SMART DATA FAST.™

VOLTDB

SOFTLAYER BENCHMARK

DATA SHEET

SOLVING THE FAST DATA PROBLEM FOR THE CLOUD INTEGRATING VOLTDB AND THE IBM SOFTLAYER CLOUD COMPUTING PLATFORM DELIVERS UP TO 5X THE PERFORMANCE OF AMAZON WEB SERVICES

VoltDB, the industry's fastest in-memory operational NewSQL database, and the global cloud computing platform from SoftLayer, an IBM company, provides users with the performance of bare metal and the Fast Data processing capabilities required to analyze and make real-time decisions to meet the needs of today's high-volume, high-velocity businesses.

OVERCOMING FAST DATA HURDLES

The cloud has become the platform of choice for accelerating Big Data application development. However, to tap the maximum value of data, you must be able to extract intelligence from it and take immediate, decisive action at the moment it arrives, rather than days or weeks later. Not all cloud platforms provide the scale, performance, reliability and fast, functional data management capabilities crucial to supporting applications that thrive on Fast Data.

The combination of VoltDB and IBM SoftLayer solves the challenge of running Fast Data in the cloud, providing the performance of bare metal and the processing capabilities needed to analyze and make real-time decisions.

VOLTDB INTEGRATION WITH IBM SOFTLAYER

Many businesses seeking the benefits of data processing in the cloud are saddled with legacy data management systems that are too slow to handle the rapid ingestion, real-time analysis and decision-making requirements of Fast Data applications.

The VoltDB in-memory NewSQL database makes traditional databases obsolete by combining high-velocity data ingestion and export, massive scalability, and real-time analytics and decision-making capabilities. SoftLayer offers unparalleled performance and control, with a full-featured API and sophisticated automation controlling a flexible unified platform that seamlessly spans physical and virtual devices, and a worldwide network for secure, low-latency communications.

Together, VoltDB and SoftLayer deliver a high-performance solution for running Fast Data applications in the cloud. Enterprises gain:

- The ability to ingest large volumes of fast-moving data and process it in real time to support decision-making;
- An automated cloud platform with the performance of bare metal, a modular, customizable infrastructure, and an API-driven, simple management system;
- A no-compromises solution based on modern approaches to data management with the power and familiarity of SQL and ACID guarantees of RDBMSs.

VOLTDB - SOFTLAYER BENEFITS

Increasingly, enterprises are realizing the benefits of using best-in-class solutions for their growing data needs. The VoltDB and SoftLayer integration provides the following:

- The ability to serve large-scale, concurrent users;
- A seamless, global cloud–based application development platform;
- The ability to select the highest-performance servers in the cloud;
- The query speed needed for interactive applications; and
- The ability to make real-time, automated, per-event decisions.

The VoltDB - SoftLayer solution is ideal for organizations in industries as diverse as telecommunications, online media, financial services, gaming, public utilities and defense. Use cases include customer-specific dynamic advertising placement and ad targeting, click-stream analysis, analysis of data from IoT-connected devices, financial tick data analysis, log analysis, analysis of real-time sensor readings in utility grids, and analysis of user interactions in online gaming to drive in-game purchases. The integration of VoltDB and SoftLayer offers users the scale and anytime, anywhere access of the cloud, up to 5.4X the performance of Amazon EC2, and a single, encompassing view of data flows in real time. Businesses are empowered to quickly act on vital information that can mean the difference between profit and loss.

THE TEST: VOLTDB - SOFTLAYER PERFORMANCE BENCHMARK YIELDS 3.4X – 5.4X THE PERFORMANCE OF AMAZON EC2

VoltDB tested three, six, nine and 12-node clusters on IBM SoftLayer using the Yahoo Cloud Serving Benchmark (YCSB), a de facto industry-standard performance benchmark for cloud databases. The YCSB defines a number of different workloads, the most widely reported of which is Workload B with 95% read/5% update.

The bare metal benchmark was run on servers using dual Intel E5-2690v2 processors on SuperMicro X9DRI with 64GB and 10GbE running 64bit CentOS 6.5. VoltDB was directly installed after provisioning; the only additional requirement was to install a JAVA virtual machine.

In comparison, the VoltDB benchmark of YCSB on AWS ran on Paravirtual (PV) c3.8xlarge instances running Ubuntu 13.10. These instances are based on Intel Xeon e5-2680 v2 (Ivy Bridge) processors with 60gB each of memory.

