Arquitectura Técnica, Innovaciones y Beneficios de Exadata

Francisco Javier Jiménez Orden
Grado de Ingeniería en Informática

Nombre Consultor
Ivan Rodero
Junio 2013

FICHA DEL TRABAJO FINAL

Título del trabajo:	Arquitectura Técnica, Innovaciones y Beneficios de Exadata
Nombre del autor:	Francisco Javier Jiménez Orden
Nombre del consultor:	Ivan Rodero
Fecha de entrega (mm/aaaa):	06/2013
Área del Trabajo Final:	Arquitectura de Computadors
Titulación:	Grado en Ingeniería Informática

Resumen del Trabajo (máximo 250 palabras):

Exadata es una máquina de Oracle que incluye servidores, discos, *switches* de comunicación, tarjetería *flash* y *softwar*e especializado, empaquetado y preinstalado, que mejora sustancialmente los tiempos de ejecución de acceso a las bases de datos Oracle para aplicaciones OLTP, *Data Warehouse* y mixtas, y que puede ahorrar costes comparado con las arquitecturas tradicionales como plataforma de consolidación de servidores de bases de datos.

Las novedades introducidas en los motores de bases de datos que se ejecutan en Exadata, y otros *appliances*, junto con su cada vez más extrecha relación con el *hardware* son de tal calado que, probablemente, nos encontremos ante la mayor evolución de los últimos años en la arquitectura de los servidores para la ejecución de bases de datos. Por otra parte, las capacidades de consolidación de Exadata, y los previsibles ahorros en costes en su adopción, parecen presuponer un fuerte despliegue en todo el mundo.

La memoria presenta las características técnicas de Exadata, resume las mejoras obtenidas en rendimiento y escalabilidad con casos de clientes y asocia las novedades tecnológicas con las mejoras de rendimiento y escalabilidad observadas. Adicionalmente, se resumen las características de otros *appliances* para bases de datos y se compara a Exadata, desde el punto de vista económico, con la arquitectura tradicional para servidores de bases de datos.

Abstract (in English, 250 words or less):

The Oracle Exadata database machine is an optimized package of servers, storage, communication switches, flash memory and specialized software that delivers extreme performance for OLTP, Data Warehouse and mixed workloads applications, where IT costs can be saved via Oracle databases consolidation.

The changes introduced in the database engine that runs on Exadata, and other appliances, along with his increasingly relationship with the hardware shows that we are facing the greatest development in recent years in the servers architecture for implementation of databases. Moreover, the Exadata consolidation capabilities, and the expected cost savings in its adoption, appear to assume a strong deployment around the world.

This work presents the technical features of Exadata, it summarizes the improvements made in performance and scalability with real cases and it links the new Exadata functionalities with the performance and scalability improvements observed in the cases. Additionally, the work summarizes the characteristics of other appliances for databases deployment and it compares Exadata, from the economic point of view, with traditional database servers architectures.

.

Palabras clave (entre 4 y 8):

Exadata appliances comparativas rendimiento inversión consolidación

Índice

Lista de figuras

1. Introducción

1.1 Contexto y justificación del Trabajo

Algunos proveedores de *hardware* y *software* para la explotación de bases de datos comerciales están ofreciendo al mercado nuevos productos, empaquetados, que incorporan mejoras de rendimiento y escalabilidad con respecto a las arquitecturas tradicionales para la ejecución de bases de datos. Estas mejoras están facilitadas por innovaciones tecnológicas que cambian la forma en que los motores de bases de datos se ejecutan, con servidores *hardware* especializados y funciones *softwar*e asociadas a los motores de bases de datos que reducen cuellos de botella en el acceso y presentación de los datos a las aplicaciones. Estos cambios son de tal calado que, probablemente, estemos asistiendo a la mayor evolución de los últimos años en la arquitectura de los motores de bases de datos comerciales. Este es el caso del producto Exadata de la compañía Oracle, formado por *hardware* y *software* especializado para ejecutar bases de datos Oracle.

Es relevante, por tanto, el describir las novedades tecnológicas incorporadas en Exadata resumiendo sus características, identificando las mejoras obtenidas en casos reales de clientes y en pruebas de escalabilidad realizadas por terceros y asociando estas mejoras con las novedades tecnológicas incorporadas. Esta descripción se complementa con un resumen de características de otros productos similares, denominados habitualmente *appliances*, y con una comparativa económica con las arquitecturas tradicionales de servidores para la ejecución de bases de datos, comparativa de utilidad para identificar las características de los proyectos de inversión en esta tecnología.

No existen pruebas estándar sobre Exadata, tipo TPC-C o TPC-H, y no es viable para un trabajo de estas características el disponer de un equipo Exadata donde realizar test de rendimiento y escalabilidad dado que este tipo de pruebas están restringidas en Oracle a actividades comerciales de venta del producto. Otras empresas ofrecen con un cargo económico sus servicios para la realización de pruebas sobre Exadata y publican poca información sobre pruebas realizadas para clientes. La información disponible en este aspecto es difícil de encontrar y, si bien la información sobre Exadata puede recopilarse libremente desde la *web*, no es sencillo agrupar y separar la información estrictamente útil, por lo que un lector de esta memoria encontrará una visión global, de fácil lectura para un profesional de informática y suficientemente documentada para un estudiante avanzado de informática, que introduce al mismo al mundo de los *appliances*.

1.2 Objetivos del Trabajo

El objetivo principal de este trabajo es el de ahorrar tiempo de estudio al lector interesado en Exadata o en los *appliances* en general, proporcionando información relevante que permita al lector entender los cambios más importantes en la arquitectura de los servidores y en los motores de bases de datos proporcionados por Exadata y otros *appliances*. Otros objetivos son:

- Permitir identificar al lector dónde es útil Exadata a través de la exposición de las mejoras obtenidas en casos de clientes para la ejecución de bases de datos en entornos OLTP, *Data Warehouse* o mixtos, expuestos públicamente por Oracle y por los clientes de Exadata.

- Resumir los resultados obtenidos en pruebas de escalabilidad presentadas públicamente por terceras empresas.

- Asociar las mejoras descritas en los casos de clientes y en las pruebas de escalabilidad con las novedades tecnológicas incorporadas.

- Presentar al lector un resumen con las características principales de otros *appliances*.

- Permitir entender al lector cómo son los proyectos de inversión en Exadata, extrapolable a otros *appliances*, comparando Exadata con la arquitectura tradicional de servidores de bases de datos en un sencillo ejemplo.

1.3 Enfoque y método seguido

Si se hubiese podido tener acceso para pruebas a un equipo Exadata, probablemente habría sido de utilidad diseñar un test que vinculase las nuevas funcionalidades aportadas con los rendimientos observados. Sin embargo, se hace difícil diseñar un conjunto sencillo de tests que permitan extrapolar conclusiones sobre la utilidad de Exadata, lo que, probablemente, es una de las razones por las que la compañía Oracle no publica test estándar sobre Exadata, y sí se presentan públicamente comparativas en casos reales de sus clientes, lo que parece más eficiente desde el punto de vista comercial. Por otra parte, Exadata es relativamente nuevo por lo que tampoco existen comparativas entre *appliances* que incluyan a Exadata lo suficientemente documentadas para, empíricamente, permitir extraer conclusiones.

La estrategia adoptada, dada la dificultad para encontrar información útil, ha sido, en primer lugar, la de recopilar información publicada en

Internet, por Oracle, por otros proveedores, por terceros y por clientes de Exadata. En segundo lugar se ha procedido a clasificar la información adecuadamente y estudiarla en detalle, concretando brevemente los aspectos más relevantes que permitían cumplir con el objetivo principal de ahorrar tiempo a un posible lector. Por último, se ha procedido a su exposición a través de esta memoria. Creo que la visión global aportada permite acercar Exadata, y los *appliances* en general, al lector de esta memoria, de una forma práctica y más completa que la exposición de los resultados de test o pruebas de concepto que hubiese podido diseñar y ejecutar.

1.4 Planificación del proyecto

Para la realización de este proyecto sólo se ha requerido mi portátil, con acceso a Internet. Las tareas ejecutadas han sido las siguientes:

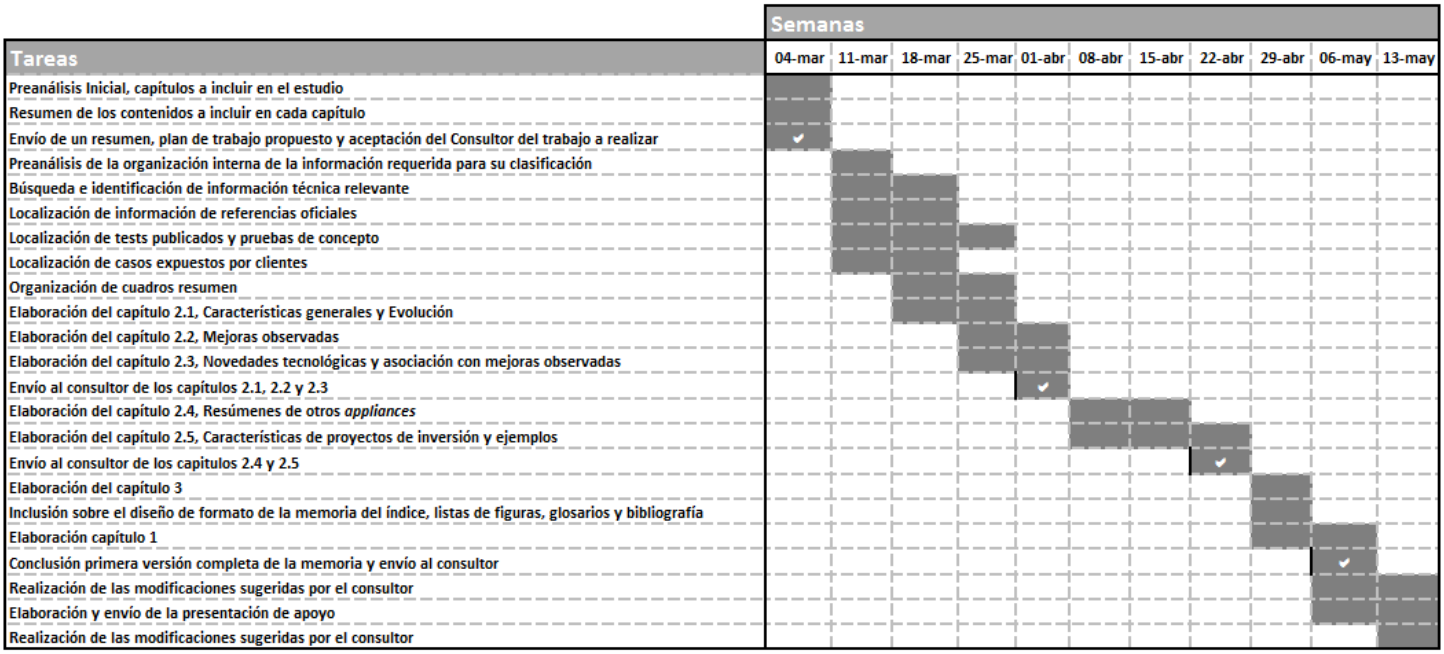

Fig 0. Listado de Tareas en la realización del TFG

1.5 Breve descripción de los capítulos de la memoria

La memoria está distribuida en tres capítulos principales, el capítulo 1 de introducción, el capítulo 2 con el cuerpo principal compuesto de 5 subcapítulos y el capítulo 3 de conclusiones. En este apartado se hace una breve descripción del capítulo 2 y de sus subcapítulos, del 2.1 al 2.5.

El capítulo 2.1 detalla las configuraciones de Exadata y describe las diferentes familias existentes, focalizándose en el conjunto de recursos *hardware* y *software* que lo componen. Existe información abundante en

Internet con descripción técnica sobre Exadata pero este punto resume la información relevante que será de utilidad para el resto de capítulos y ahorra al lector el tiempo requerido para ello. Adicionalmente, en este capítulo se describen también las configuraciones de equipos Exadata de anteriores generaciones, dado que las comparativas y pruebas de escalabilidad descritas en el resto de capítulos se han realizado sobre varias generaciones de Exadata y es útil disponer de esta información.

El capítulo 2.2 presenta de manera sumarizada y resumida los resultados obtenidos en mejoras de tiempos de ejecución de consultas, cargas de datos, y otras operaciones con Exadata, utilizando para ello las referencias oficiales publicadas por Oracle y profundizando en tres casos de clientes que han obtenido mejoras en procesos OLTP, *Data Warehouse* y entornos mixtos. Existen más de 300 ejemplos disponibles en Internet pero los casos seleccionados muestran un resumen que puede aplicarse a todos y ahorran tiempo al lector. Este capítulo incluye también pruebas de escalabilidad relevantes sobre Exadata realizada por terceros. Adicionalmente, se presentan las características que debe cumplir un *benchmark* para testear eficazmente esta plataforma.

El capítulo 2.3 presenta las novedades tecnológicas de Exadata y las relaciona con las mejoras observadas y descritas, tanto en las referencias oficiales como en otras pruebas de escalabilidad. Dado que las novedades implican un cambio importante en cuanto a la forma en la que se ejecuta el motor de la base de datos Oracle, se describen las novedades más relevantes para el lector y se enlazan con pruebas de escalabilidad realizadas y presentadas públicamente por terceros. Concluye este punto con una breve descripción de consideraciones críticas a Exadata.

El capítulo 2.4 resume otros *appliances* para la ejecución de bases de datos, como Teradata, Netezza/PureData de IBM, Greenplum de EMC y el Data Accelerator de Fusion-io. El objetivo es que el lector pueda disponer en una lectura sencilla de una visión de los componentes y arquitectura de estos *appliances*. Concluye este capítulo con unas consideraciones sobre la comparativa entre *appliances* basada en términos económicos y un apartado de consideraciones ante la disyuntiva de construir un *appliance* basado en componentes estándar.

El capítulo 2.5 compara, desde el punto de vista económico, a Exadata con la arquitectura tradicional para servidores de bases de datos, tanto para proyectos de consolidación donde se muestra un sencillo ejemplo de una empresa ficticia que tiene dos alternativas, seguir como está o cambiar a Exadata, como para otro tipo de proyectos donde la funcionalidad aportada es un plus extra. Concluye este capítulo con un resumen de los beneficios de Exadata.

2. Capítulos Principales

2.1 Breve descripción de Exadata a través de sus configuraciones, evolución del producto y características generales

2.1.1 Introducción

La empresa Oracle define Exadata como una máquina de base de datos, formada por componentes *hardware* y *software* preinstalados para la ejecución de bases de datos Oracle[1]. Este empaquetamiento de *hardware* y *software* se distribuye en diferentes configuraciones cuyos elementos comunes a todas ellas son la existencia de servidores *hardware*, discos, memoria *flash*, tarjetas y *switches* de comunicación, *software* de sistema operativo, *software* específico de Exadata y *software* de base de datos Oracle 11g. Actualmente cada configuración es distribuida con un número concreto de estos elementos y dos familias de producto, la X3-2 y la X3-8. En el apartado 2.1.2 se realiza una breve descripción de las configuraciones actuales.

Los casos de clientes de Exadata[2] y las comparativas de tipo test con otros productos, analizadas en el apartado 2.2 de esta memoria, han sido realizadas sobre configuraciones actuales X3 y sobre configuraciones en versiones anteriores del producto. Es de utilidad, por tanto, incluir una descripción de configuraciones anteriores a la X3. En el apartado 2.1.3 se describe brevemente la historia de Exadata y las diferentes configuraciones por las que Exadata ha ido evolucionando.

Esta memoria está centrada en las características técnicas de Exadata. Con el objetivo de proporcionar una visión más amplia de Exadata se incluye el apartado 2.1.4 con una enumeración de observaciones adicionales.

2.1.2 Configuraciones de Exadata

Oracle distribuye configuraciones de 1/8, 1/4,1/2, un Exadata completo o varios Exadatas interconectados entre sí en la familia X3-2 y configuraciones de un Exadata completo o varios Exadatas interconectados entre sí en la familia X3-8. La diferencia más importante entre las familias X3-2 y X3-8 es el tipo de servidores *hardware* para las bases de datos. Hay que distinguir aquí que los servidores *hardware* de cualquier Exadata pueden ser servidores de base de datos o servidores de almacenamiento, una característica de Exadata en la que profundizaremos en el apartado 2.3. El número 2 y el número 8 de las referencia X3-**2** y X3-**8** está relacionado con el número de procesadores de cada servidor de base de datos y no está relacionado con el número

de servidores de base de datos de cada modelo, hecho que puede causar confusión y que el siguiente cuadro diferencia:

X3-2 Completo	X3-8 Completo
Ocho Servidores de base de datos, con 2 Procesadores de 8 cores, cada uno, de tipo Intel® Xeon® E5-2690 (2,9 GHz) **Total 128 Cores**	Dos servidores de base de datos, con 8 Procesadores Intel® Xeon® E7-8870(2.4GHz) cada uno, 10 Cores cada procesador **Total 160 Cores**

Fig 1. Diferencia servidores de base de datos entre el modelo X3-2 completo y el modelo X3-8 completo

La diferencia más importante entre un Exadata X3-2 completo y un X3-8 completo es el número de servidores de base de datos de su configuración. El modelo X3-2 completo es un conjunto *elevado* de servidores con pocas CPUs en cada servidor, es un *cluster de servidores* de base de datos[3], mientras que el modelo X3-8 son 2 servidores con muchas CPUs cada servidor, son dos *SMP*[4] *en cluster.* Esta diferencia es importante dado que el tipo o tipos de aplicación a ejecutar, el número de accesos simultáneos requeridos, el número de bases de datos accedidas o las necesidades de rendimiento esperadas pueden hacer que una familia ofrezca mejores prestaciones que la otra.

Como veremos en el capítulo 2.3 los servidores de almacenamiento son una de las novedades introducidas en Exadata para la obtención de mejoras en el rendimiento en el uso de las bases de datos. Están formados principalmente por *hardware*, memoria *flash* y un *software* especializado denominado Exadata Storage Software que, en su conjunto, permiten reducir los tiempos de acceso a disco desde los servidores de base de datos, comparados con arquitecturas para servidores de bases de datos tradicionales. Exadata incorpora funcionalidades[5] que aceleran el proceso de ejecución del SQL y el rendimiento global del sistema, mejorando la entrada/salida a disco.

La diferencia más importante entre los servidores de una misma familia es el número de servidores que contiene, tanto de base de datos como de almacenamiento, tal y como se refleja en el siguiente cuadro resumen de la familia X3-2:

	1/8	1/4	1/2	Completo
Servidores de Base de Datos (Cada Servidor se compone de 2 Procesadores de 8 cores cada uno Intel® Xeon® E5-2690 (2,9 GHz) y 256GB Memoria)	2	2	4	8
Cores totales servidores de Base de datos	16 Cores	32 Cores	64 Cores	128 Cores
Memoria RAM Total servidores de Base de datos	512 GB	512 GB	1 TB	2 TB
Servidores de Almacenamiento (Cada servidor se compone de 2 Procesadores de 6 cores cada uno Intel® Xeon® E5-2630L y 4 tarjetas PCI Flash F40 eMLC de 0,4 TB cada una)	3	3	7	14
Cores totales servidores de Almacenamiento	18 Cores	36 Cores	84 Cores	168 Cores
Tarjetas PCI Flash Totales	6 tarjetas	12 tarjetas	28 tarjetas	56 tarjetas
TBs Flash totales	2,4 TB	4,8 TB	11,2 TB	22,4 TB
Discos de 600 GB cada uno a 15.000 rpm o discos de 3 TB cada uno a 7.500 rpm (o una configuración o la otra)	18 discos	36 discos	84 discos	168 discos

Fig 2. Familia X3-2. Características principales de los servidores de base de datos y almacenamiento

Para cada familia de Exadata existen diferentes configuraciones de fábrica para la conectividad entre los servidores de bases de datos, los servidores de almacenamiento y los discos. Cada configuración está cableada con *switches QDRs Infiniband*[6] de 40GB/segundo en cada sentido y tarjetería *Ethernet*. La conectividad Infiniband permite enlaces punto a punto entre los diferentes servidores y los discos a alta velocidad. Este detalle es importante, para las comparativas, dado que la velocidad y latencias proporcionadas por Infiniband se han de tener en cuenta en los resultados de los test. Además, las configuraciones de más de un Exadata, conectados entre sí, se realiza a través de Infiniband, por lo que la diferencia entre las diferentes configuraciones, incluyendo múltiples X3-2 o X3-8 interconectados, reside en el número de enlaces Infiniband existentes, lo que implica un número diferente de *switches* y tarjetería por configuración.

Otros componentes de los equipos Exadata pueden presentar diferencias para el servicio de base de datos, tanto en arquitecturas en cluster[7] como en SMP, con sistemas *enracados*[8] de múltiples servidores conectados entre sí o con grandes servidores *multicore.* El espacio ocupado por el equipo, el número de unidades de enracado, el consumo de potencia y otras características, pueden determinar diferencias de costes de mantenimiento anual. Estas diferencias, si bien no han sido consideradas en los casos de clientes, sí que pueden ser significativas de cara al análisis comparativo de los proyectos de inversión. En el capítulo 2.5 se contemplan los elementos a considerar desde el punto de vista de su impacto económico. El siguiente cuadro muestra una imagen de un equipo Exadata y algunas de sus especificaciones técnicas:

	X3-2 Completo
Altura	1.998 mm
Anchura	600 mm
Profundidad	1.200 mm
Peso	871,4 Kg
Consumos con configuración de discos de 15.000 rpm	
Consumo a Potencia máxima	11,9 KW
Cooling a máximo uso	40.600 BTU/hora
Consumos con configuración de discos de 7.500 rpm	
Consumo a Potencia máxima	10,9KW
Cooling a máximo uso	37.200 BTU/hora

Fig 3. Especificaciones técnicas equipo Exadata X3-2 Completo

Los servidores de bases de datos X3 pueden ejecutarse con 3 tipos diferentes de sistema operativo, seleccionables en el momento de la instalación:

- Oracle Linux 5 Update 8 con el Kernel *Unbreakable Enterprise*
- Compatible Red Hat
- Solaris 11

Exadata incorpora un conjunto de funcionalidades para proporcionar alta disponibilidad y seguridad, con protección para fallos de *hardware* y de *software*, mediante redundancia de elementos, como por ejemplo los suministros de potencia, los *switches* de comunicaciones, la protección ante problemas de *hardware*, protección ante roturas de disco, clusterización global y elementos para la configuración de Exadata para la protección ante desastres. Asimismo, Exadata incluye la administración centralizada de todos los componentes. Estas características avanzadas han de ser tenidas en cuenta en las comparaciones con otras arquitecturas y en los análisis de costes / beneficios de los proyectos de inversión.

Una de las características de Exadata más destacadas es el tratamiento de todo el *software* en su conjunto para la resolución de parches, hecho que simplifica las actividades de resolución de problemas, comparado con otras arquitecturas, y que se tendrá en cuenta en el capítulo 2.5 en la valoración económica.

2.1.3 Evolución de Exadata

Las familias X3-2 y X3-8 son la cuarta generación de Exadata. La primera generación, denominada V1, apareció en el año 2008 y estaba formada por componentes de Oracle y de la compañía Hewlett-Packard, en una única familia. Con la adquisicion de la compañía SUN Microsystems[9] Oracle cambió su estrategia y sustituyó, en 2009, los componentes de Hewlett-Packard por componentes de SUN en la segunda generación de Exadata, denominada V2, también en una única familia. En la tercera generación del año 2010 Oracle introdujo las familia X2-2, que sustituía a la familia V2, e introdujo una nueva familia denominada X2-8. Las diferencias entre las familias X2-2 y X2-8 eran similares a las expuestas en el apartado 2.1.2 entre las familias X3-2 y X3-8. Estas últimas familias forman la cuarta generación de Exadata y fueron introducidas en el año 2012.

