization allows the adoption of advanced resiliency practices such as using telemetry from the underlying system to predict impending hardware failures and then proactively migrating the workload to another host to avoid application interruption. For example, the system would detect a potential fan failure or an increase in the rate of soft memory errors and take action to make sure the workload continues despite incipient system problems.

This approach should also reduce the need for frequent checkpointing and restoration, resulting in increased overall job throughput.

Separate Workloads Means Security

Virtualization allows workloads to be compartmentalized in their own VM in order to take full advantage of the underlying parallelism of today's multicore, heterogeneous HPC systems without compromising security. This approach is particularly beneficial for organizations centralizing multiple groups on to a shared cluster or for teams with security issues — for example, a life sciences environment where access to genomic data needs to be restricted to specific researchers.

VM abstraction provides a security separation between workloads that is not available in traditional HPC environments.

Some government mandates require research organizations or companies (e.g. pharmaceutical companies) to store test results for years. Archiving a VM is an easy way to record and save the precise software environment used for the trials. The same holds true for academic and research organizations that are concerned about reproducing their scientific efforts or responding to any subsequent audits.

Virtualization permits the live migration of workloads when there is contention in the cluster.

Workload Agility

Virtualization permits the live migration of workloads when there is contention in the cluster. For example, it is often the case when several jobs are running that another job will be introduced that gobbles up more memory than anticipated. This situation crops up frequently in EDA (electronic design automation) work. In a bare metal environment, when the new job starts to consume all available memory there are only two somewhat unsatisfactory alternatives — either let the jobs continue to run very slowly, or manually intervene to kill and restart jobs to untangle the mess.

In a virtual environment, when a newly introduced job starts on a memory rampage, DRS can be used to move the offending workload to another, less loaded physical host, allowing the total environment to work as efficiently as before.

Multi-tenancy with Resource Guarantees

DRS, mentioned above, has other benefits as well. In a virtualized cluster, it can be used to enforce the fair sharing of resources between the workloads of multiple groups. This takes place below the OS layer within the virtual platform. It allows HPC environments to move beyond the less structured, trust-based approach typically found in traditional HPC environments where multiple user jobs may be scheduled on the same OS instance.

When users can be given guaranteed access to their share of compute resources, they are more likely to contribute their physical resources to a common pool.

Instead, the VM approach provides guaranteed resources to specific groups or departments as needed.

It also addresses another problem — when users with tight budgets and aggressive deadlines are able to add more computing power to the cluster, they are typically reluctant to share those resources. So other users in need of more computing power must add their own resources to the cluster and CAPEX starts to escalate. Cloud computing, which enables automated self-provisioning and policy-based resource sharing, can help. When users can be given guaranteed access to their share of compute resources, they are more likely to contribute their physical resources to a common pool. For IT this has CAPEX implications — the need to add more computing power to the cluster can be postponed or eliminated.

Exploring Accelerators

Early VMware testing of GPUs for computation (GPGPUs) in a virtual environment indicate that performance is generally within 2% of that achievable in a native, bare metal environment. Tests using the Intel Xeon Phi coprocessor show the same close alignment with native performance. Several VMware customers are exploring the use of GPUs and coprocessors in a virtualized environment for HPC.

Software Defined Data Center (SDDC)

Managing, provisioning and monitoring very high scale workloads and large datasets has been a staple of HPC environments for some time. Now, with the advent of Big Data and the advanced computational capabilities and affordability of today's hardware and software, enterprise data centers are also dealing with the problems and opportunities that this deluge of data brings with it.

By leveraging the benefits of virtualization and the power of private, public and hybrid clouds, IT departments in both the enterprise and HPC are transforming themselves into providers of access to IT as a service (ITaaS)

The software defined data center is the underlying data center architecture that allows most of an IT's compute, storage, networking, security and availability infrastructure to be abstracted and defined in software and to function as enterprise-wide resources. In an SDDC, policy-driven automation supports the provisioning and ongoing management of compute, storage and network services. This approach enables ITaaS to be delivered in a virtualized environment that allows IT to deploy and deliver applications with greater agility, speed and quality of service.

Virtualization in the Cloud

One of the best ways to realize the full performance benefits of virtualization is to make it available through a private cloud. The VMware vCloud Suite realizes operational efficiency through policy-driven operations. By providing simplified operations management, the cloud solution drives greater resource utilization and staff productivity.