As with the AWS test, the SoftLayer benchmark targeted three core YCSB workloads: Workload A (50% read/50% update), Workload B (95% read/5% update), and Workload E (95% of the paging "scan" operation of up to 100 rows/5% insert). This combination gives a representative view of the kinds of things YCSB is used to test. Running Workload B, the most popular of the three benchmarks, on the 12-node cluster, VoltDB and SoftLayer achieved from 3.4X to 5.4X the performance of Amazon Web Services. (For more information on AWS benchmark results, see this VoltDB post.)

Scenarios were tested against three, six, nine, and 12-node clusters, where each cluster was matched with 150 client threads per server node, distributed such that no client machine was managing more than 300 threads. Further, the YCSB client was implemented to assign one instance of the VoltDB Java client to each 50 YCSB threads. The purpose of this trial was to ensure that the database was running at load in order to measure peak throughput.

In the benchmark run against a 12-node cluster (the largest configuration, and the one most commonly used by VoltDB customers), client threads were incrementally added. This was accomplished by adding additional client machines, each running 150 threads. This test demonstrates how much load the cluster can handle before seeing a noticeable impact in terms of latency. Note that latencies remain essentially flat until a point just short of peak throughput. This behavior is true not only of average latencies, but also is preserved at the 99th percentile.

In all cases, the database was preloaded with 10 million records per node using the YCSB defaults of 10 fields with 100 bytes each, and the workloads ran on each setup for two independent 10-minute trials.

The source code for the client implementation of the YCSB benchmark is public, and can be found on GitHub at https://github.com/VoltDB/voltdb/tree/master/tests/test_apps/ycsb. YCSB is an open source project. As of the running of this test, the current version of YCSB is 0.1.4. Since VoltDB does not maintain YCSB, it can't be guaranteed that future revisions will not break compatibility with our driver.

RESULTS

Workload A

Workload A is a 50-50 mix of read and update operations. Throughput on SoftLayer increased from 426K transactions per second (TPS) for three nodes to 904K for 12 nodes, which is ~ 2.5X AWS.

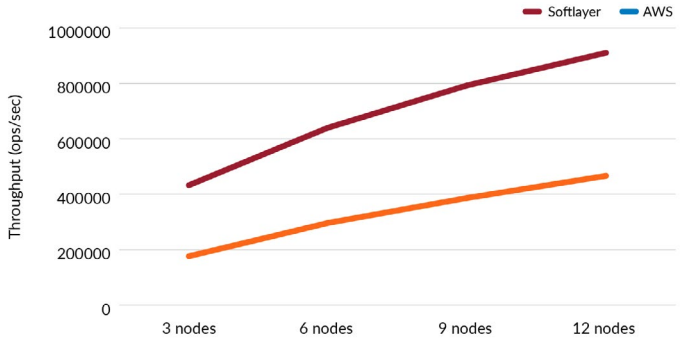

FIGURE 1.
Workload A Average Throughput

The 12-node cluster was able to sustain a high degree of responsiveness through 816K TPS.

