TPC State of the Council 2013

Raghunath Nambiar[1], Meikel Poess[2], Andrew Masland[3], H. Reza Taheri[4], Andrew Bond[5], Forrest Carman[6], and Michael Majdalany[7]

[1] Cisco Systems, Inc., 3800 Zanker Road, San Jose, CA 95134, USA
rnambiar@cisco.com
[2] Oracle Corporation, 500 Oracle Pkwy, Redwood Shores, CA 94065, USA
meikel.poess@oracle.com
[3] NEC Corporation of America, 14335 NE 24th Street, Bellevue, WA 98007, USA
andy.masland@necam.com
[4] VMware, Inc., 3401 Hillview Ave, Palo Alto CA 94304, USA
rtaheri@vmware.com
[5] Red Hat, 100 East Davie Street, Raleigh, NC 27601, USA
abond@redhat.com
[6] Owen Media, 3130 E. Madison St., Suite 206, Seattle, WA 98112, USA
forrestc@owenmedia.com
[7] LoBue & Majdalany Mgmt Group, 572B Ruger St. San Francisco, CA 94129, USA
majdalany@lm-mgmt.com

Abstract. The TPC has played, and continues to play, a crucial role in providing the computer industry and its customers with relevant standards for total system performance, price-performance, and energy efficiency comparisons. Historically known for database-centric standards, the TPC is now developing benchmark standards for consolidation using virtualization technologies and multi-source data integration. The organization is also exploring new ideas such as Big Data and Big Data Analytics as well as an Express benchmark model to keep pace with rapidly changing industry demands. This paper gives a high level overview of the current state of the TPC in terms of existing standards, standards under development and future outlook.

Keywords: Industry Standard Benchmarks, Transaction Processing Performance Council.

1 TPC a Look Back and a Look Ahead

System benchmarks have played, and continue to play, a crucial role in the advancement of the computing industry. Existing system benchmarks are critical to both buyers and vendors. Buyers use benchmark results when evaluating new systems in terms of performance, price/performance, and energy efficiency, while vendors use benchmarks to demonstrate the competitiveness of their products and to monitor release-to-release progress of their products under development. With no standard system benchmarks available for Big Data systems, today's situation is similar to that of the mid-1980s, when the lack of standard database benchmarks led

R. Nambiar and M. Poess (Eds.): TPCTC 2013, LNCS 8391, pp. 1–15, 2014.

many system vendors to practice what is now referred to as “benchmarketing,” a practice in which organizations make performance claims based on self-designed, highly biased benchmarks. The goal of publishing results from such tailored benchmarks was to state marketing claims, regardless of the absence of relevant and verifiable technical merit. In essence, these benchmarks were designed as forgone conclusions to fit a pre-established marketing message. Similarly, vendors would create configurations, referred to as “benchmark specials," that were specifically designed to maximize performance against a specific benchmark with limited benefit to real-world applications. The TPC was founded to address these issues and it continues to do so today. To keep up with rapid changes in the industry, the TPC introduced its annual international conference series on performance evaluation and benchmarking (TPCTC) in 2009.

2 TPC Benchmark Roadmap

Over the years, TPC benchmarks have raised the bar for what the computing industry has come to expect in terms of benchmarks themselves. Though the original focus has been on online transaction processing (OLTP) benchmarks, to-date the TPC has approved a total of nine independent benchmarks. Of these benchmarks, TPC-C, TPC-H, and TPC-E are currently active, and are widely being used by the industry.

TPC-V, TPC-VMC, and TPC-DI are under development. As described below, TPC-Express is another initiative from the TPC to bring out packaged benchmark kits that are easy to run and report.

The TPC-Pricing Specification and the TPC-Energy Specification are common across all the benchmark standards.

The timelines are shown in Figure 1.

