TECHNICAL BRIEF

Achieving Mission-Critical Performance with InterSystems Caché*, VMware vSAN*, and the Intel® Solid State Drive Data Center Family for PCIe*

Executive Summary

Healthcare organizations are optimizing infrastructure to improve efficiency and reduce cost. Running the InterSystems data platform on software-defined storage using VMware vSAN* and the Intel® Solid State Drive (Intel® SSD) Data Center Family for PCIe* delivers high performance with significant savings.

Benchmark testing demonstrates the viability of this approach, running the InterSystems TrakCare* 2015 electronic medical record (EMR) system on the InterSystems Caché* 2015.2 database, with loads generated using utilities provided by InterSystems. The primary goals of this work are to quantify the throughput profile, maximum throughput, and scaling of applications built on the InterSystems data platform and vSAN, running on hardware based on Intel® processors and the Intel SSD Data Center Family, while maintaining acceptable server, storage, and user response times.

The results of this testing show linear scalability of workloads, across both a single database instance and in a scale-out scenario across multiple application servers. In addition, random-read I/O performance testing demonstrates the ability to maintain acceptable disk latency under high loads.

These findings confirm that the healthcare industry can deliver excellent performance and end-user experience using a highly scalable infrastructure based on building blocks from VMware, InterSystems, and Intel. This standards-based solution enables analysis and management of large collections of patient data, setting the stage for open-ended product innovation with high capability and low cost.

Overview of Solution

The testing reported on in this brief uses a solution architecture that highlights InterSystems Caché, VMware vSAN, and the Intel SSD Data Center Family for PCIe*.

InterSystems Caché

InterSystems Caché is an advanced database management system and rapid application development environment that offers high performance, scalability, and reliability with low maintenance and hardware requirements.

This new generation of database technology provides multiple modes of data access. Data is described only once, in a single integrated data dictionary, and it is instantly available using object access, high-performance SQL, and powerful multidimensional access, all of which can simultaneously access the same data.

Caché comes with several built-in scripting languages and is compatible with popular development tools.

Authors

Murray Oldfield
Technology Architect, InterSystems

Christian Rauber
Senior Solutions Architect, VMware

Vivek Sarathy
Storage Solution Architect, Intel Corporation

VMware vSAN

VMware vSAN allows IT environments to seamlessly extend virtualization to storage with a hyper-converged solution built to interoperate with the overall VMware environment. Using this technology, IT organizations can easily mitigate risk in digital-transformation projects using existing tools, skill sets, and solutions. Fully integrated with VMware vSphere*, vSAN offers the flexibility to expand using other VMware solutions for the software-defined data center and multi-cloud environment as IT needs grow.

Embracing vSAN lowers total cost of ownership by 50 percent or more with capital and operational savings.[1] The solution reduces capital expenditures with software-defined storage that takes advantage of server-side economics and affordable flash storage. Users can avoid large capital expenses with on-demand, granular scaling of compute and storage that delivers predictable performance and costs.

Users can also decrease operational expenses with the rapid deployment of compute and storage infrastructure, while managing day-to-day operations in one tool and with one team. Additionally, the solution accelerates responsiveness to traditionally time-consuming tasks, from troubleshooting to performance tuning, with intelligent analytics, advanced monitoring, and virtual machine (VM)-level automation.

Intel SSD Data Center Family for PCIe*

The Intel SSD Data Center Family for PCIe is based on Non-Volatile Memory Express* (NVMe*), a standard that brings PCI Express* (PCIe) drives into the mainstream with industry-standard software and drivers. Compared to SATA drives, PCIe drives based on the NVMe standard can provide[2]:

- Low latency
- Improved CPU utilization
- Scalability
- Increased bandwidth
- Direct attachment to CPU, which can eliminate costs and overhead associated with the host bus adapter (HBA)

Because these enterprise-grade drives connect to the CPU through the PCIe interface, they can deliver up to six times the performance of Serial Attached SCSI (SAS)/Serial ATA (SATA) SSDs with reduced latency and improved CPU utilization[3,4]. They are designed for scalability, quality of service, and low latency.

The performance of a single drive from the Intel SSD Data Center Family for PCIe, specifically the Intel SSD DC P3700 Series, can replace the performance of multiple SATA SSDs[5]. This performance makes the Intel SSD Data Center Family for PCIe an ideal solution for the most demanding data-center workloads.

