An Oracle White Paper
February 2011

Oracle Internet Directory 11g and Oracle Exadata Database Machine in the Facebook Age

Disclaimer

Table of Figures and Graphs:

Introduction

The meteoric rise of social media services, and the recent announcement of Facebook crossing the 500 million user mark provided a welcome opportunity to kick off a benchmark with Oracle Internet Directory 11g using the latest Oracle Exadata Database Machine. Our intention was to revisit the results of our previous two billion user benchmark[1], and measure how hardware and software improvements affect performance of a 500 million user directory. At the same time we wanted to prove that the powerful combination of OID and Oracle Exadata can be the infrastructure backbone to handle challenges posed by large scale user populations.

Telecommunication, Internet Service Providers, and social media services are some areas which deal on a daily base with an ever growing number of users who were concurrently accessing their systems to:

- Register for or update service subscriptions, or update their profiles
- Buy mobile applications and services
- Search for information

The directory is in the hot spot to:

- Identify and authorize callers in the 'call path' when they place a phone call, and verify they can access a particular service e.g. send a multi media message.
- Manage large group memberships with millions of users, and requires search, add and delete operations

Providing non-disruptive, fast, reliable and scalable authentication and authorization services that can easily be maintained and integrated into an existing infrastructure is an essential business requirement to address these demands.

In the following we outline that the unique architecture of Oracle Internet Directory and Oracle Exadata Database Machine is another proof for "Hardware and Software, Engineered to work together" that provides a solution for those challenges. Before we cover the details here are some highlights of the 500 million user directory benchmark:

- **433,684** search ops/sec
- **314,861** group lookup ops/sec for groups with 10 million members
- **100,000** concurrent clients without loss of throughput

[1] See Appendix D for details

Oracle Internet Directory Details

Oracle Internet Directory (OID) implements a unique architecture which enables the directory to fully utilize the underlying server hardware, scale on any given hardware, and at the same time provide high availability.

This architecture provides the following benefits:

- Multi-threaded server processes using DB connection pooling enable each server process to open and maintain simultaneous access the DB, and at the same time minimize system resources.
- Multi-processing utilizes existing CPU's and is NUMA aware for efficient memory access.
- Multiple-instances of directory servers use multiple hardware nodes, and multiple listening endpoints for high concurrency with separated application traffic to dedicated LDAP endpoints.
- OID scales with the number of CPU's in SMP architectures.
- OID scales with the number of nodes in hardware cluster architectures.
- No replication or partitioning is required to scale horizontally.

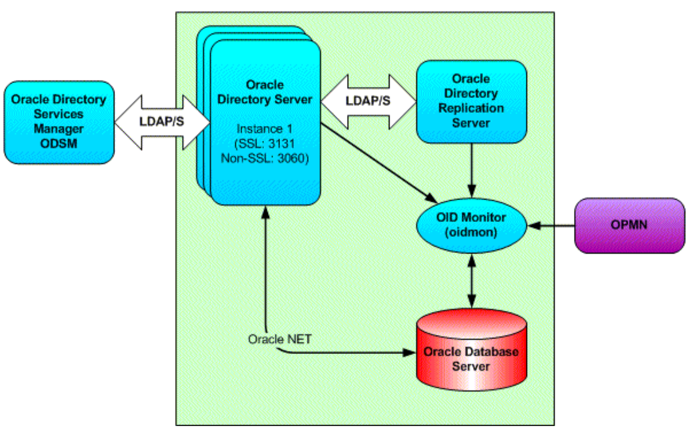

Figure 1: Oracle Internet Directory Typical Node Architecture

In a typical deployment one or more OID servers, together with the OID replication server, are usually collocated with the DB instance on the same physical host.

For this benchmark we took advantage of OID's flexible deployment options. To achieve best performance with Oracle Exadata Database Machine, the OID servers were deployed on dedicated hosts (see Figure 2) to maximize host resources, and to simultaneously provide high availability without replication.

For further details regarding Oracle High Availability options please refer to the *Oracle® Fusion Middleware High Availability Guide* and the *Oracle® Identity Management Deployment Guide for Oracle Identity Management*. Both guides outline recommended deployment architectures, and serve as a blueprint for an Enterprise deployment.

Oracle Exadata Database Machine X2-2 Overview

With only 10 days to conduct this benchmark, we knew from previous benchmarks that we would need to spend a significant amount of time on hardware configuration and DB tuning before we could even start data loading. These tasks range from Operating System tuning to database storage configuration using Solid State Disks.

The Oracle Exadata[2] Database Machine in combination with the Exadata Storage Servers, simply eliminates these tasks by replacing isolated special-purpose systems with a consolidated platform to deliver leading performance and scalability for all database applications.

The Oracle Exadata Database Machine provides an optimal solution for all database workloads, ranging from scan-intensive data warehouse applications to highly concurrent OLTP applications. With its combination of smart Oracle Exadata Storage Server Software, complete and intelligent Oracle Database software, and the latest industry standard hardware components from Sun, the Database Machine delivers extreme performance in a highly-available, highly-secure environment.

With Oracle's unique clustering and workload management capabilities, the Database Machine is also well suited for consolidating multiple databases onto a single grid. Delivered as a complete pre-optimized and pre-configured package of software, servers, and storage, the Sun Oracle Database Machine is simple and fast to implement and ready to tackle your large-scale business applications.

The Database Machine uses the Oracle Exadata Storage Servers, which are highly optimized for use with the Oracle database. Exadata delivers outstanding I/O and SQL processing performance for data warehousing applications by leveraging a massively parallel architecture to enable a dynamic storage grid for Oracle Database 11g deployments.

Exadata is a combination of software and hardware used to store and access the Oracle database. It provides database-aware storage services, such as the ability to offload database processing from the database server to storage, and provides this while being transparent to SQL processing and your database applications. Exadata storage delivers dramatic performance improvements, with unlimited I/O scalability, is simple to use and manage, and delivers mission-critical availability and reliability to Oracle Internet Directory and your enterprise

[2] See Appendix D for further technical details

Benchmark Deployment Architecture

Distributed Architecture

As outlined in the Oracle Internet Directory introduction, we used a distributed architecture.

