InComparison

IBM Smart Analytics Systems vs Oracle Exadata X2-2

An InComparison Paper by Bloor Research
Author : Philip Howard
Publish date : December 2010

When we commenced this exercise we expected to find that there were some areas in which IBM excelled and others in which Oracle did so. We have been surprised to find that that is not the case and that the IBM Smart Analytics System out-competes Oracle Exadata in almost every area we have examined

Philip Howard

Executive summary

This paper is organised into two parts: this section, together with the conclusion, is intended for executives who have no, or limited, technical knowledge or interest, and the remainder of the paper, which is intended for the more technically minded. Information in this section, in particular, is duplicated in the remainder of the paper, though we go into more detail there.

The basic theme of this paper and its companion piece (IBM pureScale Application System vs Oracle Exadata X2-2) is to provide a comprehensive comparison of IBM and Oracle's offerings for data warehousing and on-line transaction processing (OLTP) respectively.

Oracle's view of these two sets of requirements is that a single solution, Oracle Exadata, is ideal to cover both of them; even though, in our view (and we don't think Oracle would disagree), the demands of the two environments are very different. IBM's attitude, by way of contrast, is that you need a different focus for each of these areas and thus it offers the IBM pureScale Application System for OLTP environments and IBM Smart Analytics Systems for data warehousing.

In practice, IBM's approach is not quite as simple as this. In terms of data warehousing there are, in fact, two possible approaches: to license DB2 and InfoSphere Warehouse or an IBM Smart Analytics System. The Smart Analytics offerings contain the former but the systems have been pre-built and pre-tuned in conjunction with the particular hardware platform chosen, together with any optional modules such as a Cognos business intelligence module. In effect, licensing DB2 and InfoSphere Warehouse is for those preferring a more DIY approach while the IBM Smart Analytics System is for those that want a complete system ready to run. A similar concept applies to IBM's OLTP offerings.

IBM refers to its approach as "workload optimised systems". That is, these offerings, and particularly the packages, have been designed and optimised for their specific environments. These include System x (running Linux or Windows), System p (running AIX) or System z (on a mainframe, running z/OS).

The question we will address in this paper is which of these two approaches is best. Of course, one could make theoretical arguments in favour of either Oracle's or IBM's approach, in which case we could argue until the cows come home. Whether one concept is better or not from a theoretical point of view is beside the point; what counts is which is best in terms of performance, scalability, ease of administration and management, and cost.

The first point to make is that the smallest available Oracle Exadata offering has 6Tb of usable data. With compression this equates to at least 18Tb and quite possibly a lot more. Conversely, the smallest Smart Analytics System has 330Gb of usable disk space. So we can immediately conclude that the Smart Analytics System will be suitable for use by many small and medium sized organisations, as well as departments, for which Oracle Exadata will be too large and too expensive. Of course, those users might opt for Oracle Database 11g Release 2 on its own but that is outside of our consideration.

Another distinction is IBM's approach to upgrading the Smart Analytics System. Unlike IBM, Oracle is prescriptive about Exadata: you can have a ¼ rack, a ½ rack or multiples of full racks. If you want to add new processing nodes you have to add one or more additional storage servers and, in addition to the hardware costs involved, you have to pay increased license fees to run the Exadata software, which comes at $10,000 per disk (and there are 12 disks per server) plus 22% maintenance; and all this because you need some processing power even if you don't need extra disk capacity!

Conversely, IBM offers a variety of models, running on a choice of operating systems that you can upgrade in a conventional manner. You can choose between models that do or do not offer solid-state disks and, if you do opt to use this technology then you can upgrade this also. At the opposite extreme all Oracle Exadata implementations come with flash storage even though it is primarily targeted, by Oracle, at OLTP environments.

In so far as performance is concerned, making comparisons on a theoretical basis (neither company has a published benchmark based on current systems although IBM currently holds the performance record for the TPC-H benchmark at the 10TB scale factor), it is not so much a question of comparing apples with pears but more like comparing bowls of fruit: both have apples but one has Cox's Pippin and the other Granny Smiths, while one vendor has bananas and grapes to go with their apples and the other has nectarines and kiwi

Executive summary

fruit. Thus we have the native capabilities of the Exadata Storage Server plus storage indexes and hybrid columnar compression and workload management from Oracle versus piggy-back scans, tokenised compression, multi-dimensional clustering, cubing services and workload management from IBM, plus all the normal performance characteristics of a database, together with the different ways that the vendors make use of flash disk. Not to mention the fact that there are somewhat different facilities offered in the Smart Analytics System on the mainframe as opposed to those running on distributed platforms.