Testing Summary

Two configurations are tested in this study: a single database instance and a scale-out architecture using InterSystems data platform Enterprise Cache Protocol (ECP) with multiple application servers.

Note: *It is InterSystems policy to verify and release its products against various operating systems and processor types. For example, InterSystems supports applications on specific Red Hat versions, whether bare metal or a virtualized platform such as vSAN. The same is true for the InterSystems TrakCare health information systems and InterSystems HealthShare* health informatics solution sets, which are built on the InterSystems data platform and supported on selected operating systems. Further details of InterSystems support policy are available at http://www.intersystems.com/services-support/product-support/virtualization.*

Benchmark Methodology

The testing uses the InterSystems WebStress testing tool (included with the InterSystems data platform) to drive variable test-transaction rates. Test scripts are run for one hour, with metrics that include average CPU utilization, IOPS, and database accesses per second averaged over the hour. Multiple test runs are used for each transaction rate. The test scripts and transaction profile for this benchmark are derived from current performance statistics and TrakCare component usage measured on live customer systems.

The TrakCare system is used to simulate realistic application database activity. TrakCare is built on the InterSystems data platform, which incorporates the InterSystems Caché high-performance database, InterSystems Ensemble* integration platform, and InterSystems DeepSee *embedded analytics. To test scaling, the transaction rate for each one-hour test is increased in increments (1x, 1.5x, 2x, and so on) up to and beyond 100-percent guest operating system CPU utilization. A 2x test is set to be twice the workflow transaction rate of a 1x test. A cache tier based on the Intel SSD Data Center P3700 Series is used to avoid bottlenecks associated with storage throughput. Support for the NVMe interface standard on these drives provides very high data transfer rates.

Although the benchmark tests are based on usage data from live TrakCare sites and are a mixture of high-use and high-impact hospital workflows, it must be noted that TrakCare deployments have wide variation. The TrakCare application is used as a utility to drive the tests to stress the InterSystems data platform and vSAN platform with realistic transactional activity. Therefore, hardware requirements and sizing for customer sites are not to be inferred from this report, because other factors such as architecture, scalability, high availability, additional applications, and growth must also be considered. The InterSystems Technology Architect group provides TrakCare and HealthShare configurations for customers.

Testing Approach

A TrakCare database generator utility is used to create a large test database with the types of data required for the workflows and screens used in the tests. The generator allows creation of different sizes of databases, allowing realistic simulation of very large customer databases to match the throughputs being benchmarked.

A single test is made up of a set of scripts; each script simulates an application workflow moving through several screen or application components. An application screen may also be made up of one or more components. The scripts for a test are selected based on observations and metrics from live sites, taking into account high-use workflows, transactions, and components that use the highest system resources.

One or more driver VMs simulates web sessions (or users) by executing test scripts at a set transaction rate, with tests tuned to achieve an average throughput at a set target rate. A test with a rate of 1x is the baseline for the different script rates that make up a test. This approach allows test rates to be scaled up and down in increments of 1x, 1.5x, 2.x, and so on.

As shown in Figure 1, the test drivers send HTTP traffic to one or more Apache* web servers to simulate web sessions (or users) by executing application workflows using parameters made up of random data exported from the built benchmark database. The web servers connect to the database server for application processing.

