CLABBY ANALYTICS

Research Report

Contrasting IBM's Three Server Approach to Workload Optimization to Oracle's Exadata Single Server Approach

Introduction

In short, a *workload* is an application or group of applications that, when executed on an information system, completes a unit of work. There are tens-of-thousands of different types of workloads in the computer world, ranging from processing electronic mail applications, through financial/distribution/human resource/manufacturing/sales and other run-the-business applications, to industry-specific applications.

Workload optimization is all about matching workloads to systems architectures best suited to serve them — and then optimizing the hardware, systems software, applications and databases to achieve high performance and utilization.

There are three approaches to workload optimization in today's information technology (IT) marketplace:

1. "*Tuned-to-the-task*" — This approach focuses on matching workloads to the server best suited to serve them. Two vendors, IBM and Oracle, strongly focus on building information system environments that are tuned to perform specific tasks.
2. "*Software optimization*" — the use of software accelerators to streamline performance or to offload the CPU from having to process input/output (I/O), networking, and other tasks.
3. *Application performance Management* (APM) — this third approach involves a discipline known as APM. APM tools help IT managers and administrators figure out how applications are behaving within a given system or across distributed systems. APM tools and utilities are then used to tune those applications for maximum performance. Unlike the previous two approaches, APM is usually a customer-driven management activity.

In this Research Report we take a closer look at how two vendors — Oracle and IBM — tune their respective systems for certain workloads (business analytics and transaction processing). As we evaluated IBM's and Oracle's approaches to tuning-their-systems-to-the-task, we found that IBM offers three different system implementations designed to process deep analytics, operational analytics, and high-speed transaction processing. Oracle, by contrast, uses a one-size-fits all approach with its Exadata server. Which approach is better? You decide…

Background: Why Processors, Systems Designs, and Optimization Matters

At *Clabby Analytics*, we've spent the last four months writing a book on workload optimization (soon to be available on *Amazon*, entitled "*Workload Optimization for Smarter People*"). And what we learned during the course of our research is that:

1. *Microprocessor characteristics matter (a lot) when it comes to executing workloads*;

> ***We have a fundamental belief, and it is that no single microprocessor architecture or systems design does every computing job (workload) optimally. Some microprocessors excel at processing serial jobs (linear, step-at-a-time workloads). Others excel at processing parallel tasks (processing slices of an application and then gathering these slices together to produce an end result — for instance, x86-based servers are outstanding at processing parallel tasks). Still others excel at processing large volumes of data.***

2. *Systems designs are extremely important when it comes to optimizing workloads*:
    a. Some designs do an extraordinary job offloading the central processing units (CPUs) from having to handle the management of storage, memory, and I/O subsystems — and this results in significant performance improvements;
    b. There is great variability in systems designs when it comes to quality-of-service (QoS) and reliability, availability, and security (RAS); and,
    c. Issues such as contention, coherence, saturation and queuing can limit application/database performance.

> ***For instance, successive generations of Oracle's Exadata platform continue to improve in terms of I/O performance — but Exadata is based on a shared disk architectural design where resource contention can limit Exadata I/O performance. In this case, a design improvement is limited by a design flaw. System designs are important…***

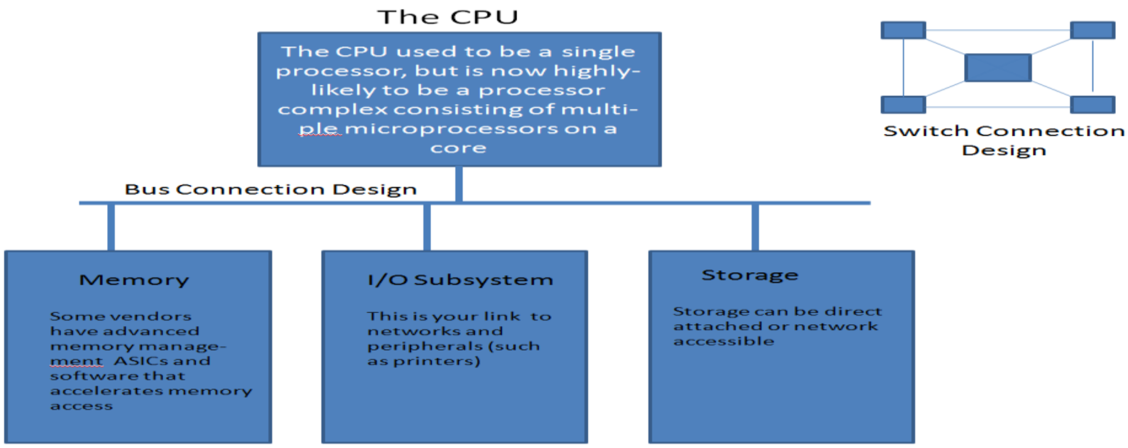
3. *"Software optimization" also plays an extremely important role in optimizing workloads (*in some cases, query processing performance can increase 100 fold!);
4. *Some vendors are changing the basis of competition in the computer industry by building pre-optimized, pre-integrated systems/software environments*:
    a. Vendors that build their own integrated software stacks (middleware, system software, etc.) — and that build their own systems platforms — have a distinct edge in terms of their ability to tune for performance over those that do not. This is because these vendors can add extensions to their products to accelerate performance — and because these vendors have deep developmental expertise, and know how, when, and where to tune their own products to achieve maximum performance).