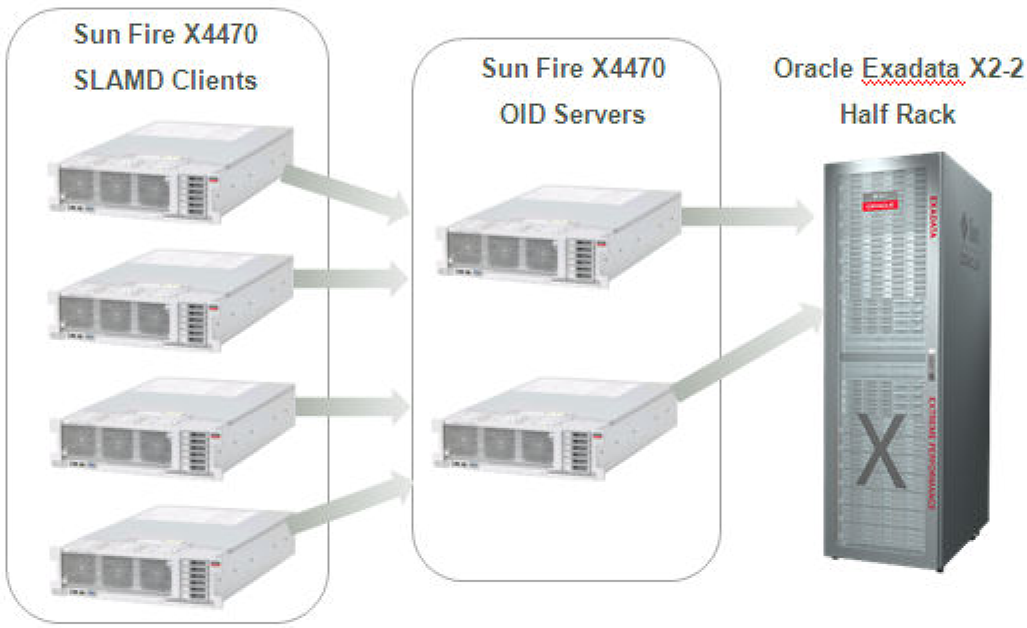

Figure 2: Oracle Internet Directory Distributed Deployment Architecture

To run the benchmark we used the SLAMD distributed load generation engine. SLAMD clients were distributed across four Sun Fire X4470, and connected to the SLAMD server. The SLAMD server distributed the load across two Oracle Internet Directory Servers, which were also deployed on two Sun Fire X4470. The Oracle Exadata Database Machine X2-2 with a half rack configuration served as a backend server. All systems were connected via a 10 Gigabit network.

A characteristic feature for OID is the use of a single directory information tree (DIT). No partitioning is required to scale with the amount of data. Just add further OID server nodes, if required, to scale horizontally, or add Exadata Database, and Exadata Storage Servers to meet increasing data volume.

Database Tuning

Out of the box configuration was retained with the following modifications:

Parameter	Value	Description
SGA_MAX_SIZE	48GB	Configurable upper bound to which sga_target_size can dynamically adjust for the System Global Area (SGA).
SGA_TARGET	48GB	Target size for the System Global Area.
PGA_AGGREGATE_TARGET	2GB	Target size for the Program Global Area.
audit_trail	none	Configures database auditing.
processes	2000	Number of database processes.
db_file_multiblock_read_count	16	Number of blocks to read in a single I/O operation during a sequential scan.
open_cursors	1000	Maximum number of cursors a database session can have.

OID Tuning

We retained the out of the box OID configuration for this benchmark. That meant that all default access controls and password policies were effective and in play. However, to best utilize the hardware resources at hand, we performed minimal tuning by changing the following OID configuration parameters.

Parameter	Value	Description
Orclserverprocs	32	One server process per CPU core. Since there were 32 cores on system, we keep one process per core.
Orclmaxcc	4	Number of worker threads per server process. This was the optimal configuration for this hardware.
Orclgeneratechangelog	0	Change log generation turned off because no replication was configured. This only affects modify and add tests.
Orclskiprefinsql	1	Skip referral processing in SQL enabled, since

		there are no referral entries in this test.
Orclmatchdnenabled	0	Matched DN enabled, this does not affect successful search.

Data Characteristics

LDAP

500 million LDAP user entries were generated using the SLAMD makeLDIF template shown in Appendix C. These entries had the following properties:

- There were 26 attributes (7 generated operational attributes), of which
  - One was a binary attribute with length 2KB
  - The others had cumulative length of 0.8KB
- Total entry size was 2.8KB

These properties were chosen to represent a real-world deployment where the user entry might contain binary information such as a photograph or cryptographic identifiers like certificates.

In addition, 50 group entries were created with the following properties:

- Each group had 10 million members each.
- Each user is a member of one group.
- Each group entry was 450MB in size.

These properties were chosen to represent the realistic grouping requirements in a deployment of this size.

Note: <u>No data partitioning</u> was necessary to accommodate this workload scenario.

Database

The database representation of this data had the following characteristics:

TableSpace Name	Size in GB
OLTS_ATTRSTORE	1032
OLTS_CT_STORE	890
OLTS_BATTRSTORE	1436
SYSAUX	1.3
OLTS_DEFAULT	0.001
OLTS_SVRMGSTORE	0.0002

Workload Scenarios

Two deployment workload scenarios were tested:

- **500 Million User Deployment With Sparse Load-**
  This scenario examines the case of a large deployment where only a subset of the users in the deployment are active at given time. For the purposes of this benchmark, the subset was stipulated to be the first 10 million users. Therefore in our tests benchmarking this scenario we targeted 10 million out of the total 500 million users deployed.
- **500 Million User Deployment With Exhaustive Load-**
  This scenario examines the case of a large deployment where all the users in the deployment are active at a given time. Therefore in our tests benchmarking this scenario we targeted all 500 million users deployed.

