Use of VMware for providing cloud infrastructure for the Grid

Robin Long\textsuperscript{1}, Matthew Storey\textsuperscript{2}

\textsuperscript{1}Physics Department Lancaster University, Lancashire, U.K., LA1 4YB;
\textsuperscript{2}ISS, Lancaster University, Lancashire, U.K., LA1 4WA

E-mail: r.long@lancs.ac.uk, m.storey@lancs.ac.uk

Abstract. The need to maximise computing resources whilst maintaining versatile setups leads to the need for flexible on demand facilities through the use of cloud computing. GridPP is currently investigating the role that Cloud Computing, in the form of Virtual Machines, can play in supporting Particle Physics analyses. As part of this research we look at the ability of VMware’s ESXi hyper-visors\cite{6} to provide such an infrastructure through the use of Virtual Machines (VMs); the advantages of such systems and their potential performance compared to physical environments.

1. Introduction

Cloud technology comes in many forms; The main forms are “Software as a Service”, whereby users are provided with a pool of computers on which they can run pre-installed software; “Platform as a Service”, in which a computing platform where software can be developed, compiled and run is provided; The final type is “Infrastructure as a Service”, this is the most basic setup, users are provided with a piece of hardware or a hyper-visor on which Operating Systems can be installed or Virtual Machines can be deployed.

The Grid already provides cloud computing in the form of “Software as a Service” and “Platform as a Service”. There is a desire to make more efficient and opportunistic use of resources as the needs placed on the Grid grow. An example of such resource use comes from the High Level Trigger (HLT). The HLT is a massive computing farm dedicated to filtering events from the ATLAS (or other experiments) triggers. In periods of no data taking, this farm is sat idle for anywhere from a few hours to years. Being able to rapidly provision new VMs onto this hardware that could convert it into a Grid site, and then re-provisioning it with new HLT VMs so that it could act as a Trigger farm again would provide a very efficient use of resources.

To be able to do this relies heavily on virtualisation technologies. This paper therefore examines the ability of VMware to provide such a virtualisation infrastructure, and how virtualisation affects performance.

2. Virtualisation

Virtualisation has been widely adopted by business to enable multiple operating systems often of differing types (Linux / Windows) to leverage a single piece of hardware. This has led to large scale reduction of hardware both in terms of servers but also in the connecting infrastructure needed to support such hardware, monitors, keyboards, racks, air-conditioning to name but a
few. The ability to run many virtual machines on a single host is attractive in business to save money through hardware but also savings in software purchases. The rapid provisioning of machines due to the removal of the need to buy a physical machine for each operating system instance is also very attractive.

There are several virtualisation platforms available and perhaps the most common two are VMware ESXi and Microsoft Hyper-V[7]. With Microsoft Hyper-V being the platform used in the Azure cloud offering[8], whilst VMware offer cloud services based on their own offering[9], these environments are now moving to enable seamless migration of workloads into the cloud should that be a need, and are being tested and improved constantly in real world environments with mixed loads on them. Both offer the ability to host multiple client operating systems on a single physical piece of hardware. This abstraction has led to the ability to perform operations such as live migration of virtual machines between physical hosts (assuming both hosts have the same vendor class[2]). This allows a workload, from the users perspective, to be active and seamlessly migrated it from one piece of hardware to another. This ability minimises downtime due to hardware failure and maximises purchasing options allowing for seamless migration to newer environments with faster / different hardware in them.

The ability to rapidly deploy through common and published interfaces even between hosts at different physical sites is very attractive in essence removing the tie between operating systems and the underlying hardware for a minimal impact of the virtualisation layer. Virtualisation also allows workload segregation and accounting charge back, or show back of resources used, enables scaling and planning for the future.

VMware is the platform of use at Lancaster for business and research use and allows high density and OS flexibility with management tools allowing the management of hundreds of operating systems from a single administrative interface (VMware vCenter).

3. VMware
VMware ESXi is a Linux-based native hyper-visor created by VMware. Each instances of VMware ESXi is administered from the VMware vSphere client. The client allows full control of each ESXi server. VMware vCenter Server is the centralised management tool for VMware. vCenter allows for the administration of multiple VMs and ESXi servers. An installation of vCenter is needed to perform certain tasks which include migration of VMs from one ESXi to another and cloning VMs. vCenter is also controlled from the vSphere client, or via the web client, CLI or VMware APIs.

3.1. Our setup
Our VMware setup consisted of one server running an instance of ESXi (version 5.1) with a Windows guest installed that has the vCenter (version 5.1) server on it, and another ESXi server used for the testing. These were controlled from a vSphere client (version 5.1) on a local machine.

The test server is a SGI Altix XE310 consisting of two sockets with an Intel Xeon CPU E5440, 16GB of RAM and a single Seagate ST3250620NS hard drive.

Converting one of our servers to an ESXi hyper-visor was simple. The installation of the hyper-visor took up roughly 150MB. There was then an additional overhead of RAM for each guest that depends on the number of virtual cores and devices assigned to it[1].

4. Analysis
We wanted to determine what effects virtualisation would have on the performance of a physical machine, and if any of these scaled with the number of machines. To do this we used standard tools designed to benchmark critical parts of a system, namely CPU, network and disk access. The tools used were:
Figure 1. Performance of (a) HEP-SPEC and (b) IPerf for the various machine configurations. Note that neither of these plots start at zero so that the differences can be seen more clearly.