La evolución de las diferentes generaciones de Exadata se ha producido a través de la incorporación de mayor capacidad de proceso global, bien sea por la introducción de diferentes tipos de procesador, mayor número de *cores* o más memoria *flash* disponible, entre otras. El siguiente cuadro resume las diferencias más importantes en las configuraciones completas de las generaciones V2, X2-2 y X3-2:

	V2 Completo	X2-2 Completo	X3-2 Completo
Servidores de Base de Datos y características principales de cada servidor	8 Servidores de base de datos Cada servidor se compone de 2 procesadores con 4 cores cada uno Intel® Xeon® X5540 (2.53 GHz) y 96GB Memoria)	8 Servidores de base de datos Cada Servidor se compone de 2 Procesadores de 6 cores cada uno Intel® Xeon® X5675 (3.06 GHz) y 96GB Memoria)	8 Servidores de base de datos Cada Servidor se compone de 2 Procesadores de 8 cores cada uno Intel® Xeon® E5-2690 (2,9 GHz) y 256GB Memoria)
Cores totales servidores de Base de datos	64 Cores	96 Cores	128 Cores
Memoria RAM Total servidores de Base de	576GB GB	768 GB	2 TB
Servidores de Almacenamiento y características principales de cada servidor	14 Servidores de Almacenamiento Cada servidor se compone de 2 procesadores de 4 cores Intel® Xeon® L5540 (2.53 GHz)	14 Servidores de Almacenamiento Cada servidor se compone de 2 procesadores de 6 cores Intel® Xeon® L5640 (2.26 GHz)	14 Servidores de Almacenamiento Cada servidor se compone de 2 Procesadores de 6 cores cada uno Intel® Xeon® E5-2630L y 4 tarjetas PCI Flash F40 eMLC de 0,4 TB cada una
Cores totales servidores de Almacenamiento	112 Cores	168 Cores	168 Cores
TBs Flash totales	5,3 TB	5,3 TB	22,4 TB
Almacenamiento Total	168 discos ->Discos de 600 GB cada uno a 15.000 rpm SAS o discos de 2 TB cada uno a 7.200 rpm SATA (o una configuración o la otra)	168 discos ->Discos de 600 GB cada uno a 15.000 rpm SAS o discos de 2 TB cada uno a 7.200 rpm SAS (o una configuración o la otra)	168 discos ->Discos de 600 GB cada uno a 15.000 rpm SAS o discos de 3 TB cada uno a 7.500 rpm SAS (o una configuración o la otra)

Fig 4. Generaciones V2, X2 y X3. Evolución de las características principales

Desde su concepción Exadata ha sido considerada una máquina de base de datos de propósito general, para la ejecución de una o múltiples bases de datos en modo *OLTP*, *Data Warehouse* o sistemas mixtos, con unos rendimientos y capacidades superiores a las arquitecturas tradicionales[10]. De una forma breve, se puede definir la evolución de las diferentes generaciones como la posibilidad de *cada vez más* ejecutar aplicaciones más complejas y con mayores requerimientos de tipo

OLTP, *Data Warehouse* o mixtas. Los servidores de almacenamiento han existido desde la primera generación, la segunda generación facilitó la ejecución de cargas mixtas al proporcionar velocidades de acceso superiores gracias al uso de memoria *flash*, la tercera generación incorporó escalabilidad adicional con la introducción de la familia X2-8 y mayores capacidades *OLTP* y *Data Warehouse* y la cuarta generación dispone de más memoria *flash* y más *cores* en los servidores de base de datos. El siguiente esquema muestra la evolución de Exadata desde el punto de vista de las capacidades aportadas por cada generación:

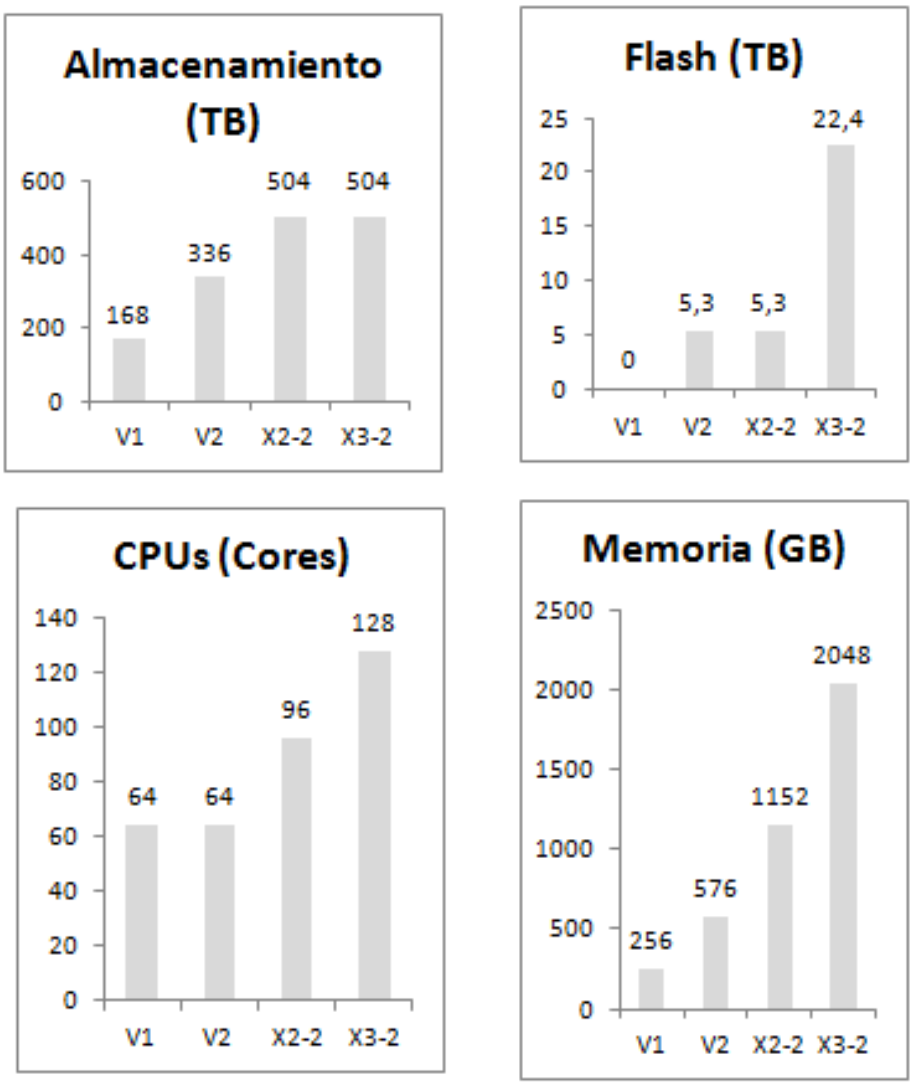

Fig 5. Almacenamiento, RAM, Memoria Flash y CPUs de las generaciones V1, V2, X2-2 y X3-2, equipos Completos

2.1.4 Características generales

Exadata es una máquina de base de datos Oracle de propósito general, por lo que cualquier tipo de aplicación que accede a una base de datos

Oracle puede utilizar Exadata para la capa de base de datos. Sin embargo, no todas las bases de datos Oracle pueden ejecutarse en Exadata dado que se requiere una versión específica, la 11gR2.

Exadata puede ser contemplado como un entorno donde el cliente puede consolidar todas o casi todas sus bases de datos. La compañía Oracle considera que Exadata permite reducir el número de servidores requeridos para la capa de base de datos, eliminando costes de sistemas de almacenamiento, costes asociados a la gestión y al mantenimiento. Por otra parte, Exadata está certificado para la ejecución de bases de datos Oracle de aplicaciones estándar del mercado, como por ejemplo SAP[11].

¿Dónde es imprescindible Exadata?

Actualmente no existe una casuística de tipo de cliente o requerimiento de servicio para la capa de base de datos que sólo pueda resolverse con Exadata, dado que no existe una funcionalidad adicional aportada por el producto de Oracle que no pueda ser satisfecha con otras arquitecturas. Exadata ejecuta bases de datos Oracle 11g de una forma eficaz y con un alto rendimiento, lo que puede ser indicado para cualquier tipo de cliente de base de datos Oracle con un volumen transaccional medio-alto o unos requerimientos de mejora de tiempo de acceso en consultas a la base de datos. Sin embargo, algunos clientes de la base de datos que tengan problemas con operaciones de tipo *batch* o con la ejecución de las operaciones de *Backup* y *Recovery* pueden encontrar en Exadata una plataforma imprescindible para conseguir reducir las ventanas de ejecución de estas operaciones a unos tiempos que permitan el servicio ininterrumpido. Otras casuísticas de requerimientos sólo satisfechos con Exadata pueden presentarse en cualquier tipo de cliente.

¿Cambia el rol del administrador de bases de datos (DBA)?

Los *appliances* como Exadata incorporan en un *box* servidores, *switches* y tarjetas de red, discos, tarjetas *flash*, *software* de sistema operativo, *software* específico y *software* de base de datos. Adicionalmente se incorporan herramientas de administración del conjunto para la monitorización y gestión y las funciones de gestión del cambio, incluyendo el parcheado, se simplifican, al proporcionarse soporte desde un único fabricante. Quizás lo más relevante en cuanto a la administración de los *appliances* es que determinadas funciones de gestión de activos, como la gestión del *storage*, pueden ser realizadas por los DBAs. Un cliente de bases de datos Oracle que opte por incluir todas sus bases de datos en un Exadata puede encontrarse con el desafío de la posible consolidación de roles de administración.

Fin del capítulo 2.1

Notas y Bibliografía de 2.1

1 Folletos de producto públicos de Exadata. ***DataSheets families X3-2 & X3-8,*** **http://www.oracle.com/us/products/database/exadata/overview/index.html [Consulta Web], Fecha de Consulta: Marzo 2013.**

La información de las características técnicas de Exadata ha sido compilada, principalmente, desde los *DataSheets* publicados por Oracle y de la información proporcionada públicamente por Oracle en su red Oracle Technology Network, accesible desde la url mostrada.

2 *Búsqueda de Casos de Clientes Exadata en la web pública de la compañía Oracle,* http://www.oracle.com/search/customers [*Consulta Web], Fecha de Consulta: Marzo de 2013.*

3 Cluster de pequeños servidores de bases de datos en una arquitectura MIMD, según la taxonomía de Flynn, de multicomputadores en cluster con memoria distribuida, débilmente acoplados.

4 SMP arquitectura MIMD, según la taxonomía de Flynn, de multiprocesadores con memora compartida, fuertemente acoplados, UMA o NUMA. Cuando en una arquitectura de servidores de bases de datos se conectan varios de estos servidores SMP, también proporcionan una memoria distribuida débilmente acoplada y gestionada por el software de cluster de la base de datos, convirtiéndose en Multicomputadores en cluster con memoria distribuida.

5 El concepto de Exadata Storage Software es intrínseco a Exadata e incorpora un conjunto de nuevas funcionalidades para la ejecución interna de los procesos de acceso a la información de la base de datos Oracle.

6 Quad Data Rate, *Infiniband,* http://en.wikipedia.org/wiki/InfiniBand [Consulta Web] Fecha de Consulta: Marzo de 2013.

7 Multicomputadores con memoria distribuida

8 Armarios para montaje de varios servidores

9 *Sun Microsystems,* http://es.wikipedia.org/wiki/Sun_Microsystems [Consulta Web] Fecha de Consulta: Marzo 2013.

10 Exclusivamente para la ejecución de bases de datos Oracle 11g

11 Chris Kanaracus, IDG (2012) *Oracle Exadata gains certification for SAP applications* http://www.pcworld.com/article/230175/article.html *[Consulta Web] Fecha de Consulta: Marzo 2013*

2.2 Casos de Clientes Exadata y pruebas de escalabilidad

2.2.1 Introducción

Oracle define Exadata como el producto de la compañía con mayor crecimiento en número de clientes de la historia[1] y enumera, en un conjunto de referencias públicas[2], las mejoras obtenidas por clientes de Exadata de todo el mundo, de cualquier industria y tamaño, para aplicaciones *OLTP*, *Data Warehouse* o mixtas. Estas referencias describen las mejoras obtenidas con altas variaciones de resultados y no describen exhaustivamente los entornos de partida, por lo que se hace difícil establecer, empíricamente, reglas de asociación entre las mejoras y la tecnología específica que las facilita. Sin embargo, una enumeración de las mejoras puede ser útil para entender el porqué del éxito del producto entre sus clientes. El apartado 2.2.2 resume y sumariza las mejoras descritas por Exadata en sus referencias oficiales.

Otras referencias[3], también proporcionadas por Oracle públicamente, realizan un análisis en mayor detalle de la situación de partida de los clientes de Exadata y de las mejoras obtenidas. El apartado 2.2.3 resume 3 casos de clientes de Exadata: El caso de la Agencia de protección de fronteras de USA, como ejemplo de entorno mixto *OLTP-Data Warehouse*, el caso de la empresa de telecomunicaciones Turkcell, como ejemplo de *Data Warehouse* puro y el caso de la empresa de equipos de navegación por GPS Garmin, como ejemplo de entornos *OLTP*.

En el apartado 2.2.4 se describen 2 resultados de tests obtenidos en una prueba de rendimiento y escalabilidad con Exadata gracias al uso de los servidores de almacenamiento y las cachés ampliadas con las tarjetas *flash*. Estas pruebas han sido realizadas por la empresa SAS y expuestas públicamente en su *web*.

2.2.2 Resumen de resultados oficiales publicados de casos de clientes Exadata

La compañía Oracle no ha publicado resultados de pruebas test de tipo estándar[4] sobre plataformas Exadata, como por ejemplo resultados TPC-C para pruebas OLTP y TPC-H para pruebas de entornos *Data Warehouse*. Por otra parte, realizar una prueba de valoración de Exadata no es una labor sencilla para un investigador dado que se requiere acceso a una plataforma Exadata y este acceso es proporcionado por Oracle para actividades comerciales. Además, si bien es cierto que es relativamente sencillo encontrar en Internet empresas de servicios que disponen de una plataforma Exadata y que ofrecen su plataforma para la realización de pruebas de valoración del producto,

estas pruebas se realizan a través de la contratación de los servicios ofrecidos[5].

Los resultados obtenidos con Exadata y publicados en Internet por Oracle permiten identificar mejoras en rendimientos de uso y acceso a las bases de datos con Exadata comparando rendimientos actuales contra rendimientos anteriores, pero estas mejoras pueden haberse obtenido no sólo por las nuevas tecnologías aportadas por Exadata, como los servidores de almacenamiento, en *hardware*, memoria *flash* y, el *software* de Exadata Storage, sino que pueden existir otros elementos que propician las mejoras entre los entornos previos del cliente y la nueva plataforma Exadata:

- Diferentes versiones de Base datos Oracle. Exadata sólo ejecuta versiones 11gR2 de la base de datos Oracle y, mientras en las referencias no se especifica la versión anterior de la base de datos, estas pueden ser versiones anteriores del producto. La versión 11gR2 puede ofrecer mejoras de rendimiento sobre versiones anteriores e influir, en parte, en las mejoras observadas.
- Procesadores no comparables. Los entornos previos de base de datos de los clientes pueden ser antiguos, con diferente rendimiento por *core* y diferente número de *cores* comparado con los entornos Exadata finalmente adquiridos.
- Diferente *sizing* de la plataforma destino, para mejorar situaciones de saturación de entornos con una nueva plataforma de cada cliente. Es posible que los entornos previos de los clientes presentasen grados de saturación de uso que hacían recomendable una plataforma mayor, con mayor número de servidores.

Aún así, una inversión en equipos Exadata y una migración hacia esta plataforma de las bases de datos tiene un coste que ha de estar justificado y las referencias descritas son casos reales de clientes, con unos requerimientos de negocio concretos, que han optado por la plataforma Exadata, por lo que la información de las mejoras, descrita en las referencias, pueden haber justificado la inversión.

Las mejoras descritas en las referencias oficiales muestran rendimientos superiores en Exadata comparados con las plataformas previas de los clientes, tanto en tiempos de acceso a datos vía SQL simples y múltiples, como en tiempos de ejecución de SQLs complejos, en carga de datos, tiempos de *backup* y *recovery* y capacidad de compresión de información. Estas mejoras en tiempos han facilitado a los clientes de Exadata procesar o realizar actividades que antes no hacían y es, probablemente, la razón más importante para su adopción. El siguiente cuadro resume los beneficios expuestos de Exadata en 10 clientes, con diferentes modelos X2-2 y X2-8:

Cliente	Sector	Queries más rápidas	Informes más rapidos	Carga de datos	Backup	Recovery	Compresión	Otros
Algar Telecom	Comunicaciones	85% mejor	Algunos informes de 2 horas a segundos	De 12 horas a 3,5 horas 2000 cargas diarias	De 26 horas a 6 horas, incluyendo copia a cinta	De 10 días a 5 horas	Compactación de 1000 tablas, 8TB, 50% espacio reducido	
Asiana Airlines	Aerolíneas	10 veces más rápido, en algunos casos hasta 338 veces más rápido	Informes de booking de 10 horas a 10 minutos	Transformaciones E-LT reducidas en 5 horas				
Autoglass	Automóviles	Algunas consultas hasta 60 veces más rápidas	De minutos a segundos en tiempos de respuesta de las aplicaciones de negocio		DB Backup de 23 horas a 5 horas			Procesos de negocio hasta 60 veces más rápidos
BNP Paribas	Banca	Media de mejoras globales de 30 segundos a 1 minuto por consulta	Informes 17 veces más rápidos	Cargas 6 veces más rápidas			De 40TB a 8TB con HCC	
Digicel Haiti	Comunicaciones	Mejora del rendimiento global en un 55%	Informes para la dirección de 10 horas a 4 horas			De 15 horas a 6 horas		Consolidación de 300 servicios
Finansbank A.S.	Banca		600.000 informes por mes con una media de ejecución de 31 segundos a 17	Refresh del dwh de 341 a 250 minutos			De 18 TB a 3,5 TB	
Garantibank	Banca	Queries 10 veces más rápido gracias a smart scan y flash	Informes de 8 horas a pocos minutos	8.000 millones de inserciones diarias, desde el mainframe al dwh en 2,5 TB a través de 1.300 transferencias de ficheros	backups de 7 tb en 3 horas		De 17 TB a 5TB	Queries intensivas que antes no se podían ejecutar
Procter & Gamble	CPG	Algunas queries hasta 30 veces más rápidas Un query que tardaba 20 horas ha pasado a ejecutarse en 3 minutos						Expansión: Más usuarios de las capacidades analíticas
SK Telecom	Comunicaciones	x10		De 2 a 3 veces más rápido			x10	Implementación de procesos antes inviables
TargetBase	Servicios	Hasta 30 veces más rápido en queries de uso común	Mejoras de hasta 9 veces más rápido				30% reducción de costes de almacenamiento	

Fig 6. Beneficios Exadata publicados por Oracle[6], subconjunto seleccionado. X2-2 y X2-8

Otros casos con resultados similares, también presentados oficialmente por Oracle, muestran la misma situación, con consultas simples o complejas más rápidas e informes más rápidos, por lo que vuelven a repetir los beneficios expuestos y no aportan información adicional para esta memoria. Aunque una búsqueda en la *web* de Oracle de casos de clientes con Exadata ofrece 359 informes[7], el subconjunto seleccionado refleja una situación similar para todos los casos.

Es difícil encontrar ejemplos en Internet donde se puedan contrastar resultados de test sobre la plataforma Exadata comparadas con otras plataformas, con *benchmarks* realizados por terceros. Además, la mayoría de los casos encontrados son de equipos X2 o V2. Exadata es relativamente nuevo y puede que en un futuro próximo, así como aumente el número de clientes, aparezcan más casos publicados y más test de evaluación realizados por terceros.

2.2.3 Casos Exadata con mayor detalle

Customs & Border Protections, USA. Caso Mixto *OLTP-Data Warehouse*[8]

Algunos clientes de Exadata presentan públicamente informes, en las *webs* de sus empresas u organizaciones, donde se resaltan beneficios obtenidos con su implementación. Un caso a resaltar es el de la Agencia de fronteras de USA o CBP (Customs & Borders Protection) del Departament of Homeland Security (DHS), donde el Sr. Kenneth M. Ritchhart, Deputy Assistant Commisioner and DCIO, a principios de 2011 resaltaba las razones por las que se adquirieron 12 Exadatas completos y los beneficios obtenidos en rendimiento y reducción de costes después de su implementación. El siguiente cuadro muestra un resumen del

volumen operacional que, a diario y según el Sr Richhart, debe gestionar la informática de la Agencia:

Pasajeros diarios	1,1 millones
Pasajeros de vuelos internacionales a diario	256.897
Pasajeros de Cruceros a diario	43.188
Camiones, trenes y contenedores a diario	70.900
Coches privados a diario	331.347
Envíos diarios aprobados de bienes	85.300
Arrestos diarios	107
Droga comisionada a diario (en libras USA)	17.420
Dinero no declarado decomisado en dólares	300.582
Vehículos	21.863
Aviones	290
Barcos	225
Rescates diarios	488
Ubicaciones o centros	16.000

Fig 7. Resumen de las operaciones diarias de la Agencia CBP, DHS, USA

El siguiente cuadro resume la descripción de algunos de los activos informáticos del CBP, expuesta por el Sr. Richhart en el mismo documento:

Funcionamiento	24x7x365
Estaciones de trabajo	65.000
Petabytes de información SAN/NAS	6,5+
Petabytes de información totales	45
Servidores	5.400
Empleados IT propios y subcontratados	5.399
Objetivo IT para 2011	**Reducir 214M$ de costes**

Fig 8. Resumen de algunos activos IT de la Agencia CBP, USA

La Agencia CBP realizó un proceso de sustitución de algunos de los servidores de base de datos Oracle por 12 Exadatas, siendo el concepto de *appliance*[9] uno de los objetivos de la Agencia para la simplificación de los activos en servidores, discos y *software*. El menor coste y el mayor rendimiento de Exadatas también formaron parte de las razones para la adopción. El siguiente cuadro resume los beneficios obtenidos, expuestos por el Sr. Richhart, una vez que los sistemas entraron en producción:

Sustitución de grandes equipos SMP por Exadatas completos V2, X2-2 y X2-8	12 Exadatas completos
Coste reducido	1/4 coste de los grandes SMP
Velocidad obtenida en consultas online	x10
Velocidad obtenidas en consultas batch	x77
Reducción de espacio y energía, en Armarios	De 120 Racks a 12 Racks
Reducción de espacio en disco SAN/NAS	-1 Petabyte

Fig 9. Resumen de los beneficios obtenidos con los Exadatas de la Agencia CBP, DHS, USA

Turkcell[10], caso Data Warehouse

Otro caso descrito en las referencias oficiales de Oracle de mejoras obtenidas con Exadata es de la compañía de telecomunicaciones turca Turkcell. Con más de 30 millones de clientes en Turquía que, a 31 de Marzo de 2011 representaba el 54% del mercado turco, y con más de 60 millones de clientes en total, incluyendo todos los países donde Turkcell opera, es la 3ra compañía de telecomunicaciones europea por número de clientes. En un día la red de Turkcell procesa 1.500 millones de registros de llamadas (*Call Detail Records* o CDRs), que forman la base que luego ha de ser procesada por las áreas de negocio de Turkcell, y que genera un volumen acumulado de 600 TB de información que se duplica cada año. Turkcell, previo a la implantación del Exadata, extraía un subconjunto de la información generada en un *Data Warehouse* resumido en las siguientes características:

Tamaño del Data Warehouse extraído	250 TB
Informes Mensuales	50000
Cabinas de Disco, en RACKs	10
Servidor (no se especifican características)	M9000
Total RACKS	11
Tiempo de ejecución informe típico	27 minutos
Informes con más de 4 horas de tiempo de ejecución	87
Usuarios analistas ejecutando informes	300
Velocidad de transferencia al Data Warehouse, en GB/Segundo	6 GB/segundo
Duración del backup	44 horas
Número de cintas de backup por cada backup	159
Metros cuadrados de espacio ocupados	33
Kwh Consumidos	50

Fig 10. Resumen de las características del Data Warehouse de Turkcell previo a la implementación con Exadata

Turkcell inició en 2009 un proceso de implementación de su *Data Warehouse* en un Exadata X2-2 completo con discos de 15.000 rpm que concluyó en 4 meses y que permitió las siguientes mejoras:

Tamaño después de la aplicación de la compresión (Advanced Compression + Hybrid Columnar Compression)	30 TB
RACKs Finales (Full Exadata X2-2, discos de 15.000 rpm)	1
Tiempo de ejecución informe típico	3 minutos
Informes con más de 4 horas de tiempo de ejecución	1
Velocidad de transferencia al Data Warehouse, en GB/Segundo	21 GB/segundo
Duración del backup	14 horas
Número de cintas de backup por cada backup	57
Metros cuadrados de espacio ocupados	3
Kwh Consumidos	10

Fig 11. Resumen de las mejoras obtenidas con el Exadata X2-2 completo con discos de 15.000 rpm

Garmin[11], caso OLTP

Garmin fabrica y comercializa equipos de navegación GPS para automóviles, aviación, marina y equipos de deporte, siendo un líder mundial en su sector. Con centros de distribución en USA, el Reino Unido, Australia y Taiwan y fábricas en USA y Taiwan dispone de 35 oficinas propias y 9.200 asociados. Garmin es cliente de Oracle para las aplicaciones críticas de negocio, como la gestión de pedidos, la gestión de la fabricación, la gestión de la cadena de suministro, la planificación del suministro de equipos navegadores y la gestión de la información *online* de clientes con dispositivos de control *fitness*. Estas aplicaciones se ejecutan sobre bases de datos Oracle y tienen unos requerimientos de servicio de tipo *OLTP*, tanto en rendimiento como en concurrencia y disponibilidad del servicio. El siguiente cuadro muestra una breve descripción de los requerimientos de Garmin previo a su migración a Exadata.