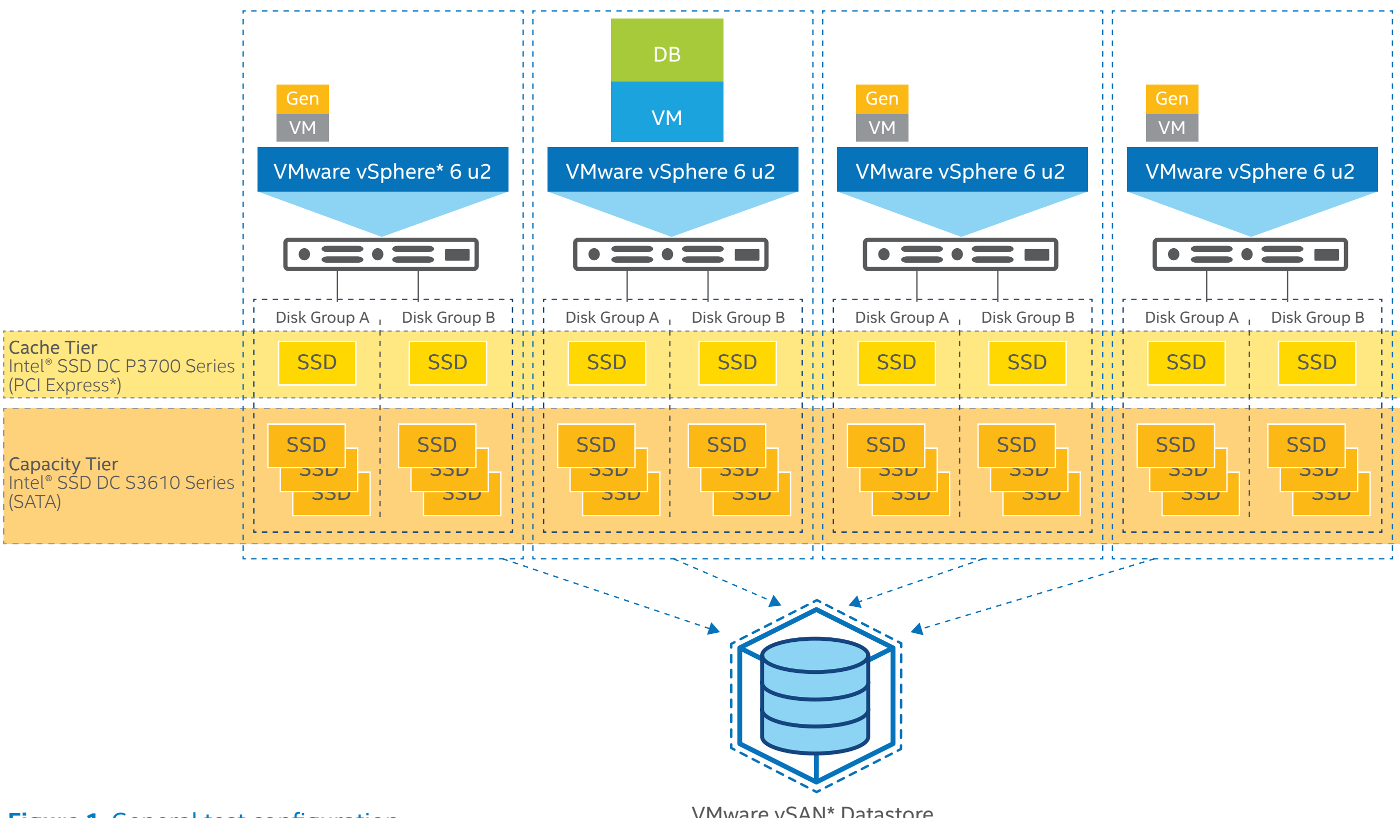

Figure 1. General test configuration.

This configuration is optimized using the combination of a very high-throughput cache tier based on the Intel SSD Data Center Family for PCIe and high-capacity SATA SSDs for a blend of high performance and capacity. For scale-out, distributed database tests with multiple application servers and one or more generators are used to drive one or more web and application servers for application processing that connects to the data server. An Enterprise Cache Protocol (ECP) test configuration is illustrated in Figure 2.

Further details of WebStress can be found in the Caché documentation.

System Configuration

The following configuration is used to compare scalability and response times:

- InterSystems Caché 2015.2
- Four servers based on Intel® Xeon® processors E5-2697 v3 (2.6 GHz) running VMware ESX 6 update 2
- Two vSAN disk groups per ESX host comprising two Intel SSD Data Center P3700 Series (PCIe) for the cache tier
- Six Intel® SSD Data Center S3610 Series (SATA) for the capacity tier
- One VM database server running Red Hat Enterprise Linux*