Our view is that, depending on the particular requirements of the customer, one or other of the vendors may have a performance advantage. To be specific, we would expect Oracle Exadata to perform well with relatively static data (because you get better compression then) and when supporting complex queries of the type usually associated with analytic applications, data mining or statistical analysis. On the other hand we prefer IBM's cubing services, compression and workload management in particular. In general, a proof of concept will be the best way of determining which of the suppliers can provide the best performance characteristics to meet your workload and requirements.

Going into pricing itself, IBM has targeted its pricing for the Smart Analytics System at Oracle environments that do not require any software licensing. That is, those customers with an unlimited license agreement (ULA), at least for the database. For these users, a small Smart Analytics System, which has approximately the same processor performance and storage capacity as a ¼ rack Oracle Exadata system, costs around the same as its counterpart whereas the medium and large configurations are less expensive than their ½ and full rack competitors. Of course, if you don't have a ULA then these price comparisons are very much in IBM's favour. Bear in mind, however, that list prices are notoriously fickle and sometimes bear little resemblance to reality.

Apart from price, perhaps the biggest differences between the two systems are in manageability and flexibility. We have already discussed the latter with respect to the configurations and disk capacities that are available while, in terms of manageability, the Smart Analytics Systems are easier to install in the first place, easier to grow, and do not require application tuning at any point (which is the case when you add a node to an Oracle cluster).

There will be some environments where Oracle Exadata can out-perform IBM and there will be some where the reverse is true. However, in all other respects the IBM Smart Analytics System is, in our opinion, superior to Oracle Exadata: it is easier to manage and tune, easier to install, more flexible and costs (at least notionally) less money.

System descriptions

Before we begin to make any sort of comparisons we need to have a clear idea about the architectures of each product offering and what is and is not included within each.

IBM InfoSphere Warehouse

IBM InfoSphere Warehouse might be thought of as just DB2. However, DB2 in the InfoSphere Warehouse is surrounded by a suite of tools to simplify the creation and management of data warehouses that have been built on DB2. The Linux/UNIX/Windows (LUW) edition includes the DB2 server while the z/OS edition does not, as most interested customers already own a license. Included within the InfoSphere Warehouse are graphical and web-based tools for the development and execution of physical data models, data movement flows (SQW), OLAP analysis (Cubing Services) and data mining (LUW edition only at this time).

As far as this paper is concerned we will not be discussing InfoSphere Warehouse, per se, as all of its elements are included within the Smart Analytics System.

IBM Smart Analytics System

The difference between the IBM Smart Analytics System and DB2 with InfoSphere Warehouse is that the latter simply consists of software that runs on any platform that you want to install it on and is not pre-configured, pre-tuned or pre-built in any way. In effect it is a do-it-yourself solution. The Smart Analytics System, on the other hand, is precisely available in a pre-configured, pre-tuned and pre-built fashion across all of IBM's operating system environments (mainframe z/OS, AIX, Windows and Linux), with a variety of different options that you can use from an entry-level system upwards. This pre-integration and tuning of the various components means that it is easier to get better performance and you can get faster time to value.

Oracle Exadata X2-2

There are actually two Oracle Exadata Database Machine products: Exadata X2-2 and Exadata X2-8. The latter is a full-rack only system, primarily intended for the largest OLTP and consolidation environments. It is not yet available and we will therefore be focusing on Exadata X2-2. This consists of Oracle Database 11g Release 2, Oracle RAC (Real Application Clusters) Database server grid, an InfiniBand interconnect, the Oracle Enterprise Linux operating system, and the Exadata Storage Server Grid using either High Performance (600Gb) or High Capacity (2Tb) disk storage, where the latter give more capacity but with lower performance.

The way that the system works is that data is stored in the Exadata Storage Server grid and the storage servers act as a sort of preprocessor for accessing data from disk in an optimised fashion, using what Oracle calls smart scans, before passing the results to the database itself. This significantly reduces the amount of data that the database has to process and is particularly efficient in data warehousing environments. In order to improve performance in OLTP environments, Oracle Exadata also includes flash storage for caching hot data.

You can have multiple databases running within an Exadata environment and you can have multiple small databases on a single RAC node or you can have larger databases that span nodes. This means that you can have an OLTP system sharing an Exadata environment with a data warehousing implementation. You cannot similarly share the Smart Analytics System with a pureScale environment. This is because Oracle uses a shared disk environment throughout whereas IBM uses shared disk for OLTP in its pureScale implementation but a shared-nothing architecture for data warehousing on Linux, UNIX and Windows. The exception is that the mainframe Smart Analytics offering provides a shared disk environment for both OLTP and data warehousing/business analytics. This is easily deployed with the logical partition option of the 9600 for the System z platform. On the downside, and admittedly it is a small point, you cannot re-purpose Exadata Storage Servers. If you decide in the future to move to some other vendor then you can re-use your RAC servers and you can re-use your IBM servers but it will not be easy to do the same thing with the Exadata Storage Servers because of their particular functional design.