Servidores dedicados para las bases de datos tipo SMP (no se especifican características)	12 servidores, entre M9000, M5000 e Intel Linux
Tipología de Discos	SAN
Requerimientos de mayor capacidad de proceso habiéndose llegado al límite en CPUs e IOPS	Altos requerimientos anuales , 50% de incremento anual en requerimientos de espacio
Tiempo medio de proceso de preenvío de equipos	35 minutos aprox
Tiempo medio de proceso de envío de equipos	35 minutos aprox
Crecimiento en visitas por hora a las páginas de fitness	De 200.000 por hora a 750.000 por hora

Fig 12. Resumen de requerimientos y características de servicio de las aplicaciones de Garmin previo a su migración a Exadata

Garmin inició un proceso de migración a 2 ½ Exadata V2, de las bases de datos que dan soporte a la operativa descrita. Previo a la migración Garmin no disponía de una arquitectura de protección ante desastres y la migración a los equipos Exadata les permitió disponer de dos entornos separados. Los resultados después de la implementación de las bases de datos Oracle en los equipos Exadata están resumidas en el siguiente cuadro:

Servidores dedicados para las bases de datos	2 1/2 Exadata V2
Picos de CPU, máximo porcentaje	50%
Porcentaje de peticiones resueltas en memoria flash	93% peticiones de páginas de fitness, 70% de peticiones aplicaciones de planificación de la demanda y fabricación
Procesos de Booking	Hasta un 30% más rápidos
Procesos batch de planificación de la demanda	Entre un 20% y un 50% más rápidos
Procesos de planificación de la demanda	Entre un 20% y un 30% más rápidos
Tiempo medio de proceso de preenvío de equipos	20 segundos aprox
Tiempo medio de proceso de envío de equipos	21 segundos aprox
Rendimiento en visitas por hora a las páginas de fitness	750.000 por hora

Fig 13. Resumen de mejoras obtenidas en Garmin con los equipos Exadata

2.2.4 Pruebas de Escalabilidad realizadas por terceros

Las mejoras de Exadata presentadas en las referencias oficiales muestran una alta variabilidad de resultados y reflejan una mayor eficiencia en las operaciones de entrada/salida. En algunos casos las mejoras son tan altas que parecen indicar un cambio sustancial en el número ejecutado de operaciones de entrada/salida por segundo. Esto no es de extrañar si tenemos en cuenta el uso de *flash*. Las capacidades de Exadata encontradas en las especificaciones, en los modelos X2-2 y X3-2, en entrada/salida por unidad de tiempo, son las siguientes:

	Discos de 600GB a 15.000 rpm	Discos de 7.500 rpm de 3 TB
Máximo ancho de banda a disco	25GB/s	18GB/s
Máximos IOPS de disco	50.000	28.000
Capacidad (raw)	100 TB	504 TB
Capacidad (utilizable)	45 TB	224 TB
Máximo ancho de banda flash	100GB/s	93GB/s
Máximos IOPS en lectura flash	1.500.000	1.500.000
Máximos IOPS en escritura flash	1.000.000	1.000.000
Capacidad Flash	22,4 TB X3-2 / 5,3 TB X2-2	22,4 TB X3-2 / 5,3 TB X2-2
Máximo ratio TB/hora de carga de datos	16 TB/hora X3-2, 12TB/Hora X2-2	16 TB/hora X3-2, 12TB/Hora X2-2

Fig 14. Capacidades de Exadata en entrada/salida por unidad de tiempo[12]

Estas capacidades muestran las diferencias en entrada/salida por segundo de los discos a 15.000 rpm o 7.500 rpm con respecto a los tiempos de las tarjetas *flash* PCI de Exadata, con bloques de 8k, tanto en acceso aleatorio (1.500.000 IOPS en lectura) como en acceso secuencial (100GB/s) Además, la memoria RAM disponible en el modelo X2-2 completo es de 1 TB (2TB en el X3-2) y los servidores de base de datos en Exadata comparten memoria gracias a la arquitectura RAC por lo que puede existir una contribución de esta memoria en las mejoras de los resultados compilados de las referencias.

Existen en el mercado tarjetas *flash* con capacidades similares[13] y sería relativamente sencillo implementar un entorno con un *blade* linux, con uno a varios *slots* PCI dedicados a tarjetas *flash*, con acceso a disco y donde se almacene una base de datos Oracle en una única instancia. Sin embargo, las posibilidades de escalabilidad estarían limitadas al número de *slots* disponibles, entre otras posibles situaciones. En ese sentido Exadata ofrece una plataforma *flash* escalable, de uso por múltiples instancias, y gestionada desde el *software* de Storage, con funcionalidad específica e integrada en el contexto de gestión de memoria llevada a cabo por la base de datos, tal y como se describe en el apartado 2.3.2.

El siguiente gráfico muestra las capacidades de escalabilidad de Exadata en la ejecución de algoritmos de regresión (*scoring* de SAS) con

acceso a 80 millones de registros con diferentes grados de paralelismo, en registros obtenidos por segundo sobre un X2-2 completo. Estos resultados fueron presentados en público por la compañía SAS[14] durante su Global Forum del año 2012:

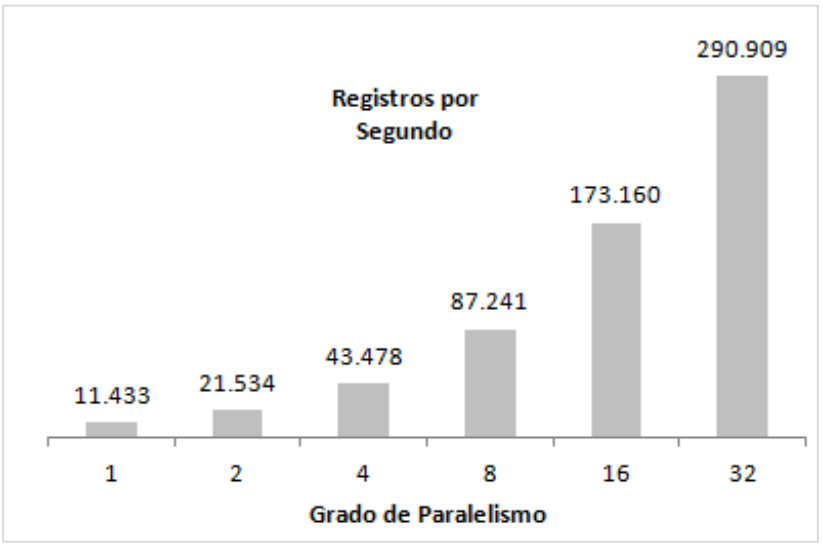

Fig 15. Registros por segundo obtenidos con diferentes grados de paralelismo: Escalabilidad Exadata, algoritmo de regresión de scoring SAS, X2-2 completo, 80 Millones de registros.

En cuanto a tiempo de ejecución, la misma prueba muestra la consistencia en la escalabilidad de Exadata:

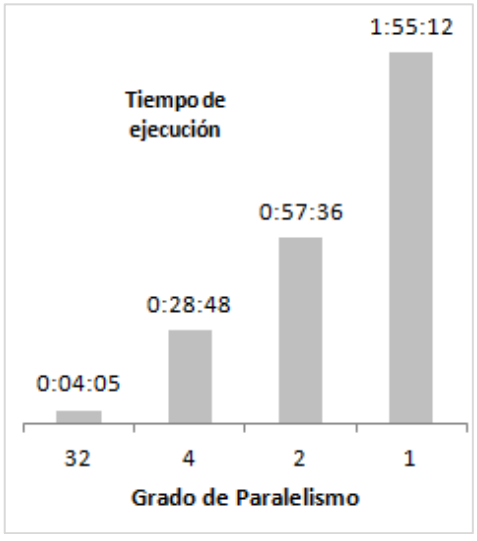

Fig 16. Tiempos de ejecución del algoritmo obtenidos con diferentes grados de paralelismo: Escalabilidad Exadata, algoritmo de regresión de scoring SAS, X2-2 completo, 80 Millones de registros.

Las mismas pruebas ejecutadas sobre 250 millones de registros y presentadas en el Global Forum de SAS mostraban tiempos de ejecución, con un grado de paralelismo de 32 procesos simultáneos, de 14 minutos aproximadamente. Sin paralelismo el tiempo de ejecución fue de 6 horas. Las pruebas presentadas no incluyen una definición del conjunto de *joins* que resuelven la operación de *score*, ni el tamaño medio de cada registro, por lo que no se puede determinar la influencia de la memoria *flash* en las pruebas de escalabilidad, pero si observarse que la escalabilidad existe.

Las operaciones de carga, *backup* y *recovery* implican acceso a disco y las referencias oficiales de Oracle muestran mejoras en esas operaciones. En este caso, la contribución de Exadata es la de proporcionar un acceso *inteligente* a los discos, como veremos en el capítulo 2.3, de tal forma que, probablemente, la combinación de *flash* junto con la compresión de datos, Infiniband y el acceso a los discos son los que permiten las mejoras globales en entrada/salida observadas.

2.2.4 Características de un *benchmark* para Exadata

Tal y como se ha comentado en otros apartados de esta memoria, Oracle no ha publicado resultados de test TPC-C o TPC-H sobre Exadata ¿Cómo puede entonces una empresa determinar si Exadata u otras soluciones cumplirá con sus expectativas? ¿Cómo comparar Exadata con otras arquitecturas para entender qué seleccionar y satisfacer unos requerimientos concretos de servicio para sus bases de datos? Estas preguntas podrían tener una fácil respuesta si existiese un *benchmark* estándar adecuado para cada tipo de cliente, configurable en cuanto al tamaño de la empresa en usuarios/procesos y uso de sus bases de datos, ejecutable desde diferentes soluciones de arquitectura, que permita comparar diferentes *appliances* y arquitecturas tradicionales de cada fabricante, pero no es el caso. Sin embargo existen alternativas que pueden ayudar a tomar las decisiones sobre qué plataforma será la idónea para la ejecución de las bases de datos de cada cliente:

- Ejecución de aplicaciones de propósito general estándar cuando éstas son ejecutadas en el cliente. Una posibilidad es la de realizar pruebas de concepto comparando las alternativas del posible cliente con la ejecución de entornos estándares del mercado como SAP, Peoplesoft, SAS, Siebel, JD Edwards, e-Business Suite u otros similares a los que ejecuta cada cliente, para la capa de base de datos y con los datos del cliente, con diferentes características de escalabilidad, concurrencia y volúmenes de datos a tratar. Esta alternativa tiene la ventaja de que el cliente conoce el entorno de ejecución de su aplicación estándar y puede extraer fácilmente conclusiones de las plataformas a probar con respecto a la ejecución de un software

conocido. La desventaja es que este tipo de pruebas pueden ser muy complejas de preparar para simular el entorno de concurrencia y escalabilidad de cada cliente.

- Pruebas de aplicaciones *core* por tipo de industria. Los clientes de un sector de la industria probablemente conozcan las características de rendimiento de las aplicaciones *core* de su sector aunque éstas no se utilicen en su empresa. Este es el caso de las aplicaciones de b*illing* en el sector *utilities*, los motores intensivos de *core banking* para el sector financiero, las aplicaciones de tratamiento de *call detail records* en el sector telecomunicaciones, las aplicaciones de *insurance & policy administration* en el sector seguros u otras similares. Un cliente de un sector concreto conoce los límites en el tratamiento transaccional de su propia empresa. Las ventajas son que este tipo de *benchmarks* pueden servir como referencia para una industria y el esfuerzo de cada fabricante en cada sector está minimizado a una prueba, mientras que la desventaja está en el grado de fiabilidad obtenido para el cliente que ha de tomar la decisión dadas las particularidades de sus propias aplicaciones en la ejecución de su operativa *core*.

- Pruebas de las aplicaciones propias, en la capa de base de datos. Esta alternativa consiste en recoger los *logs* transaccionales para el caso *OLTP* y los SQL de ejecución en los casos *Data Warehouse* propios, exportarlos a cada plataforma de pruebas y ejecutar una simulación de ejecución en cada destino. Ésta parece la estrategia de Oracle con sus pruebas de escalabilidad en cada cliente y puede resultar la más efectiva, dado que cada cliente puede comprobar las diferencias en tiempos de ejecución de sus entornos actuales y compararlos con los indicadores obtenidos en las pruebas. La ventaja es la fiabilidad de las pruebas para cada cliente pues éste se asegura unas características de servicio para su capa de base de datos antes de la adquisición, las desventajas son que en cada cliente se debe ejecutar la prueba de concepto y los resultados no son exportables a otros clientes.

- Diseño de un conjunto de *benchmarks* específicos para *applicances* o adaptación de los *benchmarks* estándar. Probablemente la mejor alternativa, puesto que la publicación de resultados genéricos pueden aportar la suficiente información para que un cliente y los fabricantes se ahorren el tener que ejecutar pruebas de concepto para cada caso.

¿Cómo debería ser un *benchmark* para *appliances*? La comparativa entre diferentes alternativas para cada cliente va a depender del uso de

la plataforma, dado que será diferente un *benchmark* para un entorno OLTP o un *Data Warehouse*. En todo caso, se han de considerar:

- En las pruebas de entornos OLTP: Pruebas de escalabilidad y saturación, con diferentes volúmenes de datos, ejecutadas en simulación de un entorno real. Es decir, con *logging* transaccional, ejecución de *backup* en tiempo real, con arquitecturas de máxima disponibilidad y protección ante desastres y con las configuraciones de cada arquitectura a probar que vayan a ser similares a la calidad de servicio esperado para la capa de base de datos.
- En las pruebas de entornos Data Warehouse: Pruebas de rendimiento de SQLs, simples y complejos, con diferentes volúmenes de datos y concurrencia de ejecución, simulaciones de carga, diferentes grados de compresión si ésta es utilizable y *backup* si es requerido.

Fin del capítulo 2.2

Notas y Bibliografía de 2.2

1 Nota de la Web de Oracle: Customers worldwide adopt Exadata Database Machinehttp://www.oracle.com/us/solutions/datawarehousing/exadata-customers-421304.html **[Consulta Web], Fecha de Consulta: Marzo de 2013.**

2 Reference *Booklet* público de Oracle sobre Casos de clientes de Exadata, http://www.oracle.com/us/products/database/exadata-reference-booklet-400018.pdf **[Consulta Web] Fecha de Consulta: Marzo de 2013.**

3 Casos publicados encontrados, *Búsqueda de Casos de Clientes Exadata,* http://www.oracle.com/search/customers **[Consulta Web], Fecha de Consulta: Marzo 2013.**

4 No encontradas pruebas TPC sobre Exadata en Febrero de 2013 en http://www.tpc.org

5 Varias empresas encontradas utilizando el buscador de Google que proporcionan servicios de benchmarking sobre Exadata. Clave de búsqueda: *Benchmark Exadata*, Fecha de Consulta: Marzo 2013.

6 Casos recopilados del Reference Booklet (nota 2)

7 Número de referencias encontradas con la palabra clave Exadata (nota 3)

8 Kenneth M. Richhart, *Transforming Information Technology Services: Reducing Cost and Improving Availability*. CBP Agency, Department of Homeland Security, http://www.actgov.org/knowledgebank/documentsandpresentations/Documents/Program%20Events/Executive%20Session%20featuring%20Ken%20Ritchhart,%20CBP%201-12-11.pdf **[Consulta Web] Fecha de Consulta: Marzo 2013.**

9 Definición de Computer Appliance, Wikipedia http://en.wikipedia.org/wiki/Computer_appliance [Consulta Web] Fecha de Consulta: Marzo 2013.

10 Turkcell, *Turkcell Accelerates Reporting Tenfold, Saves on Storage and Energy Costs with Exadata Database Machine*, http://www.oracle.com/us/corporate/customers/turkcell-1-exadata-case-study-456284.pdf [Consulta Web] Fecha de Consulta: Marzo 2013.

11 Oracle Exadata Database Machine Technical Case Study: *Garmin International Inc, White Paper* http://www.oracle.com/technetwork/database/availability/garmin-1667151.pdf [Consulta Web] Fecha de Consulta: Marzo 2013.

12 Folletos Exadata, *Datasheets Exadata X2-2, X2-8, X3-2 y X3-8,* http://www.oracle.com/us/products/database/exadata/overview/index.html [Consulta Web], Fecha de Consulta: Marzo de 2013.

13 Flash Memory, *Wikipedia* http://en.wikipedia.org/wiki/Flash_memory [Consulta Web] Fecha de Consulta: Marzo 2013.

14 Paul Kent, *Big Data SAS (2012) SAS Scoring Accelerator for Oracle and Beyond, presentación realizada y publicada en el* SAS Global Forum 2012 http://www.oracle.com/technetwork/database/.../sas/385-2012-1608576.pdf [Consulta Web] Fecha de Consulta: Marzo 2013.

2.3 Arquitectura Técnica y componentes de Exadata

2.3.1 Introducción

Esta capítulo detalla las funcionalidades y componentes de Exadata principales que facilitan el rendimiento observado y expuesto en el capítulo 2.2, tanto en las referencias oficiales como en las pruebas de escalabilidad.

El apartado 2.3.2 describe la arquitectura global de Exadata, con énfasis en la arquitectura de los servidores de almacenamiento, la distribución de tareas en la ejecución del SQL entre los servidores de bases de datos y los servidores de almacenamiento y las funcionalidades de *smart scan*, el uso de la memoria *flash*, la comprensión híbrida columnar y los índices de almacenamiento dado que sus aportaciones son clave para el rendimiento observado.

Exadata utiliza la funcionalidad de ASM[1] para la distribución y acceso a la información residente en los discos, y aunque esta funcionalidad no es exclusiva de Exadata, es de utilidad conocerla para la comprensión del uso de los discos. El apartado 2.3.3 resume ASM en el contexto de Exadata e introduce el concepto de los *Storage Indexes*, como aportadores de la reducción de la entrada/salida.

El apartado 2.3.4 describe el nuevo protocolo iDB y el uso de Infiniband[2] entre los servidores de bases de datos y los servidores de almacenamiento, también importantes contribuidores a las mejoras de rendimiento de Exadata.

En el apartado 2.3.5 se resumen las probables áreas de influencia de cada funcionalidad descrita en este capítulo con respecto a las mejoras de rendimiento presentadas por Oracle en sus referencias oficiales.

Acaba este capítulo con el apartado 2.3.6 donde se exponen algunas de las críticas a Exadata recopiladas en la *web*.

2.3.2 Componentes de Exadata y su contribución para el rendimiento observado

En el apartado 2.1.2 se describen las diferentes configuraciones de la 4ª generación de Exadata de la familia X3-2 y la familia X3-8, y las características de los servidores de bases de datos y los servidores de almacenamiento. Por ejemplo, el equipo 1/4 Exadata X3-2 se compone de dos servidores de bases de datos y 3 servidores de almacenamiento, cada servidor de almacenamiento dispone de 12 discos, y todos los discos del 1/4 Exadata deben ser de 15.000 rpm o de 7.500 rpm.

Además, los servidores de bases de datos y los servidores de almacenamiento están conectados con una red Infiniband. Gráficamente:

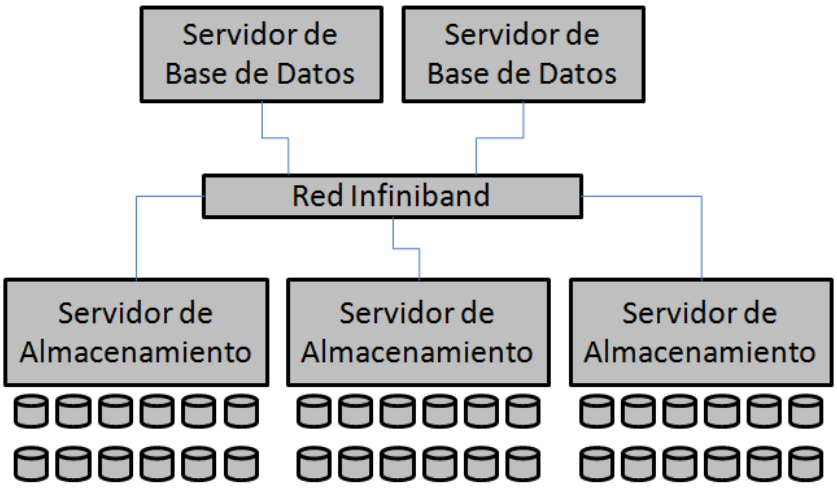

Fig 17. Esquema de servidores de ¼ Exadata X3-2

Cada servidor de base de datos de la familia X3-2[3] se compone de dos procesadores Intel Xeon E5-2690, con 8 *cores* por procesador y 256GB de RAM, mientras que cada servidor de almacenamiento[4] se compone de dos procesadores E5-2630L, con 6 *cores* cada procesador, 4 tarjetas PCI *flash* F40 eMLC de 0,4TB cada una y los 12 discos. Gráficamente, y siguiendo con el caso de 1/4 Exadata:

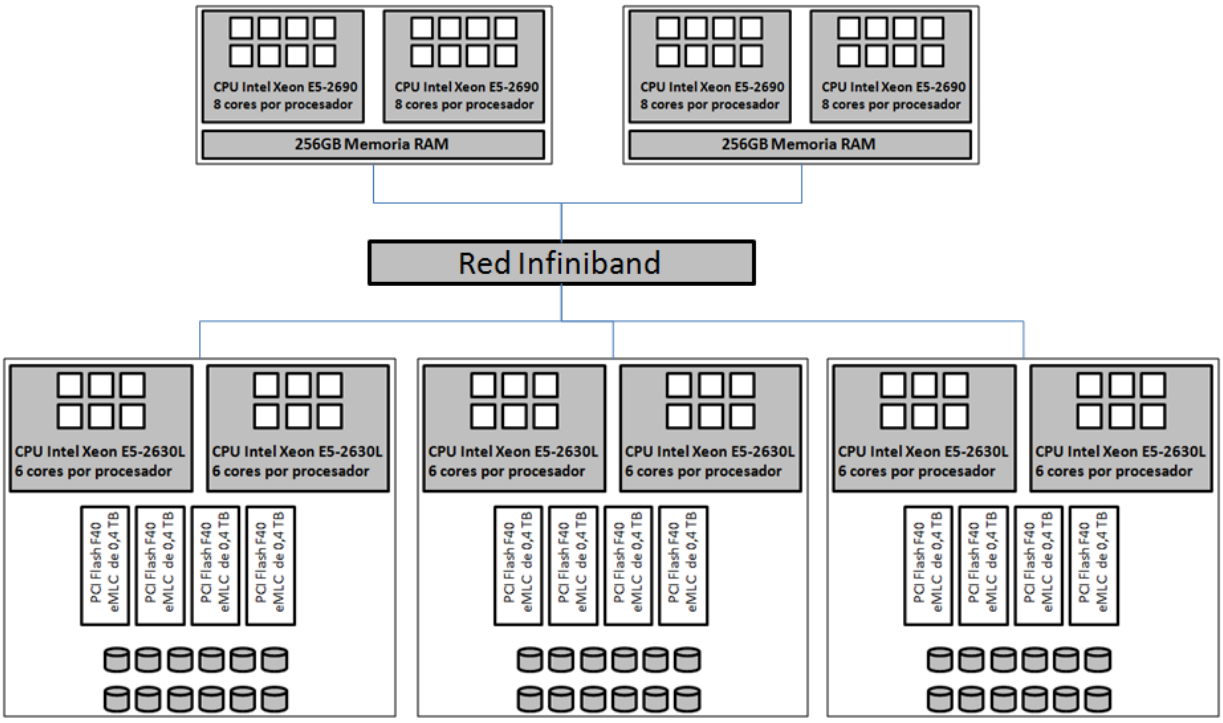

Fig 18. Esquema de servidores de bases de datos y almacenamiento de ¼ Exadata X3-2 en mayor detalle

Los servidores de almacenamiento proporcionan un conjunto de funcionalidades diseñadas para mejorar los tiempos de acceso a disco. Entre otras, las más destacadas son *Smart Scan*, *Smart Flash* Caché y la compresión híbrida columnar.