> ***Point #4 is especially interesting to us because we are now seeing three vendors (IBM, Oracle, and Microsoft) tuning their products to the extreme to create significant competitive advantage. We especially like the competition between IBM and Oracle because these vendors build their own integrated software stacks — and tune those stacks to exploit their own underlying hardware.***

Optimizing Systems: A General Discussion

One of the top priorities of systems designers is to find ways to expeditiously feed data to the CPU. The components that feed the CPU are memory (the container that is closest to the CPU that holds the most relevant data being worked); the input/output subsystem (which is the connection to the network and peripherals); and, in some cases, a direct connection to directly attached storage subsystems. Data is passed over a bus or switch architecture to the CPU (the bus or switch is the highway between the CPU and related systems resources). The relationship of these components is illustrated in Figure 1.

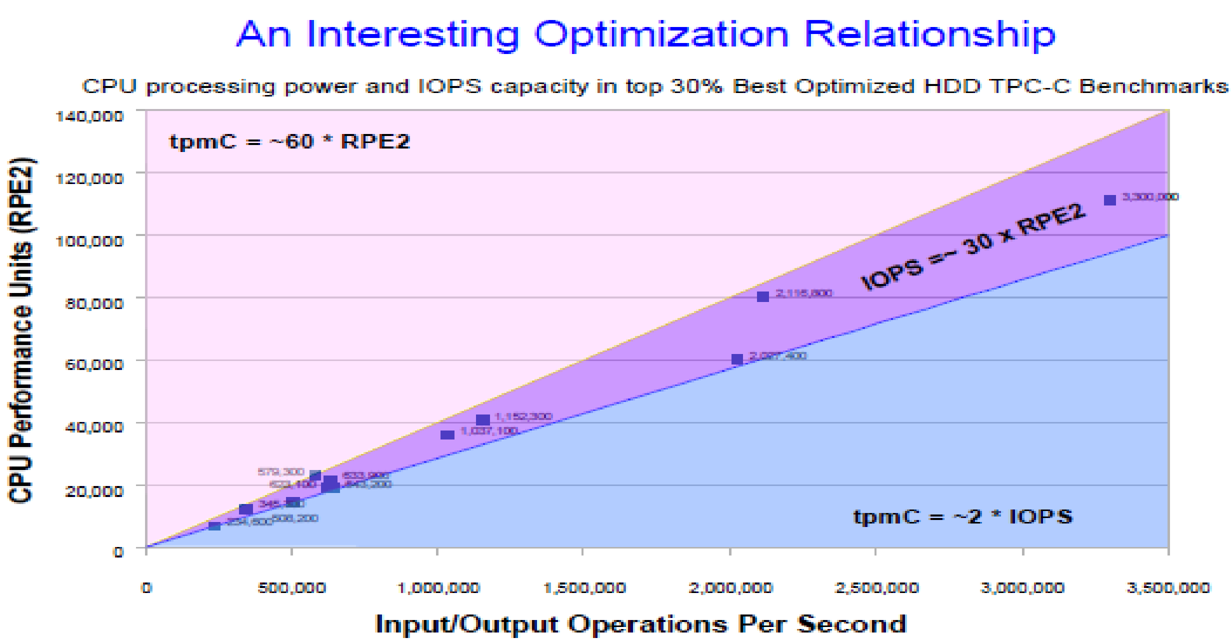
Figure 1 — The CPU and Related Service Components