Scaling your system

The implementation options for Oracle Exadata X2-2 are illustrated in Table 1.

	¼ Rack	½ Rack	Full Rack	2-8 Racks
Database Servers	2	4	8	16–64
Exadata Storage Servers	3	7	14	28–112

Table 1: Oracle Exadata 2-2 configuration options

Note that, when upgrading, these are the only options available: a ¼ rack can be upgraded to a ½ rack and a ½ rack to a full rack; you cannot have a ¾ rack, for example, and you can only have whole numbers of racks above a full rack. A quarter rack configuration holds 21Tb of raw disk capacity and a full rack (using High Performance drives) contains almost 100Tb of raw disk capacity. If you are using High Capacity drives the capacity of a full rack is 336Tb. Each Exadata Storage Server also includes 4 flash cards, with a capacity of 96Gb each, scaling up to around 5Tb on a full rack. Note that you cannot scale upwards without adding extra disk: this means that you cannot simply add new processing capacity if you have, say, a CPU bottleneck: you have to have additional storage capacity even if you don't need it. This is unlikely to be a major concern within a data warehousing environment.

In practice, of course, actual disk capacity and usable disk capacity are very different things. To begin with there is disk mirroring, which is needed for resiliency, which halves your available capacity and then there are considerations with regard to space needed for logs, temp space, indexes and so on. Oracle's own estimates are that 55% of disk capacity, before taking account of mirroring, is actually usable for storing data, which means that a ¼ rack actually provides around 6Tb of usable space, a ½ rack 14Tb and a full rack 28Tb. Of course, these figures do not take account of compression.

IBM offers six models for the Smart Analytics System:

- 1050 – an entry level system based on System x (Linux or Windows) with 3 size options, each with one data module (see below), 2 sockets, a maximum of 12 cores, a maximum of 48GB of memory, and total disk capacity of either 0.9, 4.2 or 7.2Tb with user capacities being 0.333, 1.65 and 3.3Tb respectively. Compression is not included as standard with the 1050.

- 2050 – a larger departmental model based on System x with 3 size options each, with one data module, 4 sockets, a maximum of 32 cores, a maximum of 1TB of memory and spinning disk capacity of up to 28.8Tb (13.2Tb user space). Compression not included. This system, together with the 1050, would typically be deployed in environments with less than 100 users.

- 5600 – a medium-sized configuration based on System x with 4 standard size options with 2 to 8 data modules (though you can have more than this), each containing 6 cores and 64Gb of memory. Spinning disk capacities range from 14.4 to 57.6Tb, which provide (with compression at 2.5x) total user capacities of between 24 and 96Tb. This system, together with the 5600 with SSD option, would typically be deployed for as many as a few hundred users.

Scaling your system

- 5600 with SSD Option – the largest system available on System x with 4 standard size options with 2 to 8 data modules, each (again, this not a limit) containing 12 cores and 128Gb of memory. Disk capacities range from 21.6 to 86.4Tb. In addition, the 5600 with SSD option includes between 1.28 and 5.12Tb of solid-state disk. Total usable capacities, after compression, range from 24 to 96Tb.

- 7700 – the most powerful of IBM's offerings for non-mainframe environments, this is built around Power7 servers running AIX. There are six base options with 1, 2, 3, 6, 10 or 20 data modules though both larger and in-between sizes are available. Each has 16 cores and 128Gb of memory. Spinning disk capacities range from 28.8 to 560Tb and solid-state disks from 0.7 to 14 Tb. There is also the option to add further solid-state disk capacity, up to 80.4Tb. Total user capacities go from 28.75 to 575Tb after compression is taken into account. This system would be suitable for 800 or more users, according to IBM.

- 9600 – IBM's mainframe version of the Smart Analytics System provides approximately the same performance as the 7700 but scales, in terms of disk capacity, somewhere between the 5600S and 7700 (nearer the latter). This offering includes the zEnterprise or z10 server with DS8700 or DS8800 disk drives. A big advantage is the very low latency achieved when loading transactional data from System z. Another key feature of the 9600 offering is the ability to add resources to an existing System z to support a 9600 logical partition (LPAR) on the same system. This lowers deployment costs and avoids the support problems and expense introduced when installing separate discrete servers for individual workloads. Today, Oracle is being successfully run on Linux on System z, particularly for consolidating workloads, but Exadata does not exist on System z. The 9600 can be extended with IBM's latest offering, the IBM Smart Analytics Optimizer, designed for large table scans.

It should be apparent from these details that with IBM, unlike Oracle Exadata, you can start very small: with systems involving less than 1Tb of data. While we will discuss pricing in due course this strongly suggests that Oracle Exadata does not scale down and will not be suitable, or will be overly expensive, for at least some departmental environments and for small and mid-sized organisations that do not have large amounts of data to analyse.

