

A Systematic Review Identifies a Lack of Standardization in OLAP Queries on Cloud Computing

Marcio Alexandre P. da Silva,¹ Eonassis Santos,² Paulo Caetano da Silva²

UFPE,¹ UNIFACS²

maps3@cin.ufpe.br, eonassis@msn.com, paulo.caetano@unifacs.br

Abstract. *Due to storage and processing capacity of cloud computing, OLAP (OnLine Analytical Processing) systems have found a suitable environment in order to minimize issues of data storage and processing performance. Objective: This paper proposes to identify what guideline has been standardizing the implementation of OLAP query languages applied on cloud computing. Method: A systematic review was developed with three researchers, through Internet, whose method contains research question, search strategy, inclusion and exclusion criteria, selection process, data extraction and synthesis. Result: The result points to use of different OLAP query syntaxes and mechanisms, proprietary or not, meaning the lack of standardization. Conclusion: Considering the importance of standardization in any technology field, which is spurred by several international organizations worldwide, this paper identifies the necessity of works that propose a standardization to guide the building of OLAP query languages on cloud computing.*

1. INTRODUÇÃO

A computação em nuvem é um ambiente de computação baseado na Internet que fornece recursos e dados advindos do processamento de um conjunto de computadores compartilhados. Desta forma, trata-se de um modelo para permitir o acesso sob demanda a um conjunto de recursos computacionais configuráveis (e.g. redes de computadores, servidores, armazenamento, aplicativos e serviços) [7].

OLAP é um acrônimo para processamento analítico on-line, o qual realiza análises multidimensionais de dados empresariais e fornece a capacidade de cálculos complexos, análise de tendências e modelagem sofisticada de dados. É a base para muitos tipos de aplicativos de negócios para gestão empresarial, *business intelligent* (BI). Consultas do tipo OLAP são usadas para realizar previsões, fazer análises financeiras, para descoberta de conhecimento. Ferramentas OLAP permitem aos usuários finais – por meio de mecanismos ou linguagens de consultas - realizar análises ad hoc (isoladas) de dados em múltiplas dimensões (*data warehouse*), apoiando as tomadas de decisão [8].

Entretanto, a falta de padronização tecnológica dificulta o desenvolvimento de soluções uniformes por fabricantes de software, tornando o produto de software mais caro para desenvolver. De acordo com a Organização Internacional de Normalização (ISO) [1],

as normas internacionais trazem benefícios tecnológicos, econômicos e sociais, também harmonizam as especificações técnicas de produtos e serviços, tornando a indústria mais eficiente e quebrando obstáculos ao comércio internacional.

A infraestrutura proposta pela computação em nuvens oferece capacidade de escalabilidade, elasticidade e tolerância a falhas. Além disso, ela é projetada para lidar com grandes quantidades de dados, em quase um número ilimitado de máquinas interligadas [4]. A migração dos sistemas OLAP para as nuvens trouxe uma nova maneira de trabalhar com o armazenamento de dados e linguagens de consultas multidimensionais [2][3][4][5]. Verifica-se nos trabalhos o desenvolvimento de diferentes infraestruturas de gerenciamento de dados e transações, com a finalidade de fazer processamentos analíticos mais velozes, assim como o poder de escalabilidade nesse ambiente.

Desta forma, considerando a importância da padronização tecnológica, este trabalho propõe identificar qual padrão as linguagens de consulta OLAP têm seguido quando operadas sobre a computação em nuvens.

Na seção 2, a metodologia aplicada na revisão sistemática feita neste trabalho é apresentada, cujo método é composto por questão de pesquisa, estratégia de busca, critério de inclusão e exclusão, processo de seleção (qualificação), extração e síntese dos dados. Na seção 3, os resultados obtidos nos trabalhos pesquisados são discutidos. Seguindo-se da última seção que trata da conclusão e investigações futuras.

2. METODOLOGIA

Mesmo tendo poucos trabalhos que abordam especificamente as linguagens de consulta OLAP nas nuvens, o método aplicado (o qual é detalhado a seguir) permitiu a identificação de 14 trabalhos, os quais compõem a análise deste artigo.

2.1 Questão de Pesquisa

Para iniciar as buscas de artigos, foi elaborada a seguinte questão de pesquisa: Qual o cenário atual das linguagens de consultas OLAP na computação em nuvens? Com esta pergunta buscou-se identificar o estado da arte do tópico abordado e, posteriormente, verificar o padrão utilizado por cada trabalho.