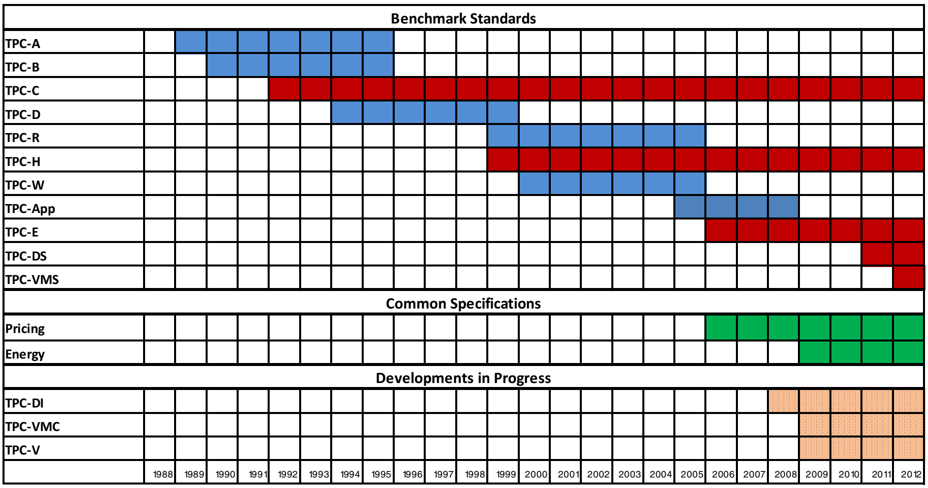

Fig. 1. TPC timeline (Color coding: blue=obsolete, red=current, green= common specifications, beige=under development)

3 TPC Development Status Report

3.1 TPC-Data Integration (TPC-DI)

Data Integration (DI), also known as '"ETL" (Extract, Transform, Load), is the analysis, combination, and transformation of data from a variety of data sources and formats into a unified data model representation. Having a performing data integration system is a key element of data warehousing, application integration, and business analytics solutions. This is especially important as the variety and volume of data are always increasing and performance of data integration systems is critical. Despite the significance of having a highly performing DI system, there has been no industry standard for measuring and comparing the performance of DI systems. Recognizing this benchmark void, the TPC established a subcommittee to develop TPC-DI, a benchmark for Data Integration. It is based on ideas first presented at TPCTC09. The release date of the benchmark is expected in 4^{th} quarter 2013.

The TPC-DI benchmark workload transforms and combines data extracted from a fictitious On-Line Transaction Processing (OTLP) system and other data sources, and loads it into a data warehouse. The source and destination data models, data transformations, and implementation rules have been designed to be broadly representative of modern data integration requirements. No single benchmark can reflect the entire range of possible DI requirements. However, using data and operation models of a retail brokerage, it exercises a breadth of system components associated with DI environments, which are characterized by:

- The manipulation and loading of large volumes of data
- A mixture of transformation types including data validation, key lookups, conditional logic, data type conversions, aggregation operations, etc.
- Fact and dimensional table building and maintenance operations
- Multiple data sources having a variety of different data formats
- Historical loading and incremental updates of the destination data warehouse
- Consistency requirements ensuring that the integration process results in reliable and accurate data
- Multiple data tables with varied data types, attributes, and inter-table relationships

The benchmark is executed in a series of phases, consisting of:

- Initialization
- Loading the data warehouse with large volumes of historical data
- Two incremental updates to the data warehouse, each representing one day of new data
- An automated audit check to verify the results

The Performance Metric reported by TPC-DI is a throughput measure, the number of source rows processed per second. The metric combines the throughputs achieved for each phase to produce the single throughput performance metric.

3.2 TPC-Decision Support (TPC-DS)

The TPC Benchmark DS (TPC-DS) is a decision support benchmark that models several generally applicable aspects of a decision support system, including queries, and data maintenance. The benchmark provides a representative evaluation of the System Under Test's (SUT) performance as a general purpose decision support system.

This benchmark illustrates decision support systems that:

- Examine large volumes of data
- Give answers to real-world business questions
- Execute queries of various operational requirements and complexities (e.g., ad-hoc, reporting, iterative OLAP, data mining)
- Are characterized by high CPU and IO load
- Are synchronized with source OLTP databases through database maintenance functions that are executed while queries are being run, a.k.a. trickle updates

A benchmark result measures query throughput and data maintenance performance for a given hardware, operating system, and DBMS configuration under a controlled, complex, multi-user decision support workload. There have not been any benchmark publications since the benchmark was introduced and only one minor revision was published to clarify wording regarding trickle updates.

3.3 TPC -Virtualization in Progress

The TPC has been working on multiple fronts to deliver benchmarks for measuring the performance of virtualized databases. This section presents a benchmark that has already been released, one that is under development and close to being released, and one that is still a little further from completion.