The IBM Smart Analytics System also differs from Oracle in the way that you can upgrade. The latter is relatively inflexible but the Smart Analytics System is built on a modular basis with each system consisting of a foundation module, one or more data modules and a variety of optional modules. These optional modules include admin/user modules, failover modules, management modules and application modules. The idea here is that if you need extra disk capacity then you license an extra data module, but if you need to support additional users then you license an extra user module, and so on. In other words you pay for what you need. This may not be the case with Oracle Exadata where, if you need more CPU capability, you have to have more disk.

Compression

Oracle uses two different types of compression: one, which is known as 'advanced compression', has been designed specifically for OLTP environments and is discussed in more detail in the companion paper to this; and what is known as 'hybrid columnar compression', which is used by Oracle in data warehousing and archival environments. The advantage of using column-based compression is that each column has its own datatype and therefore you can optimise your compression algorithms for specific datatypes. In theory, therefore, columnar compression should be optimal though in practice some algorithms work better than others. However, as its name suggests, Oracle's compression is not purely columnar but hybrid. This is because the Oracle database is not itself columnar but is row-based. This means that you have to convert rows into columns for compression purposes, which means that load times are slowed down and, as the database processes data in rows it also means that you have to decompress and convert the data back into row format for processing purposes. So, although the compression may give you significant space savings you will lose some of the performance benefit gained from the reduced I/O that compression provides. In addition, you need to bear in mind that the way Oracle's compression works is that you have a series of compression units. This is fine except that when you update a compression unit the entire unit is locked. So, if you insert a new row then this is written to a new page outside the

Scaling your system

compression unit. Conversely, if you delete a row then the relevant space in the compressed storage is not reused (until you do a reorg) so that the effectiveness of compression gradually deteriorates over time. Further, this will result in erratic query processing times because some data you are accessing may be in the original compression unit while other data is in new pages (which are not ordered in any way and therefore will likely have inferior compression ratios) and because there are empty blocks interspersed throughout the compression units.

IBM uses a completely different technique for compression: it looks for repeating patterns (which may be either within a column or across columns) and then replaces each of those strings with a symbol (token) backed by a system-wide dictionary of symbol-string equivalences. Thus if you have 5,000 instances of 'Milwaukee' in your database you store 'Milwaukee' just once plus 5,001 instances of the symbol. The same approach is used for index compression and it useful to contrast this with Oracle's approach: suppose that you are compressing a customer index then every time that 'Bloor' appears Oracle will store 'Bloor' plus a row ID. So if there are 250 entries for Bloor there will be 250 separate Bloor-Row ID pairs. IBM however, would only store Bloor once, followed by a string of 250 Row IDs. Thus IBM's approach is more efficient for compressing indexes.

Oracle's data compression, at its best, will produce more space savings than IBM's. However, as we have noted that space saving will deteriorate. In practice, Oracle claims 3 to 10 times savings while IBM claims up to 4 to 7 times. Taking averages we would expect Oracle to do better when it comes to data. On the other hand, Oracle is inferior when it comes to compressing temp space and indexes so, overall, we would expect there to be little in it or somewhat in IBM's favour. However, there is also the issue of performance—because data is compressed you have fewer I/Os required to read the same amount of data—in this case, because of the row-column-row conversions required within an Oracle environment, we would expect the performance gains (IBM estimates them at 40%) from IBM's compression to surpass those achievable in an Oracle Exadata environment. Add this in to the space saving mix and we prefer IBM's approach.

Performance

Unlike the TPC-C benchmarks for OLTP processing where both Oracle and IBM have posted results within the last 12 months, there are no comparable results for TPC-H, which is the benchmark for data warehousing. IBM does hold the current performance record for 10Tb systems but that is not with the current version of DB2 and there is no available figure for Oracle Exadata. However, that should not concern us unduly as benchmark results are no more than indicative. We will therefore proceed directly to a discussion of particular elements of each system that militate in favour of good performance. Some of these features we have already discussed, notably the advantages that IBM has in terms of compression performance.

Flash (solid-state) storage

Oracle Exadata comes with flash disk as standard but it is only standard on some IBM systems, as previously discussed. However, Oracle sees the use of flash primarily as something to speed up OLTP processing, whereas IBM uses it more generally. The IBM pureScale Application System does not come with solid-state disks though you can use them if you choose to. There are two major differences between IBM (with Easy Tier) and Oracle in their use of flash. The first is that Oracle uses a PCIe flash card as opposed to solid-state disks. The advantage of this is that you don't have a disk controller between the flash storage and processor, which can potentially slow the environment down if the disk controller has not been designed to operate at flash speeds. The other difference between Oracle and IBM is in the way that the two companies use flash. Oracle, which refers to its technology as the Exadata Smart Flash Cache, actually uses this as a read cache. That is, it copies hot data from storage into the cache. Deciding what data should be held in the flash cache is handled automatically, though users can define directives at the database table, index or segment level to ensure that specific application data is held in flash, subject to the proviso that the software is smart enough to know when data will not fit into the cache.

