LSI LOGIC

The Optimal Oracle Configuration: Using the Oracle Orion Workload Tool to Accurately Configure Storage

James F. Koopmann

Technical White Paper

25034-00, Rev. A
February, 2007

Contents

Chapter 1: THE OPTIMAL ORACLE CONFIGURATION: USING THE ORACLE ORION WORKLOAD TOOL TO ACCURATELY CONFIGURE STORAGE

Introduction 1-1
Orion Workload Tool Options 1-2
Configuration and Operation Benefits 1-3
Configuration Benefits 1-3
Operational Benefits 1-3
Using the Orion Workload Tool 1-4
Determining the Application Mix 1-4
Online Transaction Processing (OLTP) 1-4
Online Analytical Processing (OLAP) 1-5
Changing Application Mixes 1-5
Determining Oracle Workloads 1-6
Determining the Workload for an Existing Oracle Database 1-7
Setting Up the Test 1-9
Understanding the Test Results 1-9
Summary Results 1-9
Single Disk Drive 1-10
Three Disk Drives 1-10
I/O Throughput 1-10
Single Disk Drive 1-10
Three Disk Drives 1-11
Data Transfer Rate 1-11
Single Disk Drive 1-12
Three Disk Drives 1-12
Latency 1-13
Single Disk Drive 1-13
Three Disk Drives 1-14
Conclusion 1-14

Contact Information 1-15
North American Headquarters 1-15
LSI Logic KK Headquarters 1-15
LSI Logic Europe Ltd. European Headquarters 1-15

Appendix A: Test Data Tables A-1

Appendix B: References B-1

Chapter 1

The Optimal Oracle Configuration: Using the Oracle Orion Workload Tool to Accurately Configure Storage

Introduction

As a database architect responsible for configuring storage array or purchasing database storage, you must identify the aspects of storage configuration that most affect database performance. You could just base your purchasing decisions on only the amount of data that you plan to house in the Oracle database. However, to determine the optimal configuration of a storage array, you must consider the relationship among the central processing unit (CPU), memory, and disk drives as well as the database's projected size. By carefully planning and configuring the storage array, you gain a vital edge in maintaining and predicting the performance of an Oracle database and ultimately in performance. Careful planning and configuration can help you avoid these pitfalls:

- **Making assumptions about specifications** – Storage vendors, often regarded as definitive authorities, can provide impressive statistics and specifications. Impressive statistics might lead you to assume that a storage array is above industry standards and viable for serving the Oracle database.
- **Solving poor configuration problems with extra CPUs or memory** – If you purchase database storage and later the database system is CPU-bound or memory-deprived, it is easy to purchase additional CPUs or memory. With this type of purchase, you might never question configuration of the storage array.
- **Forgetting that faster CPUs or more memory can lead to more I/O requests** – Having purchased database storage, you might want to service input/output (I/O) requests faster and purchase additional CPUs or memory. Adding faster CPUs or more memory might actually lead to more I/O requests and tasking disk drives even more.

Specifications are invaluable for a wide range of purchasing decisions. Unless a storage vendor has run the targeted application with an Oracle database on the storage array, you cannot predict how an application will ultimately run. It is up to you, the purchaser, to validate the end configuration for the particular environment.

Configuring a storage array for use with an Oracle database need not be difficult. Start with proper planning to optimize the performance of the hardware. Use the Orion Workload tool to effectively develop a workload that mimics and stresses a storage array

in the same manner as the planned application with an Oracle backend database. Because the Orion Workload tool does not require running an Oracle database, you can test multiple configurations and discover the optimal configuration for reliability, stability, and scalability.

Whether a workload has been sampled and realized from a current database system, you must measure a storage array for expected performance before creating an Oracle database. Only after stressing a storage array, can you determine its capabilities and properly assess its current use or future growth.

Orion Workload Tool Options

Use the Orion Workload tool to calibrate and design a storage array that you can rely upon to meet expected performance levels. The Orion Workload tool measures Oracle performance without your having to install Oracle software or even create an Oracle database. Instead, the Orion Workload tool sends I/Os to raw disks using the same libraries that an Oracle database would use. The Orion Workload tool can simulate a variety of Oracle I/O workloads. You can use the command line options to fine tune the storage array and create the optimal storage array architecture for a particular environment without guesswork. After using the Orion Workload tool, you will have a finer understanding for the performance capabilities of a storage array and can be confident in its deployment.