Source: Clabby Analytics — January, 2012

Figure 2 illustrates why balancing a systems design is important. What this figure shows is that, if you take a look at the leading HDD-based Transaction Processing Council TPC-C benchmarks that were published before solid state disks were in use, ***the top benchmark results*** (shown in the pink section as little purple squares) ***are achieved by the systems that balance I/O operations per second with the CPU's ability to execute those operations.***

Figure 2 — Optimizing I/O with CPU Processing to Achieve Maximum Performance



Source: IBM Corporation — January, 2012

A Comparison of IBM's 3 Workload Optimized Servers Vs. Oracle's Single Server Approach

In this section, we will take a look at three distinct workloads: deep analytics, operational analytics, and transaction processing — and we'll show how IBM's approach to optimizing systems for these workloads differs from Oracle's approach.

Contrasting IBM's Three Server Approach to Workload Optimization to Oracle's Exadata Single Server Approach

First: Let's Look More Closely at Business Analytics and Transaction Processing Workloads
Let's start with the business analytics discussion. The way we see it, there are essentially two different approaches to analytics processing: deep analytics and operational analytics:

1. ***Deep analytics*** queries are usually complex queries, often involving sifting through large data sets to get answers. Another defining characteristic of deep analytics is that it tends to involve low concurrency of queries — these types of queries are so resource-demanding and have tended to take so long to execute that it is not practical to execute many at the . The focus of deep analytics is to get an analytics result as quickly as possible — so there is a big focus on achieving *rapid response time* from the system running the query. So, from an optimization perspective, you want to use a system that reduces the run time it takes to execute a particular query.

2. ***Operational analytics*** (and this includes predictive analytics) involves processing many concurrent queries from multiple users — some that may be simple, while others may be complex — using a data warehouse back-end. In this case, lots of threads of execution are taking place simultaneously. The focus when processing operational analytics is to get a lot of reports done quickly — so "*throughput*" (how many reports per second or how many queries per second can be executed) is a major focal point when processing operational analytics.

A Closer Look at Transaction Processing Workloads
Every time a user requests a service in real time from a computer, that user is initiating a transaction. (Contrast this with a computer that stores a bunch of requests and executes them at a later time — this is called batch processing). The metrics that are important when it comes to transaction processing are cost per transaction and transactions per second (in other words: price/performance).

Driving up transaction performance has long been about matching the amount of data being fed to the CPU, and the CPU's ability to process that data (illustrated in Figures 1 and 2 earlier in this report). And one of the obstacles to feeding data to the CPU has long been the amount of time it takes to read/write data to the storage subsystem. Transaction processing vendors have tackled this problem by finding ways to compress data, and by finding new technologies (such as solid state disk — SSD) that can read/write faster than mechanical hard disks (hard disk drives — HDDs).

> ***What should be clear after reading the previous two subsections is that each workload places different demands on underlying information systems. Deep analytics workloads focus on rapid response. Operational analytics focuses on throughput. And transaction processing focuses on price/performance.***

Now, lets's take a closer look at how IBM tunes and optimizes its systems for processing deep analytics, operational analytics, and transaction processing as compared with Oracle's approach.

Netezza: Purpose Built for Deep Analytics
Let's start with IBM's Netezza TwinFin appliance. IBM's Netezza appliance is a turnkey, high-performance data warehouse/advanced analytics engine. With Netezza, IBM has integrated database, server and storage components into a single server environment that

can be used to run business intelligence, predictive analytics, and other data intensive applications. This environment is based on the IBM blade architecture; it uses an Asymmetric Massively Parallel Processing (AMPP) approach to process workloads; and it uses FPGAs (field programmable gate arrays — specialized processors) as a means to filter data before it is processed.

Netezza FPGAs perform a preprocessing function where specific computational kernels are processed by FPGAs as opposed to making the CPU do all of the work. Using these FPGAs for kernel preprocessing, only relevant portions of the data being analyzed are passed on to the CPU to be processed — leading to exponential increases in analytics performance. And any spare FPGA processing cycles can be made available to perform data compression — again lightening the processing load on the CPU.