IBM, on the other hand, offers two approaches to the use of flash (solid-state disks):

1. As standard on the 5600 and 7700, it is used for temp space. That is, where you store temporary data such as the results of sort-merge joins within a much larger query. Of course, multiple queries may use up temp space concurrently and access from those queries to the temp space is effectively random rather than sequential, which is why it will benefit from being placed on a solid-state disk, because disks in data warehousing environments are better suited to sequential access.
2. With the IBM System Storage DS8700, which is available with the 9600, (and, in due course, on mid-range storage servers) it is used in conjunction with IBM's Easy Tier technology. The idea here is that some (hot) data is held on solid-state disks (SSD arrays) and the remainder on conventional hard drives, with the hottest data migrating up to SSD arrays or (when it gets colder) down to hard disks as appropriate, with the relocation of data (which can be as little as 1Gb at a time) being automatically handled by the disk system's Easy Tier feature. Easy Tier also provides the ability to dynamically relocate individual logical volumes from one tier to another (for example, from faster spinning disks to slower ones) and to change a volume's RAID type.

Easy Tier clearly has broader capabilities and can be used more widely (it is application agnostic so will be suitable for both OLTP and data warehousing environments) than the Exadata Smart Flash Cache. However, it has limited availability at present. For most companies the choice will be between the Exadata Smart Flash Cache and the use of solid-state disks for temp space. The former is targeted by Oracle at OLTP environments while temp space is especially important in warehousing environments, so we would have to expect that IBM's approach will be more useful, given the context of this paper.

Database

Bloor Research has regularly conducted performance comparisons between the DB2 and Oracle database systems, in 2003, 2005 and 2007 prior to the current paper. We do not intend to go into the sort of detail that those reports have done, otherwise we would double the length of this paper! In general there are individual features of each product that we like but these roughly balance out. The same tends to be true as the two vendors leapfrog one another as new versions of their respective database systems are released. Of those elements that pertain to data warehousing, as

Performance

opposed to OLTP, we have historically seen the two engines as more or less comparable. For example, we have preferred Oracle's indexing and preferred IBM's cubing services (to support OLAP) as well as its tuning and management capabilities. However, Oracle has made significant strides in reducing its administrative requirements so that it has narrowed this gap, but we still rate it as lagging behind IBM.

However, there are a number of elements within the two databases that require particular discussion.

Large table scans

Certain types of queries require whole tables to be read. Where these tables are very large and/or where complex queries require multiple tables to be scanned, this can have a serious impact on performance. Both IBM and Oracle have techniques that directly address this issue, as well as complementary techniques (discussed in the next section) that also help to alleviate this problem.

This is what Exadata, as an add-on to the Oracle database, was specifically designed to address. Put simply, it puts processing close to the disk to provide an MPP-like, shared-nothing approach. Data is streamed off disk and pre-processed locally before the results are passed to the database itself for final processing. This results in faster table scans and reduced traffic across the network between the Storage Server and the database. The actual technique that Oracle uses is known as Smart Scans. While in general it will significantly speed up query processing especially, but not only, for large table scans, it does have one weakness. This is that Smart Scans, which operate at a lower level than the database, are turned off if there is a write to the table in question. In other words it will be best suited to static environments or situations where it is acceptable to update data on a periodic basis. If you are operating in real-time or near real-time mode then Smart Scans will need to be reserved only for that part of your queries that access historic data.

IBM takes a different approach, depending on whether the Smart Analytics System is implemented on a distributed or mainframe environment:

- DB2 for LUW, of course, is a shared-nothing database but it does not put processing as close to disk as Oracle Exadata does. Historically, it too has suffered with performance issues with large table scans. To overcome this, it offers a facility called piggy-back scans. The idea here is that one query can 'piggy-back' onto another. Suppose queries a and b both need to scan a particularly large table. Query a starts. One minute later query b starts. For the duration of the scan initiated by query a both queries retrieve the data they need from this table. Meanwhile query b scans the portion of the table that query a scanned in the first minute. There are two net results: first, query b finishes one minute earlier than it would have done otherwise and, secondly, you are using much less I/O resources.

  Clearly, IBM's approach to large table scans is workload dependent. If you have a lot of users and queries addressing the same tables, and if they are requiring complete scans, then you will get significant benefits when using piggy-back scans. However, the number of complex and other queries requiring whole table scans may be quite low, in which the benefits of piggy-back scans will be similarly low. To sum up with respect to distributed environments: neither Oracle nor IBM has an ideal solution and each will have advantages over the other depending on the situation and workload.