Test Scenarios

Each of the following read tests was executed against both of the above workload scenarios.

Sequential LDAP Search Operations Test

This test scenario involved concurrent clients binding once to OID and then performing repeated LDAP Search operations. The salient characteristics of this test scenario was as follows:

- SLAMD ‘SearchRate’ job was used.
- BaseDN of the Search was root of the DIT. The scope was SUBTREE. The search filter was of the form “UID=<a unique value>”. DN was the required attribute to be returned.
- Each LDAP search operation matched a single entry.
- The total number of concurrent clients was 3200 and was distributed amongst 4 client nodes

- Each client bound to OID once and performed repeated LDAP Search operations. Each Search operation resulted in the lookup of a unique entry in such a way that no client looked up the same entry twice and no two clients lookups were the same.
- Test job was run for 60 minutes.

Random LDAP Search Operations Test

This test was the same as the previous test, except that the entries were looked up randomly. Additionally this test was run with 500 concurrent clients against both workload scenarios to determine low latency metrics. Furthermore, this test was run with 32,000, 64,000 and 100,000 concurrent clients in the exhaustive workload scenario to showcase OID's ability to handle large volumes of concurrent connections.

Sequential LDAP Authentication Operations Test

This test scenario involved concurrent clients binding once to OID and then performing repeated LDAP Compare operations. The salient characteristics of this test scenario were as follows:

- SLAMD 'AuthRate' job was used.
- The BaseDN of the Search was root of the user container. The scope was BASE. The search filter was of the form "UID=<a unique value>". The DN was the required attribute to be returned.
- Each LDAP search operation matched a single entry.
- All entries had the same userpassword value.
- The total number of concurrent clients was 3200 and they were distributed among four client nodes.
- Each client bound to OID once and performed repeated LDAP searches followed by bind operations. Each Search operation resulted in the lookup of a unique entry in such a way that no client looked up the same entry twice and no two clients lookups were the same.
- Test job was run for 60 minutes.

Random LDAP Authentication Operations Test

This test was the same as the previous test, except that entries were looked up randomly. Additionally this test was run with 500 concurrent clients against both workload scenarios to determine low latency metrics.

Sequential LDAP Compare Operations Test

This test scenario involved clients repeatedly executing the sequence of performing an LDAP search operation to look up a user and performing a simple bind as that user to verify its credential. The salient characteristics of this test scenario were as follows:

- SLAMD ‘CompRate’ job was used.
- Each compare assertion was done on the userpassword attribute which created additional overhead for handling hash comparisons and password state policy enforcement.
- The total concurrent clients was 3200 and they were distributed among four client nodes.
- Each client bound to OID once and performed repeated LDAP compare operations, each Compare operation performed on unique entry in such a way that no client examines the same entry twice and no two clients lookup the same.
- The test job was run for 60 minutes.

Random LDAP Compare Operations Test

The same as above, except that entries are looked up randomly. Additionally this test was run with 500 concurrent clients against both workload scenarios to determine low latency metrics.

The following large group-related tests were also performed.

LDAP Group Search Operations Test

This test scenario consisted of concurrent clients binding once to OID and then performing repeated LDAP search operations. The salient characteristics of this test scenario were as follows:

- SLAMD ‘SearchRate’ job was used.
- The total number of concurrent clients was 3200 and they were distributed amongst four client nodes.
- Each client performed a subtree search with the filter member=’RandomDN’.
- Each lookup returned one group entry.
- The required attribute was set to DN.
- The test job was run for 60 minutes.
- A search performed such that all the 50 group entries were touched randomly.

LDAP Group Modify Operations Test

This test scenario consisted of concurrent clients binding once to OID and then performing repeated LDAP modify operations. The salient characteristics of this test scenario were as follows:

- A custom JNDI script was used.
- A total of 10 concurrent LDAP clients were used.
- The uniquemember attribute value was added and deleted.
- The test job was run for 60 minutes.
- A modify was performed on large groups that had 10 million members.

The following write tests were also performed.

Sequential LDAP Modify Operations Test

This test scenario consisted of concurrent clients binding once to OID and then performing repeated LDAP Modify operations. The salient characteristics of this test scenario weres as follows:

- SLAMD 'LDAP ModRate' job was used.
- A total of 200 concurrent LDAP clients were used.
- Each client updated a unique, sequentially selected entry each time, and a total of 54 million entries were updated.
- The test job was run for 60 minutes.
- The value length was set to 11.
- The attribute that was modified was not indexed.
- All modify operations were routed to a single database instance.

Random LDAP Modify Operations Test

This test scenario was the same as the previous scenario, except that entries were looked up randomly and 43 million entries were updated.

LDAP Add Load Test

The test scenario involved concurrent clients adding new entries as follows:

- A SLAMD 'LDAP Add Rate' job was used.
- A total of 500,000 entries were added.
- A total of 500 concurrent LDAP clients were used.
- Slamd added an 'inetorgperson' objectclass entry with 21 attributes (including operational attributes).

LDAP Mixed Load Test

The test scenario involved concurrent clients performing a mix of LDAP operations. The characteristics were as follows:

- A SLAMD 'LDAP Mixed Load' job was used.
- The LDAP client operations were 65% search, 16% bind, 16% compare and 3% modify.

- There was a total of 3200 concurrent LDAP clients. They were distributed on four client nodes.
- The test job was run for 60 minutes.