The Orion Workload tool provides three run levels:

- **Simple** – Small and large random I/O are tested individually.
- **Normal** – Small and large random I/O are tested individually and combinations of small and random I/Os together.
- **Advanced** – A wide variety of options are available to fine tune the workload.

Use these Orion Workload tool options to fine-tune the sample workload and stress the storage array:

- **num_disks** – Defines the number of spindles in the storage array to be tested.
- **size_small** – Defines the size for small random I/O.
- **size_large** – Defines the size for large random or sequential I/O.
- **type** – Defines the type of large I/O (random or sequential).
- **write** – Defines the percentage of writes in the workload.
- **matrix** – Defines the mixture of workload to run.

Configuration and Operation Benefits

Orion Workload tool provides two very important benefits.

- **Ease of configuration** – The guesswork is quickly removed by performing a variety of benchmark tests. You can run tests for specific applications and for different physical structures within Oracle.
- **Long-term operational benefits in performance, scalability, and reliability** – Without the Orion Workload tool these operational benefits are often harder to realize. The Orion Workload tool lets you simulate an actual Oracle workload through particular storage configurations and empowers you to develop an optimal storage configuration.

Configuration Benefits

The Orion Workload tool provides a range of options for the storage array. Here are some of the configuration tasks possible with this tool:

- Creating benchmarks for specific database application mixes
- Extending single-node tests to multi-node Real Application Clusters (RAC) configurations
- Balancing disk drive layout against system workload
- Verifying performance levels for various I/O scenarios
- Revealing specific I/O benefits and deficiencies
- Configuring storage array before use
- Reconfiguring the storage before installing an Oracle database
- Timely fixing of hardware flaws
- Comparing benchmarks across different storage products and configurations

Operational Benefits

The Orion Workload tool provides operational benefits in the areas of performance, scalability, and reliability:

- When used to simulate a real-world Oracle workload, the Orion Workload tool directs you to configurations that provide current and future levels of performance not possible from non-benchmark solutions.
- The Orion Workload tool provides a picture of how I/O patterns scale when workloads change. Although you typically allow for growth when configuring storage arrays, the Orion Workload tool relates growth in workload to performance, giving a clear picture of the scalability of the system.

- When scalability and performance are met, you experience higher levels of reliability in the storage array. By performing Orion Workload tool benchmarks, you can, with a high level of assuredness, state that the configuration is able to meet or exceed I/O requirements of an application mix.

Using the Orion Workload Tool

It is very easy to start using the Orion Workload Tool.

1. Download the Orion Workload tool.
2. Install the Orion Workload tool by unzipping the file.
3. Create a file that contains a list of raw volumes or files to test.
4. Run the Orion Workload tool binary with workload options.
5. View the tabular output.

NOTE Visit the Orion Workload tool download page, listed in Appendix B, "References" on page B-1, and read the *Orion User Guide* to obtain the latest instructions on how to run various workload scenarios.

For any workload that is tested against the storage array, the Orion Workload tool loads the system with I/O streams of varying intensities to determine I/Os per second (IOPS), latency, and megabytes per second (MBPS). The exact load level is represented in the form of outstanding asynchronous I/Os that can continually be maintained. For instance, combining a mixed workload of two outstanding large random I/Os with four outstanding small random I/Os produced a data transfer rate of 95 MBPS.

IMPORTANT Remember, the load level represents the number of outstanding I/Os that are queued and not the number of I/Os that have been serviced.

Determining the Application Mix

Start the planning process by analyzing the workload mix placed on a database and, ultimately, the performance requirements of the storage array.

Online Transaction Processing (OLTP)

An OLTP workload has these characteristics:

- Read- and write-intensive
- Transaction-based

- Performance measured by IOPS and average latency (I/O turn-around time)
- Having typical usage of large bursts of small reads followed by inserts and updates
- Primarily driven by the database issuing small random I/Os the size of the defined Oracle database block size (db_block_size).
- Having a database block size ranging from 2 KB to 32 KB, with 8 KB being average.

IMPORTANT Even if a database has a low write percentage, database logs must also be considered. Also, certain operations, such as table or index scans, backups, and parallel queries, can issue large I/Os.