3.3.1 TPC-VMS

Performance analysts have a choice of virtualization benchmarks [5], including some that have been around for years [6]. But TPC-VMS is the first industry standard virtualization benchmark with the characteristics that have made TPC benchmarks the benchmarks of choice for enterprise-class servers:

- Includes a Price/performance metric
- Is an audited benchmark
- Has database-centric workloads
- Scales the database size with performance when running TPC-E and TPC-C workloads

The goal of the TPC-VMS benchmark was to develop a benchmark specification quickly by utilizing the existing TPC benchmark specifications. The TPC Virtual Measurement Single System Specification (TPC-VMS) leverages the TPC-C, TPC-E, TPC-H, and TPC-DS benchmarks by adding the methodology and requirements for running and reporting performance metrics for virtualized databases. The intent of

TPC-VMS is to represent a virtualization environment where three database workloads are consolidated onto one server. Test sponsors choose one of the four benchmark workloads (TPC-C, TPC-E, TPC-H, or TPC-DS) and run one instance of that benchmark workload in each of the three virtual machines (VMs) on the system under test. The three virtualized databases must have the same attributes, e.g. the same number of TPC-C warehouses, the same number of TPC-E Load Units, or the same TPC-DS or TPC-H scale factors. The TPC-VMS Primary Performance Metric is the minimum value of the three TPC Benchmark Primary metrics for the TPC Benchmarks run in the Virtualization Environment.

Several characteristics of the benchmark are worth noting:

- It models a consolidation environment of three identical databases on three virtual machines with the same workload on the same OS, DBMS, etc.
- The four possible workloads are the well-understood existing TPC-C, TPC-E, TPC-H, and TPC-DS TPC benchmarks.
- To enhance ease of benchmarking for test sponsors, the benchmark was defined such that existing benchmarking kits for TPC-C, TPC-E, TPC-H, and TPC-DS can be used to also run the TPC-VMS variants of these workloads.
- An elegant feature of the benchmark is specifying that the metrics reported are those of the VM with the lowest primary performance metric. This avoids the possibility of a test sponsor *gaming* the test by dividing system resource unevenly among the VMs, but it does so without having to resort to complicated run rules to prevent such gaming.

The TPC-VMS benchmark was adopted in August of 2012, after a very short development phase of one year, hence meeting its goal of a quick development schedule. Prototyping results [7] show that the benchmark meets its goal of exercising the virtualization management system with a complex, database-centric workload.

3.3.2 TPC-V

In 2010, the TPC formed a subcommittee to develop a new benchmark for virtualized databases. The TPC-V benchmarks aims to capture some of the most important properties of databases in the cloud:

- Multiple VMs of varying sizes and different workload types.
- Load elasticity: the benchmark poses a challenge to the hypervisor to react to unexpected changes to the load, and allocate just the right amount of resources to each VM. TPC-V specifies four groups of VMs. Although a constant overall tpsV load level is maintained throughout the run time, the proportion directed to each group changes every twelve minutes, as depicted in the Figure 2 below.
- As the processing power of the system under test (SUT) grows, TPC-V specifies more sets of VMs in each of the four groups. The minimum configuration has four groups, one set per group, and three VMs per set for a total of twelve VMs. But unlike many other virtualization benchmarks, the number of sets does not scale linearly with the power of the SUT. Using a

logarithmic scale, TPC-V specifies two sets per group for today's high end systems, and around three sets per group for the high end systems of the foreseeable future. This makes sense in the context of databases in the cloud: a server may host hundreds of application VMs, but the number of database VMs on one server, even in a cloud environment, is likely to be much more limited.

- TPC-V uses the TPC-E DDL and DML as a starting point to reduce the development time. The TPC-V specification is vastly different from TPC-E, and comparing the results would bring to mind the proverbial apples and oranges comparison. Yet, the reliance on the proven properties of TPC-E has cut years from the TPC-V development process.