> ***IBM's Netezza appliance is ideal for processing complex, scanning type queries such as those found when performing deep analytics. It preprocesses data, and then feeds that data to the CPU in an optimal, balanced fashion. It be tailored for complex analytics applications that have high predictability. Further, it is not burdened by legacy database structures and online transaction processing features — resulting in a simple code path for faster performance. (Note: Netezza customers have observed that instead of one report per day, for example, they are now submitting many reports, analyzing various views of the data, and finding new questions to ask that they did not have the capacity to ask before).***

IBM's Smart Analytics Systems

As for optimally processing concurrent operational analytics queries (such as those run against enterprise data warehouses), IBM's Smart Analytics systems are designed differently than Netezza. In this workload scenario, IBM tunes its software stack and database to process a lot of threads (threads are elements of applications).

> ***We call this type of environment a "data load ready environment". The hardware, the stack, the database, and the query engines have all been tuned for optimal performance. All that is needed is access to data and report requests — and this system will execute in optimized fashion!***

Let's look more closely at IBM's Smart Analytics System 7700. It is a POWER micro-processor -based designed for processing data intensive tasks. The Smart Analytics System 7700 provides a data partition facility that helps maximize parallel computing performance. To improve analytic query performance, IBM has built materialized query tables (MQT) and multi-dimensional clustering (MDC) — programs that accelerate table reads/writes and indexing (more specifically, MDC is designed to help OLAP processing through the use of clustering indexes). IBM also applies deep compression methods (taking up less storage space and accelerating reads) not only to tables as other vendors do, but also to indexes, temp space, log records and backups. Using these facilities, IBM's Smart Analytics System 7700 can rapidly execute simple queries typically assigned to a single processing thread — or complex queries (which can be decomposed into operations that can be executed in parallel). Further, IBM offers functionality that can rewrite complex queries automatically to significantly improve performance.

> ***All of these extensions and improvements accelerate operational analytics processing on IBM's Smart Analytics System 7700. Notice how IBM has built a processor optimized for data-intensive processing work (the POWER7 microprocessor); has integrated its software to support a complete data warehouse/business analytics solution; and has tuned this solution for operational analytics workloads.***

Transaction Processing Optimization: An IBM Power 740, DB2, EasyTier Example

As stated earlier, the big challenge in transaction processing is feeding data to the CPU in a balanced fashion. IBM attacks this problem by optimizing data storage within its DB2 database using impressive data compression techniques, by buffering data — and by using SSD technology and storage caches. Data compression and data buffering reduces the need to perform as many storage subsystem reads/writes — while SSDs and storage caches move data closer to the processor for expeditious processing.

Another approach to speed-the-feed of data to the CPU is known as cold tiering and hot tiering. In essence, algorithms can examine how data is being used by observing reference patterns — and once it hits certain thresholds, that data can be assigned to a hot tier (SSD drives where it can be read and written very quickly), or to cold tiers (mechanical drives where reads/writes take place more slowly).

In a Power 740/DB2/Easy Tier configuration, IBM can configure a system with 16 cores, 64 gigabytes of memory, as well as SSD and hard drives (using its own Storwize V7000) — and optimize its DB2 database for transaction processing — and by doing so, using Easy Tier alone, IBM is able to achieve performance improvements over tuned HDD-only servers of 3x to 9x, or even more over untuned, non-workload optimized servers!

> ***In on-line transaction processing, solid state is used as primary storage and under control of Easy Tier — which is constantly monitoring data over a significant period of time to ensure that the most frequently access data is placed on fast SSD storage. This is an important distinction when compared to Oracle's Exadata configuration (described next).***

Oracle's Single Server Exadata Database Machine

Wikipedia describes Oracle's Exadata Database Machine as a "database appliance" that supports on-line transaction processing (OLTP) *AND* on-line analytical processing (OLAP) workloads. And Oracle's own website describes its Exadata as "the only database machine that provides extreme performance for both data warehousing and OLTP applications". This website goes on to describe Exadata as "a complete package of servers, storage, networking, and software that is massively scalable, secure, and redundant". See: http://www.oracle.com/us/products/database/exadata/overview/index.html?origref=http://search.yahoo.com/search;_ylt=AlYmVVHEqAiHJc6HAjc6MOGbvZx4?p=exadata&toggle=1&cop=mss&ei=UTF-8&fr=yfp-t-701 for a closer look at these Oracle claims. Oracle's description of Exadata also states that it delivers "a faster time-to-market by eliminating systems integration trial and error, and make better business decisions in real time".