Smart Scan[5]

En una arquitectura tradicional de servidores que ejecutan bases de datos Oracle no existen los servidores de almacenamiento. Cualquier configuración de servidores, ya sean multiprocesadores o multicomputadores en *cluster*, ejecutan todos los procesos SQL y acceden a los discos para obtener los bloques de datos que han de ser procesados. Por ejemplo, un equipo multiprocesador que ejecuta una base de datos Oracle con acceso a un *array* de *storage*, que almacena los bloques de datos, seguiría los siguientes pasos para la ejecución de una instrucción SQL:

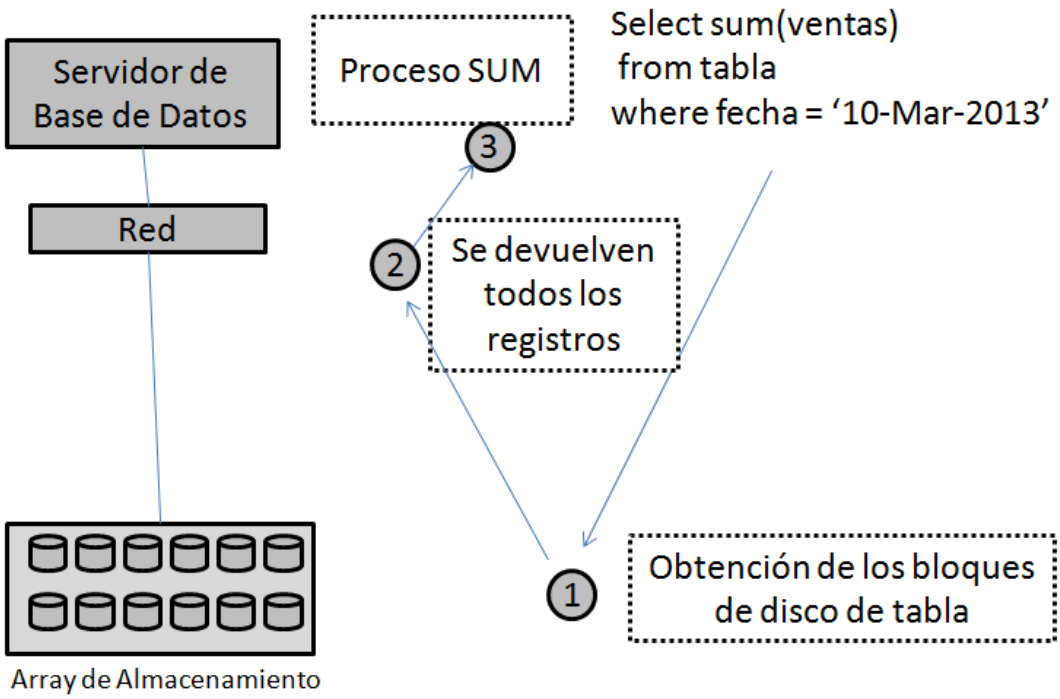

Fig 19. Arquitectura tradicional, sin servidores de almacenamiento

Con Exadata Oracle ha introducido un nuevo concepto de ejecución de las instrucciones SQL, denominado *smart scan predicate filtering*, que divide la ejecución de las instrucciones entre los servidores de base de datos y los servidores de almacenamiento. El siguiente gráfico muestra el mismo proceso del ejemplo anterior en Exadata:

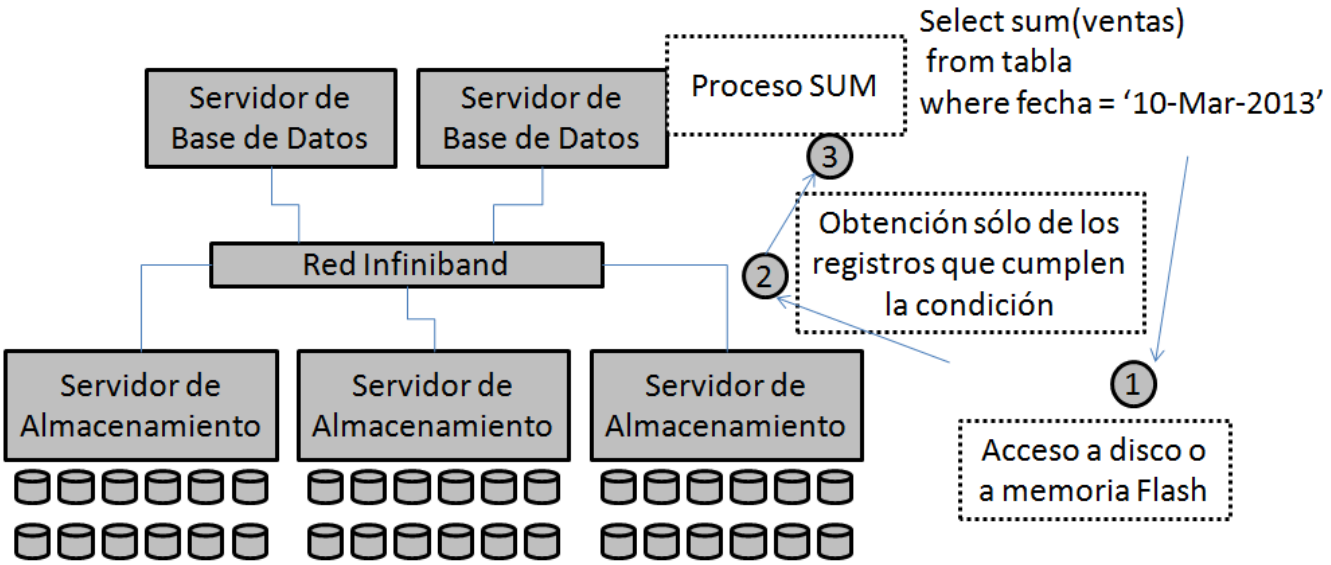

Fig 20. Distribución del proceso de ejecución de un SQL entre los servidores de base de datos y los servidores de almacenamiento en Exadata

La distribución de tareas entre el servidor de base de datos y el servidor de almacenamiento para la ejecución de SQL es transparente a los procesos de usuario y cumple con los requerimientos de consistencia estándar de la base de datos en el acceso a los registros[6]. Por otra parte, esta funcionalidad es aplicable también para SQL complejas, incluyendo subconsultas enlazadas.

La funcionalidad *Smart Scan* también permite el filtrado de columnas, denominado *Smart Scan Column Filtering*. Los servidores de almacenamiento sólo devuelven a los servidores de bases de datos las columnas solicitadas, reduciéndose el número de operaciones de entrada/salida comparado con la arquitectura tradicional, donde todos los bloques de datos y todas las columnas serán enviados a los servidores de base de datos para su proceso. Dado que la entrada/salida suele ser un cuello de botella, y en el caso de que las tablas accedidas tengan un gran número de columnas y sólo se acceda a un subconjunto de ellas, esta funcionalidad permite reducir el tráfico de red entre discos y servidores. El siguiente gráfico muestra un ejemplo del filtrado de columnas:

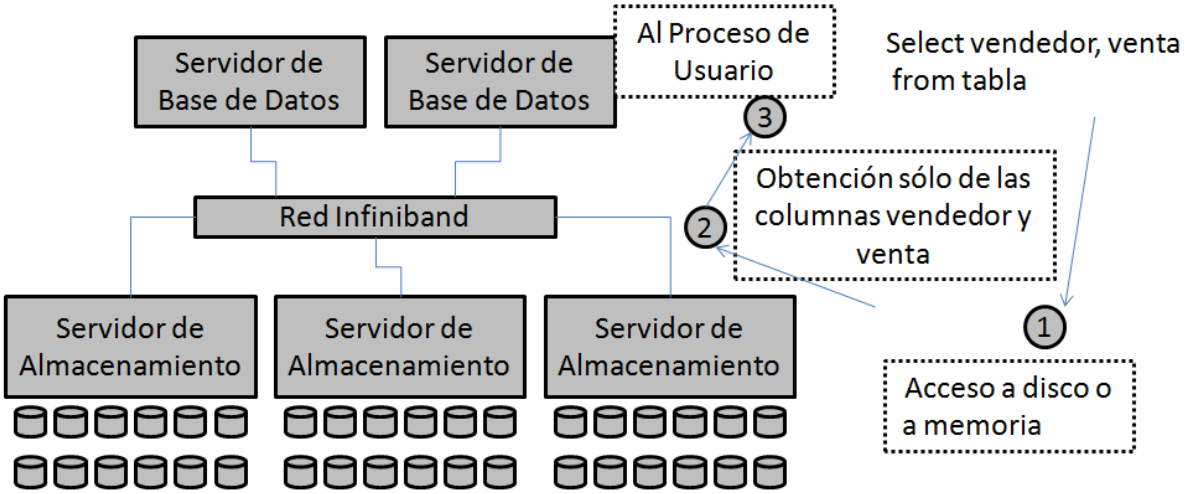

Fig 21. Distribución de tareas de ejecución de SQL en Exadata gracias a Smart Scan Column Filtering

Exadata Storage Software incluye una funcionalidad de filtrado de *joins*, denominada *Smart Scan Join Filtering*, que utiliza un mecanismo probabilístico[7] para determinar los registros que forman parte del resultado de una operación de join entre grandes tablas y pequeñas tablas de *lookup*.

La empresa Centroid[8] ha publicado en su web un estudio sobre la eficiencia de *Smart Scan* que muestra sus beneficios en reducción de operaciones de entrada/salida. El siguiente cuadro resume las pruebas y resultados obtenidos:

Número de registros tabla de test	180 millones
Tamaño de la tabla de test	4,53GB
Tiempo ejecución query de ejemplo con full scan	5,58 Segundos
Subtarea plan de ejecución query de ejemplo	Table Access Storage Full
Área elegible para Smart Scan de la query de ejemplo	4,49GB
Tamaño retornado al servidor de base de datos (área PGA de usuario)	44,5 MB
Eficiencia de Smart Scan	**99,50%**
Sin Smart Scan	*Eliminación vía Alter Session, System o Hint*
Tiempo query de ejemplo	*31,39 segundos*
Área retornada al servidor de base de datos	*4,49GB*

Fig 22. Resultados publicados empresa Centroid, test de Smart Scan

Smart Flash Caché[9]

Cada *Storage Server* de Exadata tiene 4 tarjetas PCI *flash* de 0,4 TB cada una y un equipo X3-2 completo tiene 56 tarjetas *flash*, totalizando 22,4 TB. Todo este espacio es considerado como una caché de memoria adicional y es gestionado desde el *software* de *Storage* automáticamente. Cuando un servidor de base de datos envía una petición de lectura o escritura al *software* de *Storage*[10] se envía, en la petición, información adicional que permite determinar si la información tratada debe permanecer en la memoria *flash* o no. Si la operación es una petición de lectura/escritura aleatoria normalmente la información será almacenada en la caché *flash*, mientras que una petición de tipo secuencial puede que almacene la información en la caché o no. La memoria *flash* permite la ejecución de escrituras, es persistente a los *reboots* y mantiene la consistencia de la base de datos. Esta funcionalidad se denomina *Smart Flash Cache Write Back* y su rendimiento es de hasta 20x más IOPS que los discos en la generación X3 y hasta 10x más IOPS que los discos en generaciones V2 y X2, según la empresa Oracle[11], especialmente relevante para aplicaciones de tipo OLTP.

El acceso al área de *flash* puede ser directo o de tipo agregado para accesos secuenciales. En el caso agregado, dado el carácter multiproceso de las instancias de base de datos, se accede en paralelo tanto al área de *flash* como a los discos, sobre áreas contiguas, lo que puede favorecer un mayor rendimiento global de las operaciones de acceso. La gestión de todos estos procesos la realiza el *software* de *Storage* de Exadata.

Para la incorporación o eliminación manual de objetos en *flash* existen dos tipos de métodos. Por una parte existen directivas a nivel de DDL para indicar a la base de datos si un objeto (tabla, índice o segmento) debe permanecer en la memoria *flash*. Por otra parte se puede definir un disco lógico en *flash* e incluir en este disco lógico la información frecuentemente accedida. La ventaja en este caso es que la gestión del disco lógico *flash* la realiza ASM[12], integrado con la gestión global de la entrada/salida de las bases de datos.

Adicionalmente, la caché global de *flash* tiene asignado un área para almacenar las operaciones de *log*[13] de la base de datos. Estas operaciones, que requieren de escritura a disco después del *commit*, se realizan simultáneamente sobre el área *flash* y sobre la memoria de cada controladora de disco asignado al *logging*. La primera que acaba la escritura informa a la base de datos de su finalización. Esta funcionalidad puede permitir a los sistemas OLTP obtener una mejora del rendimiento dado que la actividad de *logging* suele ser un cuello de botella.

Un test realizado y publicado por la empresa suiza Benchware[14], en Junio de 2012, muestra la escalabilidad en lectura aleatoria proporcionada por la memoria *flash* de un equipo ½ Exadata X2-2. Las pruebas publicadas se ejecutaron sobre configuraciones de 1 nodo de base de datos, 2 nodos y 4 nodos, con diferente grado de paralelismo, hasta obtener la saturación de cada prueba, desde los 64 procesos concurrentes hasta 1.000 procesos concurrentes. Lo interesante de esta prueba es que parece que el mismo test[15] se realizó en Septiembre de 2012 sobre un equipo M5000, también de la empresa Oracle, con Sun Solaris y una tarjeta *flash* F20 conectada vía un bus PCI. Aunque no son comparables los resultados, dadas las diferentes características de las tarjetas *flash* (F20 tiene 96GB) y del uso de *flash* diferenciado en Exadata vía el Storage Software, sí que muestran la escalabilidad de la arquitectura *flash* del equipo Exadata. El siguiente gráfico muestra los resultados en IOPS para lecturas aleatorias en cada caso:

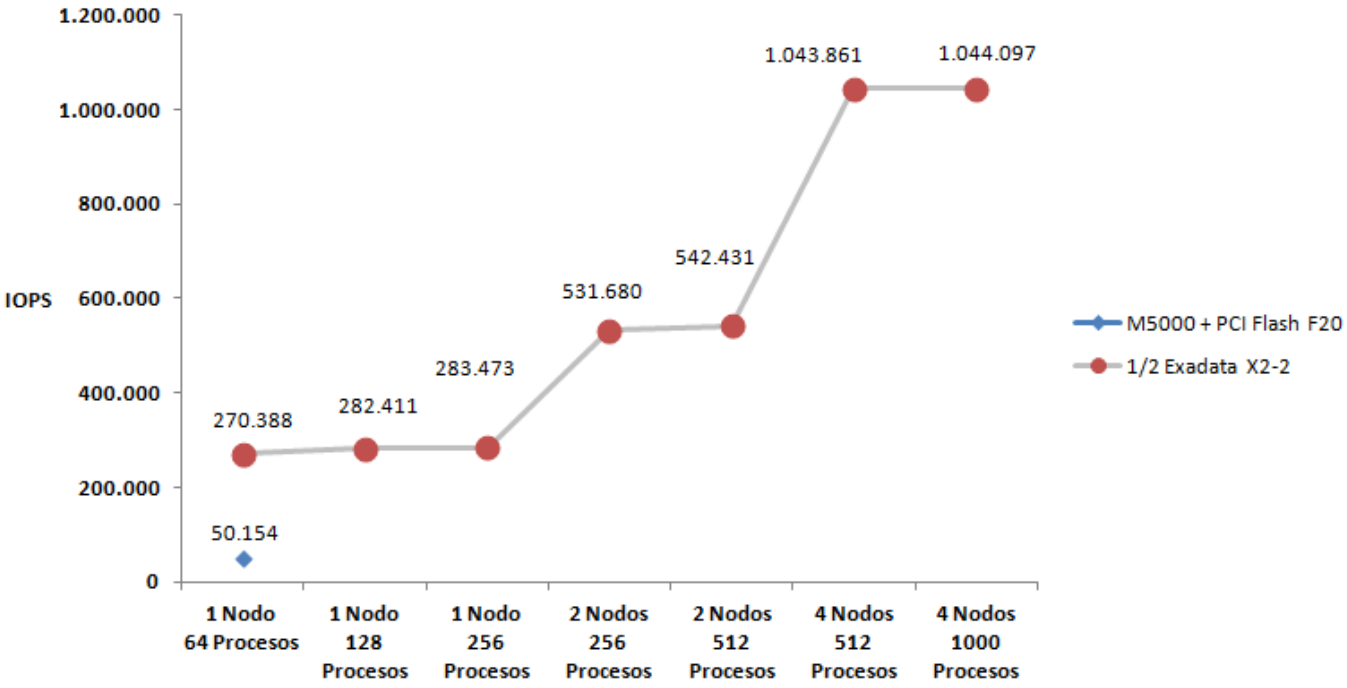

Fig 23. Resultados publicados por la compañía suiza Benchware de lecturas aleatorias sobre un M5000 con una tarjeta flash PCI F20 y ½ Exadata en diferentes configuraciones

Compresión Híbrida Columnar (*Hybrid Columnar Compression* o HCC)[16]

Tal y como se ha descrito *Smart Scan* permite reducir el número de bloques de datos que son enviados, según son requeridos, desde los discos hasta las áreas de memoria de los servidores de base de datos con un objetivo de minimización de la entrada/salida. Para reducir aún

más la circulación de datos por la red, previo a su tratamiento por los servidores de base de datos, Exadata incorpora una nueva funcionalidad de compresión denominada *Hybrid Columnar Compression*, para uso principal en entornos *Data Warehouse* y para actividades de *archiving*[17], que permite reducir el consumo de memoria, *flash* y tráfico de red.

Los bloques de datos se almacenan, en una arquitectura tradicional, registro a registro, con las columnas que forman parte de las tablas ordenadas secuencialmente. Cambiar la organización de los bloques de datos agrupando columnas, en una hipotética reorganización, puede permitir una reducción de espacio considerable con algoritmos de optimización o compresión simples, dado que, habitualmente, muchos datos de cada columna de una tabla pueden estar repetidos. Sin embargo, esta hipotética organización en agrupación de columnas dentro de cada bloque puede implicar pérdidas de rendimiento en el acceso a más de una columna, con consultas simples o complejas, al poder llegar a requerirse más operaciones de entrada/salida para acceder a columnas del mismo registro[18].

Exadata ha introducido un sistema híbrido de compresión, entre la organización de registros secuenciales en cada bloque y la organización de agrupación en columnas. Para un subconjunto de registros[19] se define una unidad lógica de compresión, denominada *Unit Compression*, formada por varios bloques de datos, donde se utiliza una organización de agrupación por columnas. Los datos son cargados en cada unidad de compresión utilizando las técnicas de carga masiva de los *Data Warehouse*, como *Create Table as Select, Direct Path* o cargas en modo *bulk load*. Una vez cargados son comprimidos utilizando un algoritmo de compresión[20]. La siguiente figura ilustra la organización en unidades de compresión por agrupación de columnas:

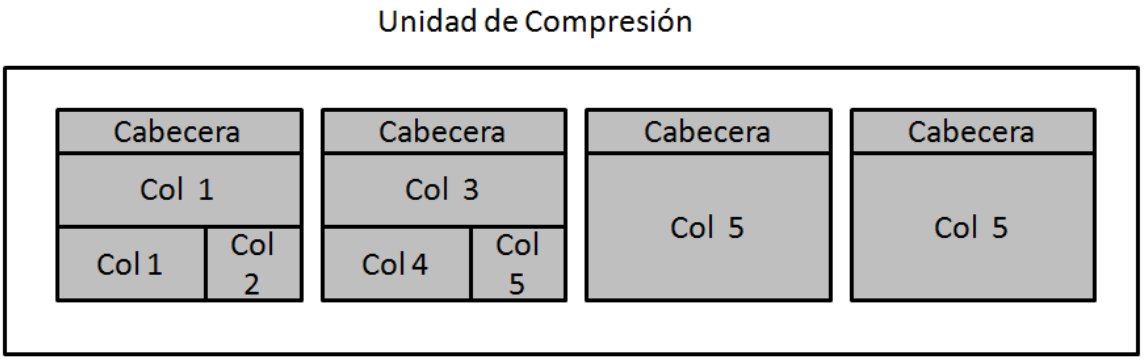

Fig 24. Organización de bloques de datos por agrupación de columnas en una unidad de compresión, HCC

El algoritmo de compresión utilizado distingue dos tipos de operaciones que Oracle denomina *Warehouse Compression* (o *Query Compression*) y *Archiving Compression*. En ambos casos, vía DDL, se puede definir un

método de compresión más agresivo (*HIGH*) o menos agresivo (*LOW*). Oracle afirma[21] que para un *Data Warehouse* típico la media de compresión obtenida con un algoritmo *High* es de cerca de 10x (ratio de compresión x 10), en el caso *Low* es 6x. Para el caso de *Archiving Compression* la media es 15x pero puede suponer pérdida de rendimiento y es recomendado su uso para datos accedidos poco frecuentemente.