- With DB2 for z/OS, IBM has recently introduced new facilities for handling the large table scan issue through the use of the Smart Analytics Optimizer working in conjunction with a zEnterprise BladeCenter Extension. This Extension combines compressed, columnar storage with a multi-parallel processing architecture. When queries are received the Smart Analytics Optimizer will intelligently route queries, or parts thereof, either to the normal row store or to the columnar Extension (the latter being particularly useful for large table scans) as appropriate. In many ways this is directly comparable with the approach taken by Oracle with Exadata.

Storage indexes and multi-dimensional clustering

There are a number of database techniques employed to try to improve performance in general and to avoid the need for large table scans. The most obvious of these is the use of indexes but these increase disk requirements,

Performance

slow down load speeds and the environment becomes more difficult to tune as you add more indexes. A second method is to use partitioning. While Oracle and IBM have comparable support for traditional indexing, historically they have had very different approaches when it comes to partitioning, with Oracle aiming to offer you a flexible approach that supports a wide variety of options and IBM focusing almost exclusively on hash partitioning (it is only in recent years that it has added range partitioning). IBM's implementation of hash partitioning in particular is very good but this must be set against Oracle's flexibility.

However, there are two features—storage indexes in the case of Oracle Exadata, and multi-dimensional clustering in the case of IBM—for which their competitor has no equivalent. Storage indexes are not indexes against the data per se but against the storage devices themselves: in effect they tell you what is (or is not) in each disk block. Thus when you are reading from disk you know which blocks to look in and which not to, thus saving a significant amount of I/O and thereby increasing performance. Multi-dimensional clustering, on the other hand, can be thought of as a type of sub-sub-partitioning. Typically, you implement hash partitioning; then range partitioning by, for example, month; and then, within each month, you cluster the data on disk around whatever dimensions are relevant. In effect, you reduce I/O depending on the extent of the cluster and improve performance in a proportionate manner.

Both of these techniques will prove useful. However, if we had to choose one or the other we would opt for storage indexes as they are more general and do not have the same level of tuning and administration.

Workload management

It is typically the case in an enterprise data warehouse that you have all sorts of different queries, of different types, that need to be run concurrently. Some of these have low priority and some of them have high priority, some of them require a lot of I/O and others very little, some of them need a lot of CPU cycles and others almost none, some of them need to be responded to within seconds or in sub-seconds while others can happily take hours. The question is how you balance these user demands in order to meet service level agreements and keep users happy. The answer is workload management but not all workload management systems are equal.

IBM uses a tiered approach to workload management whereby user requests are assigned to a service class that may be designated as high, medium or low priority and against each of these you can set thresholds as to the amount of resources that may be assigned. System requests have their own service class. Oracle, on the other hand, uses a process based on resource groups whereby applications are assigned to particular nodes. For each group you can set the maximum CPU allocation for an application, set the degree of parallelism and so on, and you can also pass directives to the Oracle I/O Resource Manager, which is used to keep disks (but not flash) performing efficiently (for example, preventing low priority tasks from flooding disks). The problem with this approach is that it is less granular than IBM's. You could define the same things and get more or less equivalent results but you would have to define (and maintain) a great many resource groups, which will add to your administrative burden.

OLTP in data warehousing environments

You will not normally run OLTP within the same environment as data warehousing though with Oracle Exadata you can dedicate one part of your cluster to OLTP and the other to data warehousing, and similar capabilities (though implemented differently) are available from IBM on the mainframe. However, that is not to say that you will not have OLTP-like applications running within a data warehouse. A classic example is master data management (MDM). If you think about MDM it is, essentially, an OLTP style system with random rather than sequential I/O and lots of look-ups, updates and inserts. For this reason we do not advocate the use of data warehouses for supporting MDM. Nevertheless there are some companies that do implement MDM in their warehouses and both Oracle and IBM support this capability. Even without MDM, however, there are environments, especially (near) real-time ones, where similar characteristics may be exhibited, often in conjunction with more traditional query processing. For example, a hedge fund might want to run complex queries against historic data in conjunction with real-time stock tick data. Of course, this can be accomplished by federating the warehouse with a front-end OLTP system or a complex event processing system but some users may prefer

Performance

to have all these capabilities in one place. Where this is the case, mixed workload management becomes especially important and Oracle's Smart Flash Cache is also likely to be of significance.

In so far as genuine hybrid environments are concerned both Oracle Exadata and the IBM Smart Analytics System 9600 represent realistic options.

Managing growth

We have already discussed the general approach to growth adopted by the two vendors, with Oracle offering rack-based growth and IBM taking a more traditional stance, allowing you to add extra capacity (memory, disk and so on) as required. Thus, expanding the capacity of the Smart Analytics System is relatively straightforward, provided you do not completely outgrow your chosen system.