Benchmark Results

Sparse Workload (targeting 10 Million of 500 Million Entries)

Test Scenario	Number of Clients	Through put (ops/sec)	Latency (msec)
Search (Sequential)	3,200	441,029	7.189
Search (Random)	3,200	441,058	7.188
- *low-latency*	500	336,506	1.481
Auth (Sequential)	3,200	163,157	19.557
Auth (Random)	3,200	161,978	19.698
- *low-latency*	500	137,573	3.629
Compare (Sequential)	3,200	285,969	11.091
Compare (Random)	3,200	286,087	11.088
- *low-latency*	500	246,316	2.026

Exhaustive Workload (targeting all 500 Million of 500 Million Entries)

Test Scenario	Number of Clients	Through put (ops/sec)	Latency (msec)
Search (Sequential)	3,200	433,684	7.313
Search (Random)	3,200	170,524	18.657
- *low-latency*	500	156,986	3.178
- *high-concurrency*	32,000	162,093	197.400
- *high-concurrency*	64,000	165,705	377.275
- *high-concurrency*	100,000	168,151	572.257
Auth (Sequential)	3,200	163,547	19.511
Auth (Random)	3,200	58,117	55.033
- *low-latency*	500	47,625	10.492
Compare (Sequential)	3,200	285,837	11.092
Compare (Random)	3,200	136,910	23.362
- *low-latency*	500	113,710	4.393
Modify (Sequential)	200	15,083	13.254
Modify(Random)	200	12,059	16.579
Mixed Load	3,200	90,170	35.478
Group Search	3,200	314,861	10.073
Group Modify	10	3,540	2.000
Add (500K)	500	1,034	482.916
Mixed	3,200	90,170	35.478

Results Analysis

Cache

In the sparse workload scenario, it was feasible to have the target set of user entries (10 million) cached in the SGA, even though OID's own entry cache was disabled. As a result the vast majority of test operations exercised in this workload benefited from significant RDBMS cache hits.

In contrast, for the exhaustive workload scenario, it was not feasible to have the target set of user entries (500 million) be cached. Furthermore, OID's own entry cache was explicitly disabled for this scenario. As a result, except for the sequential test operations that might have benefited from the sequential proximity of data, none of the other test operations were likely to benefit from any cache hits.

This indicates that OID does not need to rely on caching data for performance, and with the correct mix of hardware resources it can scale to meet virtually any deployment requirement.

Data Partitioning

None of the workload scenarios tested here involved data partitioning. The entire set of 500 million user entries was stored in a single Directory Information Tree (DIT) without compromising performance or requiring excessive memory. Consequently, no a priori knowledge of the expected workload characteristics was required to partition for maximum performance. Furthermore, no subsequent maintenance, such as repartitioning, was required to address changes in workload characteristics over time.

This illustrates how OID's ability to handle all data within a single DIT empowers administrators to load and forget their data thereby reducing administrative overhead.

Large Entry Size

This benchmark operated on user entries that were 2.8KB in size (including a 2KB binary attribute). Nevertheless, since OID's architecture allows it to operate on each individual element of an entry, no discernible performance impact was observed.

Large Group Size

Part of this exercise also involved operations on groups with 10 million members. To put things in perspective, our sparse workload scenario would be covered by the members of a single such group. Nonetheless, OID was able to service these requests without hindrance at high throughput.

Latency

The results confirm the intuitive notion that latency increases with concurrency (that is the number of concurrent clients). As the number of clients connecting with OID grows, the number of resources used and context switches that occur within the OID server also grows. It is clear that this would cause an increase in latency.

While it was feasible to reduce the number of clients connecting to OID in order to obtain better latency numbers, we believe that for a deployment of this size at least 500 concurrent clients would be expected.

Concurrency

In order to showcase OID's ability to handle large numbers of concurrent clients, we ran the Random LDAP Search Operations Test on the exhaustive workload scenario for 32,000, 64,000 and 100,000 clients, respectively. While operation latency increased as expected, throughput was only marginally affected.

It must be noted here that we did not perform tests involving even higher numbers of concurrent clients because the SLAMD load generator had issues handling such large numbers of clients, not because OID was unable to handle such a load. For example, for the 64,000 and 100,000 concurrent client tests latency had to be manually calculated using the detailed statistics collected by SLAMD since the generated report was inaccurate.

Network Connectivity

In both scenarios we were unable to maximize CPU utilization on any of the three types of nodes in the benchmark deployment topology (Exadata DB, OID and SLAMD client). Utilization topped out at around 70%. This indicates that we were bottlenecked on network capacity.

In this benchmark exercise, the network connectivity between the SLAMD clients and OID and between OID and the Exadata DB was based on a 10 Gigabit Ethernet. Consequently, we believe that we would have been able to maximize CPU utilization and obtain higher performance numbers if network connectivity had been powered by the newer 40GbE or 100GbE Ethernet.

Sustained Uniform Performance

As can be seen in the performance graphs for our benchmark tests in Appendixes A and B, OID was capable of maintaining uniform performance levels with no spikes throughout each test.

This showcases OID's ability to maintain high performance numbers over long periods of time without service disruption.

Conclusion

IT infrastructures that serve user populations in the hundreds of millions are becoming increasingly more common. Providing a scalable, powerful, and highly available solution to address the challenges exposed by these large scale deployments is a business critical priority.

As evidenced in this whitepaper, the unique architecture of Oracle Internet Directory allows it to scale virtually linearly based on the available hardware. Given that the Oracle Exadata Database Machine exhibits the same traits, the combination of these two products provides the most optimal future proof Directory solution today.

As such, with the appropriate hardware and network configuration we can expect to successfully scale to meet requirements exceeding 1,000,000 ops/sec.

Appendix A: SLAMD Performance Graphs (Sparse Workload)