Un estudio[22] publicado en Mayo de 2011, basado en una presentación realizada en vivo, vía *webcast*, sobre la funcionalidad de HCC presentada por el Sr. Diwakar Kasibhotla, de la empresa OppenheimerFunds, muestra una comparación entre los tiempos de acceso y carga con y sin HCC en un test realizado sobre Exadata:

Equipo	1/4 Exadata, no se especifica modelo
Tabla particionada	9 particiones con 10 subparticiones cada una
Índices	No
Registros	67 millones aprox
Query	Select c1,c2, sum(c3) from tabla group by c1,c2
Registros obtenidos en la query	92.633 registros

Resultados	Sin HCC	Query High	Query Low	Archiving High	Archiving Low
Tamaño Final	7,23GB	1,76GB	2,68GB	1,35GB	1,65GB
Ratio de compresión	1	4,1	2,69	5,35	4,38
Tiempo ejecución query	26,07 seg	12,04 seg	11,12 seg	13,71 seg	12,03 seg
Tiempo de carga Bulk loading (incluyendo otras tablas de hechos y dimensiones)	23,49 min	19,02 min	22,32 min	41,06 min	20,11 min

Fig 25. Resultados del test de HCC sobre la tabla de 67 millones de registros

2.3.3 Almacenamiento, ASM y *Storage Indexes*

Cada disco físico asignado a un servidor de almacenamiento se denomina *cell disk*[23] y se subdivide, como mínimo, en una entidad lógica denominada *grid disk* que puede expandirse en más de un disco físico. Los *grid disk* pueden utilizarse para separar bases de datos o para separar zonas de la misma base de datos con diferente intensidad de uso y pueden expandirse a más de un servidor de almacenamiento. Por otra parte, se reservan grupos de *grid disk* para proporcionar copias de los bloques de datos, mediante *stripping y mirroring*. El siguiente diagrama muestra estos conceptos:

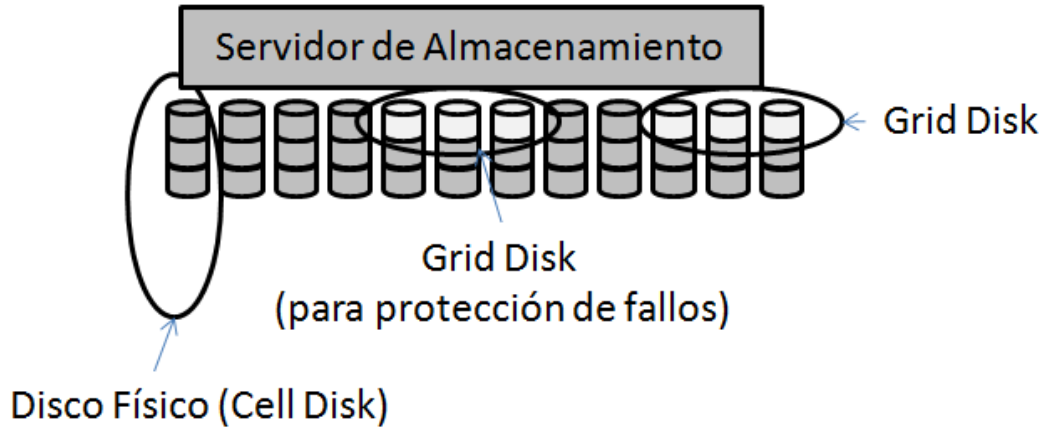

Fig 26. Cell Disk y Grid Disk en un servidor de almacenamiento

La gestión del almacenamiento se realiza a través de ASM[24], funcionalidad de la base de datos Oracle introducida en la versión 10, que automatiza procesos de gestión de los volúmenes, realiza las funciones de *stripping, mirroring*, la asignación de áreas de espacio o el movimiento de grupos para optimizar la entrada/salida. Así, los discos se controlan, asignan o deasignan desde SQL. Adicionalmente, ASM se encarga de la redistribución de los grupos lógicos al añadir o eliminar un disco físico. Desde el punto de vista de los procesos *software*, ASM se ejecuta en los servidores de base de datos:

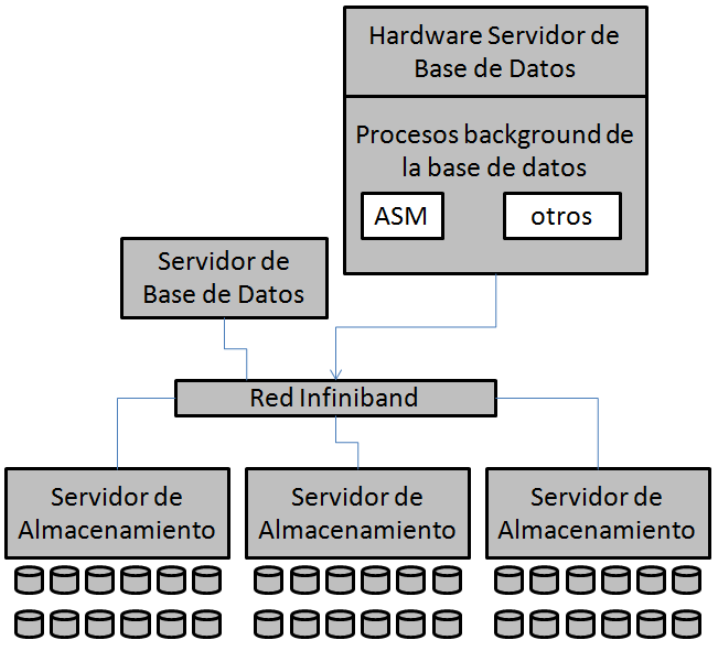

Fig 27. Los procesos de background de ASM se ejecutan en los servidores de base de datos

El proceso CELLSRV de cada servidor de almacenamiento gestiona la entrada/salida con los discos físicos, ejecuta la parte de SQL de *Smart Scan*, gestiona la priorización de las tareas en el acceso a disco y se comunica con un subproceso de los procesos de *background* de ASM denominado LIBCELL y con las instancias de la base de datos vía el nuevo protocolo iDB. Además de CELLSRV, cada servidor de almacenamiento dispone de un proceso de rearranque de tareas, denominado *Restart Service* o RS, similar al proceso SMON de la instancia de la base de datos, y de un proceso de gestión denominado *Management Service* o MS que ejecuta los comandos de administración del servidor de almacenamiento. Gráficamente:

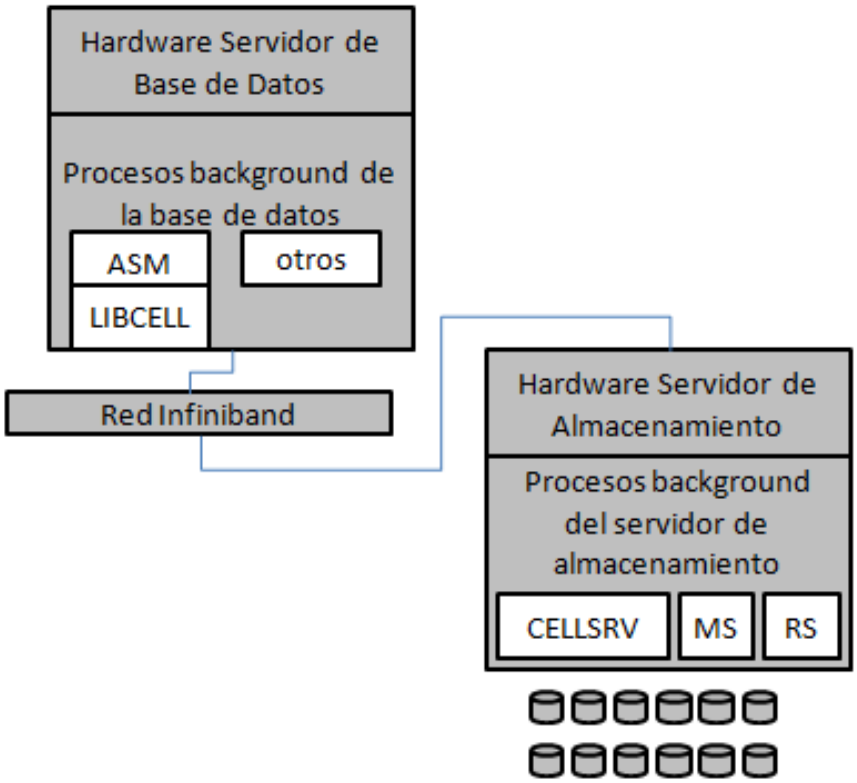

Fig 28. Procesos del servidor de almacenamiento

En un equipo Exadata se pueden ejecutar diferentes bases de datos, algunas de las cuales pueden formar parte de un mismo entorno de ejecución, como por ejemplo una aplicación que requiere de 2 bases de datos diferentes. Incluso para una misma base de datos pueden existir tareas diferentes, como por ejemplo tareas *batch* y tareas de tipo OLTP. Cada uso o sesión de la base de datos supone una carga en cuanto a consumo de CPU o entrada/salida, entre otras, y los administradores pueden asignar prioridades agrupando las sesiones por tipo de carga, agrupando determinado uso de recursos en conjuntos y asociando los grupos de sesiones a los conjuntos de uso de recursos en planes de

ejecución. Cuando estos planes involucran a más de una base de datos los administradores crean categorías de planes de ejecución.

Esta funcionalidad ya existía en la arquitectura tradicional de bases de datos Oracle (sin Exadata), donde el proceso *background* DBRM es el encargado de la distribución de recursos y la vigilancia y asignación de recursos a los grupos[25]. Con Exadata la asignación de los recursos de CPU la sigue realizando el proceso DBRM, y la asignación y priorización de recursos de entrada/salida es una tarea distribuida entre DBRM y un nuevo proceso, exclusivo de Exadata, denominado IORM. Cada servidor de almacenamiento recibe, para petición de entrada/salida, un metadato de la categoría o grupo de recursos donde está asociada la sesión, a través de varias colas de mensajes gestionadas por CELLSRV. A posteriori, CELLSRV envía al proceso IORM cada petición, se evalúa el plan de ejecución y se encola en función de la prioridad.

Para cada MB de disco existe una entrada a un índice gestionado por CELLSRV y denominada *Storage Index* que almacena en memoria el valor máximo y mínimo de cada columna, de tal forma, que cada operación SQL (la parte SQL que ejecuta el servidor de almacenamiento vía CELLSRV) determina transparentemente si un subconjunto de bloques han de ser accedidos o no. Previo a la ejecución del *Smart Scan*, cuando una *query* con cláusula *where* es ejecutada se analiza el *Storage Index* y se eliminan los accesos a los bloques de datos que no van a ser requeridos. El siguiente diagrama muestra un ejemplo del funcionamiento de los *Storage Indexes*:

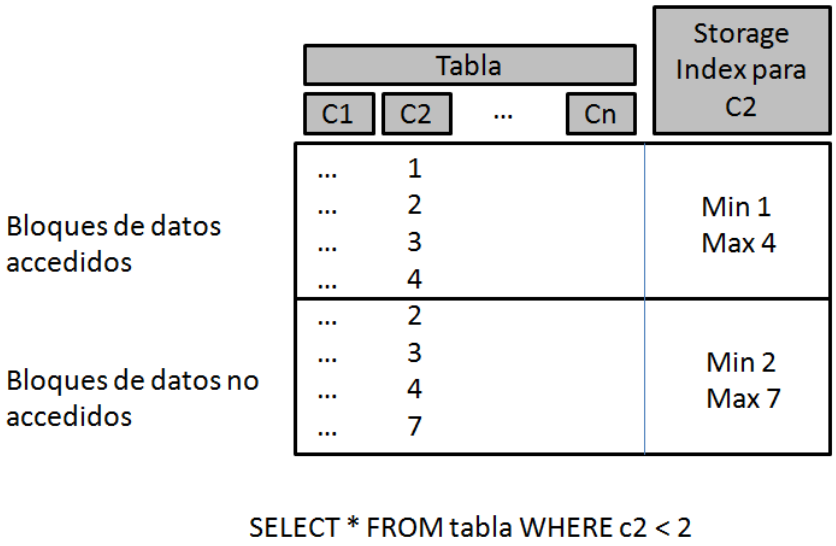

Fig 29. Ejemplo de discriminación en el acceso a bloques de datos vía Storage Indexes

La empresa Centroid también ha publicado en su *web* un estudio sobre la eficiencia de *Storage Indexes* que muestra sus beneficios en

reducción de operaciones de entrada/salida. El siguiente cuadro resume las pruebas y resultados obtenidos:

SIN Storage Index vía	Acceso a tablas del diccionario con Wild Cards en la consulta (eliminan el uso del Storage Index)
CON Storage Index	Acceso a tablas del diccionario sin Wild Cards
Registros obtenidos con la consulta	1,6 millones aproximadamente
CON Storage Index, tiempo de ejecución de la consulta y operaciones en GB IO evitadas	0,39 segundos 14GB de accesos IO evitados
SIN Storage Index, tiempo de ejecución de la consulta	4,12 segundos

Fig 30. Resumen de resultados del test de la empresa Centroid sobre la eficiencia de Storage Indexes

2.3.4 Infiniband en Exadata[26]

Exadata utiliza Infiniband para la conexión entre los servidores de bases de datos entre sí, cuando se requiere el uso de RAC, y para conectar los servidores de datos con los servidores de almacenamiento, que permite múltiples enlaces simultáneos, un alto ancho de banda en la comunicación y menor latencia que las arquitecturas tradicionales en el acceso a *arrays* de *Storage* basadas, por ejemplo, en *Fiber Channel*[27]. Exadata incluye varios *switches* SUN Datacenter QDR (Quad Data Rate) de 36 puertos, para la creación de enlaces de comunicaciones de hasta 40GB/segundo en cada sentido. En cuanto a conectividad, las diferentes configuraciones de la familia X3-2 y X3-8 disponen de los siguientes elementos:

	X3-2 1/8	X3-2 1/4	X3-2 1/2	X3-2 Completo	X3-8 Completo
Switches Infiniband SUN Datacenter de 36 Puertos QDR a 40GB/segundo	2	2	3	3	3
Puertos 10GigE de 10GB	10 (5xservidor)	10 (5xservidor)	20 (5xservidor)	40 (5xservidor)	16 (8xservidor)
Puertos Ethernet para Administración	1	1	1	1	1

Fig 31. Switches Infiniband y puertos GigE en la familia X3-2 y X3-8 de Exadata

Cada servidor (de datos y de almacenamiento) dispone de una tarjeta Infiniband *dual port* QDR PCI-E LP HCA. En las configuraciones de 3 *switches* (1/2 X3-2, Completo X3-2 o X3-8) cada servidor se conecta a

dos *switches* “hoja”, para obtener redundancia, y los switches “hoja” se conectan al tercer *switch* denominado *Core* o *Internal Spine*. En las configuraciones 1/8 y 1/4 no es requerido un *switch Core*. Cuando se requiere conectar un elemento externo a Exadata (*backup* de cinta por ejemplo) se utiliza el *Internal Spine* en las configuraciones 1/2 o completas o uno de los *switches* hoja en las configuraciones 1/8 y 1/4. Todas las conexiones vienen preinstaladas de fábrica. Gráficamente:

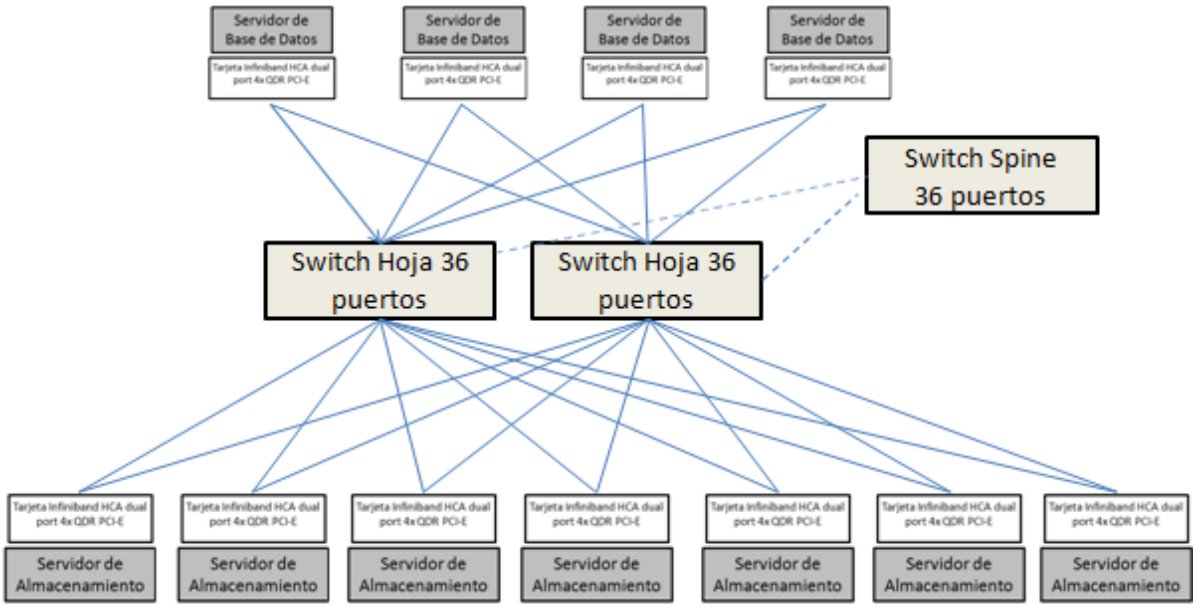

Fig 32. Conexiones con switchs hoja y switch spine en una configuración X3-2 1/2

Para la conectividad entre servidores Exadata utiliza el protocolo *Zero Loss Zero Copy Datagram* (ZDP), basado en *Reliable Datagram Sockets*, con acceso directo a memoria (*DMA*) desde los cables hasta la memoria de cada servidor. Para la conectividad con elementos externos se utiliza *Internet Protocol* sobre Infiniband (IPoIB) o los puertos 10GigE que proporciona una *ethernet* externa para las aplicaciones cliente de las bases de datos. Sobre ZDP se implementa el protocolo iDB para la comunicación entre los procesos (LIBCELL y procesos de *background* de las bases de datos y CELLSRV para los servidores de almacenamiento) Cuando es posible la ejecución de *Smart Scan*, iDB implementa una función de envío y recepción de SQL y datos del resultado obtenido, adicionalmente a la copia de los bloques de datos de disco cuando *Smart Scan* no se utiliza. Gráficamente:

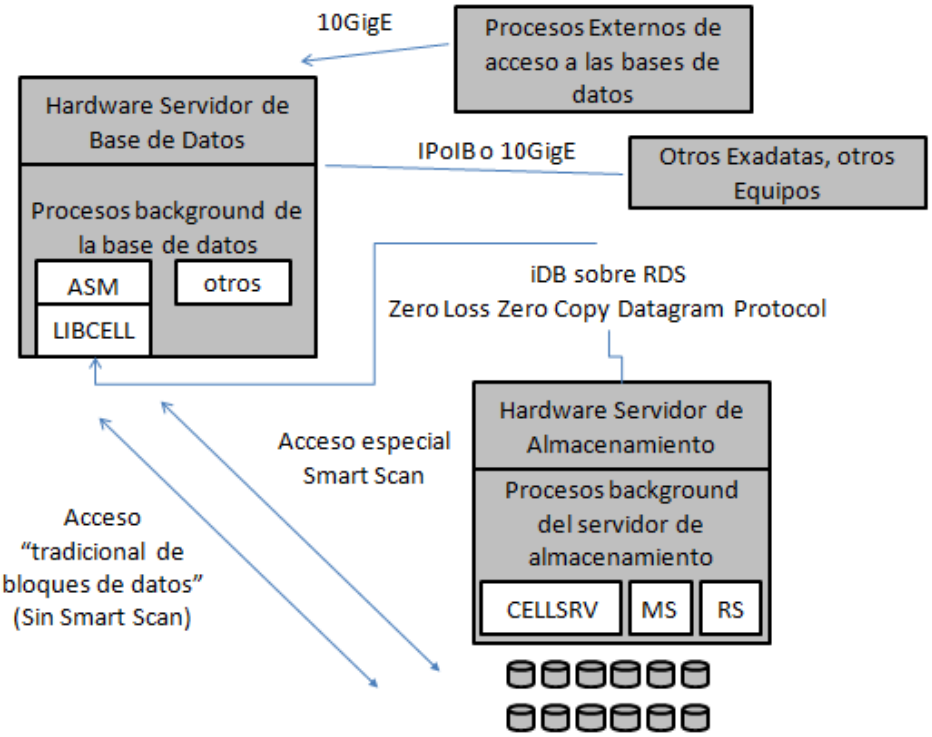

Fig 33. Protocolos y accesos internos y externos a Exadata

2.3.5 Cuadro resumen

Las mejoras de rendimiento de Exadata publicadas por la empresa Oracle en sus referencias oficiales pueden haberse debido, en su mayor parte, a la nueva funcionalidad proporcionada por Exadata y descrita en este apartado.

Funcionalidad	Mejora Principal	Tipo de Aplicación
Smart Scan Predicate Filtering	- Tráfico de red	Mixto
Smart Scan Column Filtering	- Tráfico de red	Mixto
Smart Scan Join Filtering	- Tráfico de red	Mixto
Smart Flash Cache	Velocidad de acceso a bloques de datos	Mixto
Smart Flash Cache Write Back	Velocidad de acceso a bloques de datos	OLTP
Discos lógicos en flash gestionados por ASM	Velocidad de acceso a bloques de datos	OLTP
Flash Logging	Velocidad de acceso a bloques de datos	OLTP
Compresión híbrida columnar (HCC) Warehouse Compression	- Tráfico de red	DWH
Compresión híbrida columnar (HCC) Archiving Compression	- Tráfico de red	DWH
IORM	- Rendimiento Global	Mixto
Storage Indexes	Velocidad de acceso a bloques de datos	Mixto
Infiniband	- Tráfico de red	Mixto
Zero Loss Zero Copy Datagram	- Tráfico de red	Mixto
iDB	- Tráfico de red	Mixto

Fig 34. Cuadro resumen funcionalidades Exadata y probables áreas de influencia en las mejoras de rendimiento descritas en las referencias oficiales

2.3.6 Críticas a Exadata

Exadata es un producto en evolución que, en pocos años, ha sufrido varios cambios. Además, al ser un producto asociado a la base de datos, a medida que vayan apareciendo nuevas versiones de la base de datos, probablemente, este hecho exija a Oracle esfuerzos adicionales para adaptar funcionalidades a su ejecución en Exadata. Por otra parte, la evolución del motor de base de datos hacia un funcionamiento explícito en nuevas arquitecturas distribuidas entre servidores de base de datos y servidores de almacenamiento es un camino que acaba de comenzar y que puede implicar falta de estabilidad del producto. Aún así las mayorías de las críticas a Exadata encontradas en Internet no vienen de este importante hecho sino de la adecuación de Exadata, su coste o la ineficacia de alguna de sus funcionalidades en determinados casos.

Las críticas a Exadata existentes en Internet provienen principalmente de:

- Inexistencia de *benchmarks* TPC-C o TPC-H publicados sobre Exadata[28]
- Exadata no adecuado para aplicaciones OLTP[28]
- Coste excesivo o menos económico que otras alternativas[29]
- Almacenamiento con carencia de funcionalidades comparado con vendedores de disco *Tier* 1[29]
- Vendor *lock-in*[30]
- Ineficacia de funcionalidades como *smart scan*
- Secretismo de Oracle con respecto al funcionamiento interno de alguna de las funcionalidades

En la bibliografía de este capítulo se han añadido algunos de los enlaces a comentarios y documentos críticos con Exadata para que el lector pueda entender también la visión crítica contra Exadata.

Fin de 2.3

Notas y Bibliografía de 2.3

1 No es objetivo de este TFG profundizar en funcionalidades también existentes fuera de Exadata.

2 En este capítulo se describen las configuraciones de *switchers* Infiniband y los protocolos utilizados. No es objetivo de este TFG explicar conceptos también existentes fuera de Exadata. Existen multitud de documentos en Internet que describen Infiniband. Un ejemplo:

Infiniband Wikipedia, http://en.wikipedia.org/wiki/InfiniBand **[Consulta Web], Fecha de Consulta : Abril 2013**

3 Datasheets Exadata X3-2 y X3-8, http://www.oracle.com/us/products/database/exadata/overview/index.html **[Consulta Web], Fecha de Consulta: Febrero y Marzo 2013.**

4 En las descripciones *Data Sheets* de las familias de Exadata no se especifica la memoria DRAM existente en cada servidor de almacenamiento.

5 Toda la información de este apartado se ha obtenido de la red técnica pública de Oracle (OTN):

Oracle Exadata Storage Software, http://www.oracle.com/technetwork/server-storage/engineered-systems/exadata/dbmachine-x3-twp-1867467.pdf Oracle Technical Network, {Consulta Web}, Fecha de Consulta : Abril 2013.

6 Consistencia Smart Scan: Garantía de que las operaciones SQL devuelven los registros según el modelo ACID Oracle.

7 No se encuentra de forma pública información sobre las características del algoritmo probabilístico de filtrado de joins.

8 Webinars Centroid, Centroid http://www.centroid.com/webinars/SmartScan.pdf [Consulta Web], Fecha de Consulta: Abril 2013.

9 Smart Flash Caché http://www.oracle.com/technetwork/server-storage/engineered-systems/exadata/exadata-smart-flash-cache-366203.pdf [Consulta Web], Fecha de Consulta: Abril 2013

10 Exadata Smart Flash Cache Technical White Paper, http://www.oracle.com/technetwork/server-storage/enginee...ystems/exadata/exadata-smart-flash-cache-366203.pdf [Consulta Web], Fecha de Consulta: Marzo 2013.