What we like about the Exadata configuration is that it uses Infiniband connections for high-speed; it makes use of the latest/greatest Intel cores to process vast amounts of data; and it has some good software for offloading the CPU from having to deal with storage management tasks.

> ***But can Oracle's single system approach trump IBM's three different servers in terms of performance when it comes to distinctly different workloads? There are some major design differences between IBM's approach and Oracle's approach to building pre-integrated servers that seem to indicate that IBM's different-system-for-each workload approach is more targeted and thereby, more effective.***

Consider the following: Oracle's Exadata servers are supported by Oracle Exadata Storage Servers that use a massively parallel architecture "to dramatically increase data bandwidth between the database server and storage". These storage servers use solid state drives as a read cache that front ends hard disk drives. (Read cache sits inline between servers and storage — meaning that all the data being processed passes through it — but the active data is left on the SSD such that if can be accessed more quickly than having to read from the HDD storage subsystem).

> ***Now consider this: SSDs are very good at random reads and writes, as found in OLTP workloads and very simple analytic queries. But SSDs are far less effective at the large sequential reads found in complex analytic queries, particularly when implemented as a cache. Given this situation, it can easily be argued that Exadata is not as well optimized for complex analytics queries as the IBM Netezza environment that we described earlier. For complex OLAP workloads, where the benefits of SSD on large sequential are vastly less and not as cost effective, IBM places SSD within its servers, (close to the CPU) as overflow space for temporary data generated by large queries. This approach makes Netezza more effective when processing complex analytics workloads.***

As for on-line transaction processing workloads, readers should be aware that Exadata Storage Server software costs $10,000, plus $2,200 service and support per disk drive (this pricing can be found at http://www.oracle.com/us/corporate/pricing/exadata-pricelist-070598.pdf). This software provides hybrid columnar compression and other optimizations related to analytics processing — but sits mostly idle when used for OLTP workloads.

> ***So Oracle buyers who buy the Exadata Database Machine and accompanying Exadata Storage Server for transaction processing are essentially paying for software and support (per disk) that they don't need. As we mentioned earlier, price and performance are the primary drivers in the transaction processing marketplace...***

Another Oracle Concern: The Real Application Cluster (RAC) Design

According to Oracle, its Real Application Cluster design "enables a single database to run across a cluster of servers, providing unbeatable fault tolerance, performance, and scalability with no application changes necessary". This RAC design is central to Oracle's Exadata Database Machine architecture.

When it comes to workload optimization, Oracle RAC's resource allocation practices and the resulting behaviors that occur after resource allocation are worth closer scrutiny. Oracle RAC allocates resources for query processing upon a query's arrival. If the workload at that moment is light, a query can be allocated many resources and execute quickly. If the workload is heavy, it is "downgraded", receives fewer resources and takes substantially longer. *The result is unpredictable response times.*

Also, once assigned, resources are not reallocated as the workload changes. So down-graded queries, for example, are not given additional resources as those resources become available. As a result, queries continue to execute on the originally assigned resource set, even if the rest of the system is idle. *The results here are wasted resource and needlessly slow response times.* By contrast, DB2 and Smart Analytics systems use a different method for resource allocation that does not exhibit these characteristics. As resources are freed up by finished work, resources are smoothly reallocated to running queries and, as a result, IBM systems produce far more predictable, rapid response times.

Don't Take Our Word for It — Read ITG's Report

The independent technology research and analysis firm International Technology Group (ITG) of Santa Cruz, California compared one of IBM's data warehousing solutions (IBM Smart Analytics System) to those of Oracle and Teradata (leading competitors in the data warehousing field). ITG established its own evaluation criteria based on certain workloads (ITG closely examined complex mixed workloads and queries involving large sequential table scans — and then compared acquisition costs and three year ownership costs based on "street" pricing [i.e., discounted prices paid by users]). ITG then presented its findings in a report entitled: "Management Brief: Cost/Benefit Case for Enterprise Warehouse Solutions" which is available for free at:

http://www.informationweek.com/whitepaper/Hardware/Data-Centers/itg-management-brief-cost/benefit-case-for-enterp-wp1315500836?articleID=191703376

