

VMware vSphere™ Sets New World Record TPC-H with the ParAccel Analytic Database™

The First Fully Virtualized and Audited TPC-H on vSphere

VMware's ability to run this industry standard and widely-accepted benchmark leaves no question as to VMware vSphere's ability to process even the most challenging enterprise workloads. ParAccel Analytic Database (PADB) running on VMware vSphere achieved the fastest 1TB TPC-H score on record, with 7.7 times better price/performance than the prior performance record holder.¹ Furthermore, this combination beat the previous TPC-H record using 37% fewer servers and provided a database load time 8.7 times faster than the previous performance record holder.

VMware's sponsorship of this benchmark was not only record-setting, but historic as well, being the first ever fully-audited and virtualized TPC-H on vSphere. VMware vSphere executed flawlessly under the stringent TPC run rules in the presence of a certified TPC-H auditor.

The Perfect Workload to be Virtualized

Analytic workloads are transient in nature, seeing heavy resource utilization rates — for example, when performing financial reporting at the end of month, end of quarter, or end of year. After these business events have completed, resource consumption can drop substantially. In the world of physical hardware, analytic environments must be sized to service peak workloads while, much of the time, these expensive big-iron systems will be substantially underutilized. So, sizing analytic workloads on physical hardware effectively locks in compute resources, which cannot be freed up when not in use.

Using VMware vSphere's workload management and cloning capabilities, resources can be directed to service peak analytic activities in virtual environments and then redeployed or freed to service other applications, after peak business events or cycles have completed.

The Transaction Processing Council and TPC-H Benchmark

The TPC is a non-profit corporation which supports a consortium of hardware and database software vendors devoted to defining transaction processing and database-related benchmarks. The primary goal behind TPC benchmarks is the definition of functional requirements that can be run on any database, regardless of the hardware or operating system.

The TPC Benchmark™ H (TPC-H) is a decision support benchmark. It consists of a suite of business-oriented ad hoc queries and concurrent data modifications. The queries, and the data populating the database, have been chosen to have broad industry-wide relevance. This benchmark illustrates decision support systems that examine large volumes of data, execute queries with a high degree of complexity, and provide answers to critical business questions.

Test Configuration

The diagram in Figure 1 shows the hardware selected to run the ParAccel Analytic Database (PADB). The Leader and Compute nodes were configured on forty Hewlett Packard DL360 G6 servers, each equipped with two Intel Xeon x5560 quad-core 2.8 GHz processors, 72 GB of RAM, two PCIe Dual-port Gigabit Server Adapters, and eight 300GB internal drives. The forty servers were connected to four Cisco 3750 48-port switches.

The Hewlett Packard servers ran VMware ESXi™, which hosted the virtual machines. Two virtual machines were created on each of the servers. Each virtual machine had 8 vCPUs, 32 GB of memory, and 4 virtual NICs. All of the virtual machines ran 64-bit Red Hat Enterprise Linux. There were two PADB v2.5 virtual machine compute nodes running on each ESXi host. The only exception was that one ESXi host also housed a virtual machine leader node responsible for communications external to the PADB private network.

¹ As of 4/12/2010: HP BladeSystem c-Class 128P RAC with Oracle Database 11g Release 2 Enterprise Edition; 1,166,976 QphH@1000GB, 5.42 USD per QphH@1000GB, available on 09/01/09.

