**Project References**

# Arneil, J. (2012) *My Year With Exadata* [online]. e-DBA. Available at: <http://2012.ukoug.org/

# presdisplayfile.asp?prs\_prsid=7038&filename=UKOUG%5FOCT%5FMy%5Fyear%5Fwith%5FExadata%2Eppt>.

# This source is quite useful as it gives a consumers perspective on Exadata, a summary of exactly what Exadata is, their experiences from the set up, it also goes into detail on the technical specifications of Oracles newly released Exadata X3 which improves greatly on previous versions and finally sheds light on the hardware and software maintenance needed to keep Exadata running.

# Clegg, D. (2012) *Formulating Your Big Data Strategy: It's Not About Technology - It's About Business Value; It's Not About Data - It's About Insight from Data* [online]. Acunu. Available at: <http://2012.ukoug.org/presdisplayfile.asp?prs\_prsid=7013&filename=121203UKOUG%2Epdf>.

# This source outlines what the market is using to handle their big data needs, such NoSQL, Hadoop etc and shows exactly what kind of uses they are popular for, such as on-line big data workloads, batch processing. It also shows why RDBMS are not able to handle the issue of big data and how a good analytics program will help to sift through big data to find exactly what is needed.

# Colvin, A. (2012) *Patching Exadata Demystified* [online]. Enkitec. Available at: <http://2012.ukoug.org/presdisplayfile.asp?prs\_prsid=7040&filename=patching%5Fexadata%5Fdemystified%2Epdf>.

# This source illustrates the patching that Exadata goes through in order to keep running smoothly and secure. It shows why this is an issue that must be addressed as well as several methods of doing this without outage and the pros and cons of each. This is useful as it shows that after installation of Exadata there is still much work to be done and how it can become dangerous if you do not plan for it. The source provides useful information on how Exadata was not as smooth in the early days for patching as well it has improved over time, this unbiased reference will be used in my report where I compare my artefact to Oracles system for big data.

# Howard, P., Sowerby, D. (2012) *The Next BIG Thing?* [online]. SolStonePlus. Available at: <http://2012.ukoug.org/presdisplayfile.asp?prs\_prsid=7491&filename=SolStonePlus%5F%5FUKOUG%5FConference%5FEndeca%5FPresentation%5F2012%2Epptx>.

# This source explains what Big data is and how it should be used, it goes on to show how a graphical representation of such data can help to see exactly what you are getting from the data you have collected and stored. This is useful as it provides me with background information on big data and demonstrates how a graphical representation of my artefacts results could help understand the outcome when testing my artefact.

Kutrovsky, C. (2012) *Exadata Data warehousing: Leveraging Parallel Query* [online]. Pythian. Available at: <http://2012.ukoug.org/presdisplayfile.asp?prs\_prsid=7482&filename=Christo%5FKutrovsky%

5F%2D%5FExadata%5FDatawarehousing%5FLeveraging%5FParallel%5FQuery%5F%2D%5F2012%2Epptx>.

This source shows how data moves in Exadata in order to get the end result, it also details Parallel query concepts such as the big sort, big group by and a big join which allows for the querying of very large tables in real-time, this is achieved by using a flow-through system of handling the data. This is useful as it details how querying has changed from traditional RDBMS to allow for the increased performance requirements as well as the underlying concepts used to make this a reality.

|  |  |
| --- | --- |
| Lewis, J. (2012) *Are You Sure You Need Exadata?* [online]. | JL Computer Consultancy. |

# Available at: <http://2012.ukoug.org/presdisplayfile.asp?prs\_prsid=7425&filename=pessimist%5Fexadata%2Epdf>.

# [Accessed 4 December 2012].

This source goes into much detail on Oracles relatively new hardware solution to the issue of big data, Exadata and whether or not you might need such a system. It shows the technical details of the system, useful functions as well as parts of Exadata that consumers should be wary of when considering investing into. This is useful as it shows how Oracle has transitioned from software only, to hardware and software systems in order to tackle some of the pressing issues being felt in the market today such as big data. The source is useful as it shows an existing solution to big data and so I will use it when evaluating my artefact.

Logan, P., Abbey, M. (2012) *Applying Traditional DBA Skills to Oracle Exadata* [online]. Pythian. Available at: <http://2012.ukoug.org/presdisplayfile.asp?prs\_prsid=7213&filename=exadata%

5Fukoug12%5Fpost%5Fabbey%5F%2D%5FPaul%5FLogan%2Epptx>.

This source details the difference when working on an Exadata machine compared to a traditional RDBMS, it discusses how Exadata is tuned to get the best performance out of its hardware, what should be considered when planning for backup's, it discusses the internals of the Exadata machine as well as detailing functions that help with compression, performance and storage. This is useful as it shows how Oracle is providing innovative improvements to hardware and how this differs to traditional RDBMS.

Norris, D. (2012) *Exadata X3: The Fourth Generation* [online]. Oracle. Available at: <http://2012.ukoug.org/presdisplayfile.asp?prs\_prsid=7989&filename=Exadata%5FX3%5FTechnical%5FDelta%5Ffrom%5FX2%2Ddn%2Epdf>.

# This source is useful as it shows the technical specifications of Oracles newest addition to its Exadata machines, it compares this to previous versions to show the vast improvements on hardware and software made over the years. This is useful as it shows the direction that Oracle is taking in order to provide the best performance for its customers and to resolve the issue of big data, therefore I will use this when researching database systems that handle big data as well as for evaluating my artefact.

# Poder, T. (2012) *Exadata Performance Method* [online]. Enkitec. Available at: <http://2012.ukoug.org/presdisplayfile.asp?prs\_prsid=7312&filename=Tanel%5FPoder%5FExadata%5FPerformance%5FMethod%2Ezip>.