- iperf-2.0.5-9
- bonnie++-1.96-2
- hep-spec06 (v1.1)

These tools were then used on several configurations: Bare-metal with all resources, 1 VM with all resources and then on 2, 4 and 8 VMs with the resources split equally between each guest.

These setups allowed us to compare the performance and scalability of VMware. In the multiple VM setups, the tests were run concurrently on each virtual machine.

4.1. HEP-SPEC
HEP-SPEC06[3] is the HEP-wide benchmark for measuring CPU performance and is based on the SPEC CPU2006 benchmark suite. The HEP-SPEC tests were run three times for each configuration, where more than one guest is running on the host, the scores are summed together to give a total HEPSPEC for the machine.

Figure 1(a) shows that when 1 virtual machine is assigned all of the available resources then the HEPSPEC score drops by about 2 points. This drop is slowly reversed as the resources are divided up between more guests. Once their is one core per guest the drop is about 0.2 points.

This effect is because of the management overhead of the hypervisor management agents is most visible when a single VM is using all the cores. As the cores are split between more VMs the hyper-visor is able to compensate for this management overhead leading to an increase in HEPSPEC as number of VMs increases.

4.2. Bonnie++
Disk access speed is very important in a High End Computing environment. We used bonnie++[5] to test the disk access speeds; Bonnie++ was set to write to a file four times the size of the RAM available to each VM.
Figure 2. Performance of bonnie’s write (white), rewrite (grey) and read (black) tests for the various machine configurations.

Figure 2 shows that no significant effect is seen on the write rates, although there is some fluctuation in values. It can also be seen from figure 2 that the re-write rate drops when going to more than one guest on a host. Disk access can be seen to improve in the virtual environment when moving to 4 and 8 guests. Whilst the test files were greater than the RAM available to the VMs, they were not bigger than the RAM available to the hypervisor, as such we believe efficiencies in the hypervisor are causing a bias in these results.

4.3. IPerf
The separation of CPU from storage disk on grid sites means that network bandwidth is very important. We tested this on the physical and virtual machines using IPerf[^4]. The physical machines contained Intel 80003ES2LAN Gigabit ethernet cards. As these test machines are re-purposed production machines, and have not been moved from our production room, all tests were done on a non-dedicated network. Figure 1(b) shows that bandwidth is not obviously effected by virtualisation. Any changes in value fluctuate a lot between configurations and seem to be due to contention on the network infrastructure rather than to do with the virtualisation process.

5. Conclusion
VMware appears to perform very well in the tests. Whilst a significant hit to the CPU is seen when virtualising the whole system, the hyper-visior is able to compensate for this when more VMs (with less cores are added). The IPerf results suggest that there is little effect from virtualising the NIC, although further tests need to be performed to be performed to be confident that the fluctuations are from network traffic. Bonnie demonstrated the largest difference with a significant drop in re-write performance, and some odd results for read access; however such results are non-representative of a HPC environment and should be viewed lightly.

The overhead from the hyper-visior proves to be small, our analysis suggests that making the VMs as small as possible decreases this overhead compared to larger VMs. In the example of the HLT which can vary from months to hours of idle time, we expect the overhead of utilising
a hyper-visor approach to be minimal in normal operation and provide massive flexibility and enable complete utilisation with minimal administrative overhead.

Overall VMware performed very well and should be monitored closely with advances in disk access management allowing potential performance improvements in the just released 5.5 version of ESXi.

6. Acknowledgements
We would like to thank VMWare for granting us permission to publish this paper.

References
[1] vSphere 4.1 - ESX and vCenter, Overhead Memory on Virtual Machines. (Accessed 03/02/2013);
   http://pubs.vmware.com/vsphere-4-esx-vcenter/index.jsp?topic=/com.vmware.vsphere
   resourcemanagement.doc_41/managing_memory_resources/r_overhead_memory_on_virtual_machines.html
[2] ESXi and vCenter Server 5.5 Documentation, CPU Compatibility and EVC (Accessed 03/02/2013);
   https://pubs.vmware.com/vsphere-55/index.jsp?topic=%2Fcom.vmware.vsphere.vcenterhost.doc%2FGUID-
   03E7E5F9-06D9-463F-A64F-D4EC20DAF22E.html
[3] HEP-SPEC06 Benchmark(Accessed 03/02/2013);
   https://w3.hepix.org/benchmarks/doku.php
[4] IPerf - TCP and UDP bandwidth performance measurement tool (Accessed 03/02/2013);
   http://code.google.com/p/iperf/
[5] Bonnie++ homepage (Accessed 03/02/2013);
   http://www.coker.com.au/bonnie++/
[6] vSphere Hypervisor (Accessed 03/02/2013);
   http://www.vmware.com/uk/products/vsphere-hypervisor/
[7] Magic Quadrant for x86 Server Virtualization Infrastructure (Accessed 03/02/2013);
   http://www.gartner.com/technology/reprints.do?id=1-1GJA88J&ct=130628&st=eb
[8] Windows Azure (Accessed 03/02/2013);
   http://www.windowsazure.com/en-us/documentation/
[9] vCloud (Accessed 03/02/2013);
   http://vcloud.vmware.com/