11 Exadata for OLTP, http://www.oracle.com/in/corporate/events/deliver-extreme-performance-by-oltp-1891253-en-in.pdf [Consulta Web] Fecha de Consulta: Abril 2013.

12 Automatic Storage Management. Funcionalidad proporcionada por la base de datos Oracle para el acceso y gestión directo de disco, sin necesidad de software adicional de gestión de volúmenes.

13 Redo logging, escrituras de bloques de log de memoria a disco para el mantenimiento de la consistencia de la base de datos.

14 Obtenidos de la Web de la empresa suiza Benchware, http://www.benchware.ch/benchmarks/ [Consulta Web] Fecha de Consulta: Marzo 2013.

15 En el test se utiliza la misma codificación para la numeración de la prueba

16 HCC, http://www.oracle.com/technetwork/server-storage/engineered-systems/exadata/ehcc-twp-131254.pdf [Consulta Web], Fecha de consulta: Abril de 2013

17 El archivado de logs de redo para copias de backup en caliente y recuperación total en caso de problemas de disco.

18 Un acceso vía query a dos o más columnas de una tabla puede requerir menos operaciones de entrada/salido si éstas se encuentran almacenadas contiguas dentro de cada registro, al formar parte del mismo bloque de datos.

19 No se especifica qué subconjunto de registros ni cómo se seleccionan. No se ha encontrado ningún otro documento público en Internet donde se describa esta característica.

20 No se encuentra información pública sobre las características del algoritmo de compresión de HCC.

21 Presentación Exadata, http://www.oracle.com/in/corporate/events/deliver-extreme-performance-by-oltp-1891253-en-in.pdf [Consulta Web] Fecha de Consulta: Abril 2013.

22 Kasibhotla, Diwakar (2011), *HCC*, OppenheimerFunds http://dwarehouse.files.wordpress.com/2011/05/exadata-hcc-case-study2.pdf [Consulta Web] Fecha de Consulta: Abril de 2013

23 Mirza, Fahd (2010), *Physical Disk, Cell Disk, Grid Disk & ASM Disk* http://www.pythian.com/blog/physical-disk-cell-disk-grid-disk-and-asm-disk-in-exadata/ [Consulta Web], Fecha de Consulta: Abril 2013.

24 Manual de Oracle, ASM, accedible públicamente*, Introduction to Automatic Storage Management (ASM)* *http://docs.oracle.com/cd/B28359_01/server.111/b31107/asmcon.htm* [Consulta Web] Fecha de Consulta: Marzo 2013.

25 Manual de Oracle, *Background Processes*, accedible públicamente. Oracle DB Server versión 11g *http://docs.oracle.com/cd/E14072_01/server.112/.../bg*processes*.htm* [Consulta Web] , Fecha de consulta: Abril 2013

26 Manual de Oracle, Using Infiniband Switches, accedible públicamente. *http://docs.oracle.com/cd/E18476_01/doc.220/e18478/leafswitch.htm* [Consulta Web], Fecha de Consulta: Abril 2013.

27 Mellanox Technologies, Comparative I/O Analysis*, InfiniBand Compared with PCI-X, Fiber Channel, Gigabit Ethernet, Storage over IP, HyperTransport, and RapidIO,* White Paper. http://www.mellanox.com/pdf/whitepapers/IOcompare_WP_140.pdf [Consulta Web], Fecha de Consulta: Abril 2013.

28 Davidian D., *Confronting Techno-deception*, IBM Blog, https://www-304.ibm.com/connections/blogs/davidian/tags/data?lang=en_us [Consulta Web], Fecha de Consulta: Junio 2013.

29 Floyer D., *The Limited Value of Oracle Exadata*, artículo en Wikibon http://wikibon.org/wiki/v/The_Limited_Value_of_Oracle_Exadata [Consulta Web] Fecha de Consulta: Junio 2013.

30 Mullins R., *Help! My vendor's Oracle and I'm locked-in*, Sep 2010, Artículo Networkworld http://robklopp.wordpress.com/2013/02/04/my-2-cents-oracle-exadata-1q2013 [Consulta Web] Fecha de Consulta: Junio 2013

2.4 Otros *appliances*

2.4.1 Introducción

Hasta ahora se han detallado las configuraciones de Exadata, se han resumido las mejoras observadas en las referencias oficiales de Exadata y presentadas por el Marketing de Oracle y se han descrito algunas de sus características técnicas más importantes que pueden estar directamente relacionadas con las mejoras observadas. Las plataformas que han sido sustituidas por Exadata en las referencias oficiales se han descrito como entornos OLTP, *Data Warehouse* o entornos mixtos y, en todos los casos, se trataba de clientes con bases de datos Oracle que han consolidado sus bases de datos en Exadata o han sustituido algunas de sus plataformas de bases de datos por Exadata para una función concreta, obteniendo un mayor rendimiento después del cambio. En el primer caso las alternativas de los clientes son seguir con la arquitectura *tradicional* para servidores de bases de datos, el capítulo 2.5 profundiza en las características económicas que diferencian seguir con la arquitectura tradicional o invertir en equipos Exadata. En el segundo caso existen otras alternativas de tipo *appliance*[1] que pueden permitir alcanzar mejoras de rendimiento con soluciones diferentes. Este capítulo 2.4 resume las características de otros *appliances*.

En el apartado 2.4.2 se identifican algunos de los diferentes tipos de *appliance* existentes actualmente, y se resumen sus características técnicas más importantes. Los casos a resumir son Teradata, Netezza/IBM PureData, Greenplum Data Computing Appliance de EMC y Fusion-io.

Dado que no existen *benchmarks* oficiales publicados y disponibles en Internet, de forma libre, entre los diferentes *appliances*, se hace difícil proporcionar una información adecuada sobre comparativas. La información existente es demasiado generalista, vaga, y poco empírica. Aunque el capítulo 2.5 muestra, desde el punto de vista económico, las características de los proyectos de inversión en Exadata comparadas con las arquitecturas tradicionales, no se incluyen en ese capítulo comparativas económicas con otros *appliances*. En el apartado 2.4.3 se identifican los componentes que han de ser comparados con Exadata desde un punto de vista económico sin entrar en valoraciones de diferencias de coste total de propiedad entre *appliances*.

2.4.2 Computer Appliances para bases de datos

Existen otros *appliances* en el mercado, de diferentes fabricantes *Hardware* y *Software*, orientados a la mejora del rendimiento con respecto a la arquitectura tradicional para la ejecución de bases de

datos. No es objetivo de este TFG profundizar en otros *appliances*, ni clasificarlos o compararlos entre sí, pero puede ser útil conocer, al menos, un resumen de sus características principales.

Teradata Data Warehouse Appliance

El *appliance* de Teradata para Data Warehouse[2] se compone de una familia de servidores *Hardware*, escalables y configurables en diferente número de *nodos*, software específico de Teradata, resaltando principalmente la base de datos relacional Teradata, discos atachados y *switching* de red.

Teradata se define como una arquitectura *Shared Nothing*[3], donde la base de datos, en la ejecución de las consultas, se distribuye en múltiples unidades de trabajo denominadas *Access Module Processors* (AMP) que tienen acceso exclusivo a una parte de los datos, CPU, memoria y entrada/salida. La distribución o descomposición de las consultas en unidades AMP la realiza un proceso específico de la base de datos denominado *Parsing Engine* (PE). Varios PE y AMPs pueden coexistir en el mismo nodo, ofreciendo un paralelismo *intranodo,* dado que cada nodo puede ser un equipo SMP.

Adicionalmente, la escalabilidad se consigue vía la conectividad de múltiples nodos entre sí, con una red propia denominada Bynet[4], de hasta 20GB por enlace, con conexiones bidireccionales nodo a nodo, en un único Data Warehouse, que Teradata define como *MPP*[5]. Teradata puede conectar entre sí hasta 4.096 nodos en alguna de sus configuraciones. Las diferentes configuraciones de la familia de Teradata, a diferencia de las configuraciones de 1/8 y hasta un equipo completo de las familias X3-2 y X3-8 de Exadata, están especializadas en un tipo concreto de ejecución, de mayor capacidad o volúmen de datos, o de mayor rendimiento y menor volumen.

El siguiente cuadro muestra un resumen de las configuraciones de su familia de *appliances*:

	560	1650	2690	4600	66xx
Modelo	Data Mart Appliance	Extreme Data Appliance	Data Warehouse Appliance	Extreme Performance Appliance	Active Enterprise Data Warehouse
Nodos	1 nodo con dos CPUs Intel Xeon de 6 cores cada CPU	Hasta 4.096 MPP nodos, cada nodo con 2 CPUs Intel Xeon de 6 cores por CPU	Cabina de 9 nodos con 2 CPUs Intel Xeon y 6 cores por CPU	Cabina de 9 nodos con 2 CPUs Intel Xeon y 6 cores por CPU	Hasta 4.096 MPP nodos, cada nodo con 2 CPUs Intel Xeon de 6 cores por CPU
Almacenamiento	Hasta 72TB en discos de 300GB o 600 GB	Hasta 186PB con discos de 1TB o 2 TB	Hasta 343 TB	Hasta 18TB	Hasta 92 PB
Memoria	48GB	48GB por nodo	96GB por nodo	96GB por nodo	96GB por nodo
Escalabilidad	Un nodo	Hasta 4.096 nodos	Hasta 45 nodos en 6 cabinas	Hasta 24 nodos	Hasta 4.096 nodos
Conectividad	NO	Bynet sobre 1GB Ethernet	Bynet sobre 1GB Ethernet	Bynet sobre 1GB Ethernet	Bynet v4

Fig 35. Cuadro resumen familia Teradata Data Warehouse Appliances[6]

Netezza y PureData de IBM[7]

El *appliance* de Netezza[8], dirigido también hacia los entornos *Data Warehouse*, se compone de servidores *Hardware*, *software* basado en PostgreSQL[9] y discos en un conjunto de configuraciones adaptables bajo una filosofía de no mover datos a no ser que sea imprescindible. Su arquitectura es una combinación de servidores SMP y MPPs, denominada por IBM MPP asimétrica, donde cada componente trabaja con múltiples hilos de ejecución filtrando la información requerida de la no útil tan pronto como sea posible. De estos componentes los más destacados son un tipo especial de *Field Programmable Gate Arrays*[10] (FPGAs) a los que IBM ha incorporado un software motor acelerador de ejecución de hilos denominado *Accelerated Stream* (FAST)

Los componentes de Netezza son los *Host* o servidores SMP en Activo-Pasivo para alta disponibilidad donde se ejecutan las instrucciones SQL, creando los planes de ejecución y donde se descomponen las instrucciones en subinstrucciones denominadas *snippets*, los *blades* ejecutores de *snippets* o S-Blades, compuestos a su vez de CPUs, FPGAs y memoria RAM, los discos que, al igual que en el caso Exadata pueden ser de alta capacidad o de alto rendimiento y la red entre los discos y los S-Blades que está adaptada a Netezza para ofrecer un alto rendimiento.

La información requerida viaja desde los discos a la memoria de los S-Blade, donde es cacheada con un algoritmo propio[11]. Esta información, al igual que en el resto de *appliances* puede estar comprimida. Los múltiples procesos FAST que se ejecutan en cada FPGA comprimen la información y filtran sólo la información requerida, donde las CPUs del S-Blade procesan para ejecutar su porción del *snippet* que es enviado a los equipos Host.

Los motores FAST se componen de compresores, con ratios que van de 4x a 8x, filtrado de registros y columnas similar al *Smart Scan Filtering Predicate y Column Filtering* de Exadata, y un componente denominado de “visibilidad” que permite mantener la consistencia de las SQL filtrando adicionalmente registros que no han de verse en las consultas porque aún no se ha ejecutado *commit* sobre ellos.

IBM ha mejorado y renombrado a Netezza con la familia PureData[12], donde se incorpora incluso la ejecución de funciones de *data mining* en el *appliance*. Las diferentes configuraciones de la familia N1001 se muestran en el siguiente cuadro:

IBM PureData System for Analytics N1001	Configuraciones 1 RACK (múltiples Racks disponibles)		
	N1001-002	N1001-005	N1001-010
S-Blades activos	4	7	14
Cores	32	56	112
Cores FPGA	32	56	112
TB disponibles para usuarios asumiendo compresión 4x	32	64	128

Fig 36. Cuadro resumen familia de IBM PureData Systems for Analytics N1001

Greenplum Data Computing Appliance para Bases de Datos, EMC[13]

En el caso de Greenplum de la empresa EMC, su Data Computing Appliance, también orientado a *Data Warehouse*, se compone de varias familias de productos configurables en módulos donde cada módulo tiene un conjunto de nodos servidores, disco, comunicaciones entre los nodos y el almacenamiento y el *software* especializado, basado en PostgreSQL. Su arquitectura es, al igual que en otros *appliances*, de tipo MPP *shared nothing*, donde un proceso maestro, denominado *Query Optimizer,* distribuye tareas entre segmentos independientes e incorpora en el optimizador basado en costes subplanes especiales para el *appliance*, con cálculos adicionales de costes de movimientos de planes entre *segmentos*. El siguiente diagrama muestra las características de los modulos de Greenplum Data Computing Appliance para bases de datos:

	Greenplum DB Compute	Greenplum DB Standard
Unidades Rack de cada módulo	8 Unidades (1/4 de rack)	8 Unidades (1/4 de Rack)
Software	Greenplum DB	Greenplum DB
Servidores por módulo	4	4
Cores por módulo	64	64
Memoria por módulo	256GB	256GB
Tipos de disco	Discos SAS de 300GB cada uno	Discos SAS de 900GB cada uno
Número de discos	96	96
Capacidad total física	9TB	27,5TB
Capacidad total con compresión	36TB	110TB

Fig 37. Cuadro resumen de las capacidades de cada módulo Greenplum Data Computing Appliance para Bases de Datos

La compañía EMC publica en el folleto[14] explicativo de las características de Greenplum Data Computing Appliance las capacidades de rendimiento y escalabilidad de 1 rack y de 12 racks interconectados, tanto en la familia Greenplum DB Compute como en la familia Greenplum DB Standard. El siguiente cuadro resume las capacidades, las velocidades de acceso y los ratios de carga:

	Greenplum DB Compute 1 RACK	Greenplum DB Compute 12 RACKS	Greenplum DB Standard 1 RACK	Greenplum DB Standard 12 RACKS
Número de módulos interconectados	4	48	4	48
Capacidad utilizable con compresión 4x	144TB	1728TB	440TB	5280TB
Velocidades de acceso en GB/Segundo	40	480	40	480
Velocidades de carga	16TB/hora	No especificado	16 TB/Hora	No especificado

Fig 38. Cuadro resumen de las capacidades de 1 y 12 RACKs interconectados de las familias Greenplum Data Base Compute y Standard. Escalabilidad lineal en accesos

Fusion-io Data Accelerator para Bases de Datos[15]

Los aceleradores para bases de datos de Fusion-io (Data Accelerator) son un caso diferente[16] dado que están orientados tanto a *Data Warehouse* como a OLTP y no son *shared nothing*. Estos aceleradores pueden considerarse un *appliance* pues se componen de *boxes*[17] de memoria *flash*, tarjetería de comunicaciones Fiber Channel, iSCSI e Infiniband y *software* especializado que se conectan con servidores de bases de datos y almacenamiento externo. Estos *boxes fusionan*[18] memoria y almacenamiento para las bases de datos, proporcionando mejoras, tanto en lectura como en escritura de bloques de datos, entre los procesos que se ejecutan en los servidores y el almacenamiento. Comparten memoria (y disco) entre múltiples servidores o nodos, permiten una mayor escalabilidad en el número de procesos de usuario que acceden a los datos y un mayor rendimiento.

El software especializado traslada peticiones de bloques de datos residentes en disco en direcciones del *array* de memoria *flash* NAND[19]

de Fusion-io, denominadas ioMemory, y viceversa. La traslación entre direcciones de bloques de datos e ioMemory se realiza en una capa denominada Virtual Storage Layer, y la función inversa se realiza con un software denominado *direct cache*.

El punto más interesante para este TFG de Fusion-io es que funciona y está certificado con bases de datos Oracle, RAC y ASM[20]. Además, la protección de los datos en *flash* se realiza con un RAID en memoria *flash* que permite que la arquitectura de máxima disponibilidad de Oracle basada en Dataguard también se ejecute en el *box* de memoria, lo que debe proporcionar unos rendimientos mejorados para protección ante desastres que la arquitectura tradicional[21].

La compañía Hewlett-Packard publica en su web el caso de su cliente Datalogix[22], empresa que proporciona soluciones para que organizaciones de marketing directo y captación de información para publicidad en Internet puedan incrementar la eficiencia de sus operaciones. En el caso se detalla el uso de Fusion-io, no como un *box* sino con tarjetas *flash* ioDrive añadidas a los servidores web y de base de datos. Para los servidores web, Datalogix requería una baja latencia de presentación de páginas, previo a la adquisición de Fusion-io se estaban ofreciendo páginas con una latencia de 50ms y con ioDrive esta latencia bajó a 12ms. Se instalaron 2 tarjetas por servidor. Para los servidores de base de datos se conectaron también dos tarjetas ioDrive por servidor con las siguientes características:

Servidores de Base de Datos	6x HP DL380G5
Base de Datos	Oracle 10g con RAC
Almacenamiento	SAN HP EVA 8000
Conectividad servidores de datos y almacenamiento	Fiber Channel
Tarjetas flash incluidas	2 ioDrive en cada servidor para alta disponibilidad
Almacenamiento en la memoria flash	Tablas frecuentemente accedidas y ficheros de log de la base de datos
Uso	Data Warehouse

Fig 39. Características servidores de datos del Data Warehouse de Datalogix

Los resultados obtenidos y publicados del caso Datalogix son los siguientes:

Ratio de aceleración de los servidores web	4x
Aceleración de consultas sobre tablas almacenadas en la memoria flash	Hasta 40x
Procesamiento de operativa Business Intelligence	6x
Escalabilidad	El mejor rendimiento permite menos consumo de CPU y deja más recursos para nuevas consultas
Otros Beneficios	Sencillez de implementación y gestión

Fig 40. Resumen de resultados obtenidos en Datalogix con Fusion ioDrive

Otros appliances

Existen otros *appliances* en el mercado como HANA de SAP, la combinación de HP y aceleradores SLC y MLC y otros equipos IBM PureData. Sus características principales son la de proporcionar un rendimiento adicional superior en la ejecución de aplicaciones de todo tipo comparados con las arquitecturas tradicionales.

2.4.3 *Appliances* hechos a mano y comparativas, no sólo es cuestión de rendimiento

Parece relativamente sencillo construir un *appliance* con elementos como tarjetas *Flash* o la inclusión de red *infiniband* entre los discos y los servidores de bases de datos, de tipo *blade*, con procesadores Intel y *open source* para el s.o. y la base de datos, más económicos que un *appliance* tipo Exadata y, probablemente, con buenos rendimientos dada la mejora en tráfico de red y la inclusión de la memoria *flash*. Aunque no existen o no he encontrado comparativas remarcables es probable que los costes de una arquitectura hecha a mano sean menores. Probablemente las dificultades de estas alternativas se visualicen en entornos con un alto volúmen de operaciones, usuarios y datos, con requerimientos de servicio 24x7 o necesidades de reduccion de ventana *batch* donde se requieran entornos sólidos desde el punto de vista del servicio y con el soporte y la garantía de un fabricante.

Por otra parte, no existen o no se han encontrado en Internet, de acceso público, comparativas entre los diferentes *appliances* que sean remarcables para este TFG, aunque cada fabricante expone, de manera similar a las referencias oficiales de Exadata, casos de clientes donde se han obtenido mejoras de rendimiento en accesos, carga de datos u otros

factores de mejora, como por ejemplo la escalabilidad de las diferentes configuraciones de las familias a las que pertenecen los *appliances*. Aunque pueden existir muchas razones por las que seleccionar un *appliance* u otro o construirse uno a medida, cada cliente puede optar por realizar pruebas de concepto, donde se visualicen los rendimientos concretos obtenidos en cada solución, realizar comparativas de costes entre los diferentes proyectos de inversión y seleccionar la solución adecuada teniendo en cuenta:

- La relación existente con el proveedor, la existencia de acuerdos previos o contratos de uso
- La fortaleza del proveedor
- La estabilidad del producto
- La base instalada, en casos comparables, del mismo sector
- Los conocimientos previos de la empresa u organización
- El soporte local y global
- Los conocimientos de la empresa

El capítulo 2.5 detalla los elementos a considerar, desde el punto de vista económico, en proyectos de inversión de Exadata comparado con la arquitectura tradicional. Sin embargo, cuando se comparan los costes de diferentes proyectos de inversión entre diferentes *appliances* o con soluciones hechas a mano hay que considerar, al menos:

- Costes de adquisición de las plataformas y previsión de nuevas inversiones futuras para asumir el crecimiento en función de la estrategia de cada organización.
- Costes de formación y adaptación del personal propio.
- Costes de servicios de terceros, como instalaciones, gestión de las plataformas, alojamiento en caso de *Housing, Hosting* y *Outsourcing*.
- Costes de mantenimientos anuales de las plataformas.
- Costes de gestión propios.
- Costes de adaptación y reprogramación asociados, como por ejemplo migraciones hacia 11g de las bases de datos en el caso Exadata.
- Costes de *software* del *appliance* y de *software* requerido de terceros.
- Costes de integración del *appliance* con el resto de sistemas de la organización.
- Costes de *facilities*, como adaptación de espacio en los *data centers*, energía suministrada y conectividad.
- Costes financieros.

Aunque para requerimientos similares se han de considerar los diferentes costes de cada solución, las negociaciones con cada proveedor pueden estar condicionadas a la existencia de una relación

previa, a contratos existentes, descuentos corporativos u otras circunstancias donde el coste no es el factor determinante.

Fin de 2.4

Notas y Bibliografía de 2.4

1 El término *appliance* se refiere a un conjunto de *hardware y software* preconfigurado para ejecutar una función específica. Una definición se puede encontrar en Wikipedia: *Computer Appliance, Wikipedia* http://en.wikipedia.org/wiki/Computer_appliance [Consulta Web] Fecha de Consulta, Abril 2013.

2 Toda la información de este apartado ha sido obtenida de enlaces de la web de Teradata, http://www.teradata.com

***Características de Teradata*, http://www.teradata.com [Consulta Web] Fecha de Consulta: Abril 2013**

***Higgins, D. Take a quantum leap with super-linear scalability. Artículo Teradata Magazine Online (2007)* http://apps.teradata.com/tdmo/v07n03/Viewpoints/WhyTeradata/QuantumLeap.aspx [Consulta Web] Fecha de Consulta: Abril 2013**

***Teradata Data Warehouse Appliance,* http://www.teradata.com/data-appliance/#tabbable=0&tab1=0&tab2=0&tab3=0 [Consulta Web] Fecha de Consulta: Abril 2013**

3 *Teradata is a shared nothing architecture, Coffing Data Warehousing web* http://www.coffingdw.com/Teradata_Basics/teradata_basics.htm#chapter_9___advanced_topics_that_you_will_be_tested_on/teradata_is_a_shared_nothing_architecture.htm [Consulta Web] Fecha de Consulta: Abril 2013

4 *Teradata Study Guide, Bynet* *http://lakshmikishore.blogspot.com.es/2010/09/bynet.html* *[Con*sulta Web], Fecha de Consulta: Abril 2013.

5 *At the core of the Teradata Platform*, http://apps.teradata.com/tdmo/v07n03/pdf/AR5387.pdf [Consulta Web] Fecha de Consulta: Abril 2013

6 *Krishnan, Krish, Sixth Sense Advisors, Inc,* The Teradata Data Warehouse Appliance, Technical Note on Teradata Data Warehouse Appliance vs Oracle Exadata http://www.teradata.com/white-papers/The-Teradata-Data-Warehouse-Appliance [Consulta Web] Fecha de Consulta: Abril 2013.