This source is useful because it shows the performance you can acquire from Exadata, details into how Exadata functions and how to create queries that work best for Exadata using the SQL monitoring report. It also shows how a function called smart scan makes large queries that bring back huge amounts of data much faster than before. This is useful as it shows some of the more interesting and helpful details of Exadata in order to tackle the issue of big data. Therefore I will use this to compare my artefact to, when evaluating it.

Powell, R. (2011) *Incorporating Big Data into an Enterprise Information Architecture. BeyeNetwork* [online]. 24 October 2011 [Accessed 12 December 2012]. Available at:

<http://www.b-eye-network.com/view/15480>.

This source is useful as it shows some of the most talked about topics in the database industry, such as big data analytics which allows for more in depth and faster analysis of data. Also grid computing which is a form of distributed computing where many grids loosely come together to perform large tasks. It also talks about how Oracles Exadata system is spearheading the innovations directed towards big data solutions into one system providing the performance needed to make use of big data. This is useful because it shows what big data is, it details the tools being crafted to handle such large scale data and how Oracles Exadata is dealing with the issue of big data.

Schmutz, G. (2012) *NoSQL Databases for Implementing Data Services – Should I Care?* [online]. Trivadis. Available at: <http://2012.ukoug.org/presdisplayfile.asp?prs\_prsid=7424&filename=

nosql%2Ddatabase%2Dfor%2Dimplementing%2Ddata%2Dservices%2Dshould%2DI%2Dcare%2Epdf>.

This source is useful as it provides another perspective on NoSQL databases, including a background into NoSQL and why you might use it instead of traditional RDBMS, it describes the different NoSQL database types and compares them to the RDBMS. It also describes a different way of programming called Polyglot Programming which is used to take advantage of many different languages advantages when tackling different problems and how this can be applied to database systems. This is useful as it shows different stratagems with the task of most efficiently resolving the big data issue.

# Wayner, P. (2011) *First look: Oracle NoSQL Database. InfoWorld.* [online]. 16 November 2011 [Accessed 06 December 2012]. Available at: <http://www.infoworld.com/d/data-explosion/first-look-oracle-nosql-database-179107>

This source details the developments taking place in terms of NoSQL databases with Oracle releasing its version which provides many of the functions that pioneers have been experimenting with, while keeping a professional solid database finish and the performance needed to handle big data. It also describes testing of this new database, when compared to other NoSQL databases it fared very well and shows how ACID (Atomicity, Consistency, Isolation, Durability) is clearly very useful when trying to get the most performance and reliability out of a database.

# CUBRID (2012) [*What is NoSQL for?*](http://www.cubrid.org/blog/dev-platform/what-is-nosql-for/)[online] Korea: NHN Search Solutions [Accessed 30 December 2012]. Available at: <http://www.cubrid.org/blog/dev-platform/what-is-nosql-for/>.

# This goes into great detail on the fundamentals of NoSQL as well as comparing this to relational database management systems (RDBMS) as well as a list of the popular NoSQL storage systems. This will be useful should I choose to use NoSQL when developing my artefact as a greater understanding of the underlying systems will greatly improve the successfulness of the finished system.

# LINUX For You (2003) *NewSQL — The New Way to Handle Big Data* [online] India: Electronics For You (EFY) [Accessed 02 January 2013]. Available at: <http://www.linuxforu.com/2012/01/newsql-handle-big-data/>.

# This source details on the alternatives to traditional OLTP systems when dealing with big data, such as NoSQL and goes on to show the benefits and drawbacks of using such a system. It also makes it clear why such systems are needed now more than ever and introduces NewSQL a collective name for alternative database systems, it goes on to illustrate the features and the uses for such systems. This will be useful when choosing what systems I plan to use for my artefact and will help me make a better decision and therefore a better artefact.

Transaction Processing Performance Council (TPC) (2001) *Transaction Processing - OLTP* [online] San Francisco: TPC. [Accessed 27 December 2012]. Available at: <http://www.tpc.org/tpcc/results/tpcc\_perf\_results.asp>.

This source details the benchmarks for a wide variety of hardware and software configurations which are purposefully made to provide the best performance in online transaction processing (OLTP). This will be invaluable to me when testing my artefact in order to get context from the existing systems businesses use today, I will compare my artefacts testing results to these and evaluate it with this information in mind.

# 451 Research (2000) NoSQL, *NewSQL and Beyond: The drivers and use cases for database alternatives* [online] New York: 451 Group [Accessed 21 December 2012]. Available at: <https://451research.com/report-long?icid=1651>.

# This source briefly shows the current market developments for database systems and gives an insight into the database systems emerging in order to handle the issue with the ever higher requirements needed from today's database systems. This will be useful when deciding what systems I intend to use in order to create my artefact and in doing so will help me build the most efficient system possible.

# Carr, J. (2012) *Big data’s little secret: Hadoop isn’t the end-all-be-all. VentureBeat.* [online]. 18 December 2012 [Accessed 04 January 2013]. Available at: <http://venturebeat.com/2012/12/18/hadoop/>.

# This source details the history of big data, claiming it started in 2007 with Google's Map Reduce framework being integrated into Apache Hadoop, it goes on to illustrate how Apache Hadoop and other big data solutions has led to commercialisation of the big data trend. It goes on to list these big data solutions such as Hive, Zookeeper etc, showing that Apache Hadoop is part of a large group of innovations in big data solutions. It also gives its prediction of what is next for big data and how this could be quite different from the open source solutions currently available. The source provides a decent amount of information on big data, the current solutions and its view of the future innovations, therefore this is a useful source that will help further my research and ultimately help improve the artefact, it will be included in the report for a more in depth evaluation.