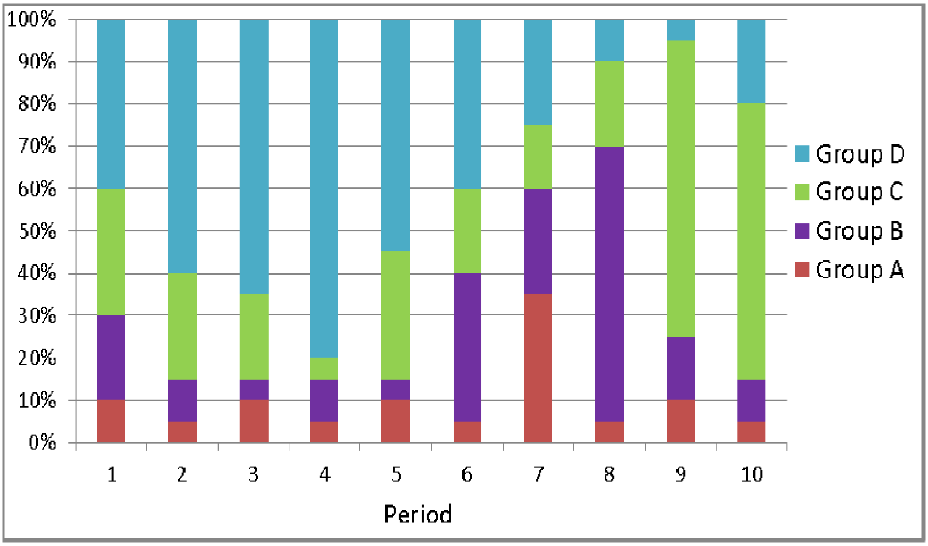

Fig. 2. Elastic load variation of TPC-V

Perhaps the most unique differentiator for TPC-V is that unlike previous TPC benchmarks, the TPC will provide a complete end-to-end, publicly available benchmarking kit along with the paper functional specification. The kit is written in Java and C++. The first implementation uses the open source PostgreSQL database. A full status update paper on TPC-V has been submitted to the TPCTC 2013.

3.3.3 TPC-VMC

TPC-VMS is the first virtualized database benchmark, and models consolidation. TPC-V takes that one step further by modeling heterogeneous workload types, varying load levels among the VMs, elasticity of load, and a VM count that scales with performance. But to truly model cloud environments, one has to include other properties such as:

- Multiple servers
- Load balancing among servers
- Migration of VMs between servers
- Deployment of VMs and applications

The TPC formed a working group to study the feasibility of such a benchmark [8]. The working group considered a number of proposals, and came up with the following requirements:

- The benchmark cannot become a test of *deep pockets*. In other words, if the number of servers is allowed to grow without bounds, a test sponsor can achieve any arbitrary performance level by simply assembling a configuration with just the right number of nodes. Note that this would be trivial for a benchmark such TPC-VMC since the application environment we are simulating is one of independent databases. So one can increase performance by simply adding more nodes. Surprisingly, one can use this very property to limit the number of servers in the configuration. A minimum set of servers can characterize the performance of a large number of servers in a large cloud environment. Therefore, the working group settled on no more than two or four servers.
- In keeping with the success of TPC-VMS in employing existing TPC benchmarking kits, the working group explored options that would not require modifications to existing kit.
- The benchmark proposal outlines a *choreographed* sequence of VM deployments and migrations, as depicted in Figure 3.

The working group has submitted its findings to the TPC, and is presently in hiatus. The TPC expects that once the TPC-V benchmark is released, the working group will resume and consider whether the TPC-V kit can be used to run a benchmark that includes migrations, deployment, etc. If the TPC-V kit proves to be well-received by the industry, extending it to simulate the properties required by TPC-VMC is only a small incremental step since the benchmark already deals with multiple VMs, elasticity, and load balancing within a server.

4 TPC-Express – A New Model for Benchmark Delivery

Traditionally, TPC benchmarks have been delivered in the form of a specification, allowing great flexibility in the way the benchmark application is implemented to satisfy the business case defined by the benchmark. This model worked well in times when customized application development was commonplace and when the various database products in the market delivered function in a wide variety of ways. The TPC considers this more traditional approach to its benchmarks as the “Enterprise” model. Compelling reasons to use the existing "Enterprise" benchmark model remain when the optimal application is developed to satisfy a functional specification.

Today, however, most database management products offer a suite of functions that are largely compatible for most database applications, and most commercial applications are purchased from an application provider. It makes good sense, then, to offer benchmarks that emulate these off-the-shelf products with the delivery of working benchmark applications in a downloadable benchmark kit, rather than requiring the development of the benchmark application by the implementer. This represents an exciting step for the TPC and those using TPC benchmarks. This new "Express" model will provide a kit that includes routines to build the database, run the benchmark application, report the results, and provide a level of validation for result compliance. This means that implementation of a benchmark can be accomplished much less expensively, with a higher confidence that the results are compliant and comparable.

A quick comparison of the two models is summarized in Table 1

Table 1. Express vs. Enterprise Models

	Express	Enterprise
Execution	Kit based (enhanced by specification)	Specification based (with some code)
Implementation	Out of the box	Customized
Audit Requirements	Mostly self validation	Full audit
Pricing	Not required	Required
ACID properties	ACI at most	Full ACID
Pricing model	License sales and benchmark registration	Benchmark registration
Expected volume	High	Limited
Cost to run the benchmark	Low	High
Time to run the benchmark	Short	Longer

Where the existing Enterprise benchmarks were typically only published by computer manufacturers, the TPC expects that the Express class of benchmarks will appeal to a wider audience that includes computer and software manufacturers, academic researchers as well as individuals interested in running test environment workloads to validate data center system changes.