7 *IBM PureData Systems for Analytics N1001, powered by Netezza.* IBM Data Sheet (2012) http://public.dhe.ibm.com/common/ssi/ecm/en/imd14400usen/IMD14400USEN.PDF [Consulta Web] Fecha de Consulta: Abril 2013.

8 *IBM Netezza Data Warehouse Appliances*, http://www-01.ibm.com/software/data/netezza/ [Consulta Web] Fecha de Consulta: Abril 2013.

9 ***Pricket T.,*** **Netezza to bake analytics into appliances, artículo. The Register (Febrero 2010) http://www.theregister.co.uk/2010/02/24/netezza_data_analytics/ [Consulta Web] Fecha de Consulta: Abril 2013.**

10 ***The Netezza Data Appliance Architecture.*** **A Platform for High Performance Data Warehousing and Analytics. IBM Redbook (Junio2011) http://www.redbooks.ibm.com/redpapers/pdfs/redp4725.pdf [Consulta Web] Fecha de Consulta: Abril 2013.**

11 No se ha encontrado información sobre las características del algoritmo de compresión utilizado

12 ***Enzee Community de usuarios de Netezza*****.** ***Netezza is now PureData for Analytics*****, IBM, http://www.enzeecommunity.com/thread/3544 [Consulta Web] Fecha de Consulta: Abril 2013.**

13 ***Greenplum Data Computing Appliance*****, EMC http://www.greenplum.com/products/greenplum-dca [Consulta Web] Fecha de Consulta: Abril 2013.**

14 ***Greenplum Data Computing Appliance Unified Analytics Platform Edition, EMC http://www.greenplum.com/sites/default/files/2013_0129_dca_ds_1.pdf*** **[Consulta Web] Fecha de Consulta: Abril 2013.**

15 ***Fusion-io Database Accelerators*** **http://www.fusionio.com/database/ [Consulta Web] Fecha de Consulta: Abril 2013.**

ION Data Accelerator, DataSheet **http://www.fusionio.com/data-sheets/ion-data-accelerator-data-sheet/ [Consulta Web] Fecha de Consulta: Abril 2013.**

16 Dado que en los casos anteriores los appliances están orientados hacia sólo Data Warehouse y Fusion-io puede instalarse con ordenadores de múltiples fabricantes.

17 Entendiendo como b*oxes* a unidades o *appliances* empaquetados.

18 De ahí el nombre de Fusion-io.

19 Compton, B. Blog. Fusion-io Senior Director Product Management, New Extended Memory Software Makes NAND a Seamless Extension to DRAM http://www.fusionio.com/blog/new-extended-memory-software-makes-nand-an-alternative-to-dram/ [Consulta Web] Fecha de Consulta: Abril 2013.

20 ***Fusion Oracle Accelerator*** **http://www.fusionio.com/overviews/oracle-acceleration-with-fusion-io/ [Consulta Web] Fecha de Consulta: Abril 2013.**

21 Oracle recomienda una arquitectura denominada de máxima disponibilidad o *Maximum Availability Architecture* (MAA) donde no se realiza réplica de cabinas sino copia de logs entre centros de proceso de datos. En el caso de Fusion-io esta copia se realizaría de memoria *flash* a memoria *flash*, a mayor velocidad que de disco a disco.

Orenstein, G. **Blog Fusion-io SVP of Products Implementing Replication for Availability with Fusion-io http://www.fusionio.com/blog/implementing-replication-for-availability-with-fusion-io/ [Consulta Web] Fecha de Consulta: Abril 2013.**

2.5 Consideraciones Económicas para comparativas entre Exadata y la arquitectura tradicional

2.5.1 Introducción

En los capítulos anteriores se han expuesto las configuraciones y familias de Exadata, se han descrito las mejoras obtenidas por clientes de Oracle a través de sus referencias oficiales, se han identificado y detallado las características técnicas y funcionales más importantes de Exadata que podrían explicar las mejoras descritas por el marketing de Oracle y se han resumido las características de otros *appliances*. Con esta información es plausible poder comparar arquitecturas tradicionales de clientes de base de datos Oracle *versus* Exadata, caso por caso, pero es muy difícil poder establecer un mecanismo de comparación con otros fabricantes de *appliances* o similares dado que, tal y como se ha comentado en el capítulo 2.4, no hay *benchmarks* oficiales publicados ni información de pruebas de validación publicadas que, empíricamente, permitan establecer comparativas de rendimiento. No es el objetivo de esta memoria profundizar en las razones por las que no se publican *benchmarks*[1] de *appliances* pero puede ser útil describir, al menos desde el punto de vista económico, los elementos a considerar al comparar Exadata con la arquitectura tradicional.

En el apartado 2.5.2 se compara desde el punto de vista económico una situación *As Is,* basada en un cliente tipo con bases de datos Oracle susceptible de migrar a una situación *To Be* basada en una configuración tipo de Exadata para un proyecto de consolidación de bases de datos donde el objetivo es una reducción de costes de propiedad. En este caso no existirá un condicionante asociado a un requerimiento de un *mayor rendimiento* de las plataformas destino, sino que se realizará una comparación estrictamente económica. Dado que existen multitud de alternativas se utilizará un ejemplo sencillo, no real, que permita distinguir los condicionantes económicos existentes y entender cómo valorar proyectos de inversión de *appliances*.

En el apartado 2.5.3 se comparará un caso donde la funcionalidad del *appliance* puede ser determinante, como por ejemplo requerir una mejora del rendimiento de un *Data Warehouse* o atender a un crecimiento en el número de usuarios y/o procesos que han de satisfacerse. Este caso mostrará los elementos a considerar más allá de las comparativas de coste total de propiedad y se utilizará el caso de la empresa Robi Axiata, publicado por Oracle.

2.5.2 *As Is* versus *To Be* para proyectos de consolidación de bases de datos Oracle

Tal y como se ha descrito en el apartado anterior, probablemente la comparación más sencilla es la de valorar económicamente el no hacer nada *versus* migrar a Exadata, dado que en esta situación no se valoran las hipotéticas mejoras de rendimiento sino la comparación de costes entre plataformas.

El tipo cliente *As Is* seleccionado es el de una organización que utiliza varias bases de datos Oracle, con diferentes versiones, con servidores independientes o en cluster alojados en uno o varios centros de procesos de datos, uno o varios subsistemas de almacenamiento SAN o NAS[2] y con varios administradores DBAs y SAs[2]. El siguiente cuadro muestra las características de este cliente tipo:

Cliente Tipo	As Is
Sector	Cualquiera
Data Centers	Uno o varios, propios o en hosting/housing/outsourcing en un tercero
Bases de Datos	Varias BBDDs Oracle en diferentes versiones. Puede que decenas o centenares de bases de datos
Servidores de Base de Datos	Varios servidores, puede que decenas o centenares, de todo tipo
Subsistemas de almacenamiento	Uno o varios subsistemas de cualquier fabricante SAN o NAS
Administradores de Bases de Datos	Uno a Varios
Administradores de Almacenamiento	Uno o varios

Fig 41. Cuadro caraterísticas As Is

En cuanto a la valoración de los costes para la comparativa, en la situación *As Is*[4] se deberán contemplar las previsibles inversiones y los gastos anuales existentes y derivados de las nuevas inversiones, durante un periodo determinado de varios años, tanto si existen nuevos proyectos que incluyan más plataformas o simplemente se considera un refresco de equipamientos por obsolescencia. Las circunstancias de cada caso determinarán la complejidad de las previsiones de inversiones y gastos. En cuanto a las inversiones interesa conocer la previsible adquisición de servidores de bases de datos, la adquisición de software asociada a los nuevos servidores y las ampliaciones o nuevas inversiones en sistemas de almacenamiento. En cuanto a los gastos anuales interesa conocer los costes de mantenimiento *hardware* y *software*, los costes de servicio y los costes de *facilities*. El siguiente cuadro muestra estos conceptos:

Partida	Inversiones y Gastos As Is
Servidores	Sustitución de los servidores existentes por nuevos servidores por reducción de obsolescencia
Servidores	Adquisición de nuevos servidores previstas por nuevos proyectos (Por ejemplo, implementación de una nueva aplicación)
Almacenamiento	Adquisición o sustitución de cabinas, ampliación de las existentes o costes TB/Año asignados a las bases de datos consolidables
Software de bases de datos	Licencias previstas de adquirir en base al crecimiento de nuevos proyectos
Mantenimientos	Mantenimientos anuales de los equipos
Mantenimientos	Mantenimientos anuales del software
Costes de Gestión propios	Costes propios de gestión de las plataformas
Costes de Gestión propios	Costes de administración de las bases de datos
Costes de Gestión propios	Costes de gestión del almacenamiento
Costes de terceros	Si existen, costes de Housing/Housing/Outsourcing de la plataforma a consolidar
Costes de terceros	Si existen, Costes de gestión
Facilities	Costes de espacio anuales, switching/routers de equipos a consolidar y energía

Fig 42. Partidas de Inversión y Gasto As Is de las plataformas consolidables en Exadata

Las partidas de inversión y gasto de la consolidación de bases de datos Oracle en Exadata deben incluir los costes esperados en el mismo periodo de adquisición, costes de *software* adicional, mantenimientos, instalaciones, formación al personal, integración, *facilities*, gestión, migraciones y costes financieros detallados en el apartado 2.4.3. Adicionalmente, en la comparativa con la situación *As Is*, puede darse el caso de que se requiera incluir como partida la recompra de activos pues algunos de ellos pueden no haberse amortizado aún. La comparativa de costes *As Is* versus *To Be*[5] en el periodo seleccionado mostrará el interés por la realización del proyecto de inversión. En todo caso, las características más importantes asociadas a los posibles ahorros de Exadata con respecto a la situación *As Is* descrita son:

Partida	Ahorros principales en Exadata
Almacenamiento	Almacenamiento: Exadata incluye discos por lo que puede producirse una importante reducción de costes por este concepto
Reducción de Licencias Software	Un coste importante es el coste de adquisición de software Oracle por core. Si se reducen el número de cores que, a futuro, el cliente tipo pueda requerir puede producirse una importante reducción de costes por este concepto
Reducción de costes de Mantenimiento	Un coste importante para cada cliente es el coste de mantenimiento software. Depende del tipo de contrato existente es posible una reducción de costes en este concepto
Reducción de costes de servidores As Is	Aunque actualmente los costes de adquisición de equipos es más económica, la suma de mantenimientos anuales y los costes de adquisición pueden ser importantes en la situación As Is
Costes de Gestión	Sustituir decenas o centenares de equipos por uno o varios Exadatas completos tiene un impacto directo en la reducción de costes de gestión, sobretodo si estos costes están asociados al número de activos como en el caso de Outsourcing, Housing o Hosting
Costes de software de terceros	Software habitual como los Volume Managers pueden ser eliminados dado que Exadata incluye vía ASM la gestión de los volúmenes
Facilities	Estos costes cada vez más importantes pueden ser reducidos

Fig 43. Hipótesis de los ahorros más importantes de Exadata versus la arquitectura tradicional

Otras consideraciones en la comparativa con las arquitecturas tradicionales en la consolidación de plataformas:

- Cada caso depende del grado de consolidación propuesto por el equipo Exadata y de la situación prevista para cada cliente en crecimiento de servidores y licencias requeridas asociadas a los nuevos servidores. Un cliente tipo que no vaya a crecer en número de servidores o en licencias de *software* puede encontrar beneficios económicos en una migración a Exadata sólo con el ahorro de costes de almacenamiento, gestión y *facilities*.
- Si se considera Exadata como plataforma de bases de datos para bases de datos no Oracle hay que incluir los costes de migración a Oracle y formación del personal

Caso sencillo *tipo*, ejemplo de consolidación

As Is

Para facilitar la comprensión de las características de los proyectos de inversión para consolidacion de servidores en Exadata utilizaré un ejemplo, el caso de la empresa A[6]. La empresa A dispone de 40 servidores de bases de datos Oracle, en diferentes versiones, con diferente grado de obsolescencia. La empresa A amortiza sus equipos en 4 años y sustituye los mismos cada 5 años[7], siendo la previsión de sustitución de servidores de bases de datos Oracle para los próximos años la siguiente:

	Año 1	Año 2	Año 3	Año 4	Año 5
Sustitución de servidores por obsolescencia	5	4	10	5	16

Fig 44. Caso de ejemplo, no real. Sustitución de servidores por obsolescencia

Hardware, S.O., Storage y Facilities

La empresa A, en la sustitución de servidores de bases de datos, planea adquirir equipos de bajo coste *Blades* Linux de 2 CPUs y 8 cores por CPU cada equipo. Actualmente, puede adquirir esos equipos con tarjetería de comunicaciones incluida por 4.500€ cada servidor. Adicionalmente, puede hacer uso de un sistema operativo de Linux por 200€ por servidor y año. Los 3 primeros años desde el momento de la adquisición tienen el coste de mantenimiento incluidos, y a partir del 4º año los costes de mantenimiento de cada servidor se ha estimado en 300 €. El total del espacio ocupado por las bases de datos, incluidos copia de seguridad para alta disponibilidad y protección ante desastres es de 50TB, y se utilizan dos cabinas de almacenamiento *Tier* 1 SAN para estas bases de datos[8]. La empresa A ha estimado un coste anual por TB de 3.500€ para los próximos 5 años, incluyendo en estos costes adquisiciones, *drivers* y mantenimientos.

Los servidores actuales de base de datos de la empresa A le suponen un coste de mantenimiento *hardware* anual de 8.000€ por servidor. Estos equipos, de diferentes fabricantes y características, ocupan junto con los sistemas de almacenamiento 10 RACKs completos de 40 unidades por RACK. Además, se encuentran alojados en un proveedor externo que les factura 10€ por unidad de RACK al mes por el espacio en sus CPDs y 1 € por Kw/hora de consumo energético[9]. Se prevee que con los nuevos equipos Blades Linux no existirá un cambio en estos costes de espacio. Los equipos actuales, incluyendo los sistemas de almacenamiento, suman una potencia de 70Kw, están funcionando todos los días del año, a todas horas y los nuevos equipos son un 50% más eficientes en términos energéticos. El siguiente cuadro muestra la previsión de inversiones y gastos en la situación *As Is* para los conceptos *Hardware, Storage y Facilities*:

	Año 1	Año 2	Año 3	Año 4	Año 5
Sustitución de servidores por obsolescencia	5	4	10	5	16
Costes de Adquisición de nuevos Servidores	22.500	18.000	45.000	22.500	72.000
Costes de s.o.	1.000	1.800	3.800	4.800	8.000
Costes de Mantenimientos de nuevos Servidores (a partir del 4º año)				1.500	2.700
Costes de Mantenimientos de los antiguos servidores	280.000	248.000	168.000	128.000	0
Costes de Storage	175.000	175.000	175.000	175.000	175.000
Costes Alojamiento Espacio	48.000	48.000	48.000	48.000	48.000
Costes Energía	65.625	62.125	53.375	49.000	35.000
Total Inversiones y Costes previstos Hwd As Is	592.130	552.929	493.185	428.805	340.716

Fig 45. Caso de ejemplo, no real. Inversiones y Gastos previstos Hardware, Storage y Facilities

Servicios y Migraciones

La empresa A dispone de un proveedor externo que le gestiona sus sistemas, a 5.000€ por servidor y año, y no espera un cambio de costes con su estrategia de *Blades* Linux. Este proveedor externo también factura por la administración de las Bases de Datos en 100.000€ anuales totales. Por otra parte, cada sustitución de un equipo antiguo por un nuevo *Blade* está previsto que incluya una modernización del *software* de base de datos, en proyectos de corta duración, siendo su coste de 20.000€ por servidor[10]. Suponemos que cada servidor ejecuta una base de datos Oracle y que todas necesitarán modernizarse a una versión actualizada. La previsión de costes de servicios en *As Is* sería:

	Año 1	Año 2	Año 3	Año 4	Año 5
Sustitución de servidores por obsolescencia	5	4	10	5	16
Costes de Administración Hardware	200.000	200.000	200.000	200.000	200.000
Costes de Administración Bases de Datos	100.000	100.000	100.000	100.000	100.000
Costes de instalaciones y migraciones	100.000	80.000	200.000	100.000	320.000
Total Costes de Servicios y Migraciones	400.005	380.004	500.010	400.005	620.016

Fig 46. Caso de ejemplo, no real. Inversiones y Gastos previstos Servicios y Migraciones

Software

La empresa A tiene un contrato con Oracle que le facilita hacer uso de licencias por *core* de bases de datos y de RAC medido en base al número de *cores* existentes. Actualmente ese contrato le permite utilizar el número de *cores* actuales pero la estrategia *As Is* implica un crecimiento en el número de *cores*. Por simplificación se supone que cada nuevo equipo Linux Blade requiere adquirir 1 licencia adicional de *software* Oracle, a 30.000 € por licencia[11] y un 22% de mantenimiento anual adicional por licencia del que ahora se dispone:

	Año 1	Año 2	Año 3	Año 4	Año 5
Sustitución de servidores por obsolescencia	5	4	10	5	16
Costes de Licencias Oracle	150.000	120.000	300.000	150.000	480.000
Costes de Mantenimiento adicional software Oracle	33.000	59.400	125.400	158.400	264.000
Total Costes de Licencias Oracle Software y Mantenimiento en *As Is*	183.005	179.404	425.410	308.405	744.016

Fig 47. Caso de ejemplo, no real. Inversiones y Gastos previstos Licenciamiento Software

El resumen de inversiones y gastos en la alternativa As Is queda:

Total Inversiones y Costes previstos Hwd As Is	526.505	490.804	439.810	379.805	305.716
Total Costes de Servicios y Migraciones	400.005	380.004	500.010	400.005	620.016
Total Costes de Licencias Oracle Software y Mantenimiento en As Is	183.005	179.404	425.410	308.405	744.016
Total Inversiones y Gastos As IS	1.109.515	1.050.212	1.365.230	1.088.215	1.669.748

Fig 48. Caso de ejemplo, no real. Inversiones y Gastos previstos en As Is

To Be

La empresa A determina analizar una alternativa a la previsible situación *As Is* con *appliances* Exadatas para consolidar sus bases de datos. Para ello realiza pruebas de concepto y obtiene un diseño alternativo de su arquitectura tecnológica basado en dos equipos de tipo "X" de Exadata, cuyo coste de adquisición total, incluyendo *hardware* Exadata y *software* de Exadata Storage Software es de 2.000.000€[12]. El coste de mantenimiento anual de este *hardware* y *software* es de 180.000€. Con estos equipos no se requieren adquirir licencias de *software* de base de datos adicionales. El coste anual estimado de alojamiento y energía de los equipos Exadata es de 40.000 €[13] anuales, el coste de migraciones e instalación se estima en 300.000€. Por el hecho de sustituir los 40 servidores se requiere incluir como pérdidas los equipos aún no amortizados, estimándose que estas pérdidas se deben reflejar el primer año con un importe total de 200.000€. Adicionalmente, se estiman los mismos costes de servicios para la administración de bases de datos que en la alternativa *As Is*, de 100.000€[14] por año y la administración de los equipos Exadata se estima en 35.000€ por año. El siguiente cuadro resume las inversiones y gastos en la alternativa Exadatas o *To Be*:

	Año 1	Año 2	Año 3	Año 4	Año 5
Adquisición Hardware / Software Exadata	2.000.000				
Instalaciones y Migraciones	300.000				
Costes de Mantenimiento	180.000	180.000	180.000	180.000	180.000
Costes de Energía y Espacio	40.000	40.000	40.000	40.000	40.000
Pérdidas por Amortización	200.000				
Administración de bases de datos	100.000	100.000	100.000	100.000	100.000
Administración de Exadatas	35.000	35.000	35.000	35.000	35.000
Total Costes inversiones y Gastos To Be	2.520.000	220.000	220.000	220.000	220.000

Fig 49. Caso de ejemplo, no real. Inversiones y Gastos previstos en To Be con Exadatas

Indicadores proyecto de inversión *To Be* vs *As Is*

La empresa A estima un coste de capital del 10%, por lo que el flujo de caja en la comparativa entre las dos alternativas para la empresa A y el valor actual neto[15] de la alternativa de Exadatas queda:

	Año 1	Año 2	Año 3	Año 4	Año 5
Beneficios To Be (Exadatas)	-1.410.485	830.212	1.145.230	868.215	1.449.748

Coste del Capital	**10%**
Valor Actual Neto a 5 Años	1.757.477

Fig 50. Caso de ejemplo, no real. Flujo de Caja To Be (Exadatas) a 5 años y VAN, coste de capital al 10%

La comparativa de costes de propiedad entre *As Is* y *To Be* (Exadatas) a 5 años queda:

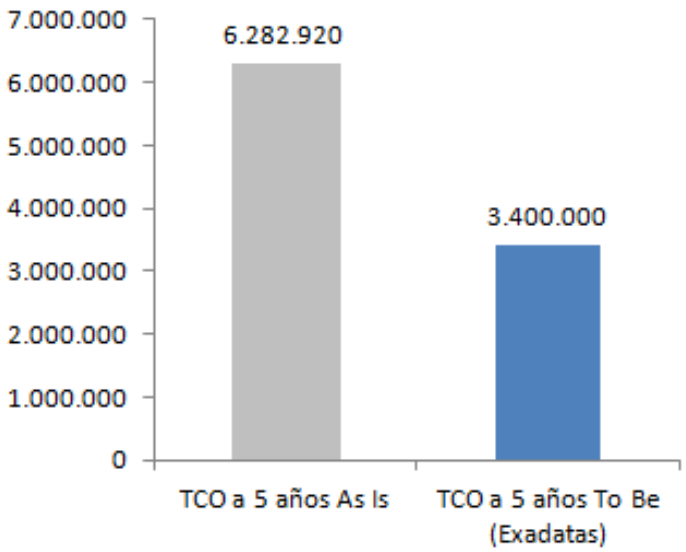

Fig 51. Caso de ejemplo, no real. Comparativa de TCOs

Todos los costes han sido *inventados* con la idea de facilitar la comprensión de las características de los proyectos de inversión para consolidación de servidores de bases de datos y permitir a un posible lector de este TFG realizar sus propios cálculos. Pueden existir otras partidas aquí no incluidas por simplificación.

2.5.3 As Is versus To Be para proyectos donde la funcionalidad aportada es un beneficio tangible

Otros casos donde la consolidación de bases de datos no sea un requerimiento del cliente *tipo* pueden estar asociados a una mejora del rendimiento de la plataforma. Este es el caso, por ejemplo, de un *Data Warehouse*, donde otros *appliances* pueden ser competencia directa de Exadata. Para este tipo de casos hay que considerar en la evaluación del proyecto de inversión la medida económica proporcionada por el rendimiento extra.

Caso Robi Axiata publicado por Oracle[16]

Robi Axiata es una empresa de Telecomunicaciones con sede en Bangladesh, proveedora de servicios GSM, que pertenece al grupo Axiata desde 1997. Axiata Group tiene 168 Millones de clientes en 10 países, mientras que Robi Axiata tiene 15,2 Millones de clientes. El interés principal de Robi es el de atraer y retener clientes, diferenciándose de su competencia en la calidad del servicio y de producto y no en precio. Para ejecutar esta estrategia Robi diseñó y puso en marcha una infraestructura de *Data Warehouse* con el objetivo de analizar las características de sus clientes potenciales y de los activos, dirigir productos específicos hacia los clientes que les podrían aportar un mayor beneficio y responder con rapidez a los cambios del mercado. Su servicio e infraestructura de *Data Warehouse* se fue quedando pequeña, dado que incluía el análisis de más de 100 millones de CDRs(*Call Detail Records* o detalles de llamadas) diarios y estas CDRs diarias fueron creciendo con los años. Esta infraestructura incluía 3 servidores HP con bases de datos Oracle 9i y diversas herramientas de carga de datos y análisis. Robi cambió a 1/2 Exadata X2-2 y se observaron las mejoras típicas ya descritas en el capítulo 2.2 de cargas más rápidas y mejoras del rendimiento en consultas.

