

# Oracle9i OLAP Scalability Using Real Application Clusters

*An Oracle White Paper*

*June 2002*

# Oracle9i OLAP Scalability Using Real Application Clusters

## EXECUTIVE OVERVIEW

Organizations today need access to rich sets of analytical information. For years, traditional on-line analytical processing (OLAP) servers have been the mainstay for accomplishing that goal. But today, because intranet and extranet applications have introduced a new, broader class of users—both inside and outside the enterprise—organizations have an explosive need for analytical information that cannot be met by traditional, stand-alone database products.

Oracle9i Release 2 diverges from the traditional architecture of stand-alone multidimensional databases. Oracle9i Release 2 stores multidimensional data *directly in the Oracle database*. It executes all multidimensional calculations on this data directly inside the Oracle database engine. This unique degree of integration brings, for the first time, the benefits of RDBMS security, scalability, and resilience to OLAP users.

To explore the benefits of this unique new architecture, Oracle conducted the only massive, concurrent multiuser OLAP test using Real Application Clusters (RAC). This scalability test was designed to demonstrate how Oracle9i Release 2 OLAP Option benefits from the strengths of the Oracle database, providing levels of user scalability well beyond the reach of any competitive OLAP server.

The results of the study demonstrate that

- **Oracle9i Release 2 OLAP Option in a RAC environment can support a huge number of concurrent users** (5,000 in this study), simultaneously requesting a large volume of analytically sophisticated work
- **Oracle9i Release 2 OLAP/RAC combination scales economically and practically linearly**, thereby guaranteeing that a suite of analytical applications deployed in this environment will never run out of the power to support ever-increasing user populations as the needs of the enterprise grow
- **The Oracle solution far outperforms that of a prominent competitive provider of OLAP servers** in the area of user scalability

Oracle9i Release 2 diverges from the traditional architecture of stand-alone multidimensional databases. Oracle9i Release 2 stores multidimensional data *directly in the Oracle database*—bringing, for the first time, the benefits of RDBMS security, scalability, and resilience to OLAP users.

One of the primary requirements for analytic servers is the ability to support large numbers of concurrent users under a heavy workload of complex calculations. To date, there is no normative standard benchmark designed specifically to evaluate the user-scalability aspects of OLAP servers.

A new benchmark definition for OLAP servers was needed, one that would demonstrate enterprise scalability. This paper focuses on three key measurements: concurrent users, query throughput, and scale factor.

## INTRODUCTION

Traditionally, multidimensional calculation engines in the form of stand-alone database products have been used as on-line analytical processing (OLAP) servers. Examples of these types of multidimensional databases include Oracle Express Server, Microsoft SQL Server 2000 Analysis Services, and IBM DB2 OLAP Server. One of the primary requirements for analytic servers is the ability to support large numbers of concurrent users under a heavy workload of complex calculations.

To date, there is no normative standard benchmark designed specifically to evaluate the user-scalability aspects of OLAP servers. The known, publicly available tests that include concurrent, multiuser scenarios are these:

- In November 1999, Oracle published the results of the Analytical Processing Benchmark (APB) designed by the OLAP Council, now disbanded. In this test, Oracle launched 300 concurrent, constantly busy users (users with 0 wait time between queries), all running on four 300 MHz Sun UltraSparc II machines.
- In April 1999, Hyperion and IBM jointly published the same APB test. Their results show 200 concurrent, constantly busy users running on a 4CPU IBM RS/6000.
- In February 2001, Microsoft published a data-scalability benchmark that included multiuser runs. This test showed only 50 concurrent users, with 30-second wait intervals between queries. No information can be derived from this test that would indicate the maximum server query-throughput rates achievable for 50 concurrent, constantly busy users.

## Measuring the Enterprise Scalability of OLAP Servers

A new benchmark definition was needed, one that would demonstrate enterprise scalability. This paper focuses on three key measurements: concurrent users, query throughput, and scale factor.

- **Concurrent users:** The practical size of a sustainable community of concurrent, constantly active users. This was shown to be at least 5,000 on a 4-node RAC cluster.
- **Query throughput:** The maximum query-throughput rates a constantly busy server can achieve under the workload generated by these users. This was shown to be 665 queries per second.
- **“Scale up” factor:** The ratio of the combined throughput resulting from adding additional RAC nodes and a proportional number of constantly busy users. This was shown to be equal to x3.76 (276-percent growth), from 8 to 32 CPUs—an almost perfectly proportional increase.