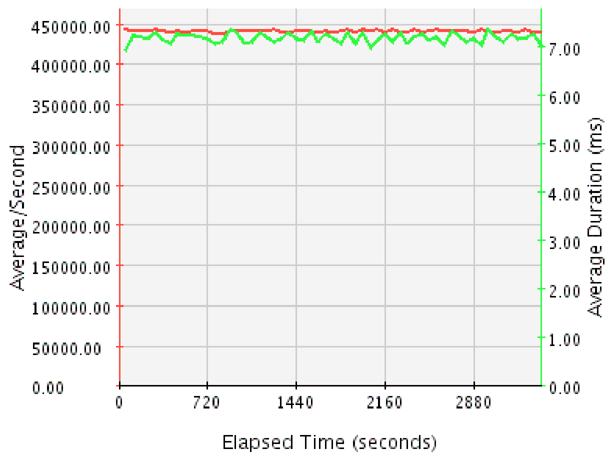

Figure 4: Sequential Search 10m Entries

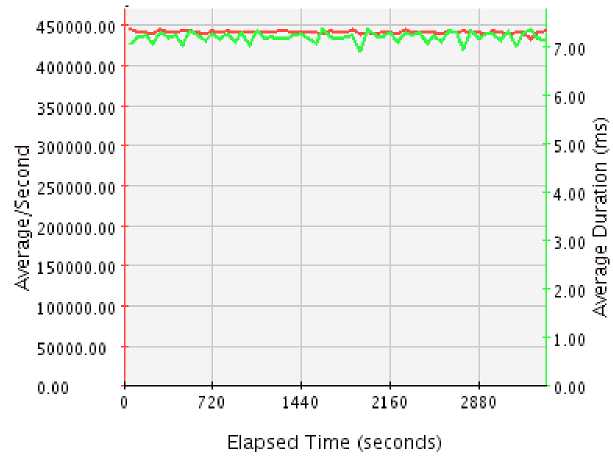

Figure 3: Random Search 10m Entries

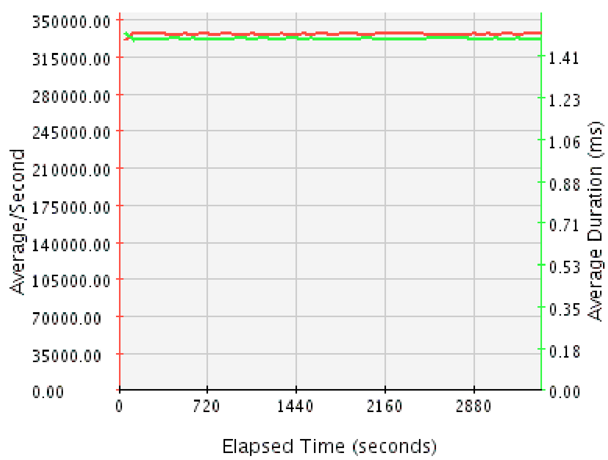

Figure 8: Low Latency Random Search 10m Entries

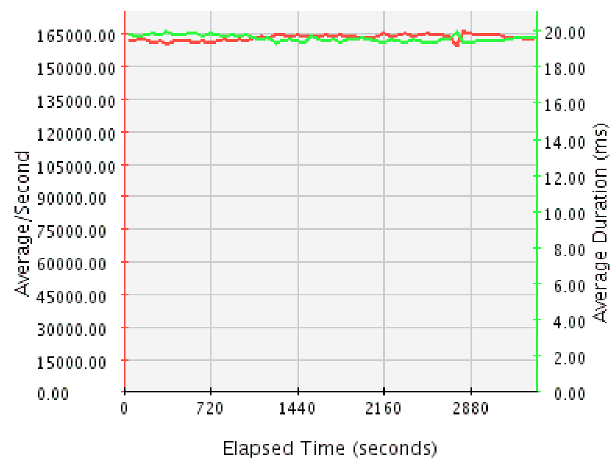

Figure 6: Sequential Authentication 10m Entries

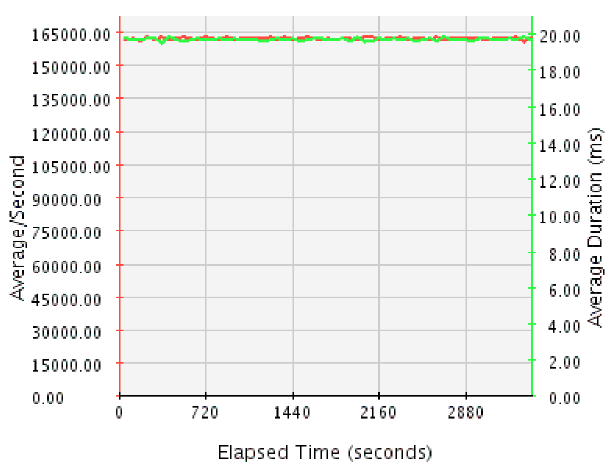

Figure 7: Random Authentication 10m Entries

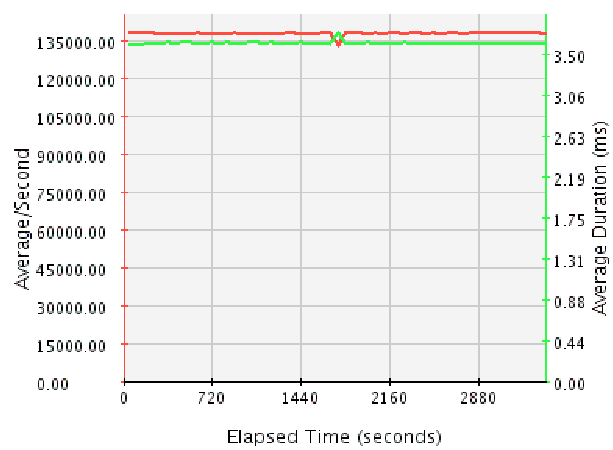

Figure 5: Low Latency Random Authentication 10m

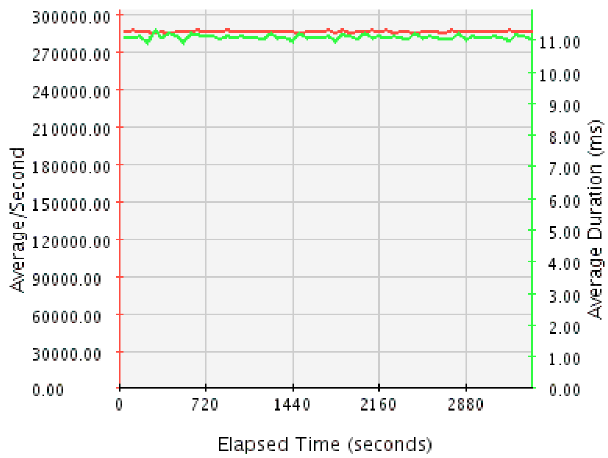

Figure 10: Sequential Compare 10m Entries