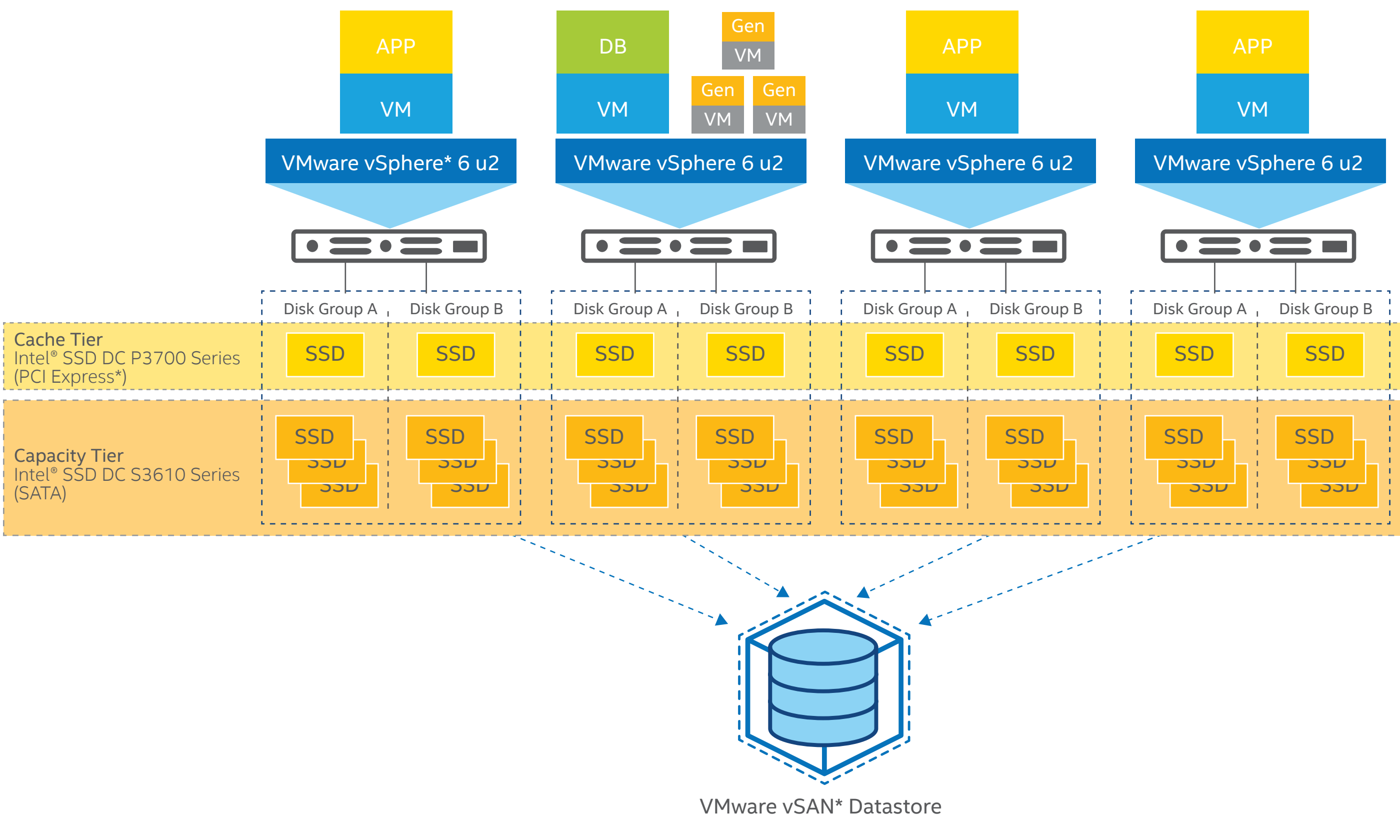

Figure 2. Enterprise Cache Protocol (ECP) test configuration.

Results

Single Database Instance Baseline Tests

Benchmark configuration for the single database baseline is as follows:

- Single database VM
- Single TrakCare database instance
- VMs for load generators and database server are on separate host servers

To test maximum throughput, the database VM is sized to consume as close to 100 percent of host CPU resources as possible. The database VM is sized at 32 vCPUs and 96 GB RAM.

Transaction throughput (global references, or glorefs) is linear, as shown in Figure 3. Test rate is driven by simple incremental steps from transaction generators. CPU utilization starts to peak at 100 percent for 2x. A gloref indicates the amount of work that is occurring on behalf of the current workload; although glorefs consume CPU time, they do not always require physical reads because of the buffer pool.