Cloud-based automatic provisioning enables on-demand deployment of IT services with full control over critical IT and business policies. Infrastructure resources can be continually modified to meet workload requirements and business needs. This leads to greater user satisfaction and reduces the risk of islands of compute being established outside the centralized IT organization.

vCloud Suite also delivers the highest levels of application uptime while providing IT with greater control over the cost of accessing and provisioning IT services.

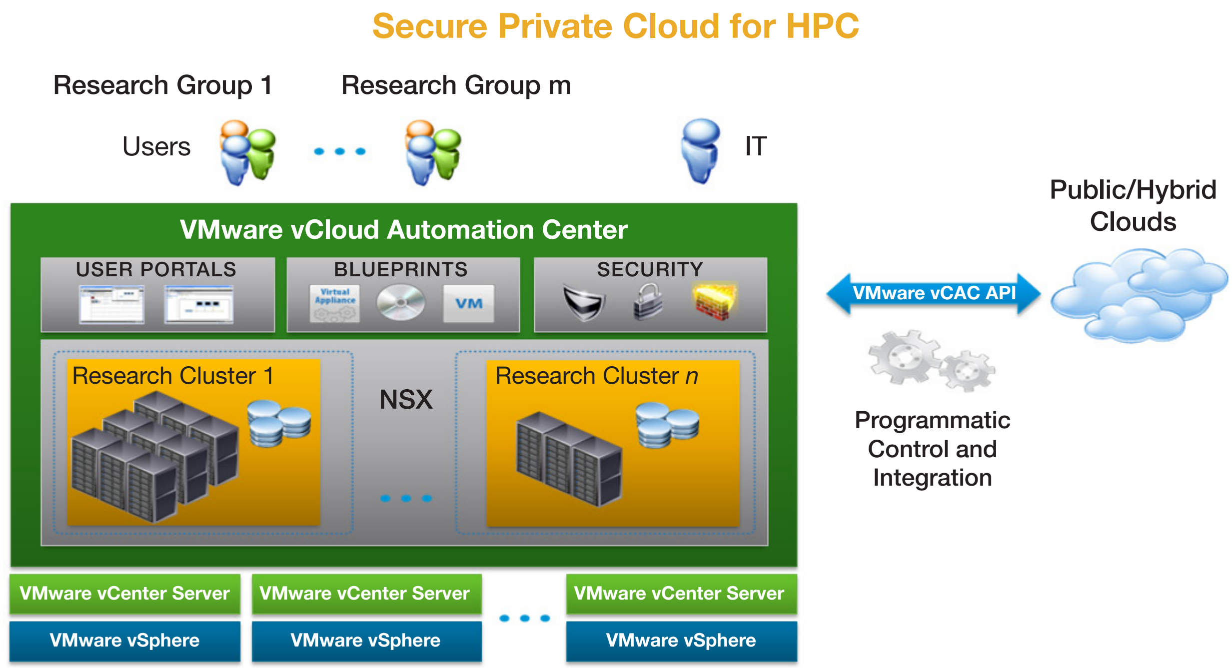

Figure 1: *Secure private cloud for HPC*

Convergence of HPC and the Enterprise

The HPC community has had a long-standing interest in creating scale-out environments for running throughput-oriented and parallel distributed workloads. Both large-scale environments (for example, cloud computing facilities) and scale-out workloads (such as Big Data) are becoming more important in the enterprise. In fact, with the rise of Big Data, the advent of affordable, powerful clusters, and strategies that take advantage of commodity systems for scale-out applications, these days the enterprise computing environment is looking a lot like HPC.

Both the enterprise and the HPC communities share a number of common concerns and requirements:

- Scale-out management
- Reduced power and cooling costs and lower data center density
- Need for dynamic resource management
- Need for high utilization, high availability
- Parallelization for multicore and heterogeneous HPC systems
- Big Data analytics
- Application resiliency
- Low latency communications
- Interest in cloud computing

VMware's approach to HPC and enterprise virtualization adds a level of flexibility and agility that cannot be achieved in bare metal environments.

The combination of virtualization and cloud computing provides value to both the end users and IT providers in HPC and enterprise environments. And, once created, these private clouds can be burst to a hybrid cloud to create seamless and secure extensions of the organization's on-premise infrastructure.

Performance is the key. VMware's approach to HPC and enterprise virtualization adds a level of flexibility and agility that cannot be achieved in bare metal environments.

The HPC community can realize significant benefits from adopting enterprise-capable IT solutions grounded in proven virtualization and cloud technology. And conversely, as business IT environments become increasingly compute-intensive, lessons learned by the scientists and engineers working with HPC can be transferred to their counterparts in the enterprise. It's a win-win situation.