With the help of an independent consulting and technology company, Oracle ran a limited number of multiuser experiments with a competitive OLAP product. The “scale up” factor from 1 to 4 constantly busy CPUs was 378-percent growth by Oracle and only 81-percent growth for the competitor.

*This drastic lack of user scalability in the lowest population ranges for the competitor’s product made further study of user scalability in the upper ranges totally impractical and was subsequently abandoned.*

## Oracle Versus a Competitive OLAP Product

As an adjunct activity, Oracle also ran a limited number of multiuser experiments with a competitive OLAP product. An independent consulting and technology company specializing in business intelligence, OLAP, analytical applications, and data warehousing was commissioned by Oracle Corporation to help with this portion of the study.

Specifically, the company was asked to deliver an optimally tuned OLAP implementation using the competitive product. It was also asked to observe some of the comparative runs, in order to verify and attest to these factors:

- Integrity of the results for both the competitor and Oracle solutions
- Fundamental compatibility of the hardware platforms used

In this segment, the sample query set was shown to execute over 8 times faster by Oracle than by the competitor on comparable hardware, and over 4 times faster by Oracle than the competitor on a variety of client platforms.

More revealing was the user-scalability aspect of the test: On a mere 16 concurrent, constantly active users, Oracle executed queries with over 22 times the competitive throughput. The “scale up” factor from 1 to 4 constantly busy CPUs was x4.78 (378-percent growth) by Oracle, and a very disappointing x1.81 (or 81-percent growth) for the competitor. **This drastic lack of user scalability in the lowest population ranges for the competitor’s product made further study of user scalability in the upper ranges totally impractical and was subsequently abandoned.**

## ENTERPRISE SCALABILITY TEST CONFIGURATION

The test consisted of a standard 2-Tier Client Server Topology. No end-user query tools were used on the client machines except to log throughput results.

### Test Data

The schema used for this study was a typical Sales and Marketing database with the following dimensions: PRODUCT (10,000 members, 7 levels), CUSTOMER (1000 members, 3 levels), CHANNEL (10 members, 2 levels) and TIME (34 members, 3 levels representing 18 months of historical data). This is intended to be a reasonable representation of a departmental OLAP application.

Below is a table describing cubes, actual measures created for the test, their dimensionality, and the representative row counts.

**Table 1: Enterprise scalability test cube descriptions**

Cube	Measures	Dimensions	Aggregated row count	Leaf row count
Actual	Sales, Units	Product,	377,477,604 (Sales)	65,610,000 (Sales)

Cube	Measures	Dimensions	Aggregated row count	Leaf row count
		Customer, Channel, Time	377,477,604 (Units)	65,610,000 (Units)
B & F	Budget, Forecast	Product, Customer, Channel, Time	64,636,866 (Budget) 12,816 (Forecast)	31,180,500 (Budget) 7,680 (Forecast)
Inventory	Inventory	Product, Customer, Time	100,546,236	46,770,750
Product Cost	Cost	Product, Time	36,000	21,600
Shipping Cost	Cost	Customer, Time	360,000	216,000

### Enterprise-Scalability Test Queries

Competitors, whether unable or unwilling to publish APB results, complained that the APB query set represents much too simplistic a set of queries to be considered representative of a true OLAP user environment. Accepting the challenge, Oracle enlarged the APB query set for this test by introducing more analytically comprehensive queries into the mix, in order to produce a more fitting simulation of a true OLAP environment.

The full query set consisted of 20 query types—10 of which existed in APB specification—and 10 new query types. A list of query types used in the test is included as an appendix to this paper.

A random number generator was used to create 11,053,200 query samples, of which no 2 would have the same type/predicate set. The 11,053,200 unique query samples were then broken up into 5,000 user streams, which provided the input to the test. The streams were further subdivided into three groups: 4,000 streams of 2,500 queries each, representing “casual users” and consisting of quicker and simpler queries; 800 streams of 1,254 queries each, representing “managers” and consisting of more-complicated queries that also select more data; and 200 streams of 250 queries each, representing “analysts” and consisting of more-complex queries that select large volumes of data. Additional information on the query mix is available in a supplemental technical paper, “Oracle9i OLAP Scalability Technical Description,” to be published shortly.

The hardware used in the enterprise-scalability test is described in the table below.