IBM's Astounding Cost Advantages

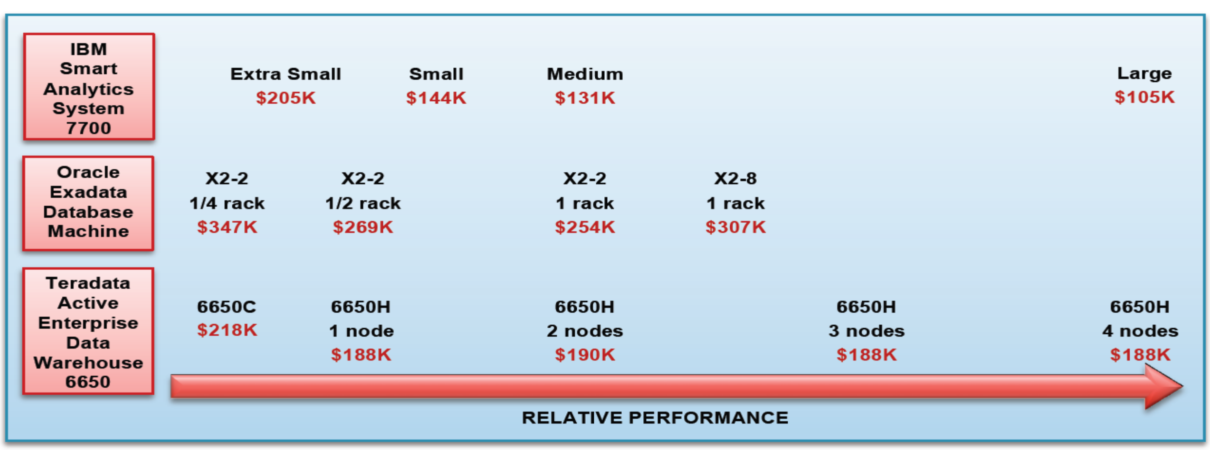
ITG's research showed that, in certain configurations offering approximately the same performance, "initial Smart Analytics System 7700 costs – including hardware and software acquisition, and installation – are 11 percent less than for Oracle Exadata Database Machine and 16 percent less than for Teradata Active EDW 6650 equivalents".

> ***However, disparities are significantly larger if three-year costs – including maintenance and support, database administration (DBA) personnel and facilities (energy and computer room occupancy) – are compared. Three-year costs for Smart Analytics System 7700 are 43 and 40 percent less than those for Oracle and Teradata systems respectively.***

ITG attributes this large disparity in three-year costs to "higher Oracle and Teradata pricing for maintenance and support. Higher personnel and facilities costs also contribute".

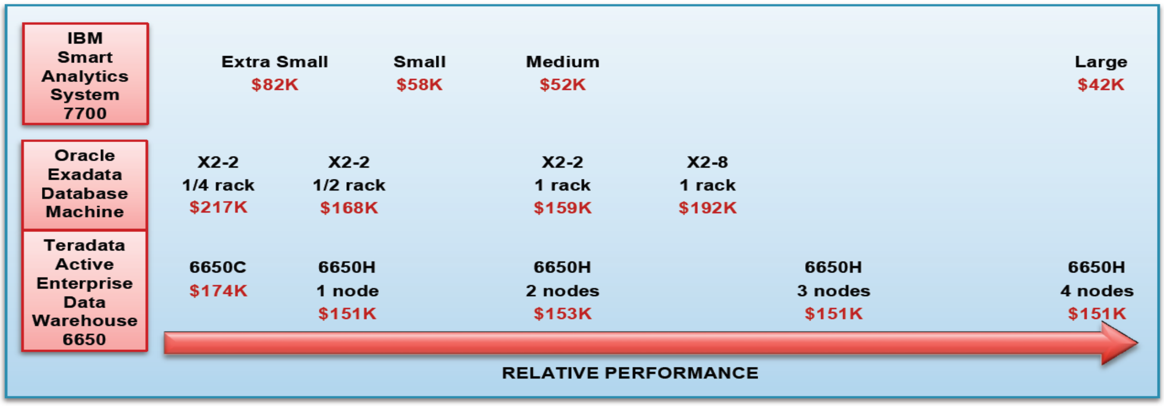
ITG also compared cost per terabyte for user data (both compressed and uncompressed) — and found the following (uncompressed comparison Figure 4, compressed Figure 5). The reason that these compressed/uncompressed data comparisons are important is because they show how each vendor handles the processing of data (the amount of space available for the customer data net of indexes, log files, temporary spaces and other structures) — illustrating how efficient each architecture is, and also illustrating capacity constraints.