Identify the percentage of reads to writes as this is a critical ratio for finding storage configurations. The typical ratio is 70:30.

Online Analytical Processing (OLAP)

An OLAP workload has these characteristics:

- Throughput-based
- Performance-based on moving large amounts of data in megabytes per second (MBPS)
- Driven by read only databases except during the load stage
- Having large multi-block read/write I/O streams containing multiple 1-MB I/Os

Changing Application Mixes

The following list captures a few of the many sources of application change.

- A simple OLTP database could be cloned to a different system and could begin to be used with an OLAP application. The data is the same, but the I/O pattern is very different.
- A new group of power users could turn an OLTP database into an OLAP reporting engine without notification.
- The application mix might change throughout the day.
  - Database backups that run during the day affect the type of read activity of the database
  - Data loads or extraction, transformation, and load (ETL) processing could switch an OLAP environment into an OLTP environment

Determining Oracle Workloads

When you have identified the application mix, translate the mix into a database workload. Keep in mind that a single storage array, if shared between two different workload types, does not perform equally effective to a single storage array configured for a single workload type. However, by understanding all workload types and properly testing the storage array, you can often produce a better solution. Just as there are basically two application mixes, there are also two primary workload types: small random I/O and large sequential I/O. As databases are interrogated, additional workload types might be seen.

- Small random I/O
  - OLTP applications as described previously
  - Transaction-based (IOPS and latency)

NOTE High levels of IOPS require more backend drives.

- Large sequential I/O
  - OLAP applications as described previously
  - Throughput-based (MBPS)

NOTE High levels of MBPS can often be satisfied by fewer backend drives.

- Large random I/O
  - Because of various database traffic or configurations, large sequential I/O might be turned into large random I/O at the disk drive level of the storage array. Here are some examples:
    - Large sequential I/O caused by data warehouses, reporting applications, or utilities, such as backups, can be broken up by normal small random I/O of an OLTP system.
    - Stripping can spread large sequential I/O across many disk drives, causing the storage array to experience large random I/O.
- Mixed I/O
  - A database might exhibit one or more of those workloads continually and thus fall into the category of a mixed I/O workload.

Determining the Workload for an Existing Oracle Database

If an existing Oracle 10g Release 2 (10gR2) database is servicing the application mix intended to be reconfigured on a storage array, use this SQL command to determine the workload.

```
set linesize 100
set head off
SELECT 'Number of Small Reads :'||
       sum(decode(name,'physical read total IO
requests',value,0)-
           decode(name,'physical read total multi block
requests',value,0)),
       'Number of Small Writes:'||
       sum(decode(name,'physical write total IO
requests',value,0)-
           decode(name,'physical write total multi block
requests',value,0)),
       'Number of Large Reads :'||
       sum(decode(name,'physical read total multi block
requests',value,0)),
       'Number of Large Writes:'||
       sum(decode(name,'physical write total multi block
requests',value,0)),
       'Total Bytes Read      :'||
       sum(decode(name,'physical read total bytes',value,0)),
       'Total Bytes Written   :'||
       sum(decode(name,'physical write total bytes',value,0))
  FROM gv$sysstat;
```

```
Number of Small Reads :205903
Number of Small Writes:106883
Number of Large Reads :40298
Number of Large Writes:2791
Total Bytes Read      :4188587008
Total Bytes Written   :2009381888
```

The statistics contained in the GV$ views are cumulative, so obtain samples from the start and the end of a peak I/O cycle. Subtract the starting statistics from the ending statistics to obtain the true reads and writes that have occurred. Use this figure to determine the read-to-write ratios and also to classify the database application mix and the workload type.

For instance, if you use the previous SQL command again in 10 minutes, you might produce the following statistics.

```
Number of Small Reads :500211
Number of Small Writes:123474
Number of Large Reads :142981
Number of Large Writes:8010
Total Bytes Read      :22232604961
Total Bytes Written   :5586081648
```

Calculate the IOPS for small and large I/O, the percentage of reads to writes, and the MBPS throughput.

```
Small Read IOPS  =(500211-205903)/(10*60) = 490 IOPS
Small Write IOPS =(123474-106883)/(10*60) =  27 IOPS
Total Small IOPS =(294308-16591)/(10*60)  = 517 IOPS
I/O Percentage of Reads to Writes = 94:6
Large Read IOPS  =(142981-40298)/(10*60) = 171 IOPS
Large Write IOPS =(8010-2791)   /(10*60) =   8 IOPS
Total Large IOPS =(102683+5219) /(10*60) = 179 IOPS
I/O Percentage of Reads to Writes = 96:4
Total MBPS Read   =((22232604961-4188587008)/(10*60))/1048576=
28 MBPS
Total MBPS Written =((5586081648-2009381888) /(10*60))/1048576=
5 MBPS
Total MBPS        =((18044017953+3576699760)/(10*60))/1048576=
34 MBPS
```

You can see from the previous numbers, the particular system being tested is an OLTP system with many small I/Os.