### Table 2: Enterprise-Scalability Test Hardware

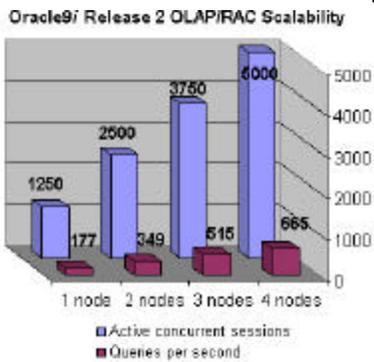
Database Tier	Client Tier
4 x HP N-Class, each with: 8 x 550 MHz 32 GB RAM 51 GB HP FC10 local disk 5.36TB HP FC60 shared disk storage among all four database servers HP UX 11.0, December 2000 Patch bundle ServiceGuard OPS Edition A.11.09 PHNE_26551 Patch for HMP Interconnect for RAC: HMP (Hyper Messaging Protocol) over HyperFabric II HyperFabric 9000 B.11.00.12 Oracle9i Release 2 64bit, OLAP and Real Application Clusters options	1 x HP N-Class, each with: 8 x 550 MHz CPU 32 GB RAM 34 GB HP FC10 local disk HP UX 11.0, June 2000 Patch bundle

### ENTERPRISE SCALABILITY RESULTS

The table below summarizes the results for each node.

**Table 3: Oracle9i Release 2 OLAP/RAC Enterprise Scalability**

Nodes	Users	User Type Mix	Average Queries/Sec.	Scale Factor	Scale % difference
1 Node	1250	1000+200+50	177		
2 Nodes	2500	2000+400+100	349	1.97	97 %
3 Nodes	3750	3000+600+150	515	2.91	191 %
<b>4 Nodes</b>	<b>5000</b>	4000+800+200	<b>665</b>	<b>3.76</b>	<b>276 %</b>



Oracle9i Release 2 OLAP Option in a RAC environment supports 5,000 concurrent, constantly busy users requesting a large volume and variety of analytically sophisticated work.

Oracle9i Release 2 OLAP/RAC combination scales practically linearly.

### TESTING DEPARTMENTAL SCALABILITY AGAINST A COMPETITIVE PRODUCT

As mentioned earlier, Oracle also conducted a test of a competitive OLAP product in an attempt to evaluate its user-scalability characteristics, such as “scale up” factors and the maximum sustainable concurrent user population. A test was done on a much smaller scale, involving only one type of user stream (analyst), and only up to 16 concurrent, constantly busy users.

Both Oracle and the competitive vendor solution were tested in a single user (1 query stream) and multiuser (16 query streams) scenario, with the intention of measuring the itinerant server query throughput and average query response times on increasing loads.

In each case, each query stream consisted of 1,360 queries whose predicates were selected randomly so no two queries would be the same. The queries were submitted with zero waits. Each stream had the exact same percentage mix of queries, as summarized in the following table.

**Table 4: Competitive Scalability Test Query Description**

Query Type	Number Queries
Random Cell Select	100
3 Month Moving Average	100
Simple Percentage Ranking	145
Percent Change Percentage Ranking	145
Market Share Percentage Ranking	145
Product Share Percentage Ranking	145
Simple Top 10 Ranking	145
Percent Change Top 10 Ranking	145
Market Share Top/Bottom 10 Ranking	145
Product Share Top/Bottom 10 Ranking	145
<b>Total queries</b>	<b>1360</b>

**Competitor's configuration:** The competitor's OLAP server was installed and configured on the server host system, a Dell PowerEdge 6300. Each client was run on a separate Dell PowerEdge 6300 server and servers from other manufacturers. The client query driver includes a flag that assigns the client to a particular set of CPUs, and in this test the clients were started so that each driver on each server dedicated one CPU to each query stream. Windows NT 4.0 EE was the operating system used for both client and server platforms.

**Oracle configuration:** Oracle9i Release 2 was installed and configured on the server host system, a Sun Microsystems E450 4x400MHz, 4GB RAM server. The clients that produce the query load were hosted on a single Sun Microsystems Ultra 60 2x450Mhz 1GB RAM workstation. Solaris was the operating system used at Oracle for both client and server platforms.

## Vendor Implementations and Optimization

Both the Oracle9i Release 2 and competitor implementations were optimized using a range of techniques. An expert, outside consulting firm was hired to optimize the competitor's implementation. In this instance, they were asked to write queries using best-known techniques, and to investigate all other possible best practices involving advanced calculations, and all sorts of server- and client-tuning options.