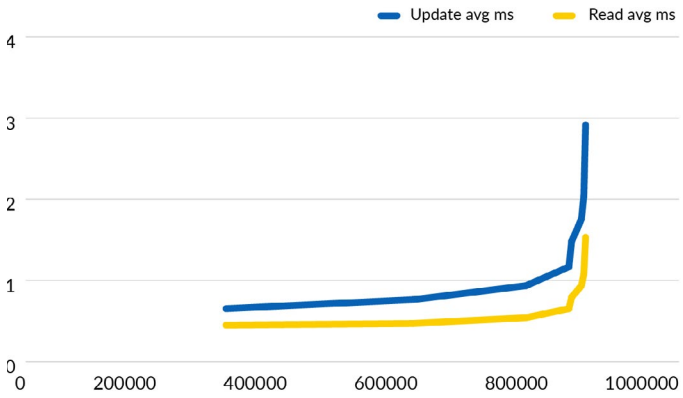

FIGURE 2.
Workload A Average Latency vs. Throughput

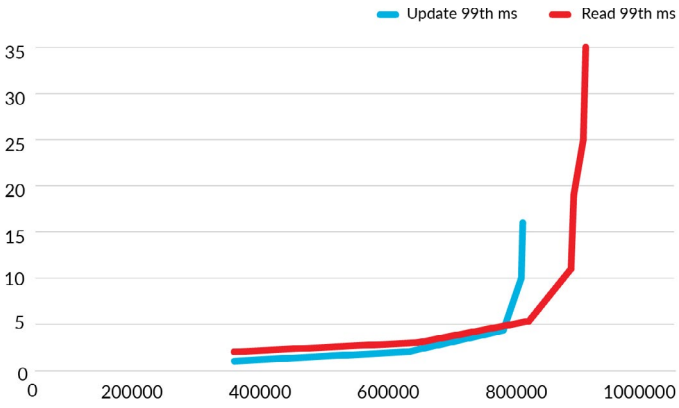

FIGURE 3.
Workload A 99th Percentile Latency vs. Throughput

Workload B

Workload B is a mix of 95% reads with 5% updates. Throughput scales in a linear fashion for this workload from 1,018K TPS for three nodes, 5.4X the AWS throughput, to 2,438K TPS for 12 nodes, 3.4X the AWS throughput.

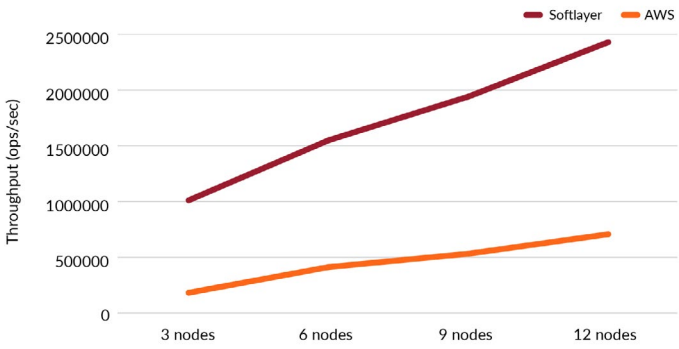

FIGURE 4.
Workload B Average Throughput

Workload B shows very favorable latency results, with low latencies holding nearly constant up until quite close to the peak throughput seen above. Up through almost 2,000K TPS the latency curve is essentially flat.

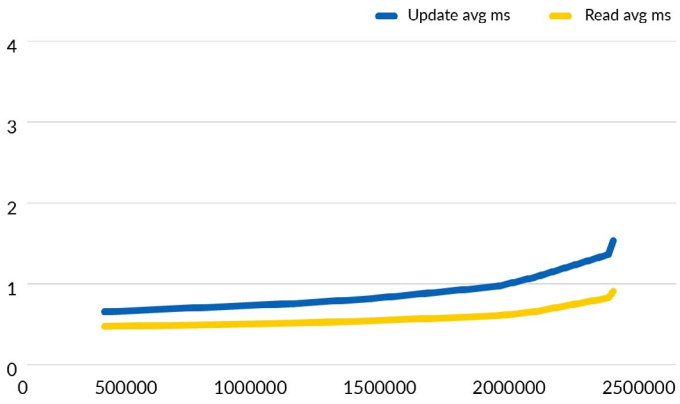

FIGURE 5.
Workload B Average Latency vs. Throughput

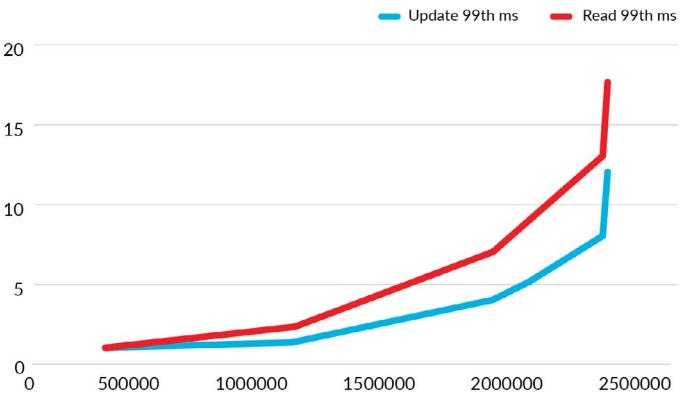

FIGURE 6.
Workload B 99th Percentile Latency vs. Throughput

Workload E

Workload E of the YCSB benchmark implements the more complicated paging "scan" operation, which can request up to 100 rows at a time – and could potentially require data from more than one partition in a variation on the "run everywhere" strategy. A full 95% of operations in this workload are this scan, with inserts filling out the rest. Despite needing to synchronously carry out this much larger operation, which has the definite potential to become network bound, the 12-node SoftLayer cluster produced 201K TPS, 30% higher than AWS. This reflects the fact that throughput for Workload E is limited by I/O with less of the processing advantage showing through as in the other workloads.