Setting Up the Test

The example test that follows is not an exhaustive evaluation of the Orion Workload tool, but a simple test to demonstrate the usefulness of the tool. In this example, you are testing two different storage configurations: one disk drive and three disk drives. The test will stress the storage array to determine the optimal configuration for an Oracle storage array.

For this test, you select an advanced run level so that a proper workload can be used for the expected storage array.

Here are example specifications:

- The number of disk drives to test
- A read-only test
- The test name
- Mixture of reads (small random with large random and sequential reads)

You can select from the many options available with the Orion Workload tool to construct a viable workload for the particular storage array purpose.

```
./orion_linux_em64t -run advanced -testname orion1 -num_disks 1
-write 0 -simulate concat -matrix detailed
./orion_linux_em64t -run advanced -testname orion3 -num_disks 3
-write 0 -simulate concat -matrix detailed
```

Understanding the Test Results

A summary file, `<testname>_summary.txt`, is created from the Orion Workload tool, which contains a description of the options chosen and some high-level statistics collected for the test.

Summary Results

For this test, you immediately see that the three-disk drive configuration gives higher performance numbers for IOPS and MBPS, while keeping I/O latency down. Note that the numbers cluster around the maximum value seen for each I/O type. These numbers provide the best case scenario for the storage configuration. Further investigation of additional Orion Workload tool statistics proves or disproves the usefulness of these values in determining scalability of one disk drive to three disk drives.

Single Disk Drive

```
Maximum Large MBPS=64.43    @ Small=0 and Large=2
Maximum Small IOPS=586      @ Small=4 and Large=0
```

Three Disk Drives

```
Maximum Large MBPS=95.38    @ Small=0 and Large=6
Maximum Small IOPS=1021     @ Small=15 and Large=0
Minimum Small Latency=3.47 @ Small=2 and Large=0
```

I/O Throughput

In the file, *<testname>*_iops.csv, Orion Workload tool produces comma-separated values representing the IOPS for small random workloads. Imposed upon the small random I/O load are varying loads of large random I/O. Along the x-axis is the small I/O load level. The lines represent the large I/O load level, and the y-axis represents the intersection of those load levels as realized IOPS.

Single Disk Drive The single-disk drive configuration shows that it is difficult to service both large and small I/Os concurrently. The only benefit for this particular single-disk drive configuration, if IOPS is of concern, is for small random I/Os only. See Appendix A, "Test Data Tables" on page A-1 for test data tables.

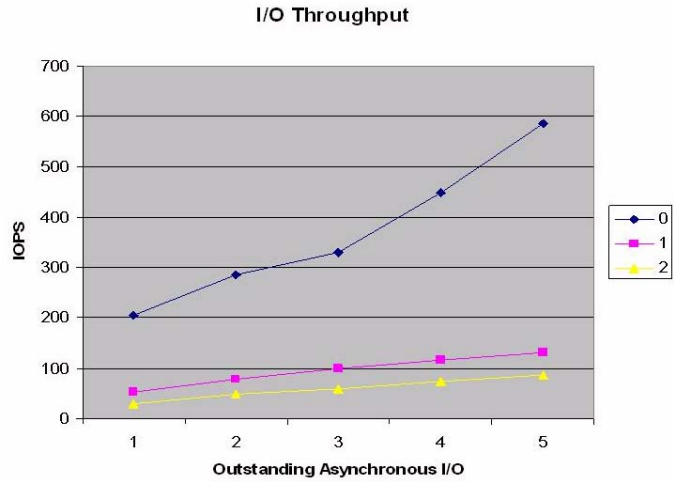

Figure 1-1 I/O Throughput – Single Disk Drive

Three Disk Drives When you move to a three-disk drive configuration, both large and small I/Os are serviced concurrently. Performance degrades quickly after more than one outstanding large I/O is added to the mix. The three-disk drive configuration produces more IOPS, at the same load levels, than the single-disk drive configuration. The three-disk drive configuration for just small random I/Os provides nearly twice as much IOPS with 15 outstanding I/Os than does the single disk drive system does with five outstanding I/Os.