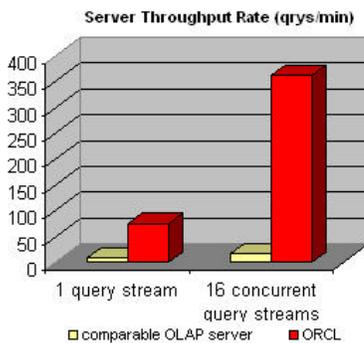
Both hardware systems had similar configurations (the same quantity of RAM, CPU, and clock speed). Although different server operating systems were used, both environments were optimized at multiple levels (query syntax, database partitioning, execution location, data retrieval, and so on) according to expert and industry-standard best practices.

For more information on the test queries and optimization best practices, refer to the supplemental technical paper titled, "Oracle9i OLAP Scalability Technical Description," to be published shortly.

## Competitive Server-Throughput Results

As seen below, Oracle delivers better query throughput than the competitor. In the single user test, Oracle's sustained query throughput is more than 8 times faster than comparable OLAP server sustained query throughput. As user load scales up to 16 concurrent streams, **Oracle's sustained query throughput is more than 22 times faster than a comparable OLAP server's sustained query throughput**.

**Table 5: Single and Concurrent User Sustained Server-Throughput Results**



Product	Concurrent Query Streams	Elapsed Time on Query Test (1 <sup>st</sup> user start – last user finish in secs)	Total Queries executed	Sustained Server Throughput (qrys/min)
Comparable OLAP server	1	9183	1360	<b>8.886</b>
	16	80895	21760	<b>16.139</b>
Oracle9i Release 2	1	1079	1360	<b>75.626</b>
	16	3611	21760	<b>361.562</b>

Although the large difference in one user run is surprising, and possibly revealing of architectural differences between Oracle and the product being tested, the result more relevant to the present test of user scalability is the apparently poor "scale up" factor exhibited by the competitive product.

**Table 6: Competitive Scalability Test Scale Factors**

	Users	Average queries per sec	Scale factor	Scale % difference
Comparable OLAP server	1	0.15		
	16	0.27	1.81	81 %
Oracle9i Release 2	1	1.26		
	16	6.03	4.78	378 %

Both systems had 4 CPUs available for utilization. As the number of concurrent, constantly active sessions went from 1 to 16, one would expect, from a well-architected server product, that the combined throughput should be increasing near 4-fold (as happened with Oracle software). The competitive product displayed very disappointing user scalability indeed.

**Client Platform Effects**

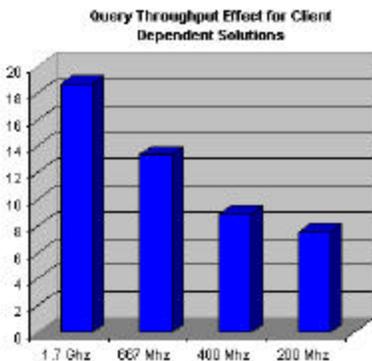
As a sidebar experiment on the competitor’s solution, different client configurations were run in the single user test in order to quantify the effect on the overall system. Four client configurations were chosen representing configurations found in office settings. They were

- 200 MHz CPU (low-end)
- 400 MHz CPU (low-mid-range)
- 667 MHz CPU (mid-range)
- 1.7 GHz CPU (high-end)

The results clearly show a significant relationship of client configuration to system-performance characteristics.

**Table 7: Query Test (Single User) – Execute 1360 queries in a single stream**

Comparable OLAP server	Elapsed Time (secs)	Queries/Sec.	Queries/Minute
1.7 Ghz CPU Client	4359	0.312	<b>18.720</b>
667 Mhz CPU Client	6105	0.223	<b>13.336</b>
400 Mhz CPU Client	9183	0.148	<b>8.886</b>
200 Mhz CPU Client	10825	0.126	<b>7.538</b>



It is interesting to note that despite being equipped with a top-of-the-line client platform, the competitor's solution still returns a single user run that is slower than Oracle's multiuser run. **Oracle's sustained query throughput (75.626 queries/minute) is more than 4 times faster than the best-sustained query throughput of a comparable OLAP server's query throughput on various client platforms.**

## CONCLUSION

The results speak for themselves. Oracle9i Release 2 delivers significantly faster query-throughput rates and sustains that high performance as user load scales up. Also, in a discussion of deployment implications of the respective architectures, it is important to understand that Oracle's true server solution is significantly more conducive to Web implementations because it efficiently handles high concurrent connection volumes and provides a more predictable user experience.