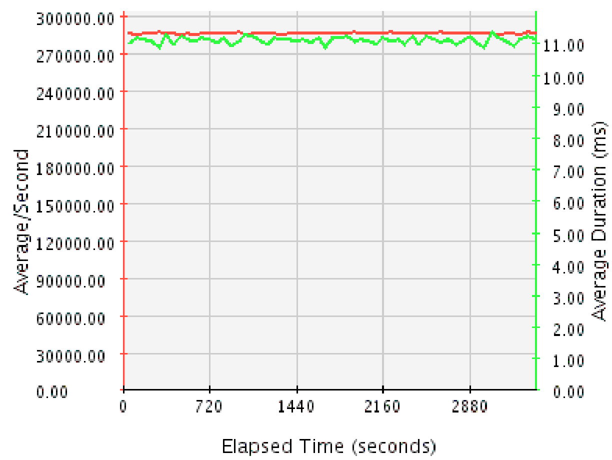

Figure 9: Random Compare 10m Entries

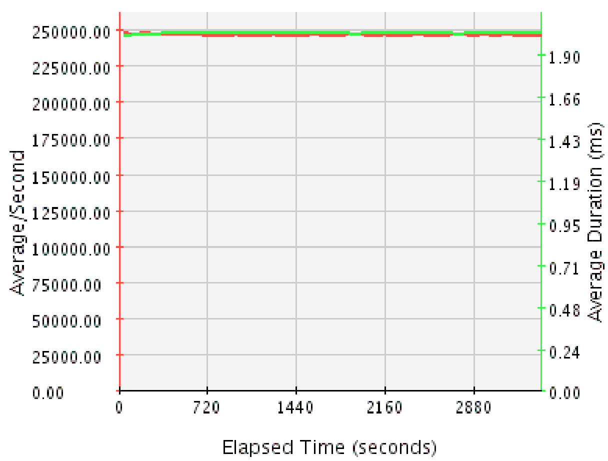

Figure 11: Low Latency Random Compare 10m Entries

Appendix B: SLAMD Performance Graphs (Exhaustive Workload)

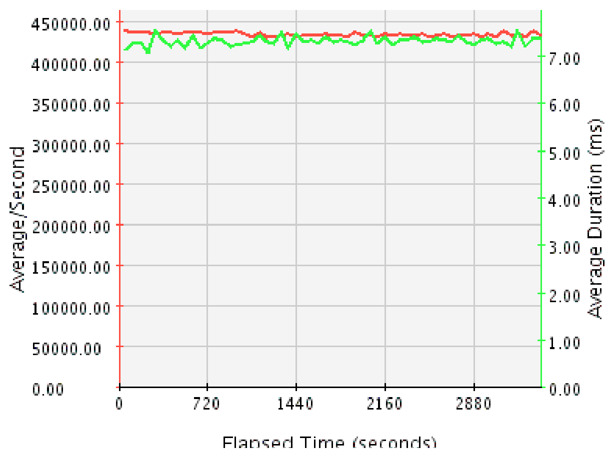

Figure 13: Sequential Search 500m Entries

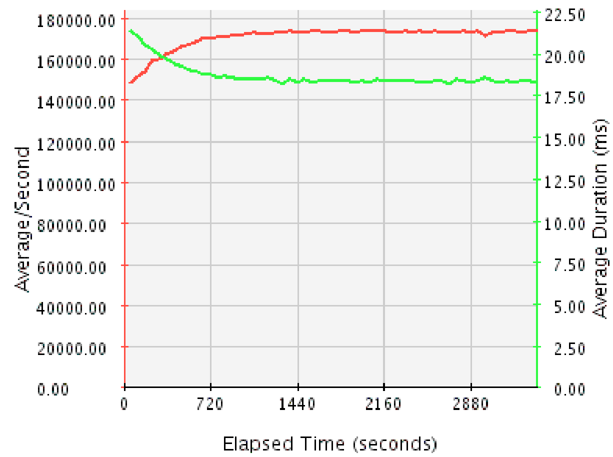

Figure 12: Random Search 500m Entries

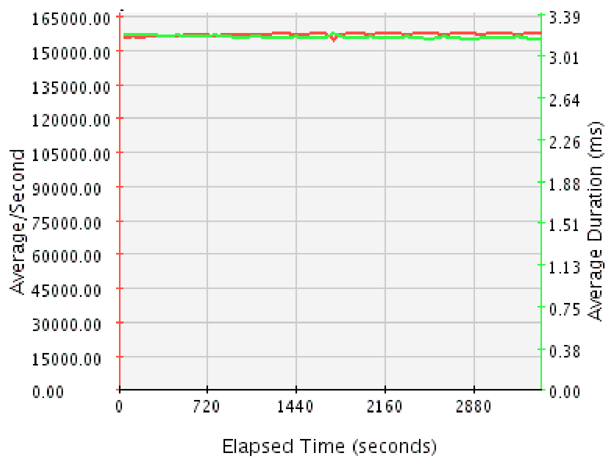

Figure 14: Low Latency Random Search 500m Entries

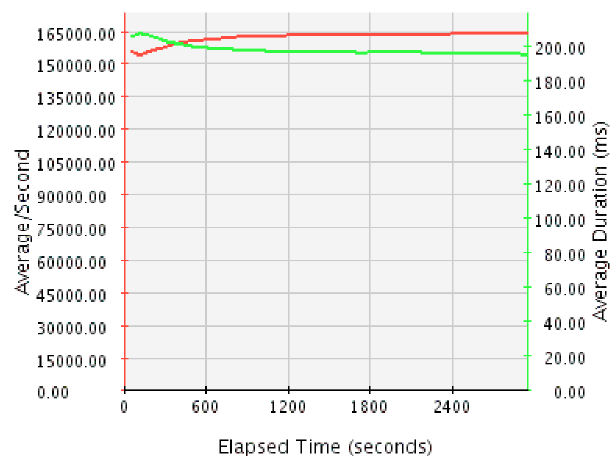

Figure 17: High Concurrency Random Search 500m (32,000 concurrent clients)