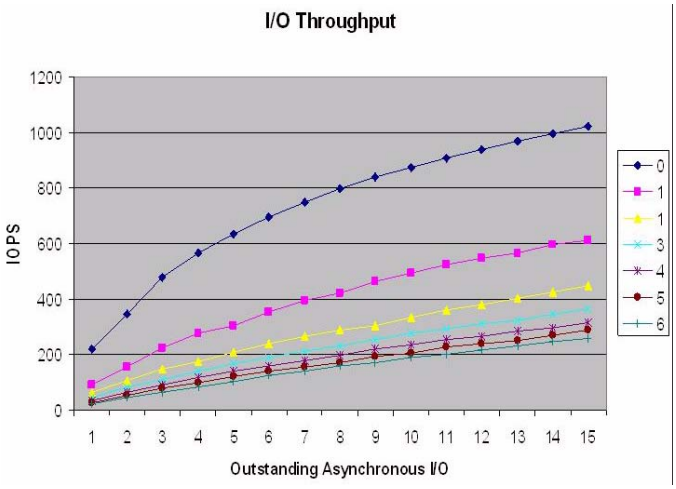

Figure 1-2 I/O Throughput – Three Disk Drives

Data Transfer Rate

In the file, `<testname>_mbps.csv`, Orion Workload tool produces comma-separated values representing the MBPS for large random/sequential workloads. Imposed upon the large random I/O load are varying loads of small random I/O. Along the x-axis is the large I/O load level. The lines represent the small I/O load level and the y-axis represents the intersection of those load levels as realized MBPS.

Single Disk Drive A single-disk drive configuration reaches the MBPS limit when there are only two outstanding large I/Os queued. As you might expect, performance declines as small random I/Os enter the test.

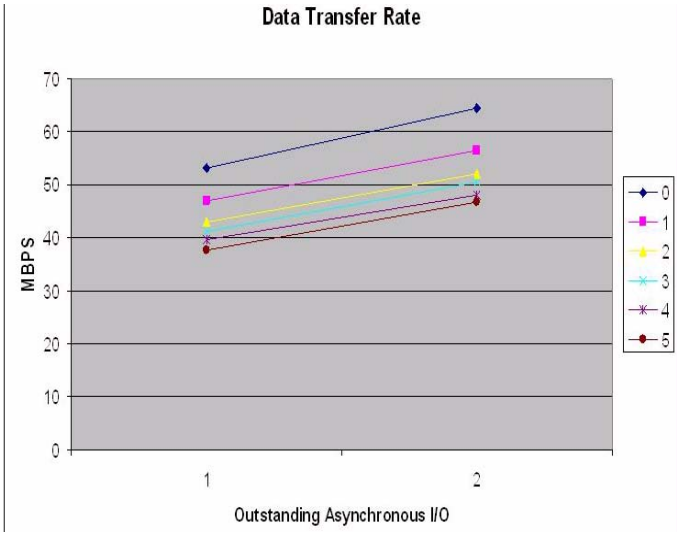

Figure 1-3 Data Transfer Rate – Single Disk Drive

Three Disk Drives A three-disk drive configuration produces more MBPS as you might expect. Under the IOPS test, the storage array was able to deliver twice the throughput when fully loaded, but the MBPS test did not do as well, indicating that IOPS is much more disk sensitive than MBPS.