Figure 4 — Cost Per Terabyte of User Data (Uncompressed)



Source: International Technology Group, Santa Cruz, California, June, 2011

Figure 5 — Cost per Terabyte of User Data (Compressed)



Source: International Technology Group, Santa Cruz, California, June, 2011

According to ITG: "if user data capacities are measured in non-compressed data, IBM Smart Analytics System 7700 configurations show the lowest three-year cost per terabyte… If, however, capacities are measured in compressed data, the gap between Smart Analytics System 7700, and Exadata Database Machine and Active EDW 6650 models widens. ***Smart Analytics System 7700 costs are two to four times less than for competitive systems***". Further, ITG goes on to state that IBM's "Smart Analytics System 7700 three-year costs per terabyte for non-compressed data average 57 percent less than for Exadata Database Machines, and six percent less than for 6650 systems. ***For compressed data, three-year costs per terabyte [for Smart Analytics System 7700] average 72 and 54 percent less respectively***".

ITG's Conclusions

ITG's conclusions after comparing IBM, Oracle, and Teradata specialty server solutions were separated into four categories: 1) workloads; 2) compression; 3) scalability; and, 4) automation.

From a workload perspective, ITG concluded that the capabilities of Exadata Database Machines are significantly different from Smart Analytics System 7700. According to IDG, Oracle's "Exadata Database Machines deliver their best performance for workloads characterized by large sequential table scans. Such workloads are generated by applications that are structurally simple, but require a great deal of processing power; e.g., identifying and collating specific variables in large volumes of records". ITG goes on to say that "the entire Exadata Database Machine architecture is geared to these types of applications and workload".

In the area of compression, ITG points out that the extent that systems can compress data without unacceptable performance degradation has a major impact on system capacities… [by using compression techniques properly] system processing may also be accelerated, and I/O throughput times reduced".

> ***According to ITG: "IBM DB2 9.7, which is employed by Smart Analytics System 7700, features one of the industry's most effective across-the-board implementations of data compression. Compression extends to***

> ***rows as well as indexes, temporary tables, log files, large objects and other data structures. Users have routinely experienced overall compression levels of 55 percent to more than 85 percent. Oracle employs two compression technologies. Database Machines implement Advanced Compression, which is a feature of Oracle Database 11g. Although higher levels have beenachieved, users have found that unacceptable performance degradation typically occurs above approximately 25 percent".***

From a scalability perspective, ITG points to the key differences in the design points of IBM's Smart Analytics Systems versus Oracle's Exadata environment based upon shared everything versus shared nothing architectures. IBM's shared nothing Smart Analytics System 7700 is capable of scaling to very large configurations (in the real world, seven production systems report the ability to process between 40TB to more than 120TB of user data). Because Oracle Database Machines implement RAC architecture, which is built around a "shared everything" model where components such as memory are shared, there is more opportunity for resource contention which can lead to scalability limitations.

IDG's final area of comparison has to do with the amount of management automation that vendors provide with their systems. ITG concludes that "IBM Smart Analytics Systems and Teradata systems typically require fewer DBAs and other IT personnel than Oracle equivalents". IDG also points out that the impact of automation when databases undergo frequent changes (less effective automation tools mean more labor is involved).

Summary Observations

In this *Research Report* we have tried to show how systems can be designed differently to support specific workloads. In IBM's case, the company looked very closely at the characteristics of business analytics workloads and transaction processing workloads. And IBM determined that deep analytics workloads have different processing requirements than operational analytics workloads. (Deep analytics workloads focus on rapid response time; while operational analytics focus on throughput in order to generate many reports). As for transaction processing, IBM determined that the focus of a transaction processing workload is price/performance. And, as a result of examining the characteristics of these workloads, IBM built three separate configurations, each designed to deliver maximum performance for each specific workload.

We also took a look at Oracle's Exadata Database Machine environment. Oracle has positioned this machine as a single architecture designed to serve all three of the workloads described above. Can this architecture process these workloads more optimally than IBM? Exadata makes use of high-speed Infiniband networking technology, it uses the fastest CPUs that Intel offers, and it has some sophisticated software for offloading storage management tasks from the CPU. But is it better at handling complex analytics, operational analytics, and transaction processing than the solutions offered by IBM?

As we said at the outset, this report articulates the differences we see in how each vendor approaches workload optimization. Which approach is better? You decide…

Clabby Analytics
http://www.clabbyanalytics.com
Telephone: 001 (207) 846-6662

March, 2012