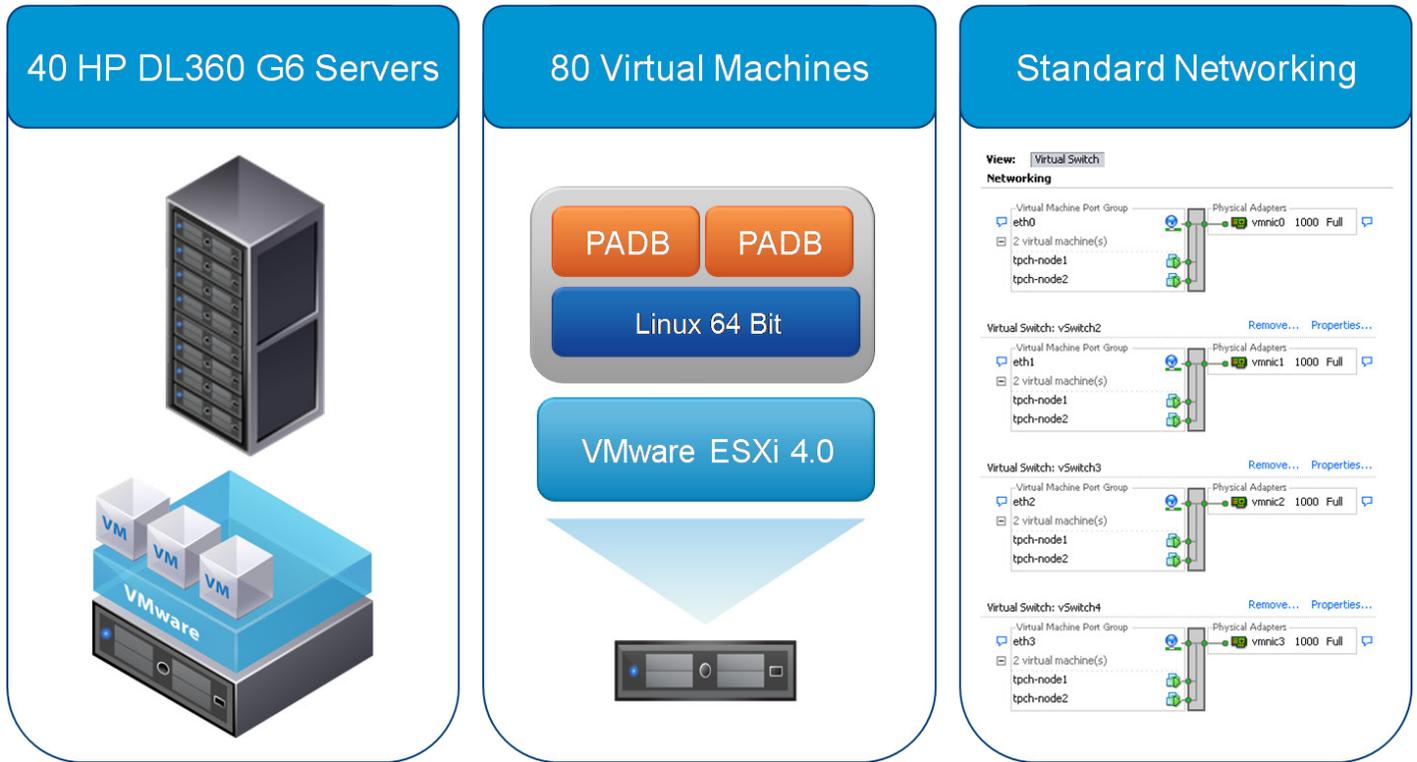


Figure 1. ParAccel Analytic Database on VMware vSphere

Faster Than Physical

Barry Zane, Chief Technology Officer of ParAccel

“We set up each server with two virtual machines and configured each virtual machine so they used half the memory of the physical server, four of the eight cores, and half of the disk I/O.

...We took measurements a couple of times and confirmed that the virtual machine configuration was actually faster than the physical server environment; the reason appears to be that VMware does a better job of mapping memory to CPU resources than Linux does. Also, VMware is more NUMA-aware than Linux is; the virtual CPUs bound themselves to the physical CPUs and to the memory connected to them.”

Conclusion

VMware sponsored the TPC-H benchmark and, along with the ParAccel Analytic Database, produced a world record for the price/performance metric — which proves even the most challenging workloads can be successfully virtualized on vSphere. The benchmark also clearly established VMware vSphere as an excellent platform for running analytic applications.

BENCHMARK PERFORMANCE HIGHLIGHTS	
Database Size	1,000GB
Composite Query Per Hour Metric	1,316.882 QphH
Price/Performance (\$/QphH)	\$.89 USD
Database Load Time	16 Minutes, 23 Seconds

The benchmark executive summary and full disclosure reports are available online at www.paracel.com or www.tpc.org. The TPC Benchmark™ H (TPC-H) specification is available at www.tpc.org. TPC and TPC-H are trademarks of the Transaction Processing Performance Council (TPC).

The performance metric reported by TPC-H is called the TPC-H Composite Query-per-Hour Performance Metric (QphH@Size), and reflects multiple aspects of the capability of a system to process queries. These aspects include the selected database size against which the queries are executed, the query processing power when queries are submitted by a single stream, and the query throughput when queries are submitted by multiple concurrent users. The TPC-H Price/Performance metric is expressed as \$/QphH@Size.