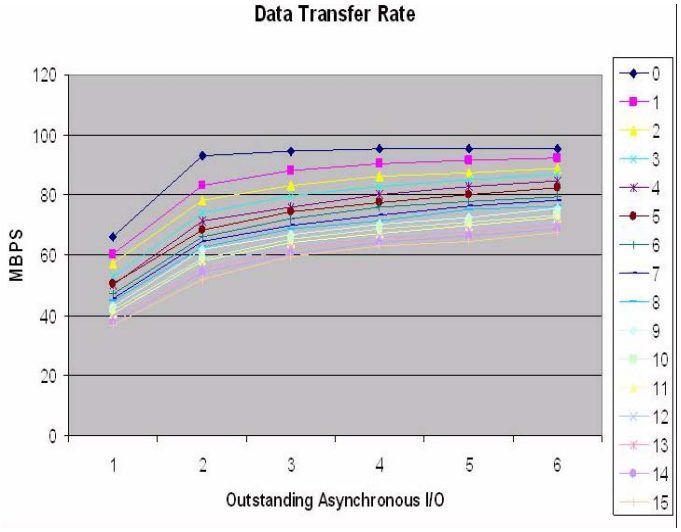

Figure 1-4 Data Transfer Rate – Three Disk Drives

Latency

In the file, *<testname>*_lat.csv, Orion Workload tool produces comma-separated values representing the latency for small random workloads. Imposed upon the small random I/O load are varying loads of large random I/O. Along the x-axis is the small I/O load level. The lines represent the small I/O load level, and the y-axis represents the intersection of those load levels as realized latency in milliseconds (ms). Latency tests are directly related to the IOPS test and, as such, have the same load levels and number of data points.

Single Disk Drive As you might expect, as the load level increases, the latency also increases.

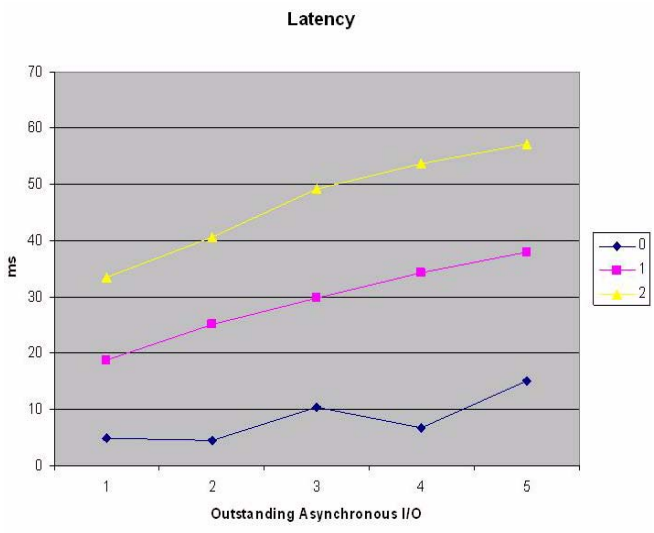

Figure 1-5 Latency – Single Disk Drives

Three Disk Drives Just as the three-disk drive configuration produced more IOPS at the same load levels as the single-disk drive, the latency is less for the three-disk drive system than for the single-disk drive system at those load levels.

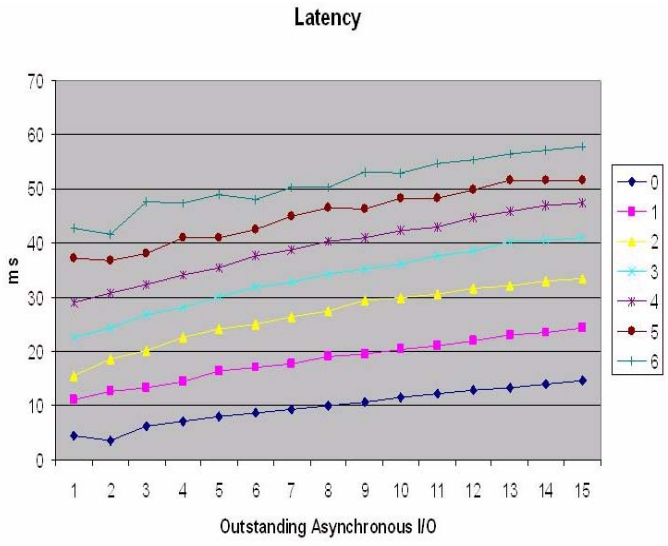

Figure 1-6 Latency – Three Disk Drives

Conclusion

If you run an Oracle database and plan to reconfigure a system or purchase a new storage array, you must stress the system and benchmark it using the Orion Workload tool. Only after running this tool can you see how the storage array performs under an actual Oracle workload and a properly configured storage array.

Understanding real-world performance levels of the systems you manage is critical to successfully running the database applications on a particular storage array. By using the Orion Workload tool statistics to project increasing levels of Oracle workloads, you can effectively provide guidance for a current storage configuration and plan and provide for future expansion.

