

DevForce Business Object Server Performance

Overview

The DevForce Business Object Server (BOS) is designed to operate in high-volume, transactional enterprise applications. It is effectively stateless, which gives it excellent scalability, fault tolerance, and load balancing characteristics. Along with performance-enhancing features in the DevForce infrastructure (such as client-side caching, span queries, and data compression), the Business Object Server enables DevForce applications to deliver high performance, scalability, and security in production environments serving thousands of concurrent users over the web.

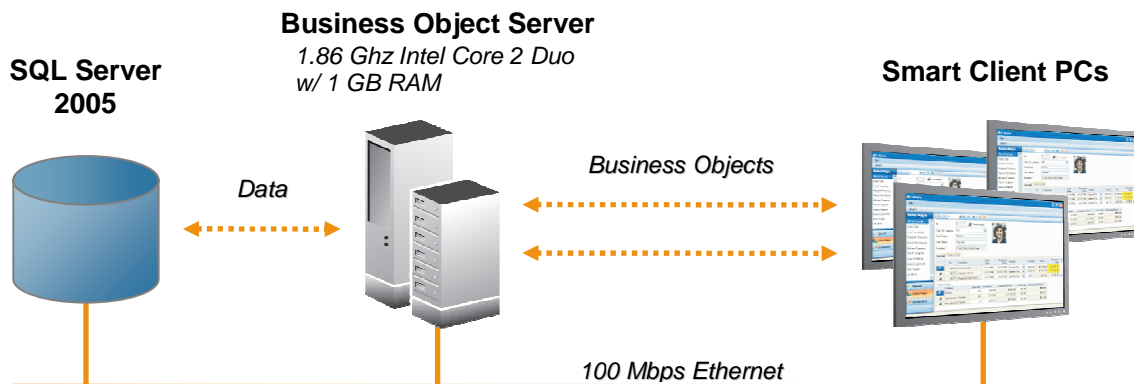
Performance testing of the Business Object Server was designed to stress the bottleneck of the server at two extremes:

- 1) When the server is object bound
- 2) When the server is query bound

The first condition occurs when clients are submitting queries that return large result sets and the server is bound by the serialization of large object graphs for transmission over the internet. The second condition occurs when clients are submitting large numbers of queries returning only one or a few objects and the query overhead dominates the processor time.

Test Setup

Performance testing of these extremes was done using the following setup:



As indicated, the hardware requirements for the Business Object Server are modest, requiring just a typical consumer desktop processor and only 1 GB of RAM. The Business Object Server can be deployed on Windows 2000 and above, but a server-grade operating system (such as Windows Server 2003) is required for high volume applications.

Performance Tests

For each of the tests, a cluster of smart client PCs was instructed to send queries to the Business Object Server as quickly as possible, simulating the effect of thousands of users concurrently working on an enterprise application. The number of PCs was increased until the load on the server reached 100% and the throughput of the server was recorded.

Bulk Query Test

This test determines the limit of server performance when it is object bound. Each query returned approximately 1,000 objects on average with each object containing about 19 database columns with representative data from an order tracking application. Under this scenario, the Business Object Server was able to process 23,500 objects per second.

Fine-grained Query Test

This test determines the limit of server performance when it is query bound. Each query returned approximately 3 objects on average with about 15 database columns per object (again with representative data from an order tracking application). Under this scenario, the Business Object Server delivered 350 queries per second.

Analysis

High Bulk Object Throughput

In scenarios where high volumes of data need to be transmitted to the client, the server is able to deliver about 23,500 objects/s/processor. Although this throughput is more than enough for most needs, developers can avoid this data transfer completely by performing the bulk processing near the data source on the server. The client application can initiate a server-side process using a DevForce Remote Procedure Call (RPC), which can perform the processing and optionally report the results back to the client.

Supports Thousands of Concurrent Users

The number of concurrent users supported by the Business Object Server will be dependent on a specific application's workflow. However, we can make some assumptions about typical application usage and derive an estimate for the number of supported users under different scenarios.

Most enterprise applications respond to user input and are waiting for the user to edit/enter data or analyze the results of the previous data before performing the next operation. The workflow of an application also tends to manipulate a contained subset of the data before moving on to the next set of data. This lends itself well to caching, and a number of objects during the workflow will be found in cache as opposed to re-querying the server.

We have found that a typical enterprise application may have only 25% of its queries actually contacting the database, with the remainder being satisfied from cache. Using this figure, and assuming a highly productive user is able to generate a query every 5



seconds, we estimate one true database query every 20 seconds per user or 0.05 queries/s/user. Given that the Business Object Server can process at least 350 queries/s/processor, we are able to estimate that the server will be able to support at least 7,000 users/processor for this application. Of course many applications will have different usage patterns; the above analysis can be repeated with whichever numbers make sense for each developer's particular application.

Scalability

Since the Business Object Server is effectively stateless, the server scales very well and there are a number of good options if additional performance is required:

- Use a high-end processor - Since the server is not memory-bound, the throughput scales well with increasing clock speed.
- Use more processing units – Moving to a quad core processor or a multi-socket configuration will increase throughput by about 80% for every doubling of processing units.
- Add more servers – This effectively adds more processing units but scales almost linearly because the memory subsystem is not shared among cores.

Summary

The Business Object Server scales well and has been proven in high-volume enterprise applications serving large numbers of concurrent users. Its fault tolerance and load balancing capabilities combine with caching and compression features in the DevForce toolset to enable enterprises to build distributed transactional applications with confidence.

#