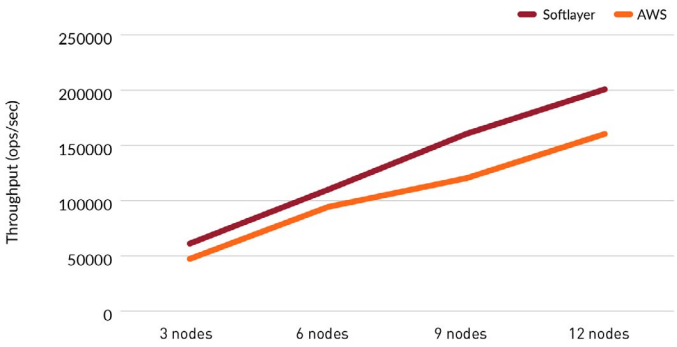

FIGURE 7.
Workload E Average Throughput

Given the nature of this workload and dependence on I/O, latencies with Workload E are, as expected, somewhat higher than the other workloads. Nevertheless VoltDB performed extremely well. The latencies shown in the figures below are not large, and are quite steady until right up against peak throughput.

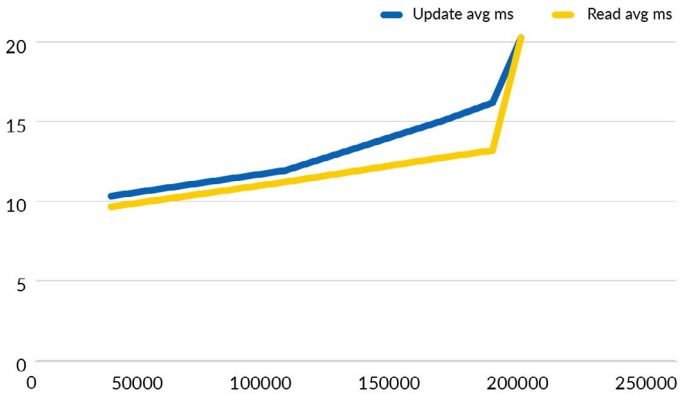

FIGURE 8.
Workload E Average Latency vs. Throughput

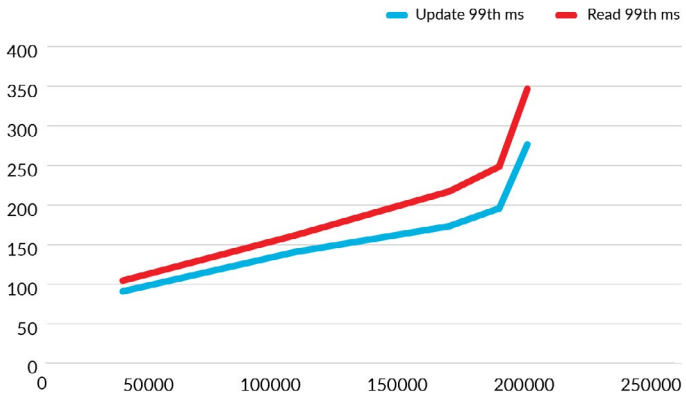

FIGURE 9.
Workload E - 99th Percentile Latency vs. Throughput

RESULTS ECLIPSE EXPECTATIONS

VoltDB's performance running the YCSB benchmark on SoftLayer reveals significant advantages compared to AWS.

For each workload, VoltDB displays essentially linear scaling behavior. For Workloads A and B, even pushing well into hundreds of thousands of TPS, 99th percentile latencies remain below 4 or 5 milliseconds, while for Workload E – where each operation is handling up to 100 times more data – that number increases to only 16 milliseconds. Against Workload B, which is the most commonly reported result, VoltDB achieves what we believe to be the world's record for YCSB at 2,438,000 average TPS on the 12-node cluster.

As good as these results are, in VoltDB's world they're not exceptional. For example, going all the way back to VoltDB version 1.1, SGI[1] ran VoltDB's "voter" benchmark against their Rackable C1001-TY3 cluster. This benchmark application does a bit more than YCSB does in each of its transactions, and with 10 server nodes they achieved a rate of about 1.2 million TPS. With 30 nodes, it scaled linearly all the way to 3.37 million TPS. More recently, running a simple, key-value workload with a 90% read/10% write mix against just a three-node, k-factor 1 cluster of Dell R510s, a maximum throughput of about 950,000 TPS was observed, with consistently good latencies up to 800,000 TPS.[2] These results showcase the strength of VoltDB's modern scalable architecture. Want to know more? Try it for yourself – download VoltDB now.

[1] https://www.sgi.com/pdfs/4238.pdf

[2] https://voltdb.com/blog/voltdb-3-x-performance-characteristics/