By including the Orion Workload tool benchmarking statistics for various load levels and configurations, the vendor community can provide added value. Although vendors cannot run every application, the Orion Workload tool can be run with relative ease, and it is good business to provide these statistics to end users. Orion Workload tool statistics give the purchaser of storage equipment additional information and planning power.

Contact Information

For more information and sales office locations, visit the LSI Logic web sites at:

http://www.lsi.com/cm/ContactSearch.do

North American Headquarters

Milpitas, CA

Tel: 866 574 5741 (within U.S), 408 954 3108 (outside U.S.)

LSI Logic KK Headquarters

Tokyo, Japan

Tel: 81 3 5463 7165

LSI Logic Europe Ltd. European Headquarters

United Kingdom

Tel: 44 1344 413200

Appendix A

Test Data Tables

Figure A-1 I/O Throughput – Single Disk Drive

Large/Small	1	2	3	4	5
0	206	285	329	448	586
1	53	79	100	116	131
2	29	49	60	74	87

Figure A-2 I/O Throughput – Multiple (3) Disk Drives

Large/ Small	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	221	346	478	564	634	696	749	798	841	875	909	937	970	995	1021
1	90	157	225	277	303	352	394	420	462	492	523	545	565	596	613
2	64	106	148	176	207	239	266	290	305	334	359	379	403	424	447
3	44	82	111	141	166	188	213	233	255	276	291	310	323	344	366
4	34	64	92	116	141	159	180	198	219	236	255	267	283	297	316
5	26	54	78	97	121	141	155	171	194	206	227	240	252	271	290
6	23	47	63	84	102	124	139	159	169	188	200	216	230	245	259

Figure A-3 Data Transfer Rate – Single Disk Drive

Large/Small	0	1	2	3	4	5
1	53.16	47.07	42.92	41.17	39.7	37.69
2	64.43	56.53	52.12	50.41	48.05	46.69

Figure A-4 Data Transfer Rate – Multiple (3) Disk Drives

Large/Small	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	66.07	60.46	57.33	52.94	50.24	50.39	47.01	45.71	44.88	42.74	41.73	40.47	39.44	38.95	37.43	36.8
	93.06	83.12	78.33	73.9	71.31	68.5	66.14	64.44	62.56	61.87	58.82	58.07	56.85	55.58	54.16	52.15
	94.46	88.17	83.33	79.82	76.05	74.48	72.26	70.05	68.62	66.93	65.26	64.36	63.21	62.34	60.19	59.66
	95.17	90.2	86.23	82.66	80.28	77.3	75.9	73.3	71.48	70.28	68.88	67.13	66.57	65.65	64.32	63.09
	95.19	91.5	87.18	84.97	82.85	79.96	77.73	76.38	75.3	72.35	71.12	69.82	68.44	68.09	66.34	64.69
	95.38	92.18	88.83	87.02	84.83	82.48	79.33	78.34	76.43	75.47	73.41	72.57	71.08	70.09	68.65	67.68

Figure A-5 Latency – Single Disk Drive

Large/Small	1	2	3	4	5
0	4.85	4.45	10.5	6.82	15.15
1	18.74	25.05	29.89	34.34	38.01
2	33.42	40.58	49.24	53.61	57.16

Appendix B

References

Oracle Orion downloads

http://www.oracle.com/technology/software/tech/orion/index.html

ORION: Oracle I/O Numbers Calibration Tool

http://download.oracle.com/otn/utilities_drivers/orion/Orion_Users_Guide.pdf

Best Practices for Creating a Low-Cost Storage Grid for Oracle Databases

http://www.oracle.com/technology/deploy/availability/pdf/ora_lcs.pdf

Resilient Low-Cost Storage

http://www.oracle.com/technology/deploy/availability/pdf/lcs_OW.doc.pdf

Document Description

This paper will assist the to use Oracle Orion Workload Tool to statistically measure workload and optimize configuration.

Ownership of Materials

Trademark Acknowledgments

Performance Information

Performance tests and ratings are measured using specific computer systems and/or components and reflect the approximate performance of LSI products as measured by those tests. Any difference in system hardware or software design or configuration may affect actual performance. Buyers should consult other sources of information to evaluate the performance of systems or components they want to purchase.