2.2 Estratégia de busca

Como estratégia de busca dos artigos, duas técnicas foram utilizadas conjuntamente: (i) a procura em motores de busca por meio de string de pesquisa, a qual deve estar alinhada à questão de pesquisa [9] e (ii) snowballing, o qual se trata de pesquisar nas referências bibliográfica outros trabalhos que serviram de base teórica [10]. Na string de pesquisa, houve a junção dos termos “OLAP” + “query” + “language” + “cloud” + “computing”, os quais foram colocados nas buscas avançadas fontes de trabalhos acadêmicos da IEEE [11], ACM [12], IADIS [13], Scopus [14], Scholar Google [15]. Os motores de buscas com acesso pago ou que não tem convênio com a Universidade Federal de Pernambuco [16], Universidade Salvador (UNIFACS) ou CAPES [17] não foram considerados neste trabalho.

2.3 Critérios de inclusão e exclusão

Como critério de inclusão, buscaram-se artigos abordando o uso de sistemas OLAP aplicados na computação em nuvens, propondo algum teste ou experimento através do qual a linguagem de consulta pudesse ser identificada. Não foram considerados artigos disponíveis em plataformas pagas, artigos não escritos em inglês ou português, não acessíveis na web, palestras, relatórios de workshop, documentos incompletos, drafts, slides de apresentação, white papers ou abstracts estendidos. Trabalhos puramente teóricos ou cuja abordagem não permite identificar a linguagem de consulta utilizada, também foram excluídos.

2.4 Processo de seleção

Trinta e três artigos foram selecionados após avaliação pelos critérios de inclusão e exclusão. A avaliação dos artigos foi feita em pares. O arquivo contendo os dados da avaliação foi armazenado na Internet porque os integrantes do grupo de pesquisa estavam localizados diferentes localidades (i.e. Salvador/BA e Recife/PE). Nesta etapa do processo foram observados quais trabalhos estariam de acordo com os critérios de inclusão, assim como sua qualidade.

Para averiguar se os trabalhos estavam de acordo com os critérios de inclusão, foram lidos e observados os títulos e abstracts, os testes ou experimentos apresentado no artigo, assim como a conclusão. No arquivo foram listados todos os artigos informando um identificador, título, autores, ano, link, análise dos artigos em pares e o status final, conforme ilustrado na Figura 1. Na coluna de Análise dos Artigos em Pares, os avaliadores classificaram em sim, concordando com a inclusão do artigo, ou não, excluindo o artigo, sempre justificando o motivo pelo qual a inclusão do artigo foi negada. Para melhor visualização do grupo de trabalho, caso havendo (i) total concordância (i.e. quando todos os avaliadores classificaram como sim) a cor da linha fica em amarelo, (ii) divergência fica verde, exigindo uma discussão mais detalhada entre os avaliadores e uma definição final (acordada entre os avaliadores) para o status daquele trabalho, e, finalmente, (iii) total discordância, a linha fica na cor cinza.

A	B	C	D
1	Incluído		
2	Divergência		
3	Excluído		
4			
id	Title	Authors	Ano
p01	Query Processing on OLAP System with Cloud Computing Environment	Nam Hun Park, Kii Hong Joo	2014
p01.1	Cloud-based Decision Support Systems and Availability Context: The Probability of Successful Decision Outcomes	S. Russell, V. Yoon, G. Forgonne	2010
p01.2	Improving Performance of Declarative Query Execution in DHT-based System	S. Arnedo, M. del, P. Villamil, R. I.	2010
p02	Accessing and steering the elastic OLAP cloud	Sicen Ye; Peng Chen; Ivan Janciak; Peter Brezany	2012
p02.1	An Elastic OLAP Cloud Platform	Peter Brezany; Yan Zhang; Ivan J.	2011

E	F	G	H
Link	Análise dos Artigos em Pares		Status
	autor 1	autor 3	
dropbox	sim	sim	ok
dropbox	não (não fala de OLAP tecnicamente somente do uso nas nuvens)	sim	não
dropbox	não aborda olap somente consultas em sistemas de hash table	não (somente hashtable)	não
http://ieeex	não (OMML? Não se aplica)	sim (modelagem genérica XML)	não
dropbox	sim	sim	resumo ok

Figura 1. Avaliação em Pares dos Critérios de Inclusão

A qualidade de cada trabalho foi mensurada por meio das respostas de seis perguntas: (i) Baseado em Pesquisa? (ii) A linguagem é testada no artigo? (iii) Tem consulta OLAP explícita? (iv) Aborda arquitetura nova nas nuvens? (v) Aborda desempenho de execução? (vi) Apresenta algum experimento? Três possibilidades de avaliação foram utilizadas: positivo (pontuando o artigo em um), parcial (pontuando em meio) e negativo (não pontuando). Para saber o *score* final de cada artigo, divide-se a soma total da pontuação (dada pelos avaliadores) com o número de avaliadores, conforme ilustrada na Figura 2.