The TPC is actively working to produce a first benchmark within the Express model. This will likely be a revision of an existing Enterprise benchmark, adjusted in ways to satisfy the needs of the Express model. The results will not be comparable with the parent Enterprise benchmark. In parallel, the TPC intends to produce a guide for other Express benchmark proposals that are both in the areas of traditional TPC benchmarks and in newer areas, such as database in the cloud, Big Data, Business Analytics, in-memory databases, and so on. The TPC welcomes proposals from within and outside of the TPC membership, and invites those who would like to participate in this development process to become active members.

5 TPC Technology Conference Series (TPCTC)

The information technology landscape is evolving at a rapid pace, challenging industry experts and researchers to develop innovative techniques for evaluation, measurement and characterization of complex systems. The TPC remains committed to developing new benchmark standards to keep pace, and one vehicle for achieving this objective is the sponsorship of the Technology Conference on Performance Evaluation and Benchmarking (TPCTC). Over the last four years we have held TPCTC successfully in conjunction with VLDB.

Table 2. TPCTC at a glance

TPCTC	VLDB	Location	Date	Keynote	Proceedings
TPCTC 2009	35th Int'l Conference	Lyon, France	August 24-28	Michael Stonebraker[1], M.I.T.	http://www.springer.com/978-3-642-10423-7
TPCTC 2010	36th Int'l Conference	Singapore	September 13-17	C. Mohan[2], IBM	http://www.springer.com/computer/communication+networks/book/978-3-642-18205-1
TPCTC 2011	37th Int'l Conference	Seattle, WA	Aug 29 – Sep 3	Umesh Dayal[3], HP Labs	http://www.springer.com/computer/communication+networks/book/978-3-642-32626-4
TPCTC 2012	38th Int'l Conference	Istanbul, Turkey	August 27-31	Michael Carey[4], UC Irvine	http://www.springer.com/computer/communication+networks/book/978-3-642-36726-7
TPCTC 2013	39th Int'l Conference	Trento, Italy	August 26-30	Raghu Ramakrishnan[5], Microsoft	

The TPC Technology Conferences have had direct effect on the TPC's direction and activities:

- The formation of TPC's Virtualization working group (TPC-V) was a direct result of papers presented at TPCTC 2009. Proposals such as dependability aspects are under consideration for future benchmark enhancements.
- Several new benchmark ideas, enhancements to existing benchmarks and lessons learnt in practice were presented at TPCTC 2010 that had a direct

[1] Adjunct Professor, Massachusetts Institute of Technology, Cambridge, MA.
[2] IBM Fellow at the IBM Almaden Research Center, San Jose, CA.
[3] ACM Fellow and Chief Scientist of the Information Analytics Lab at HP Labs, Palo Alto, CA.
[4] Donald Bren Professor of Computer and Information Sciences, University of California, Irvine, CA.
[5] Technical Fellow, Microsoft, and Professor of Computer Sciences at the University of Wisconsin, Madison, WI.

impact to the TPC and the industry, e.g. a proposal for a generic data generator.

- Papers presented at TPCTC 2011 included new benchmark ideas in the area of Event Bases Systems, Mixed Workload Benchmarks, and Dependability Benchmarks. There were also various papers on enhancing existing TPC workloads, such as an enhancement to TPC-H and a dbgen implementation for TPC-H using the generic data generator PDGF. Some more theoretical papers included analytical models of benchmarks.
- Papers presented at TPCTC 2012 included new benchmark ideas in the area of big data, energy efficiency, Cloud, ETL, and virtualization.

With the 5th TPC Technology Conference on Performance Evaluation and Benchmarking (TPCTC 2013) proposal, the TPC strives to exceed the success of previous workshops by encouraging researchers and industry experts to present and debate novel ideas and methodologies in emerging performance evaluation and benchmarking areas. Authors are invited to submit original, unpublished papers that are not currently under review for any other conference or journal. The TPC also encourages the submission of extended abstracts, position statement papers and lessons learned in practice. The accepted papers will be published in the workshop proceedings, and selected papers will be considered for future TPC benchmark developments. Topics of interest include, but are not limited to:

- Big Data
- Cloud Computing
- Social media infrastructure
- Business intelligence
- Complex event processing
- Database optimizations
- Green computing
- Disaster tolerance and recovery
- Energy and space efficiency
- Hardware innovations
- Hybrid workloads
- Virtualization
- Lessons learned in practice using TPC workloads
- Enhancements to existing TPC workloads

6 Major Areas of Focus for 2014 and Beyond

6.1 Big Data

The last five years have seen a huge change in the industry landscape: Platforms that can handle Big Data workloads have become mainstream. Big Data refers to data sets that are too large and too complex to store and process in a cost effectively and timely manner using traditional tools like scale-up systems and relational management systems. Emerging from the Web 2.0 challenge, solutions are now available to provision and manage very large workloads, including Hadoop and NoSQL. Without doubt, enterprises see the value of Big Data and Big Data analytics across all major sectors, including health care, retail, education, and government, due to two main reasons. First is an increased number of people constantly connected to the internet and second there is an increased number of devices connected to the Internet. While there were 15 billion devices connected to the Internet in 2011 it is predicted that by year 2020 there will be 50 connected billion devices connected.

