  
MAESTRIA PROFESIONAL EN TECNOLOGÍA DE BASES DE DATOS

Curso: MBD-801 Proyecto de Investigación Aplicada 1

**Evaluación de bases de datos en memoria dentro de ambientes tolerantes a fallos utilizando un clúster de alta disponibilidad de ordenadores de placas reducidas**

Estudiantes:

Douglas Murillo

Bryan Quiros S

Profesor: Prof. Luis Naranjo Zeledón

Agosto, 2015

# Tabla de Contenido

[Tabla de Contenido 2](#_Toc427156253)

[Resumen Ejecutivo 3](#_Toc427156254)

[Palabras clave 3](#_Toc427156255)

[Antecedentes 4](#_Toc427156256)

[Planteamiento del problema 5](#_Toc427156257)

[Justificación 6](#_Toc427156258)

[Objetivos 7](#_Toc427156259)

[Taxonomía 7](#_Toc427156260)

[Objetivo General 7](#_Toc427156261)

[Objetivos Específicos 7](#_Toc427156262)

[Alcances 8](#_Toc427156263)

[Limitaciones 8](#_Toc427156264)

[Marco Institucional 9](#_Toc427156265)

[Estado de la cuestión 10](#_Toc427156266)

[Ordenadores de placas reducidas (OPRs) 10](#_Toc427156267)

[Clúster de Alta Disponibilidad 11](#_Toc427156268)

[Base de Datos en Tiempo Real 12](#_Toc427156269)

[Marco Metodológico 14](#_Toc427156270)

[Características del Hardware 14](#_Toc427156271)

[Características del Software 15](#_Toc427156272)

[Dinámica de evaluación 15](#_Toc427156273)

[Tipos de Evaluaciones 15](#_Toc427156274)

[Apéndice 16](#_Toc427156275)

[Cronograma 16](#_Toc427156276)

[Referencias 20](#_Toc427156277)

# Resumen Ejecutivo

Las nuevas tecnologías obligan a una constante actualización de los sistemas y los planes de aplicación de productos tecnológicos para la solución de problemas y el mejoramiento del desempeño de los sistemas. Este trabajo pretende crear una guía de evaluación de tres motores de bases de datos en un escenario muy particular: La utilización de un clúster de alta disponibilidad utilizando hardware de bajo costo y una bases de datos con soluciones en memoria. La idea de utilizar y evaluar el desempeño de esta configuración particular se propone como el punto de partida para la utilización a gran escala de esta combinación de hardware y software como solución alternativa a soluciones propietarias de infraestructura, de alto costo y subutilizadas. Este trabajo explora las ventajas y desventajas de cada elemento involucrado en la configuración (clúster de alta disponibilidad, base de datos en tiempo real y ordenadores de placas reducidas), desarrolla una guía de buenas prácticas, recomendaciones y presenta los resultados de una evaluación basada en los tiempos de consultas a la base de datos.

Se pretende definir una base simple para la eventual implementación de este tipo de configuraciones como soluciones empresariales, de alto desempeño y disponibilidad para resolver problemas específicos y críticos para la organización.

# Palabras clave

Bases de datos en memoria, MariaDB, Altibase, memSQL, VoltDB, ordenadores de placa reducida, desempeño, consultas, clúster, alta disponibilidad, tolerancia a fallos

# Antecedentes

No existen estudios previos paralelos acerca de una implementación similar. Es necesario tomar en cuenta que el escenario de evaluación es muy específico y las variables no corresponden a configuraciones populares de sistemas empresariales o soluciones difundidas comercialmente. Los 3 conceptos principales: Ordenadores de Placas Reducidas (OPR), Bases de Datos en Memoria (BDM) y Clústeres de Alta Disponibilidad (CAD) son todas tecnologías de reciente concepción dentro de la era computacional.

# Planteamiento del problema

Las bases de datos en memoria (BDM) poseen características muy importantes para minimizar la latencia al realizar consultas y procesar los datos. Sin embargo presentan una desventaja inherente a la tecnología en memoria, la volatilidad. Los datos almacenados en memoria dependen de energía y por ende, refrescamiento continuo. Sin niveles de energía suficientes los datos se pierden y se corrompen.

Existen soluciones de memoria no volátil, que básicamente, conectan una batería recargable a los módulos de memoria. En el evento que el sistema pierda la energía, la batería seguirá proveyendo al controlador de la energía necesaria para mantener los datos en memoria mientras se restablece la fuente principal.