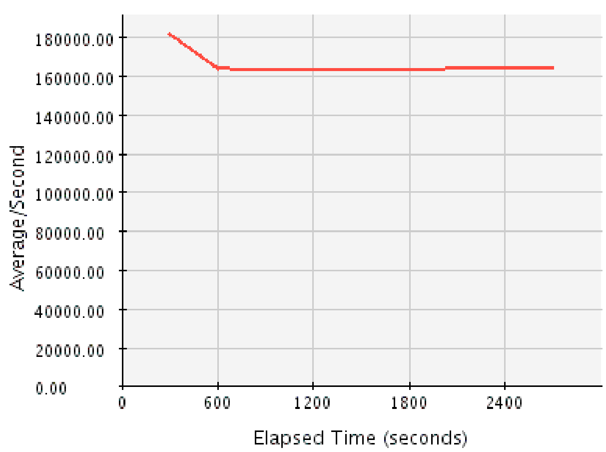

Figure 15: High Concurrency Random Search 500m (64,000 concurrent clients)

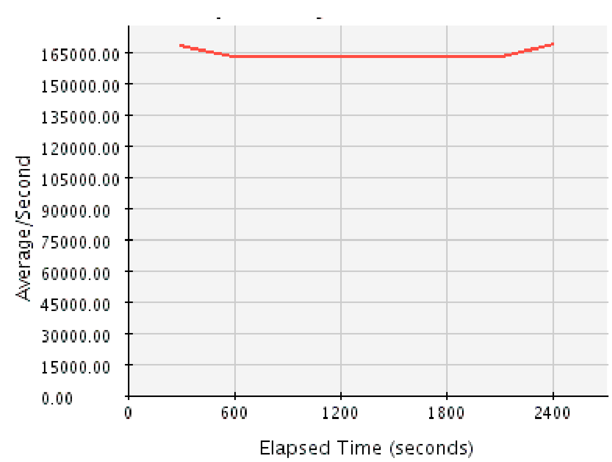

Figure 16: High Concurrency Random Search 500m (100,000 concurrent clients)