Adding a new node to an Oracle RAC implementation, on the other hand, means provisioning the new node, installing CRS, installing the RAC software, adding LISTENER to the new node, adding the database software, manually adding an ASM instance and manually adding a database test instance. In addition, with Oracle RAC, applications need to be cluster-aware in order to optimise the performance benefits associated with the cluster, which means that applications may need to be tuned as the environment scales up (or down for that matter).

Administration and management

A major feature of DB2 is its support for Oracle environments. You can import Oracle schemas directly into a DB2 database and DB2 has native (not emulated) support for Oracle concurrency control (but DB2 does it in a different way in order to avoid the locking issues that cause performance degradation in Oracle environments), SQL, PL/SQL, packages, built-in packages, OCI (Oracle call interface), JDBC, on-line schema changes and SQL*Plus scripts, amongst other features.

What this all means is that the vast majority of applications, stored procedures and other constructs written to run against an Oracle database will run unchanged, possibly with better performance because of the improved locking, against a DB2 database. According to IBM it has tested more than 750,000 lines of PL/SQL and it has achieved an average compatibility of 98.43%. This is truly impressive.

We have already commented on the fact that although Oracle has made significant strides in reducing its administrative requirements in recent releases, we still do not believe that its autonomic and self-tuning capabilities match those of IBM. Apart from this, perhaps the biggest differences are in implementation and high availability.

Implementation

While we do not have figures for installing an Oracle Exadata system, we do have such figures for an Oracle Database 11g Release 2 implementation on a 4 node cluster: according to independent research conducted by the Winter Corporation there are 208 steps involved in installing such a system. By way of contrast, an IBM Smart Analytics System is built, tested and the software installed on IBM premises before being delivered to you. It is taken apart during the physical shipment process and then the engineer will cable it up for you. At that point you can start loading data. In other words 208 steps for RAC versus none for the Smart Analytics System, and probably more than that for Exadata. Similar (though not so extreme) differences apply when it comes to upgrades and fixes, which is a single installation across all software components in the case of IBM.

High availability

Oracle relies on software-based disk mirroring: when a disk fails, this is automatically detected by the software and it recreates (and rebalances) new mirrors on other good disks. In addition, Oracle offers a high redundancy option whereby you can set up three copies of the data so that double failures cannot cause a problem.

More generally, Oracle claims that it has no single source of failure. Technically, this is true. However, there is only a single disk controller in each Exadata Storage Server, which means that you have to failover to another Storage Server if a disk controller fails.

High availability is different in the Smart Analytics System depending on the platform. On System x the architecture is such that you designate high availability groups with a spare server to act as a failover device. This is reasonable given the relatively low cost of System x servers. On the 7700, on the other hand, while you can adopt this approach if you wish, it is more typical to have servers in groups with four nodes plus a spare that will act as the failover server for the rest of the group. The 9600, on the other hand, provides an SMP architecture with shared disk, so all processors can access all the data to provide a highly available environment. The 9600 provides workload management capabilities that allow you to run utilities (backup, reorgs, updates and so on) concurrently with query workloads. Of course, not all data warehouses necessarily require high availability, which is why IBM's failover modules are optional, even where high availability is part of the standard offering. In general, IBM ensures redundancy throughout its servers, with dual adapters, dual controllers, dual cables and so on. From a software, as opposed to a hardware, perspective, IBM also offers mirrored logs, which will further enhance high availability.

Costs

Oracle has an unbundled pricing structure for Exadata while IBM has a bundled approach for the Smart Analytics System. Thus, in the case of the former, you have to separately license the database itself, RAC, partitioning, advanced compression and the tuning and diagnostic packs. Table 2 illustrates the list price for different Exadata configurations, including these additional components as well as first year maintenance and support, against comparable figures for Smart Analytics Systems configurations with similar server performance characteristics and storage capacity. Note that the disk capacities quoted are for usable, uncompressed capacities, based on Oracle's stated assumption of 55% capacity after allowing for RAID formatting. The IBM capacities vary between the 5600 and 7700 models, as shown. The Oracle pricing does not include OLAP, data mining or text mining all of which are included within the Smart Analytics offering.

These IBM prices look very attractive by comparison to Oracle's and it is worth bearing in mind that IBM has set its prices to be competitive with, or less expensive than, Oracle's even if you have an Unlimited License Agreement (ULA) so it is hardly surprising that they significantly undercut Oracle when software licenses are included. This theme continues with other models: for example, a 2050 with disk capacity equivalent to that of a ¼ rack system has a list price of just $164,394, though that doesn't include compression.

Of course the proviso must be made that these are only list prices and are subject to potentially substantial discounts.

Two other points must be borne in mind when considering costs. Firstly, upgrade costs. With Oracle you can only upgrade from a ¼ rack to a ½ rack to a full rack and then by adding additional racks. You cannot, for example, add additional database servers separately from storage servers (or vice versa) and, with storage server licensing costing $10,000 per disk drive (plus 22% maintenance), this is an expensive option if you only need extra compute power. Conversely, you can upgrade a Smart Analytics System without such limitations. You do, however, have the downside that any particular system will only expand so far, so there may come a point at which you need to change from, say, a 2050 to 5600.