Como alternativa de solución a este problema se encuentran los clústeres de alta disponibilidad. Sin embargo este concepto está asociado a configuraciones de múltiples servidores y tecnologías de alto costo de inversión inicial y mantenimiento.

Este proyecto pretende brindar una solución que permita la implementación de una BDM en un clúster de alta disponibilidad, contrarrestando el riesgo de perder los datos si uno de los nodos falla. Así mismo, y de la mano con la evaluación se realizarán mediciones del desempeño en el procesamiento de los datos, concretamente, en la elaboración de consultas y el tiempo de respuesta.

Tanto los resultados como las experiencias derivadas de la evaluación serán recopilados como parte del proyecto.

# Justificación

Este proyecto de evaluación pretende unificar tres conceptos distintos y de amplio uso en soluciones tecnológicas relacionados con bases de datos y servir como fundamento para desarrollar una solución general que cuente con las ventajas de los elementos que la componen.

Existe una necesidad creciente por sistemas computacionales que representen un menor costo de inversión inicial y mantenimiento. Las necesidades energéticas de un OPR son hasta un 80% menor que las de un sistema convencional. La configuración de clúster en Alta Disponibilidad brinda una seguridad adicional para la replicación de datos, sobretodo en soluciones donde la información es almacenada en memoria volátil.

El precedente de este proyecto radica en utilizar OPR como una solución formal y alternativa a la infraestructura de servidores común.

Si bien es cierto que ninguna de las tecnologías aplicadas es nueva, la configuración potencial abre las puertas de una nueva opción para que micro, pequeñas y medianas empresas u organizaciones puedan adoptar los beneficios de bases de datos en memoria sin incurrir en grandes costos de inversión por adquisición de infraestructura, renovación o mantenimiento.

Este proyecto sienta la base para incursionar en aplicaciones aún más complejas (tales como las operaciones en tiempo real) con mayor cantidad de nodos y en arquitecturas alternativas tales como SPARC o ARM.

# Objetivos

## Taxonomía

Para la definición de objetivos se utilizó la taxonomía de Bloom Revisada (Anderson & Krathwohl, 2001) por ser una versión más moderna y con cambios para la aplicación de un modelo cognitivo en Tecnologías de Información.

## Objetivo General

Evaluar el desempeño de consultas de tres bases de datos en memoria en un clúster de alta disponibilidad creado a partir de Ordenadores de Placas Reducidas

## Objetivos Específicos

* Documentar en una guía el proceso de elaboración del clúster de alta disponibilidad y la creación de las bases de datos en memoria
* Describir los parámetros comparativos a utilizar para la evaluación del desempeño.
* Implementar las configuraciones de hardware a utilizar para la conformación de los clústeres de prueba.
* Investigar posibles fuentes de datos de prueba y tipos de consultas que puedan generar los resultados esperados.
* Proponer las pruebas y mediciones del desempeño de las base de datos en tiempo real de acuerdo a la configuración y consulta aplicada.
* Contrastar los resultados obtenidos para cada una de las bases de datos evaluadas

# Alcances

* Se documentará una base de conocimiento para la aplicación práctica de BDTR en clústeres de alta disponibilidad en distintos escenarios.
* Formulación de un modelo de proyección de desempeño para cada configuración basado en la cantidad de nodos.
* Se redactará una guía que contenga las mejores prácticas de configuración de hardware para cada uno de los escenarios de prueba.
* Se recopilarán las lecciones aprendidas y limitaciones de software a través de las distintas configuraciones.

# Limitaciones

* Se utilizarán un total de 4 nodos (OPRs) para las pruebas.
* Solo el tiempo para completar la consulta será el factor para medir el desempeño.
* La cantidad de memoria y la capacidad del procesador estarán definidas por el modelo de la OPR seleccionada.
* Solo se utilizarán versiones de prueba o libres del software

# Marco Institucional

Este proyecto es desarrollado y financiado por los ingenieros Douglas Murillo Arias y Bryan Quiros Santana como Trabajo Final de Graduación (TFG). El TFG se encuentra descrito como requisito para optar por el grado académico de Maestría en Tecnologías de Bases de Datos impartido por la Universidad CENFOTEC en Costa Rica.

Otras leyes de la República de Costa Rica, normativas instituciones que regulan la educación superior y organismos a los que se encuentre asociados la Universidad aplican.

# Estado de la cuestión

Tal y como se expuso en los antecedentes no existe estudios previos que soporten una configuración similar. Por lo que, alternativamente, se propone explorar las investigaciones con respecto a los diferentes componentes que conforman el escenario de evaluación.