To face the challenges associated with the amount of data produced by the increased number of users and their devices, hardware and software infrastructure technologies have also evolved from traditional scale-up and client/server systems to massive scale-out clusters and clouds. Hadoop and NoSQL systems have become cost–effective, scalable platforms for handling massive amounts of structured, semi structured and unstructured data. Many of these technologies were a contribution of Web 2.0-era companies. Enterprises are also considering the use of Hadoop and NoSQL, realizing that storing and mining large data sets can help optimize their business processes, improve the customer experience, uncover strategic and competitive opportunities, and thereby gain a competitive advantage. With this new Big Data landscape, and multiple technologies to choose from, there is a need for industry standards so users can see fair and unbiased comparisons of technologies and solutions.

With no standard system benchmarks available for Big Data systems, today's situation is similar to that of the middle 1980s, when the lack of standard database benchmarks led many system vendors to practice what is now referred to as "benchmarketing," a practice in which organizations make performance claims based on self-designed, highly biased benchmarks.

Some of the existing TPC benchmarks like TPC-H and TPC-DS can easily be extended for use in large structured datasets. For example, current TPC-H and TPC-Ds benchmarks support scale factors of 100GB, 300B, 1TB, 3TB, 10TB and 30TB. This can be extended to larger scale factors like 100TB, 300GB, 1TB, 3PB and more, following the log scale, using existing data generation tools and queries. There is work in progress to extend TPC-DS to handle unstructured data also. There are initiatives like WBDB (Workshop on Big Data Benchmarking), which is intended developed brand new workloads. TPC-H and TPC-DS contain a diverse set of structured data, which makes them a suitable candidate for a Big Data benchmark.

As reported in Big Data Management, Technologies, and Applications, one of the outcomes of the first workshop on Big Data Benchmarking is BigBench _Ref355098135. BigBench is an end-to-end, Big Data benchmark proposal. It is based on TPC-DS. Hence, its underlying business model is a product retailer. In addition to TPC-DS, it proposes a data model and synthetic data generator that address the variety, velocity and volume aspects of Big Data systems containing structured, semi-structured, and unstructured data. The structured part of BigBench's data model is adopted from TPC-DS. It is enriched with semi-structured and unstructured data components. The semi-structured part captures registered and guest user clicks on the retailer's web site. The unstructured data captures product reviews submitted online.

The data generator, which was designed for BigBench, provides scalable volumes of raw data based on a scale factor. BigBench's workload is designed around a set of queries against the data model. From a business prospective, the queries cover the different categories of Big Data analytics proposed by McKinsey. From a technical prospective, the queries are designed to span three different dimensions based on data sources, query processing types and analytic techniques. In the SIGMOD paper, the authors further illustrate the feasibility of BigBench by presenting an implementation on Teradata's Aster Database. The test includes generating and loading a 200 Gigabyte BigBench data set and testing the workload by executing the BigBench queries (written using Teradata Aster SQL-MR) and reporting their response times.

BigBench's data model focuses on volume, variety, and velocity. The variety property of BigBench is illustrated in Figure 4. The structured portion of BigBench's data model is adapted directly from TPC-DS' data model, which also depicts a product retailer _Ref355098116 [13]. BigBench adds a table for prices from the retailer's competitors to the portion of TPC-DS that contain store and online sales data. TPC-DS structured part is enriched with semi-structured and un-structured data shown in the lower and right hand side of Figure 4. The semi-structured part is composed by clicks made by customers and guest users visiting the retailer's web site. The design assumes the semi-structured data to be in a key-value format similar to Apache's web server log format. The un-structured data in the new model is covered by product reviews that can be submitted by guest users or actual customers.