With the full integration of its multidimensional technology into its class-leading database server, Oracle brings the analytic power of OLAP to enterprise-scale applications. More importantly, it does so without sacrificing any of the performance characteristics of departmental applications within which OLAP technology has its roots. What this means in real terms is that businesses can now seamlessly manage and analyze data among all departmental functions within the business enterprise—and do it with an infrastructure that facilitates consolidation and summary management at governing executive levels.

## APPENDIX

Below is a list and description of the query types used in the enterprise-scalability test.

#	Query Type	Description
1	Random Cell Select	Select actual sales for randomly selected lists of products, customers, channels and time periods, all from random mix of hierarchical levels
2	3 Month Moving Average	Select actual sales and three month moving average of sales for randomly selected lists of products, customers and channels, and for all month level time periods, products, customers and channels all coming from a random mix of hierarchical levels
3	Simple Percentage ranking	Select actual sales for top 10% of products at CLASS level, based on actual sales, for each tuple from among randomly selected lists of time periods, customers and channels, all coming from a random mix of hierarchical levels
4	Simple Top 10 Ranking	As above, except top 10 products at CLASS level is being selected
5	Market Share Percentage Ranking	Select Market Share of each product CLASS for top and bottom 10% of product classes in terms of their share, for each tuple from among randomly selected lists of time periods, customers and channels, all coming from a random mix of hierarchical levels
6	Market Share Top/Bottom 10 Ranking	Same as above, except top and bottom 10 instead of 10% is being asked for
7	Product Share Percentage Ranking	Select market share of product CLASS jointly held by top 10% of customer RETAILERS in terms of their total year-to-date actual sales, for top and bottom 10% of product classes in terms of this share, for each tuple from among randomly selected lists of time periods and channels, all coming from a random mix of hierarchical levels
8	Product Share Top/Bottom 10 Ranking	Same as above except top and bottom 10 instead of 10% were used
9	Percent Change Percent Ranking	Select sales percentage change from the prior period for top 10% of product GROUP in terms this change
10	Percent Change Top 10 Ranking	Same as above, except the selection is for top 10, not 10%
11	Channel Sales Analysis	Select actual units sold, sales, and Average Price for a randomly picked single channel, and for a selection of products, customers, and channels each consisting of all children of a randomly picked member
12	Customer Margin Analysis	Select actual sales, cost and margin for a randomly selected single customer, for the sum of all channels, for a randomly selected time period and for all the children of a randomly selected product
13	Product Inventory Analysis	Select actual sales, cost and inventory for a given product sold through all the channels, for all children of a randomly picked customer, and for a bunch of months in a specific range of dates

#	Query Type	Description
14	Time Series Analysis	Select actual sales and a 6-month moving average of smoothed sales for a given Customer, children of a randomly picked product, children of a randomly picked time period, and a list which includes all channels plus the combined total of all channels
15	Customer Budget	Select budgeted units, actual sales, average price, cost, and margin for all twelve months in a given year, for a randomly picked customer for all products
16	Product Budget	Select budgeted units, actual sales, average price, cost, and margin for all quarters in a given year, for randomly selected product for all customers.
17	Forecast Analysis	Select forecasted units, actual sales, average price, cost, and margin for a randomly picked time period and all children of a randomly picked product and a randomly picked customer
18	Budget Performance	Select budget vs. actual (sales, difference and percent difference), as well and this year vs. last year-to-date sales (plus difference plus percent difference) for the given month and given as-of month, for a randomly picked customer and product.
19	Forecast Performance	Select forecast vs. actual (sales, difference and percent difference), as well and this period vs. previous period (sales, difference plus percent difference) for the given month and given as-of month, for a randomly picked customer and product.
20	Random cell select of various measures	Select one cell off a randomly picked measure from among the available stored or calculated measures, for a randomly picked product, customer, channel and time period



Oracle9i/OLAP Scalability Using Real Application Clusters

June 2002

Authors: Doug Liles, Karlene Jensen

Contributing Authors: Gregory Dorman, Grace Mateos, Charles Sperry

Oracle Corporation

World Headquarters

500 Oracle Parkway

Redwood Shores, CA 94065

U.S.A.

Worldwide Inquiries:

Phone: +1.650.506.7000

Fax: +1.650.506.7200

[www.oracle.com](http://www.oracle.com)

Oracle is a registered trademark of Oracle Corporation. Various product and service names referenced herein may be trademarks of Oracle Corporation. All other product and service names mentioned may be trademarks of their respective owners.

Copyright © 2002 Oracle Corporation

All rights reserved.