## Ordenadores de placas reducidas (OPRs)

Los ordenadores de placas reducidas (OPRs) son computadoras completas construidas en una sola placa de circuitos, con procesador, memoria y otras características que requieren los sistemas para funcionar como un computador. Se diferencian de un ordenador común en que sus placas no cuentan con puertos de expansión. (Burckle 2011)

Los diseños y aplicaciones de las OPRs varían según su objetivo. Desde simples sistemas que se utilizan como controladores hasta sistemas especializados que funcionan en conjuntos y configuraciones que permiten ahorrar espacio.

El primer OPR apareció a mediados de 1976 y desde entonces su desarrollo ha venido avanzando, enfocado en aplicaciones específicas y un amplio rango de soluciones comerciales, desde el hogar hasta empresariales. (Winn 1999)

A pesar de su gran utilidad, la industria ha limitado el desarrollo de los OPRs como soluciones empresariales exceptuando el campo de productos embebidos. Las últimas publicaciones sobre clústeres de OPRs se limitan a soluciones con procesadores basados en en la arquitectura RISC (en español, Computador con Conjunto de Instrucciones Reducidas), siendo más una exploración hacia los posibles resultados de utilizar nuevas tecnologías y no el tradicional CISC (en español, Computador con Conjunto de Instrucciones Complejas) del cual se derivan la mayor parte de soluciones de software empresarial. (Kiepert 2013)

Si bien es cierto, podría implementarse un emulador de CISC en un sistema construido con procesadores RISC, el cambio en arquitectura y emulación, sin mencionar los posibles fallos por incompatibilidad crearían una brecha comparativa que resultaría impráctica de analizar.

Los últimos avances en OPRs auguran un nuevo impulso por parte del mayor fabricante de procesadores de la arquitectura CISC, Intel. Con la presentación de la línea de OPRs Next Unit Computing (en español, Próxima Unidad de Computación) la compañía espera sobrepasar a las soluciones ARM (RISC) presentes en el mercado (Altavilla 2014).

## Clúster de Alta Disponibilidad

Un clúster de alta disponibilidad (CAD) es un grupo de computadoras interconectadas con el fin de minimizar el tiempo no operacional mediante la creación de grupos con recursos redundantes. Lógicamente, se pueden ver 1 o más sistemas dependiendo de la configuración pero su objetivo es el mismo, garantizar la continuidad de una aplicación a pesar de una falla en el equipo físico. (Bader 1996)

Esta configuración de clúster es indicada para aplicaciones críticas donde los fallos del hardware ponen en riesgo la continuidad del sistema y las operaciones. Uno de los campos donde más se utilizan es en las bases de datos.

Las diferentes configuraciones en grupos nodales o relaciones entre los sistemas permiten una flexibilidad distributiva para ambientes de ejecución. Así mismo, favorece la migración y expansión en los sistemas sin poner en riesgo la continuidad operativa del software. (Winn 1999)

Uno de los factores que facilitan la implementación de Ordenadores de Placas Reducidas (OPRs) en configuraciones de CAD es su configuración simple y la disminución en el consumo total de energía. Es imperativo considerar que la razón principal de un CAD es considerar el fallo eventual del equipo y garantizar la continuidad de la ejecución por lo que un remplazo del sistema completo facilita el restablecimiento del clúster. Sumado a las características de bajo costo de los OPR, el costo de mantenimiento se reducirá considerablemente comparado con soluciones empresariales. (Centos.org 2007) (Cubieboard.org 2013) (Kiepert 2013)

Bases de Datos en Memoria

Una base de datos en memoria (BDM) es un sistema de gestión de base de datos que se basa principalmente en la memoria principal para el almacenamiento de datos informáticos. Se pone en contraste con los sistemas de gestión de base de datos que emplean un mecanismo de almacenamiento en disco. Estos sistemas son más rápidos que las bases de datos de disco ya que los algoritmos de optimización internos son más simples y menos costos al ejecutar instrucciones de la CPU. El acceso a los datos en la memoria disminuye el tiempo de búsqueda cuando se consultan los datos, lo que proporciona un mayor rendimiento comparado con tecnologías de acceso al disco. (Vizard 2012)

En aplicaciones donde el tiempo de respuesta es crítico, tales como equipos de redes de telecomunicaciones y redes de publicidad móvil, a menudo se usan las bases de datos de la memoria principal. Las BDM han ganado mucha popularidad, especialmente en el espacio de análisis de datos, a partir de mediados de la década de 2000, debido principalmente a la disminución de costos. (Khan 2012)

TimesTen de Oracle es una BDM con persistencia y capacidad de recuperación. Originalmente diseñado e implementado en los laboratorios de Hewlett–Packard, TimesTen se creó en 1996 y fue adquirida por Oracle Corporation en el 2005.