La empresa Mainstay Partners realizó un estudio donde se calculaban los beneficios del proyecto, que incluía ahorros de IT, en línea con las características de los ahorros mostrados en 2.5.2 en cuanto a los capítulos de ahorro en costes evitados de *hardware, software* y almacenamiento. Lo más importante, para este apartado, es que Mainstay Partners también estimó los beneficios funcionales proporcionados por mejoras en el *reporting*, la productividad de los analistas y la mejora del conocimiento de la rentabilidad de sus posibles clientes. Los beneficios totales estimados fueron de 2,1M$ en 3 años, con un retorno de la inversión del 60% menor a 1 año, componiéndose estos beneficios de los ahorros IT, valorados en 800.000$ en 3 años y en ahorros funcionales de 1,3M$ en 3 años. Los ahorros funcionales estaban calculados en base a la mejora de la productividad de los 50 analistas de Robi, que ejecutaban una media de 6.000 consultas por día y donde se esperaba un ahorro de 3 horas diarias por analista gracias al

1/2 Exadata. Este ahorro suponía poder dedicar más tiempo a realizar análisis de datos u otras tareas, lo que fue valorado por Mainstay Partners en 1,3M$ de ahorro en 3 años.

Fin de 2.5

Notas y Bibliografía de 2.5

1 No se han encontrado casos de acceso público en Internet.

2 Network Attached Storage.

3 System Administrators.

4 Por convención, en comparativas entre nuevos proyectos y seguir sin cambios se suele denominar a la segunda opción "*As Is*"

5 Denominación de la alternativa de nuevo proyecto.

6 Es una empresa que no existe. Se utiliza en este ejemplo para permitir al lector del TFG la correspondencia entre partidas de costes y su distribución anual.

7 Amortización es un concepto contable y sustitución por reducción de obsolescencia es la previsión de sustitución del equipo o vida útil. No tienen porqué coincidir ambos conceptos.

8 Una forma sencilla de asignar costes globales de Storage. Se suele asignar un coste por TB o GB anual a un *Tier* diferente de Disco. Los discos más económicos serían los de *Tiering* más alto y los menos económicos serían *Tier* 1.

9 La fórmula de cálculo es muy sencilla, conociendo el coste medio de Kw/h. Si los sistemas van a estar funcionando las 24 horas del día, todos los días del año:

Coste de Energía *Power*= 24x365xCoste Kw/hxKw de Servicio

Una vez identificado este coste hay que añadir el coste de la energía de los equipos de *Cooling.* Una forma sencilla de hacerlo es utilizar un ratio adicional sobre el consumo de los equipos. Una empresa eficiente energéticamente puede tener un ratio de entre 1,5 y 1,8 veces la energía consumida.

Por lo que Coste Total: RatioxCoste de Energía *Power*

Miller R.,* Uptime Institute. *The* Average PUE is 1.8. Datacenter Knowledge (Mayo 2011) *http://www.datacenterknowledge.com/archives/2011/05/10/uptime-institute-the-average-pue-is-1-8/
[Consulta Web] Fecha de Consulta: Abril 2013.

10 Incluido en costes *As Is* para simplificar el ejemplo. Puede ser más adecuado incluir sólo los costes de modernización de Base de Datos en la alternativa *To Be* dado que Exadata obliga a una versión específica de la Base de datos (ver capítulo 2.1) y la alternativa As Is podría seguir funcionando con las versiones actuales de la empresa A.

11 Costes completamente inventados.

12 Exadata Hardware y Exadata Storage Software se licencian por separado.

13 ½ Exadata tiene un uso de potencia típico de 4,7Kw, según especificaciones del folleto del producto y depende del tipo de disco incluido.

Datasheets Exadata X3-2 y X3-8, **http://www.oracle.com/us/products/database/exadata/overview/index.html [Consulta Web], Fecha de Consulta: Febrero y Marzo 2013.**

14 El coste de administración de la base de datos no varía dado que suele ser habitual que los proveedores de servicios cobren un cargo por base de datos, aunque habría que considerar los ahorros que supone el que algunas funciones de administración se simplifiquen dado que el parcheado se realiza sobre un único punto.

15 Por simplificación, se calcula el valor actual neto con un coste de capital determinado sobre el proyecto a 5 años.

16 Robi Axiata, Reference Booklet Casos de clientes Oracle en Exadata, Caso Robi Axiata http://www.oracle.com/us/products/database/exadata-reference-booklet-400018.pdf [Consulta Web] Fecha de Consulta: Febrero y Marzo 2013.

3. Conclusiones

El *appliance* Exadata para la ejecución de bases de datos de Oracle está suponiendo una revolución tecnológica con respecto a las arquitecturas tradicionales de los servidores de bases de datos, con unas mejoras en rendimientos para aplicaciones tipo OLTP, *Data Warehouse* y mixtas y capacidades de consolidación, que pueden permitir reducir los costes de propiedad de los servidores de bases de datos y mejorar la productividad de los usuarios de las aplicaciones que acceden a las mismas. En esta memoria se han revisado las mejoras de rendimiento descritas en las referencias oficiales de Oracle y se han resumido las novedades tecnológicas que, previsiblemente, sean la causa de las mejoras. Adicionalmente, se han identificado características similares en otros *appliances* para la ejecución de bases de datos, con novedades tecnológicas que minimizan el acceso a disco, cuello de botella tradicional de las bases de datos. Ha concluido esta memoria con un ejemplo de comparativa de costes entre Exadata y la arquitectura tradicional cuyo objetivo principal es el de permitir al lector identificar los capítulos de análisis de costes asociados a los proyectos de inversión de Exadata.

Desde mi punto de vista las novedades tecnológicas encontradas en Exadata y en los *appliances* descritos han abierto un camino de evolución que está ahora comenzando. Creo que las arquitecturas de bases de datos evolucionarán para minimizar cuellos de botella, con una integración cada vez mayor del *software* con el *hardware*, y esta evolución tendrá éxito a medida que los costes ofrecidos con los *appliances* sean menores que las arquitecturas tradicionales. Sin embargo, existe el importante reto de mantener dos versiones de producto de base de datos, una con las mejoras integradas en *appliances* y otra dando servicio en arquitecturas tradicionales.

No ha sido posible ejecutar un test propio sobre Exadata, ni era el objetivo de esta memoria. Mi idea inicial era la de obtener los resultados publicados de pruebas de rendimiento de Exadata, y de acceso libre en Internet, para poder, empíricamente, relacionar adecuadamente las novedades del *appliance* con los resultados, basándome, al menos, en datos proporcionados por terceros. Lo que he encontrado es lo que se ha utilizado y esta información no ha podido ser contrastada con otras fuentes.

El plan de trabajo seguido ha consistido en recopilar, analizar y depurar información de libre disposición en Internet sobre las materias de estudio para construir esta memoria, tareas a las que he dedicado cerca de 270 horas. Una mayor dedicación habría permitido construir una visión más global, incluyendo información adicional y profundizando en la información expuesta, pero la visión sobre Exadata obtenida es suficiente, desde mi punto de vista, para entender Exadata. Por otra

parte, la mayor dificultad encontrada ha sido la de intentar separar la información puramente de marketing, obtenida en las *webs* de los fabricantes, y seleccionar adecuadamente las fuentes de información. No ha sido fácil.

Han quedado pendiente temas importantes como el describir la integración de Exadata en el ecosistema del resto de servidores de una organización o la integración de Exadata en las arquitecturas de máxima disponibilidad. También me hubiese gustado poder profundizar en las implicaciones de Exadata en la consolidación de roles de administración, dado que los *appliances*, en apariencia, pueden ser administrados por menos personas comparadas con la arquitectura tradicional. Otro estudio podría profundizar en estos aspectos.

4. Glosario

Los términos más importantes han sido descritos en las notas de cada capítulo. Los acrónimos más importantes utilizados son:

AMP : Access Module Processor

ASM : Automatic Storage Management

BTU : British Thermal Unit

CBP : Customs And Border Protection Agency

CDRs : Call Detail Records

DBAs : Database Administrators

DBRM : Database Resource Management

DDL : Data Definition Language

DMA : Direct Memory Access

FAST : Field Programmable Gate Arrays Accelerated Streams

FPGA : Field Programmable Gate Arrays

HCA : Host Channel Adapter

HCC : Hybrid Columnar Compression

iDB : Protocolo entre base de datos y storage Server

IORM : Input Output Resource Management

IOPS : IOs per Second

IPoIB : Internet Protocol over Infiniband

iSCSI : Internet SCSI

IT : Information Technology

MLC : Multi-Level Cell

MPP : Massively Parallel Processors

MS : Management Service

NAS : Network Attached Storage

OLTP : Online Transaction Processing

PCI : Peripheral Component Interconnect

PE : Parsing Engine

RAC : Real Application Cluster

RAID : Redundant Array of Inexpensive Disks

RDS : Reliable Datagram Sockets

rpm : Revoluciones por Minuto

RS : Restart Service

SAN : Storage Area Network

SAs : System Administrators

SLC : Single-Level Cell

SMP : Symmetric Multiprocessors

SMON : Oracle System Monitor

TPC-C : Transaction Processing Council, test OLTP

TPC-H : Transaction Processing Council, test Data Warehouse

QDR : Quad Data Rate

ZDP : Zero Loss Zero Copy Datagram

5. Bibliografía

Por orden alfabético

Benchware, Exadata, *Web de la empresa suiza Benchware*, http://www.benchware.ch/benchmarks/ [Consulta Web] Fecha de Consulta: Marzo 2013. *Referenciado en Capítulo 2.3., nota 14.*

Bynet. *Teradata Study Guide, Bynet http://lakshmikishore.blogspot.com.es/2010/09/bynet.html [Con*sulta Web], Fecha de Consulta: Abril 2013. *Referenciado en Capítulo 2.4., nota 4.*

Características de Teradata, http://www.teradata.com [Consulta Web] Fecha de Consulta: Abril 2013. Referenciado en Capítulo 2.4., nota 2.

Caso Exadata de Turkcell, *Turkcell Accelerates Reporting Tenfold, Saves on Storage and Energy Costs with Exadata Database Machine*, http://www.oracle.com/us/corporate/customers/turkcell-1-exadata-case-study-456284.pdf [Consulta Web] Fecha de Consulta: Marzo 2013. *Referenciado en Capítulo 2.2., nota 10.*

Casos publicados, *Búsqueda de Casos de Clientes Exadata,* http://www.oracle.com/search/customers [Consulta Web], Fecha de Consulta: Marzo 2013. *Referenciado en Capítulo 2.2., nota 3.*

Clientes Exadata*, Búsqueda de Casos de Clientes Exadata en la web pública de la compañía Oracle,* http://www.oracle.com/search/customers [*Consulta Web], Fecha de Consulta: Marzo de 2013. Referenciado en Capítulo 2.1., nota 2.*

Chris Kanaracus, IDG (2012) *Oracle Exadata gains certification for SAP applications* http://www.pcworld.com/article/230175/article.html *[Consulta Web] Fecha de Consulta: Marzo 2013. Referenciado en Capítulo 2.1., nota 11.*

Coffin Data Warehouse*, Teradata is a shared nothing architecture,* http://www.coffingdw.com/Teradata_Basics/teradata_basics.htm#chapter_9___advanced_topics_that_you_will_be_tested_on/teradata_is_a_shared_nothing_architecture.htm [Consulta Web] Fecha de Consulta: Abril 2013. *Referenciado en Capítulo 2.4., nota 3*

Compton, B. Blog. Fusion-io Senior Director Product Management, New Extended Memory Software Makes NAND a Seamless Extension to DRAM http://www.fusionio.com/blog/new-extended-memory-software-makes-nand-an-alternative-to-dram/ [Consulta Web] Fecha de Consulta: Abril 2013. *Referenciado en Capítulo 2.4., nota 19.*

Computer Appliance, Wikipedia http://en.wikipedia.org/wiki/Computer_appliance [Consulta Web] Fecha de Consulta: Marzo 2013. *Referenciado en Capítulo 2.2., nota 9 y en capítulo 2.4, nota 1*

Davidian D., *Confronting Techno-deception*, IBM Blog, https://www-304.ibm.com/connections/blogs/davidian/tags/data?lang=en_us [Consulta Web], Fecha de Consulta: Junio 2013.

Enzee*, Enzee Community de usuarios de Netezza*. *Netezza is now PureData for Analytics*, IBM, http://www.enzeecommunity.com/thread/3544 [Consulta Web] Fecha de Consulta: Abril 2013. *Referenciado en Capítulo 2.4., nota 12.*

Exadata for OLTP, http://www.oracle.com/in/corporate/events/deliver-extreme-performance-by-oltp-1891253-en-in.pdf [Consulta Web] Fecha de Consulta: Abril 2013. *Referenciado en Capítulo 2.3., nota 11.*

Exadata Smart Flash Cache Technical White Paper, http://www.oracle.com/technetwork/server-storage/enginee...ystems/exadata/exadata-smart-flash-cache-366203.pdf [Consulta Web], Fecha de Consulta: Marzo 2013. *Referenciado en Capítulo 2.3., notas 9 y 10.*

Flash Memory, *Wikipedia* http://en.wikipedia.org/wiki/Flash_memory [Consulta Web] Fecha de Consulta: Marzo 2013. *Referenciado en Capítulo 2.2., nota 13*

Floyer D., *The Limited Value of Oracle Exadata*, artículo en Wikibon http://wikibon.org/wiki/v/The_Limited_Value_of_Oracle_Exadata [Consulta Web] Fecha de Consulta: Junio 2013.

Folletos de producto públicos de Exadata. *DataSheets families X3-2 & X3-8,* http://www.oracle.com/us/products/database/exadata/overview/index.html [Consulta Web], Fecha de Consulta: Marzo 2013. *Referenciado en Capítulo 2.1., nota 1 y en capítulo 2.5 nota 13.*

Folletos Exadata, *Datasheets Exadata X2-2, X2-8, X3-2 y X3-8,* http://www.oracle.com/us/products/database/exadata/overview/index.html [Consulta Web], Fecha de Consulta: Marzo de 2013. *Referenciado en Capítulo 2.2., nota 12 y en el capítulo 2.3, nota 3*

Fusion-io Database Accelerators http://www.fusionio.com/database/ [Consulta Web] Fecha de Consulta: Abril 2013. *Referenciado en Capítulo 2.4., nota 15.*

Fusion Oracle Accelerator http://www.fusionio.com/overviews/oracle-acceleration-with-fusion-io/ [Consulta Web] Fecha de Consulta: Abril 2013. *Referenciado en Capítulo 2.4., nota 20.*

Greenplum Data Computing Appliance, EMC http://www.greenplum.com/products/greenplum-dca [Consulta Web] Fecha de Consulta: Abril 2013. *Referenciado en Capítulo 2.4., nota 13.*

Greenplum Data Computing Appliance Unified Analytics Platform Edition*, EMC http://www.greenplum.com/sites/default/files/2013_0129_dca_ds_1.pdf* [Consulta Web] Fecha de Consulta: Abril 2013. *Referenciado en Capítulo 2.4., nota 14.*

HCC, http://www.oracle.com/technetwork/server-storage/engineered-systems/exadata/ehcc-twp-131254.pdf [Consulta Web], Fecha de consulta: Abril de 2013. *Referenciado en Capítulo 2.3., nota 16.*

Higgins, D. *Take a quantum leap with super-linear scalability. Artículo Teradata Magazine Online (2007)* http://apps.teradata.com/tdmo/v07n03/Viewpoints/WhyTeradata/QuantumLeap.aspx [Consulta Web] Fecha de Consulta: Abril 2013. *Referenciado en Capítulo 2.4., nota 2.*

IBM Netezza Data Warehouse Appliances, http://www-01.ibm.com/software/data/netezza/ [Consulta Web] Fecha de Consulta: Abril 2013. *Referenciado en Capítulo 2.4., nota 8.*

Infiniband Wikipedia, http://en.wikipedia.org/wiki/InfiniBand **[Consulta Web], Fecha de Consulta : Abril 2013.** ***Referenciado en Capítulo 2.3., nota 2.***

IBM PureData Systems for Analytics N1001*, powered by Netezza.* IBM Data Sheet (2012) http://public.dhe.ibm.com/common/ssi/ecm/en/imd14400usen/IMD14400USEN.PDF [Consulta Web] Fecha de Consulta: Abril 2013. ***Referenciado en Capítulo 2.4., nota 7.***

IBM*, The Netezza Data Appliance Architecture.* A Platform for High Performance Data Warehousing and Analytics. IBM Redbook (Junio2011) http://www.redbooks.ibm.com/redpapers/pdfs/redp4725.pdf [Consulta Web] Fecha de Consulta: Abril 2013. ***Referenciado en Capítulo 2.4., nota 10.***

ION Data Accelerator, *DataSheet* http://www.fusionio.com/data-sheets/ion-data-accelerator-data-sheet/ [Consulta Web] Fecha de Consulta: Abril 2013. ***Referenciado en Capítulo 2.4., nota 15.***

Kasibhotla, Diwakar, (2011), *HCC*, OppenheimerFunds http://dwarehouse.files.wordpress.com/2011/05/exadata-hcc-case-study2.pdf [Consulta Web] Fecha de Consulta: Abril de 2013. ***Referenciado en Capítulo 2.3., nota 22.***

Kenneth M. Richhart, *Transforming Information Technology Services: Reducing Cost and Improving Availability*. CBP Agency, Department of Homeland Security, http://www.actgov.org/knowledgebank/documentsandpresentations/Documents/Program%20Events/Executive%20Session%20featuring%20Ken%20Ritchhart,%20CBP%201-12-11.pdf [Consulta Web] Fecha de Consulta: Marzo 2013. ***Referenciado en Capítulo 2.2., nota 8.***

Krishnan, Krish*, Sixth Sense Advisors, Inc,* The Teradata Data Warehouse Appliance, Technical Note on Teradata Data Warehouse Appliance vs Oracle Exadata http://www.teradata.com/white-papers/The-Teradata-Data-Warehouse-Appliance [Consulta Web] Fecha de Consulta: Abril 2013. ***Referenciado en Capítulo 2.4., nota 6.***

Manual de Oracle de Automatic Storage Management, ASM. Accesible públicamente*, Introduction to Automatic Storage Management (ASM) http://docs.oracle.com/cd/B28359_01/server.111/b31107/asmcon.htm* [Consulta Web] Fecha de Consulta: Marzo 2013. ***Referenciado en Capítulo 2.3., nota 24.***

Manual de Oracle, *Background Processes*, accesible públicamente. Oracle DB Server versión 11g *http://docs.oracle.com/cd/E14072_01/server.112/.../bg*processes.*htm* [Consulta Web] , Fecha de consulta: Abril 2013. ***Referenciado en Capítulo 2.3., nota 25.***

Manual de Oracle, Using Infiniband Switches, accedible públicamente. *http://docs.oracle.com/cd/E18476_01/doc.220/e18478/leafswitch.htm* [Consulta Web], Fecha de Consulta: Abril 2013. ***Referenciado en Capítulo 2.3., nota 26.***

Mellanox Technologies, Comparative I/O Analysis*, InfiniBand Compared with PCI-X, Fiber Channel, Gigabit Ethernet, Storage over IP, HyperTransport, and RapidIO,* White Paper. http://www.mellanox.com/pdf/whitepapers/IOcompare_WP_140.pdf [Consulta Web], Fecha de Consulta: Abril 2013. ***Referenciado en Capítulo 2.3., nota 27.***

Miller R., Uptime Institute. *The Average PUE is 1.8.* Datacenter Knowledge (Mayo 2011) *http://www.datacenterknowledge.com/archives/2011/05/10/uptime-institute-the-average-pue-is-1-8/*
[Consulta Web] Fecha de Consulta: Abril 2013. *Referenciado en Capítulo 2.5., nota 9.*

Mirza, Fahd (2010), *Physical Disk, Cell Disk, Grid Disk & ASM Disk* http://www.pythian.com/blog/physical-disk-cell-disk-grid-disk-and-asm-disk-in-exadata/ [Consulta Web], Fecha de Consulta: Abril 2013. *Referenciado en Capítulo 2.3., nota 23.*

Mullins R., *Help! My vendor's Oracle and I'm locked-in*, Sep 2010, Artículo Networkworld http://robklopp.wordpress.com/2013/02/04/my-2-cents-oracle-exadata-1q2013 [Consulta Web] Fecha de Consulta: Junio 2013

Oracle, Nota de la Web de Oracle*: Customers worldwide adopt Exadata Database Machine* http://www.oracle.com/us/solutions/datawarehousing/exadata-customers-421304.html [Consulta Web], Fecha de Consulta: Marzo de 2013. *Referenciado en Capítulo 2.2., nota 1.*

Oracle Exadata Database Machine Technical Case Study: *Garmin International Inc, White Paper* http://www.oracle.com/technetwork/database/availability/garmin-1667151.pdf [Consulta Web] Fecha de Consulta: Marzo 2013. *Referenciado en Capítulo 2.2., nota 11.*

Oracle Exadata Storage Software, Oracle Technical Network(OTN) http://www.oracle.com/technetwork/server-storage/engineered-systems/exadata/dbmachine-x3-twp-1867467.pdf [Consulta Web], Fecha de Consulta : Abril 2013. *Referenciado en Capítulo 2.3., nota 5.*

Orenstein, G. Blog Fusion-io SVP of Products. *Implementing Replication for Availability with Fusion-io* http://www.fusionio.com/blog/implementing-replication-for-availability-with-fusion-io/ [Consulta Web] Fecha de Consulta: Abril 2013. *Referenciado en Capítulo 2.4., nota 21.*

Paul Kent, *Big Data SAS (2012) SAS Scoring Accelerator for Oracle and Beyond, presentación realizada y publicada en el* SAS Global Forum 2012 http://www.oracle.com/technetwork/database/.../sas/385-2012-1608576.pdf [Consulta Web] Fecha de Consulta: Marzo 2013. *Referenciado en Capítulo 2.2., nota 14.*

Performance Exadata, http://www.oracle.com/in/corporate/events/deliver-extreme-performance-by-oltp-1891253-en-in.pdf [Consulta Web] Fecha de Consulta: Abril 2013. *Referenciado en Capítulo 2.3., nota 21*

Pricket T., Netezza to bake analytics into appliances, artículo. The Register (Febrero 2010) http://www.theregister.co.uk/2010/02/24/netezza_data_analytics/ [Consulta Web] Fecha de Consulta: Abril 2013. *Referenciado en Capítulo 2.4., nota 9.*

Quad Data Rate, *Infiniband,* http://en.wikipedia.org/wiki/InfiniBand [Consulta Web] Fecha de Consulta: Marzo de 2013. *Referenciado en Capítulo 2.1., nota 6.*

Reference *Booklet* público de Oracle sobre Casos de clientes de Exadata, http://www.oracle.com/us/products/database/exadata-reference-booklet-

400018.pdf [Consulta Web] Fecha de Consulta: Marzo de 2013. *Referenciado en Capítulo 2.2., nota 2.*

Robi Axiata, Reference Booklet Casos de clientes Oracle en Exadata, Caso Robi Axiata http://www.oracle.com/us/products/database/exadata-reference-booklet-400018.pdf [Consulta Web] Fecha de Consulta: Febrero y Marzo 2013. *Referenciado en Capítulo 2.5., nota 16.*

Sun Microsystems, *Wikipedia,* http://es.wikipedia.org/wiki/Sun_Microsystems [Consulta Web] Fecha de Consulta: Marzo 2013. *Referenciado en Capítulo 2.1., nota 9.*

Teradata, *At the core of the Teradata Platform*, http://apps.teradata.com/tdmo/v07n03/pdf/AR5387.pdf [Consulta Web] Fecha de Consulta: Abril 2013. *Referenciado en Capítulo 2.4., nota 5.*

Teradata Data Warehouse Appliance, http://www.teradata.com/data-appliance/#tabbable=0&tab1=0&tab2=0&tab3=0 [Consulta Web] Fecha de Consulta: Abril 2013. *Referenciado en Capítulo 2.4., nota 2.*

Webinars Centroid, Centroid http://www.centroid.com/webinars/SmartScan.pdf [Consulta Web], Fecha de Consulta: Abril 2013. *Referenciado en Capítulo 2.3., nota 8.*