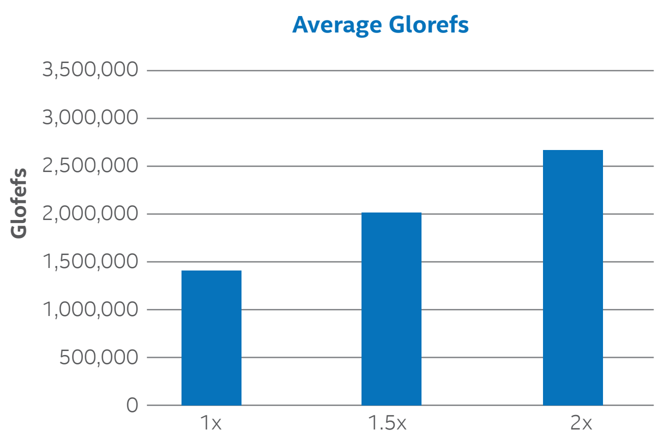

Figure 3. Single database instance results.

Scale-Out Application Server Baseline Tests

Scale-out architecture using Caché's ECP was tested using multiple application servers to drive a higher workload and utilize multiple VMs and hosts in the vSAN cluster.

ECP is a means to distribute processing load across multiple servers for massive horizontal scaling. It is a distributed data caching architecture that enables a scale-out deployment using multiple (up to 255) application servers. Data and code are stored remotely on a data server, but cached locally on application servers to provide efficient access for user sessions. The data server manages database reads and writes to persistent storage on disk, while multiple application servers perform most of the application processing. Communication between the tiers on separate servers is conducted using standard TCP/IP.

Similar to the single database instance tests, the scale-out architecture VMs are sized to consume as close as possible to 100 percent of host CPU resources. Each of the application server VMs is sized at 32 vCPUs and 96 GB RAM.

For benchmark testing, each application server runs the same transaction workload as in the single-database-server 1.5x test. Each application server added to the test increments the throughput by a further 1.5x. Three application servers were able to be added on this four-host vSAN cluster to a total of 4.5x total throughput, demonstrating efficient horizontal scaling, as shown in Figure 4.

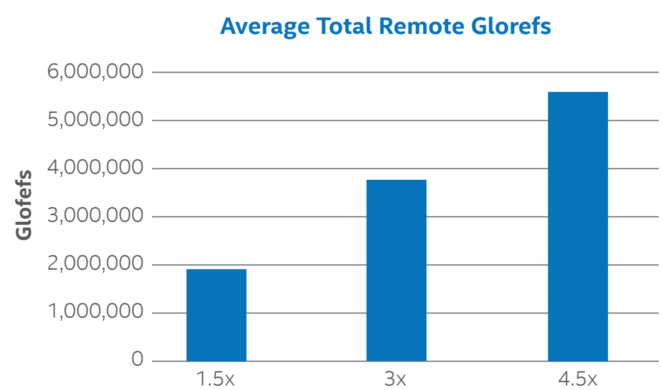

Figure 4. Scale-out application server results.

Random Read I/O Performance

Read I/O is very important for good application performance and end-user experience. This test uses Random Read IO Storage Performance Tool (RANREAD), a Caché database utility, to generate random read I/O against a Caché database, driving a set number of processes to achieve target random-read IOPS. Response time is monitored and is used to review whether acceptable disk latency is sustained by the storage system and at what I/O rate latency exceeds acceptable limits. The test does not update or write the blocks back to storage.

A 200 GB database is created before testing begins. RANREAD processes read random 8 KB blocks in the database while keeping a log of response times. As shown in Figure 5, the result is excellent, with more than 30,000 IOPS served with less than 0.7 ms latency. Please refer to InterSystems RANREAD test guidelines for further background.