Todos los datos dentro de una base de datos TimesTen se encuentra en la memoria física (RAM), lo que significa que no se requieren operaciones en el disco. TimesTen proporciona a las aplicaciones tiempos de respuesta cortos y un muy alto rendimiento en cargas de trabajo intensivas hacia la base de datos. (Oracle)

Como la memoria es mucho más rápida que el disco duro, TimesTen es muy usada en aplicaciones donde los acuerdos de nivel de servicio requieren tiempos muy bajos y de respuesta predecible, tales como las telecomunicaciones, aplicaciones de comercio de servicios financieros en tiempo real, equipos de red y aplicaciones web de gran tamaño. Otra característica fundamental es que funciona como una base de datos relacional con todas las funciones de acceso utilizando las API estándar como ODBC, JDBC, etc. y proporciona la funcionalidad del lenguaje de consulta SQL. (Oracle)

## Base de Datos en Tiempo Real

Las bases de datos en tiempo real (BDTR) es un sistema de base de datos que utiliza el procesamiento en tiempo real para manejar cargas de trabajo cuyo estado está en constante cambio. Utilizan las limitaciones de tiempo que representan un cierto rango de valores para los que los datos son válidos. A esto se le llama validez temporal. Una base de datos convencional no puede trabajar en estas circunstancias porque las inconsistencias entre los objetos del mundo real y los datos que los representan son incompatibles para realizar modificaciones sencillas. Un sistema efectivo debe ser capaz de manejar las preguntas sensibles dentro del rango de tiempo, devolver sólo los datos válidos temporalmente, y apoyar la programación de prioridad. Para introducir los datos en los registros, a menudo un sensor o un dispositivo de entrada controla el estado del sistema físico y actualiza la base de datos con información nueva para reflejar el sistema físico con más precisión. En el diseño de un sistema de base de datos en tiempo real, se debe considerar la forma de representar el tiempo válido, cómo los hechos se relacionan con el sistema en tiempo real. (Oracle)

# Marco Metodológico

Dada la naturaleza empírica de la investigación propuesta y de acuerdo a Dobles, Zúñiga y García (1998) en donde se asume que la fuente del conocimiento proviene de este tipo de naturaleza, se puede afirmar que esta investigación tiene carácter positivista; por lo tanto se propone una investigación de tipo evaluativa con un alcance descriptivo fundamentado en:

* La evaluación de múltiples escenarios de prueba: Estos permitirán recopilar gran cantidad de datos que posteriormente serán agregados y presentados para respaldar las conclusiones de la evaluación.
* Generación de datos a partir de funciones aleatorias: Lo que permite una mayor objetividad en las pruebas y resultados, los datos aleatorios facilitan el análisis de un conjunto de datos que no siguen un patrón y puedan verse afectados por algoritmos de detección de tendencias.
* La evaluación propuesta se basa en datos cuantitativos lo que permite respaldar las conclusiones con gráficos y cuadros comparativos. La utilización de sistemas de computación y la carencia de otros estudios similares generan la necesidad de establecer medidas que sirvan de referencia a futuras investigaciones dentro de la misma configuración.

## Características del Hardware

* 4 Nodos Foxconn AT-5250
  + Intel Atom D2550 @ 1.86GHz
  + 2GB DDR3 @ 1066
  + 4GB USB 2.0
* 1 Router DLINK DIR-610
* 1 Toshiba S855-S5254
* Cable UTP Cat 6

## Características del Software

* Sistema Operativo: RedHat Enterprise Linux 5 (32 bits)
* Software de clúster: Clusterware versión 11.2.0.1.0 para Linux x86 (32 bits)
* Base de datos: Timesten versión 11.2.2.6.0 para Linux x86 (32 bits)
* DTM Data Generator
* DataMaker
* Oracle Timesten® In-memory Database 11g Monitorin Plug-in For Oracle Enterprise Manager (Herramienta de evaluación)

## Dinámica de evaluación

1. Se generarán datos de pruebas utilizando las herramientas DTM Data Generator y DataMaker
2. Los datos se alimentarán a la configuración nodal directamente o a través de un equipo externo.
3. Se medirá el tiempo que consume los procedimientos sobre los datos según el rango de refrescamiento elegido (1, 0.5, 0.1 y 0.01 segundos)
4. Recopilación y procesamiento de datos.