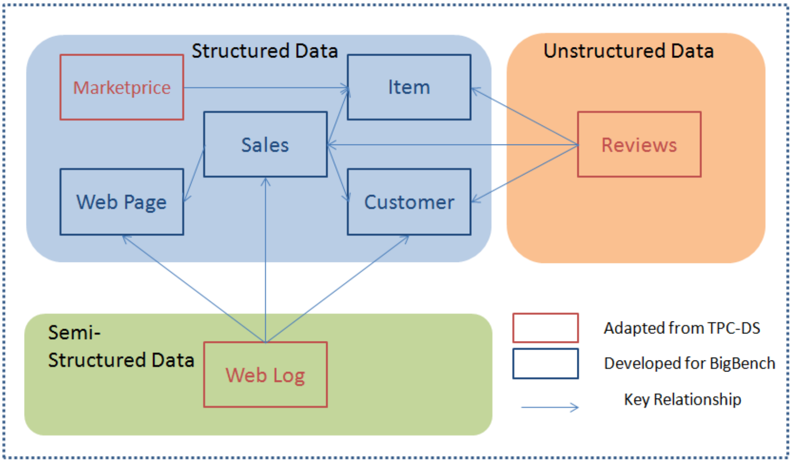

Fig. 4. Logical Data Model BigBench (Adapted from _Ref355098135)

6.2 OpenStack

The term cloud computing has different meanings depending on the target environment. There are three main types of services provided by cloud environments:

Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS), and Software-as-a-Service (SaaS). The difference between these various cloud environments has mostly to do with how much of the solution stack the user can control. For instance, with SaaS the user has access to a piece of software running in the cloud but has no control over what operating system it is running on whereas in IaaS the user has control over many aspects of the solution stack.

OpenStack (www.openstack.org) is an open source project to define and build a highly scalable common cloud computing platform for public and private clouds. OpenStack would be defined as an IaaS cloud service. Over 150 companies are participating in some aspect of the OpenStack development effort, including many of the TPC member companies.

The OpenStack project has an extremely active development community. The first OpenStack release was in October of 2010, and there have been six releases over the course of the following two and a half years. With two to three releases a year, the pace of development is very rapid. This fast development cadence is necessary, since many vendors want to implement cloud environments starting immediately rather than some time in the future.

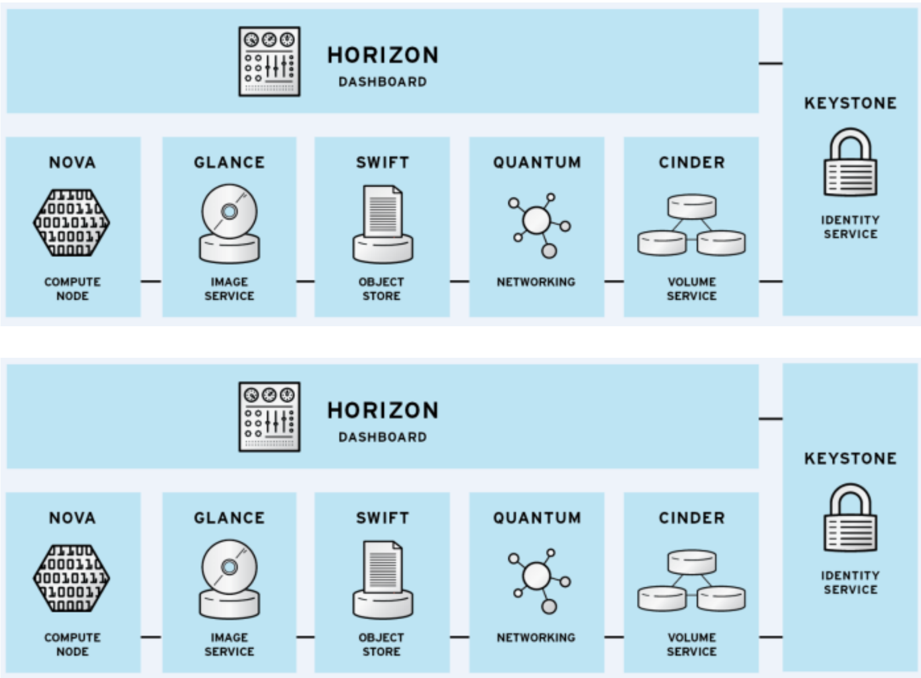

Fig. 5. OpenStack Architecture (http://www.redhat.com/products/cloud-computing/openstack/)

The design of OpenStack contains components in the areas of compute, networking, and storage. The parts of OpenStack that deal with the compute aspect of a cloud have project names such as Nova, Glance, and Horizon. Nova is a framework for providing virtual servers on demand in an OpenStack environment. Nova does not provide virtualization functionality, but can be hooked into various virtualization technologies via an API. Glance provides a way to create a catalog of virtual disk images for the compute framework to reference and use. Horizon will be the most recognizable part of OpenStack to users since it is the GUI management interface for OpenStack.