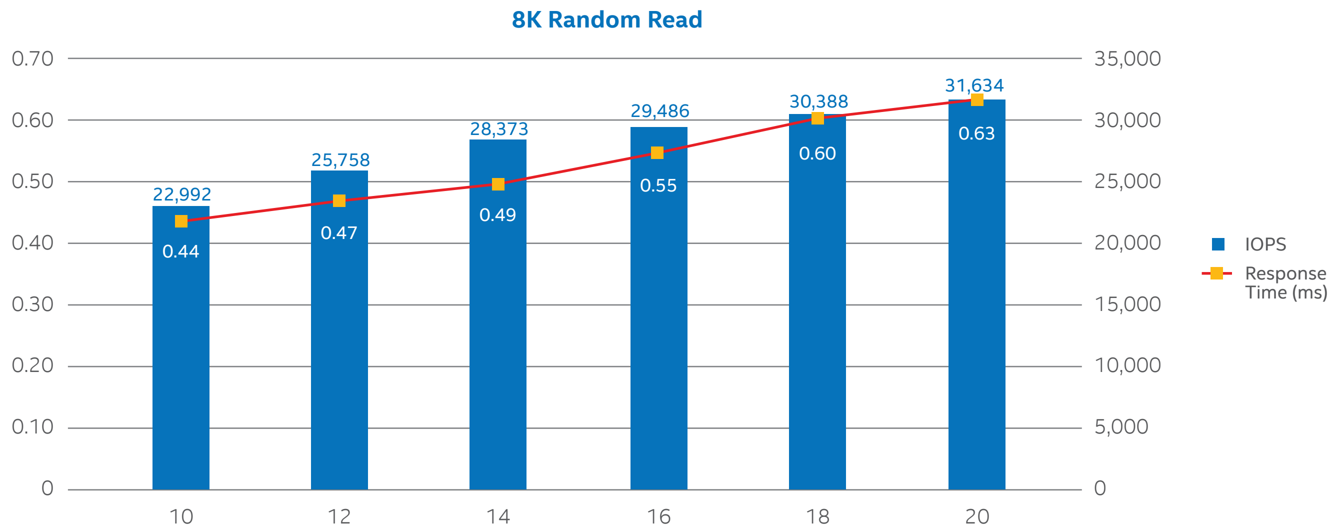

Figure 5. Random-read I/O performance.

Conclusion

In conjunction with vSAN, Caché provides a high-performance, scalable, reliable platform for patient-care applications. The Intel SSD Data Center Family for PCIe adds further to the quality of the solution, with sub-second response times, high dependability, and long-term durability. Using this solution stack, the healthcare industry worldwide is positioned to deliver forward-looking software products for analyzing and managing patient records while controlling costs.

Email us at **HealthcareCache@vmware.com** to discuss your needs, and be eligible to receive a **free vSAN assessment and TCO report**

[1] Analysis performed by VMware, November 2016. Published on VMware.com (http://www.vmware.com/products/virtual-san.html). Comparison between four-node, all-flash vSAN and EMC XtremeIO* all-flash array.

XtremeIO data based on published Lab Review, http://xtremioblog.emc.com/xtremio-tco-story-the-sequel.

[2] NVM Express, Inc. www.nvmexpress.org/wp-content/uploads/NVMe_Overview.pdf.

[3] The Intel® SSD Data Center Family for PCIe* is capable of reading data up to 2.8 GB/s and 460K IOPS and writing up to 2.0 GB/s and 175K IOPS. Source: Intel. "Consistently Amazing: Make the Switch to Faster Data." 2015. www.intel.com/content/www/us/en/solid-state-drives/intel-ssd-dc-family-for-pcie-brief.html.

[4] Configuration: Performance claims obtained from data sheet: Intel® SSD DC P3700 Series (2 TB), Intel SSD DC S3700 Series, Intel® Core™ processor i7-3770K CPU at 3.50 GHz, 8 GB of system memory, Windows Server 2012*, and Iometer*. Random performance is collected with four workers, each with 32 QD configuration for latency: Intel® Server Board S2600CP, two Intel® Xeon® processor E5-2690 v2, 64 GB DDR3, Intel SSD DC P3700 Series (400 GB), LSI 9207-8i*, and Intel SSD DC S3700 Series.

[5] Intel. "Intel Solid State Drive Data Center Family for PCIe*." https://www-ssl.intel.com/content/www/us/en/solid-state-drives/intel-ssd-dc-family-for-pcie.html. Configurations: Performance claims obtained from data sheet, sequential read/write at 128K block size for NVMe* and SATA, 64K for SAS. Intel® SSD DC P3700 Series, 2 TB, SAS Ultrastar* SSD1600MM, Intel SSD DC S3700 Series, SATA, 6 Gbps. Intel® Core™ i7-3770K processor, 3.50 GHz, 8 GB of system memory, Windows Server 2012*, Iometer*. Random performance is collected with four workers each with 32 QD.