## Tipos de Evaluaciones

* Desempeño comparativo en diferente configuración nodal
* Comparación entre configuraciones ante la pérdida de uno o más nodos
* Comparación entre configuraciones ante el restablecimiento de uno o más nodos
* Tiempo de comunicación entre clúster lógico y computador externo según la configuración nodal
* Evaluación de diferentes esquemas de replicación:
  + Clusterware
  + Timesten
  + Sistema operativo

# Apéndice

## Cronograma

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Identificación** | **Nombre** | **Tiempo**  **(en horas)** | **Fecha de inicio** | **Fecha de fin** |
| 1 | **Anteproyecto** | | | |
| 1.1 | Revisión de anteproyecto | 3 | 24/02 | 25/02 |
| 1.2 | Ejecutar cambios de contenido | 5 | 26/02 | 27/02 |
| 1.3 | Revisión de cambios | 2 | 28/02 | 28/02 |
| 1.4 | Revisión de estructura | 2 | 01/03 | 01/03 |
| 1.5 | Revisión con tutor de anteproyecto | N.A. | N.D. | N.D. |
| 1.6 | Presentación del anteproyecto oficial ante la universidad | N.A. | N.D. | N.D. |
| 1.7 | Revisión y posible aplicación de cambios por parte de la universidad | 3 | N.D. | N.D. |
| 2 | **Coordinación de lectores y tutores** | | | |
| 2.1 | Recopilar información sobre tutores y lectores disponibles | 4 | 15/03 | 22/03 |
| 2.2 | Contactar a los candidatos | 2 | 24/03 | 28/03 |
| 2.3 | Enviar el anteproyecto a los interesados | 1 | 24/03 | 28/03 |
| 2.4 | Discutir, revisar y documentar las observaciones | 6 | 24/03 | 04/04 |
| 2.5 | Coordinar las reuniones siguientes | 2 | 24/03 | 04/04 |
| 3 | **Hardware** | | | |
| 3.1 | Adquirir el hardware necesario según lo descrito en el marco metodológico | 360 | 31/03 | 21/04 |
| 3.2 | Verificar la conectividad de los puertos de red y el estado funcional de los nodos | 4 | 21/04 | 25/04 |
| 3.3 | Documentación de los resultados y procesos involucrados | 3 | 21/04 | 25/04 |
| 3.4 | Revisión y retroalimentación por parte de los lectores / tutores | N.A. | 28/04 | 02/05 |
| 3.5 | Corrección de documento | 2 | 02/05 | 05/05 |
| 4 | **Software** | | | |
| 4.1 | Instalar el sistema operativo en los nodos | 25 | 05/05 | 16/05 |
| 4.2 | Pruebas de verificación de conectividad entre los nodos y el enrutador | 15 | 10/05 | 23/05 |
| 4.3 | Configurar el clúster de alta disponibilidad | 4 | 16/05 | 23/05 |
| 4.4 | Verificar la conectividad entre los nodos, la capacidad ante fallos y la comunicación con el equipo externo | 2 | 23/05 | 23/05 |
| 4.5 | Instalar la base de datos | 8 | 24/05 | 28/05 |
| 4.6 | Configurar la base de datos | 16 | 24/05 | 28/05 |
| 4.7 | Instalar el software de generación de datos | 2 | 28/05 | 30/05 |
| 4.8 | Configurar el envío de datos automatizados a la base de datos según los periodos definidos | 16 | 29/05 | 04/06 |
| 4.9 | Instalar y configurar Clusterware | 8 | 30/05 | 08/06 |
| 4.10 | Documentación de los resultados y procesos involucrados | 6 | 08/06 | 11/06 |
| 4.11 | Revisión y retroalimentación por parte de los lectores / tutores | N.A. | 11/06 | 16/06 |
| 4.12 | Corrección de documento | 3 | 16/06 | 18/06 |
| 5 | **Pruebas** | | | |
| 5.1 | Pruebas de desempeño y recopilación de resultados con el clúster de alta disponibilidad al 100% | 8 | 19/06 | 23/06 |
| 5.2 | Pruebas de desempeño y recopilación de resultados con el clúster de alta disponibilidad al 75% | 8 | 23/06 | 26/06 |
| 5.3 | Pruebas de desempeño y recopilación de resultados con el clúster de alta disponibilidad al 50% | 8 | 26/06 | 30/06 |
| 5.4 | Pruebas de desempeño y recopilación de resultados utilizando la configuración redundante nativa de la base de datos | 16 | 30/06 | 05/07 |
| 5.5 | Pruebas de desempeño y recopilación de resultados utilizando Clusterware | 16 | 05/07 | 11/07 |
| 5.6 | Procesamiento de los resultados de pruebas | 4 | 12/07 | 15/07 |
| 5.7 | Generación de los reportes y gráficas con los resultados agregados | 6 | 15/07 | 19/07 |
| 5.8 | Documentación de los resultados y procesos involucrados | 4 | 19/07 | 23/07 |
| 5.9 | Revisión y retroalimentación por parte de los lectores / tutores | N.A. | 23/07 | 30/07 |
| 5.10 | Corrección de documento | 2 | 30/07 | 31/07 |
| 6 | **Documento Final** | | | |
| 6.1 | Integración de los documentos y resultados | 5 | 01/08 | 07/08 |
| 6.2 | Análisis y conclusiones | 3 | 07/08 | 09/08 |
| 6.3 | Aplicación de formato y estilo | 3 | 07/08 | 09/08 |
| 6.4 | Revisión preliminar por parte de tutores y lectores | N.A. | 09/08 | 16/08 |
| 6.5 | Corrección del documento final | 4 | 17/08 | 20/08 |
| 6.6 | Revisión filológica | N.A. | 20/08 | 27/08 |
| 6.7 | Corrección del documento final | 4 | 27/08 | 30/08 |
| 6.8 | Presentación del documento final | 2 | 01/09 | 30/09 |