The network and storage area is addressed by the Quantum, Swift, and Cinder projects. Quantum provides network connectivity as a service and interfaces with many different types of networking technologies. The Swift and Cinder projects both deal with storage, but different aspects of storage. Swift provides for object storage, while Cinder can provide persistent block storage to the virtual machines deployed in OpenStack.

A key aspect to any cloud environment is security. In the OpenStack environment authentication and authorization are handled by the Keystone project. Cloud infrastructure security must be both robust and efficient.

Not all of the OpenStack projects mentioned above are needed for every OpenStack use case, but all of the projects are designed to function together to provide a complete and scalable cloud infrastructure. How efficiently they function and scale is where benchmarking comes in.

6.2.1 Benchmarking OpenStack

Of course the main question facing industry consortia focused on performance like the TPC is how to measure the performance of a cloud infrastructure like OpenStack. The first step would be to realize that the performance of OpenStack should not be focused on the virtualization technology. There are already industry standard benchmarks such as SPECvirt_sc2010, SPECvirt_sc2013, TPC-VMS, and the under-development TPC-V that are focused on measuring the performance of virtualization technologies. It is also possible to take a currently available benchmark from any industry consortia and run it in a virtualized environment to try and measure virtualization performance.

The performance of a cloud environment is heavily dependent on the infrastructure used to build the cloud. Therefore, cloud benchmarks should focus on measuring this infrastructure performance while as the same time measuring overall cloud environment performance.

Below are some interesting performance questions related to a cloud environment as well as the parts of OpenStack that would most affect the answer.

How fast can a virtual machine image be deployed? Nova, Glance, Swift, Cinder

- Do my tasks take longer to run in a cloud than if I was just using virtualization? Nova, Glance, Quantum, Swift, Cinder
- What kind of performance slowdown does the security of the cloud cause? Nova, Keystone

- Do the answers to any of the previous questions change as the cloud environment scales? All OpenStack projects

The use of a cloud environment for providing compute resources to a specific set of customers revolves mainly around the ability to meet particular response time criteria for those customers. If a cloud environment cannot meet a customer's response time needs then dedicated hardware would have to be deployed instead. Therefore, any cloud benchmark must be designed around response time requirements and have it built into every aspect of the benchmark.

Because there are many different aspects to a cloud infrastructure like OpenStack, a benchmark designed to test such an environment would have to have many aspects as well. Potentially a suite of tests will be required with each designed to put stress on a particular aspect of the OpenStack environment to see how it performs. The challenge to having a benchmark that is made up of multiple tests is normalizing multiple data points into a single metric score. For an industry standard benchmark to be successful, one main metric is ideal. Multiple secondary metrics could be defined, but they should be rolled up into a single main metric.

7 Conclusion

In an environment of rapid and pervasive change, the TPC remains committed to serve the industry with benchmark standards that are relevant and up to date. While the TPC's traditional, Enterprise benchmarks continue to be the gold-standard for large database workloads, the organization has several new benchmarks in process. TPC-DI, TPC-V, and the new TPC-Express model are such initiatives that cover workloads as diverse as data integration, virtualization, and an entire new approach to benchmarks The TPC is also exploring ideas and methodologies to create benchmarks for Big Data and OpenStack. The organization also strongly supports benchmarking innovation through the TPC Technical Conference (TPCTC) and looks forward to incorporating innovative ideas from the 5^{th} TPCTC.

Acknowledgements. The authors thank the past and present members of the TPC for their contribution to the specifications and documents referenced in this paper.

References

1. Nambiar, R., Poess, M. (eds.): TPCTC 2012. LNCS, vol. 7755. Springer, Heidelberg (2013)
2. Nambiar, R., Poess, M. (eds.): TPCTC 2011. LNCS, vol. 7144. Springer, Heidelberg (2012)
3. Nambiar, R., Poess, M. (eds.): TPCTC 2010. LNCS, vol. 6417. Springer, Heidelberg (2011)
4. Nambiar, R., Poess, M. (eds.): TPCTC 2009. LNCS, vol. 5895. Springer, Heidelberg (2009)
5. SPEC Virtualization Committee: http://www.spec.org/virt_sc2010/, http://www.spec.org/virt_sc2013/
6. VMware, Inc., http://www.vmware.com/products/vmmark/overview.html
7. Smith, W.D., Sebastian, S.: Virtualization Performance Insights from TPC-VMS, http://www.tpc.org/tpcvms/tpc-vms-2013-1.0.pdf
8. Smith, W.D.: Characterizing Cloud Performance with TPC Benchmarks. In: Nambiar, R., Poess, M. (eds.) TPCTC 2012. LNCS, vol. 7755, pp. 189–196. Springer, Heidelberg (2013)