Secondly, if we are correct in our assertion that DB2 is more easily manageable than Oracle and the Smart Analytics System environment than Exadata, then we would expect the latter to require additional administration over and above that needed by IBM. This in itself represents an expense.

Oracle system	Oracle list	IBM configuration	IBM 5600	IBM 5600 with SSD	IBM 7700
¼ rack (6Tb)	$2.32m	3 data modules (10/11Tb)	$1.15m	$1.74m	$2.68m
½ rack (14Tb)	$4.73m	6 data modules (20/22Tb)	$1.56m	$2.65m	$3.73m
Full rack (28Tb)	$9.30m	12 data modules (30/33Tb)	$2.14m	$3.68m	$4.69m

Table 2: comparison of list prices for equivalent configurations

Conclusion

When we commenced this exercise we expected to find that there were some areas in which IBM excelled and others in which Oracle did so. We have been surprised to find that that is not the case and that the IBM Smart Analytics System out-competes Oracle Exadata in almost every area we have examined. The exception is performance. Putting processing close to disk for pre-processing purposes, combined with Oracle's use of storage indexes, means that there will be certain types of query and, therefore, certain types of environment, for which Oracle will out-perform IBM on distributed platforms. However, these environments are relatively specific and focused on complex analytics. [Note: although it is outside of the scope of this paper, IBM has recently acquired Netezza. We would expect Netezza to out-perform Oracle, by a significant margin, in precisely these areas].

On the mainframe things are potentially different: with the Smart Analytics Optimizer and the zEnterprise BladeCenter Extension we see no reason why Oracle should have any performance advantages.

Moreover, in other types of environment and where there are mixed workload requirements, we do not see the issue of performance as nearly so clear cut and there will be a number of areas in which IBM out-performs Oracle. In other words, on a generalised basis, we cannot make a realistic judgement between the two vendors: it will depend on your specific circumstances.

In all other respects, from scalability to flexibility, through ease of use and high availability, to cost (at least at list prices), IBM appears to offer significant advantages.

Further Information

Further information about this subject is available from http://www.BloorResearch.com/update/2069

Bloor Research overview

Bloor Research is one of Europe's leading IT research, analysis and consultancy organisations. We explain how to bring greater Agility to corporate IT systems through the effective governance, management and leverage of Information. We have built a reputation for 'telling the right story' with independent, intelligent, well-articulated communications content and publications on all aspects of the ICT industry. We believe the objective of telling the right story is to:

- Describe the technology in context to its business value and the other systems and processes it interacts with.
- Understand how new and innovative technologies fit in with existing ICT investments.
- Look at the whole market and explain all the solutions available and how they can be more effectively evaluated.
- Filter "noise" and make it easier to find the additional information or news that supports both investment and implementation.
- Ensure all our content is available through the most appropriate channel.

Founded in 1989, we have spent over two decades distributing research and analysis to IT user and vendor organisations throughout the world via online subscriptions, tailored research services, events and consultancy projects. We are committed to turning our knowledge into business value for you.

About the author

Philip Howard
Research Director - Data

Philip started in the computer industry way back in 1973 and has variously worked as a systems analyst, programmer and salesperson, as well as in marketing and product management, for a variety of companies including GEC Marconi, GPT, Philips Data Systems, Raytheon and NCR.

After a quarter of a century of not being his own boss Philip set up what is now P3ST (Wordsmiths) Ltd in 1992 and his first client was Bloor Research (then ButlerBloor), with Philip working for the company as an associate analyst. His relationship with Bloor Research has continued since that time and he is now Research Director. His practice area encompasses anything to do with data and content and he has five further analysts working with him in this area. While maintaining an overview of the whole space Philip himself specialises in databases, data management, data integration, data quality, data federation, master data management, data governance and data warehousing. He also has an interest in event stream/complex event processing.

In addition to the numerous reports Philip has written on behalf of Bloor Research, Philip also contributes regularly to www.IT-Director.com and www.IT-Analysis.com and was previously the editor of both "Application Development News" and "Operating System News" on behalf of Cambridge Market Intelligence (CMI). He has also contributed to various magazines and published a number of reports published by companies such as CMI and The Financial Times.

Away from work, Philip's primary leisure activities are canal boats, skiing, playing Bridge (at which he is a Life Master) and walking the dog.

Copyright & disclaimer

2nd Floor,
145–157 St John Street
LONDON,
EC1V 4PY, United Kingdom

Tel: +44 (0)207 043 9750
Fax: +44 (0)207 043 9748
Web: www.BloorResearch.com
email: info@BloorResearch.com