**\* N.D.: No definida**

**\*\* N.A.: No aplica**

Metodologías



Sistema operativo:

Ubuntu Server 14.04.03 x64

# Referencias

Altavilla, Dave (2014). Intel's Low Power Bay Trail Takes Residence In Tiny, Inexpensive 'NUC' DIY Computer. Recuperado de: <http://www.forbes.com/sites/davealtavilla/2014/01/24/intels-low-power-bay-trail-takes-residence-in-tiny-inexpensive-nuc-diy-computer/>

Bader, David; Robert Pennington (1996). "Cluster Computing: Applications". Georgia Tech College of Computing. Recuperado de: <http://www.cc.gatech.edu/~bader/papers/ijhpca.html>

Burckle, Robert (2011) The Evolution of Single Board Computers. Recuperado de: <http://www.winsystems.com/whitepapers/SBC_Evolution.pdf>

Centos.org (2007) Configuring and Managing a Red Hat Cluster. Recuperado de: <http://www.centos.org/docs/5/pdf/Cluster_Administration.pdf>

Cubieboard.org (2013) Hadoop(High-availability distributed object-oriented platform) on Cubieboard. Recuperado de: <http://cubieboard.org/2013/08/01/hadoophigh-availability-distributed-object-oriented-platform-on-cubieboard/>

Khan, I. (2012) Falling RAM Prices Drive In-Memory Database Surge. Recuperado de: <http://blogs.sap.com/innovation/big-data/ram-prices-drive-in-memory-surge-020097>

Kiepert, J. (2013) Creating a Raspberry Pi-Based Beowulf Cluster. Recuperado de: <http://coen.boisestate.edu/ece/files/2013/05/Creating.a.Raspberry.Pi-Based.Beowulf.Cluster_v2.pdf>

Krathwowhl, D. (2002). A Revision of Bloom's Taxonomy: An Overview. Recuperado de: <http://www.unco.edu/cetl/sir/stating_outcome/documents/Krathwohl.pdf>

Oracle Corporation. Configuring Replication Between TimesTen Databases. Recuperado de: <http://download.oracle.com/otn_hosted_doc/timesten/1122/quickstart/html/admin/rep.html>

Oracle Corporation. Oracle® Clusterware Installation Guide. Recuperado de: <http://docs.oracle.com/cd/B28359_01/install.111/b28263/toc.htm>

Oracle Corporation. Why is Oracle TimesTen In-Memory Database fast? Recuperado de: <http://docs.oracle.com/cd/E21901_01/doc/timesten.1122/e21631/overview.htm#TTCIN120>

Pfister, Gregory (1998). In Search of Clusters (2nd ed.). Upper Saddle River, NJ: Prentice Hall PTR.

Vizard, M. (2012). The Rise of In-Memory Databases. Recuperado de: <http://slashdot.org/topic/datacenter/the-rise-of-in-memory-databases/>

Winn, R. (1999) Hardware Bible Fifth Edition, Que