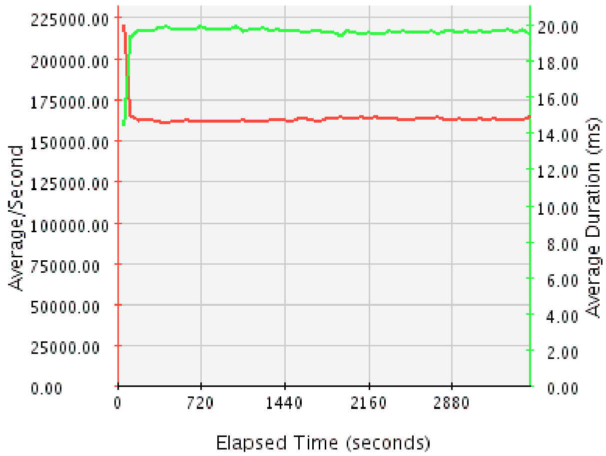

Figure 18: Sequential Authentication 500m Entries

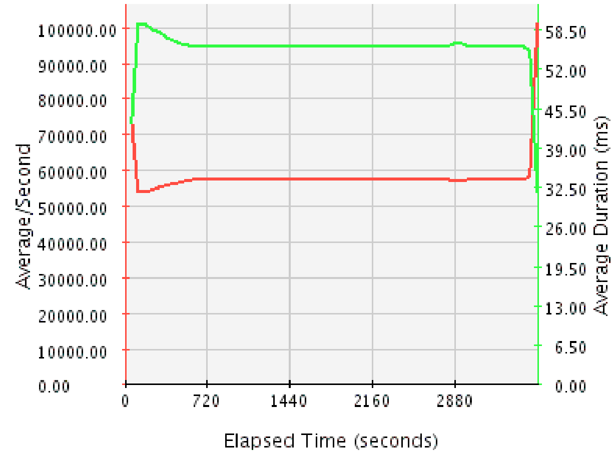

Figure 19: Random Authentication 500m Entries

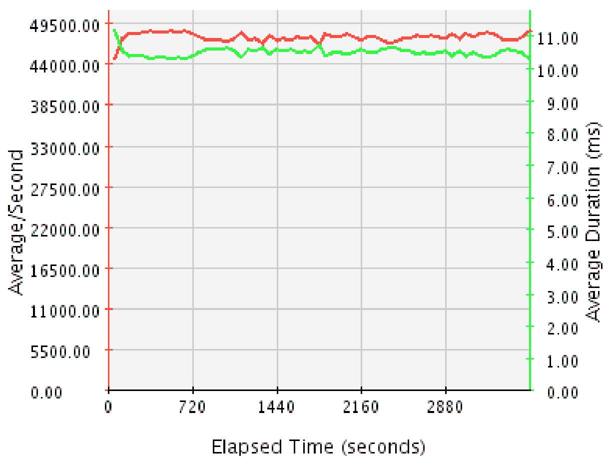

Figure 23: Low Latency Random Authentication 500m Entries

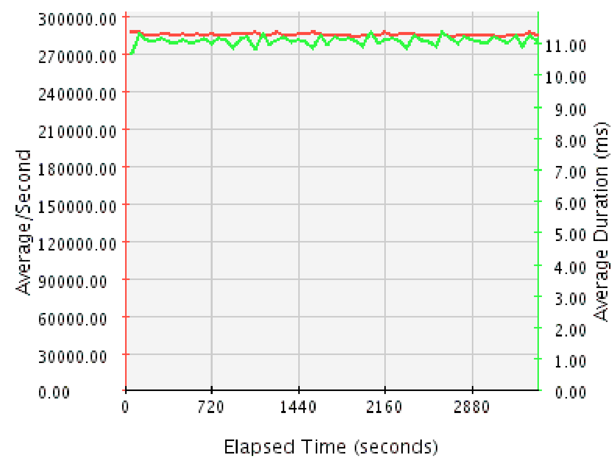

Figure 22: Sequential Compare 500m Entries

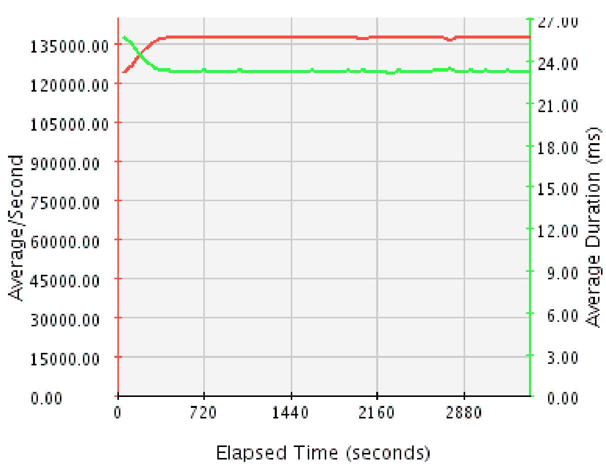

Figure 20: Random Compare 500m Entries

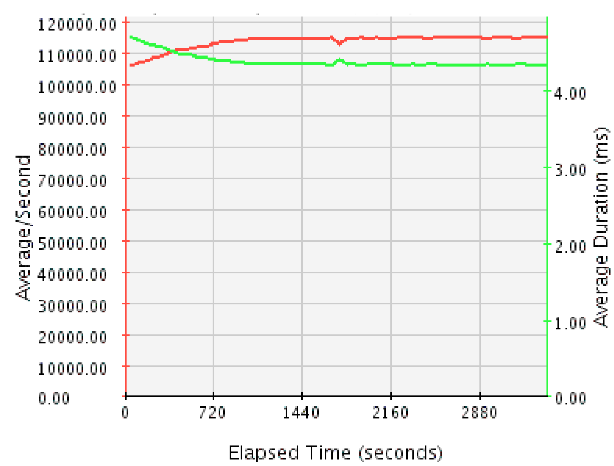

Figure 21: Low Latency Random Compare 500m Entries

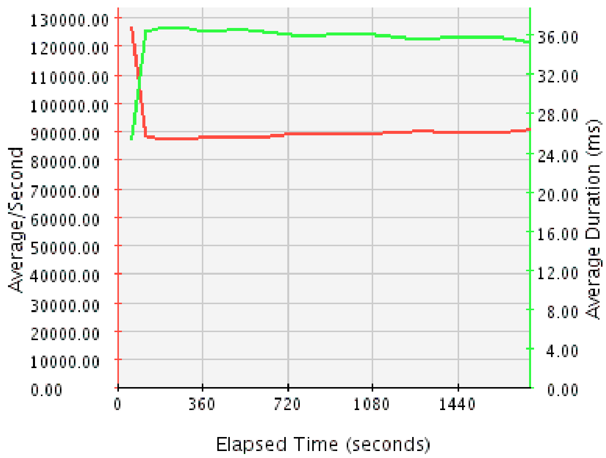

Figure 24: Random Mixed Load 500m Entries

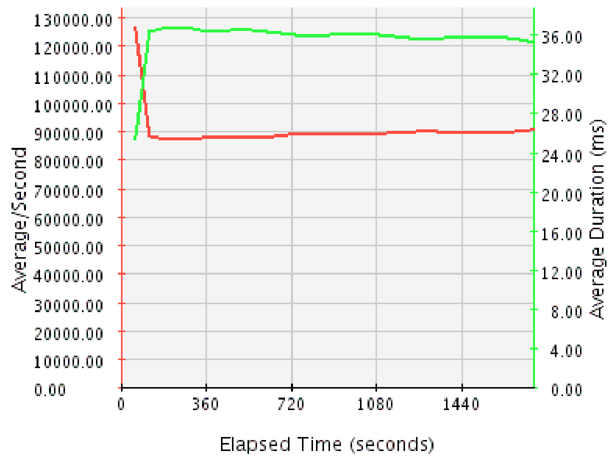

Figure 25: Random Group Search 500m Entries

Appendix C: Template LDAP User Entry

SLAMD template file used to generate data:

```
define suffix=dc=com
define NumUsers=500000000

branch: ou=oracle,[suffix]
subordinateTemplate: Users:[NumUsers]

template: Users
rdnAttr: uid
objectClass: top
objectClass: person
objectClass: inetorgperson
uid: user.<sequential:1>
cn: user.<sequential:1>
mail: user.<sequential:1>@oracle.com
photo:: <random:base64:2048>
description: <random:alphanumeric:80>
sn: <random:alphanumeric:20>
telephoneNumber: <random:numeric:6>
homePhone: <random:numeric:6>
pager: <random:numeric:6>
mobile: <random:numeric:6>
employeeNumber: <random:numeric:6>
street: <random:alphanumeric:25>
l: <random:alphanumeric:5>
st: <random:alphanumeric:2>
postalCode: <random:numeric:5>
postalAddress:  <random:alphanumeric:100>
```

Appendix D: Additional Information

- 2 Billion User Benchmark (Oracle Internet Directory 10.1.4.0.1)
- Oracle Internet Directory on the Oracle Technology Network
- Oracle Exadata Database Machine on Oracle Technology Network
- 3-D Demo: Oracle Exadata Database Machine X2-2
- Sun Fire X4470 Server
- SLAMD Load Generation Engine
- Oracle® Fusion Middleware Enterprise Deployment Guide for Oracle Identity Management 11g Release 1 (11.1.1)
- Oracle® Fusion Middleware High Availability Guide 11g Release 1 (11.1.1)
- Oracle® Documentation

ORACLE®
Oracle Internet Directory in the Facebook Age
February 2011
Author: Buddhika Kottahachchi, Ramaprakash Sathyanarayan
Contributing Authors: Ajay Keni, Olaf.Stullich@oracle.com

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

oracle.com