

INTERSYSTEMS CACHÉ BENCHMARK: ACHIEVING MILLIONS OF DATABASE ACCESSES PER SECOND INEXPENSIVELY

Abstract

In a recent benchmark test of an application based on InterSystems Caché®, a sustainable rate of 8.9 million database accesses/second, with peaks of 16.9 million database accesses/second, was achieved. These results were from a test performed on a connected system of eight application servers, using Intel® Xeon® 5570 processors, and running Linux as the operating system.

This benchmark shows that:

- 1) Caché can achieve unheard of levels of performance for an object database. It provides full persistence of data at speeds that are normally only reached by in-memory databases.
- 2) Caché demonstrates this high performance while running on inexpensive servers.
- 3) Caché offers excellent horizontal scalability, with throughput scaling linearly as application servers are added.

ACHIEVING MILLIONS OF DATABASE ACCESSES PER SECOND

Introduction

In order to compete in the marketplace, modern software applications must be able to read and write large amounts of data at high speed. As applications are scaled up, the rate of database accesses required can reach several millions/second.

In-memory databases can, with enough expensive RAM, achieve that kind of throughput, but because they don't persist data, they lack the reliability required by "run-the-enterprise" applications.

As persistent relational databases scale to handle large volumes of data, their performance often degrades. This is largely because relational databases have to battle the "impedance mismatch" between modern object-oriented development technologies and the rows-and-columns data structures used on disk. The processing overhead required to "map" between the complex data types typically used by today's sophisticated software applications and a relational database tends to limit throughput.

The usual solution is to increase throughput by adding hardware – switching to larger, more powerful machines, and connecting many machines together in a "grid". The drawbacks to this approach are increased system complexity and expense. Large distributed systems can be difficult to manage, and large computers are not cheap.

The best solution is to use InterSystems Caché® as the underlying persistent database for applications. Caché is an object database, so it is an ideal match for modern object-oriented development technologies. (Caché is seamlessly accessible via ODBC and JDBC as well, so reporting tools that rely on SQL can be used without mapping.) The benchmark results presented in this paper will show that, even when running on inexpensive servers, Caché can provide throughput of millions of database accesses/**second** – enough to support the most demanding and large-scale software applications.

Methodology

We tested Caché's throughput using a commercially available connected hospital information system (HIS) based on Caché. This Web-based solution manages clinical and administrative functions within a hospital or group of hospitals. It is entirely object-based, with a very rich electronic patient record at its core. The application has been implemented at over 350 sites in 25 countries. The size of these installations varies, with the largest site serving a population of over six million people.

The benchmark utilized scripts that mimic typical use patterns for the HIS. The scripts were derived from observed use patterns, and the database was populated with scrambled patient data from one of the larger installations. The number of simultaneous users (i.e.: number of simultaneously running scripts) was increased while monitoring both throughput and response times until response times exceeded a pre-determined limit. At that point, both the sustained and peak rates of database activity were recorded.

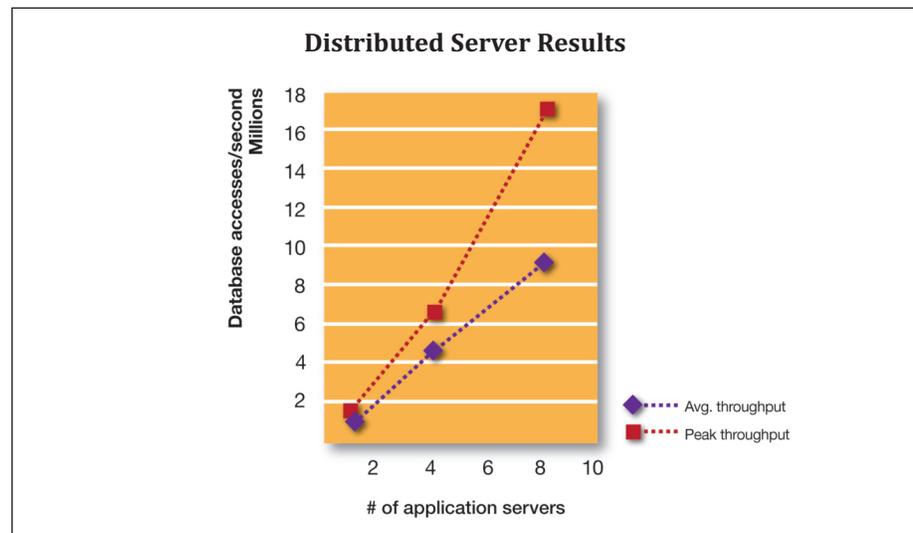
The benchmark was conducted at the Hursley IBM Test Center in the United Kingdom from July 13 – 24, 2009. The test used IBM HS22 blade servers, each with two quad-core Intel® Xeon® 5570 processors, for a total of eight cores. Both single server and various client/server configurations (ranging from one to eight applications servers) were tested. The operating system was 64-bit RedHat LINUX 5.3. Caché version 2009.1 was used throughout.

Results

The tables below shows the sustained, and peak, database access rates measured for several different system configurations:

Single Server Results		
Configuration	Avg. Throughput (database accesses/sec)	Peak Throughput (database accesses/sec)
8-core Database Server	1,190,000	2,021,000

Distributed Server Results (using Caché's ECP technology)		
Number of 8-core Application Servers	Avg. Throughput (database accesses/sec)	Peak Throughput (database accesses/sec)
1	992,000	1,690,000
4	4,100,000	6,165,000
8	8,900,000	16,900,000



Conclusion

Although it is a fully persistent object database, Caché demonstrated the extraordinarily high level of throughput normally associated with in-memory databases. In the tests of distributed systems, it scaled linearly from one to eight application servers. Furthermore, it achieved these results while running on relatively inexpensive hardware.

InterSystems Caché can provide the performance and scalability required to run complex, enterprise-wide software applications, without the need for lots of expensive hardware.

InterSystems Corporation

World Headquarters

One Memorial Drive

Cambridge, MA 02142-1356

Tel: +1.617.621.0600

Fax: +1.617.494.1631

InterSystems.com

INTERSYSTEMS