Análise de Qualidade		
Artigo:	p04. OLaaS: OLAP as a Service	
Avaliador	Baseado em Pesquisa?	A linguagem é testada?
Autor 1	Positivo	Positivo
Autor 3	Positivo	Positivo
Autor 2	Positivo	Parcial
Avaliador	Tem consulta explícita?	Aborda arquitetura nova?
Autor 1	Parcial	Positivo
Autor 3	Positivo	Positivo
Autor 2	Parcial	Positivo
Avaliador	Aborda desempenho?	Apresenta Experimento?
Autor 1	Negativo	Positivo
Autor 3	Negativo	Positivo
Autor 2	Negativo	Positivo
	Score	
	(4,5 + 5,0 + 4,0) / 3 = 4,5	

Figura 2. Avaliação da Qualidade dos Artigos em Pares.

Finalmente, os artigos com *score* acima de quatro foram escolhidos para próxima fase (i.e. extração dos dados), a qual é explicada próxima seção.

2.5 Extração dos dados

Para a extração dos dados, identificou-se no texto dos artigos a resposta para cada uma das perguntas da fase anterior, conforme ilustrado na Figura 3. Neste processo, o grupo de pesquisa optou em tirar imagens do texto original do Abstract, da Linguagem utilizada no trabalho e da Conclusão, disponibilizando-os no arquivo de planilha compartilhada entre os autores.

	Baseada em:																									
	MDX																									
	Abstract:																									
<p><i>Abstract</i>—Online Analytics Processing (OLAP) is utilised to develop multidimensional operations enabling queries and visualisation for Business Intelligence (BI). Most of the OLAP systems come with a tightly integrated user interface for querying and visualisation of data without the core OLAP operations exposed as an API. Advanced BI applications can be developed and composed to create complex workflows if the OLAP operations are available as an API. In addition, a Web Service based API would enable applications to use Service Oriented Architecture for Big Data Analytics and also easily be deployed on a Cloud. This paper documents the design and prototyping of an OLAP based Platform as a Service, termed OLAP as a Service (OLAPaaS).</p>																										
os	p01	p02.1																								
	p04	p04.2.1																								
		p05																								
		p06																								
<p>Linguagem:</p> <pre> Connection rConn = (OlapWrapper) DriverManager.getConnection("jdbc:xmla:" + "Server=" + server_name ", ) OlapConn oConn = rConn.unwrap(OlapConn.class); </pre> <p>In the above snippet the connection is casted as OlapWrapper. This allows access to the OLAP features from the oConn object. The next statement creates the actual query:</p> <pre> Query myQuery = new Query(<QueryName>, <CubeName>); </pre>																										
<p>Conclusão:</p> <p style="text-align: center;">VI. CONCLUSION</p> <p>This paper presents the design and prototyping of a framework for exposing OLAP operations as a set of RESTful Web Services. Web Services were developed for Upload Cube, Slice, Dice, Drilldown and Retrieve Data OLAP operations.</p> <table border="1"> <thead> <tr> <th>Call No</th> <th>Slice</th> <th>Dice</th> <th>DrillDown</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>500 (300)</td> <td>410 (275)</td> <td>400 (320)</td> </tr> <tr> <td>2</td> <td>500 (280)</td> <td>410 (275)</td> <td>440 (320)</td> </tr> <tr> <td>3</td> <td>492 (105)</td> <td>110 (100)</td> <td>111 (110)</td> </tr> <tr> <td>4</td> <td>178 (105)</td> <td>110 (100)</td> <td>111 (110)</td> </tr> <tr> <td>5</td> <td>178 (105)</td> <td>110 (100)</td> <td>111 (110)</td> </tr> </tbody> </table> <p style="text-align: center;">TABLE I API RESPONSE TIME PERFORMANCE (MS). VALUES IN THE BRACKETS INDICATE THE MDX QUERY PROCESSING TIMES USING SAIKU ANALYTICS ENGINE</p>			Call No	Slice	Dice	DrillDown	1	500 (300)	410 (275)	400 (320)	2	500 (280)	410 (275)	440 (320)	3	492 (105)	110 (100)	111 (110)	4	178 (105)	110 (100)	111 (110)	5	178 (105)	110 (100)	111 (110)
Call No	Slice	Dice	DrillDown																							
1	500 (300)	410 (275)	400 (320)																							
2	500 (280)	410 (275)	440 (320)																							
3	492 (105)	110 (100)	111 (110)																							
4	178 (105)	110 (100)	111 (110)																							
5	178 (105)	110 (100)	111 (110)																							

Figura 3. Extração de Dados dos Artigos

3.6 Síntese dos dados

Após a extração dos dados averiguou-se que todos os trabalhos abordam mecanismos de melhoramento de desempenho das ferramentas OLAP em arquitetura distribuídas (incluindo nas nuvens). Verificou-se, também, a ausência de um padrão para as consultas OLAP, considerando que os trabalhos em análise fazem uso de diversos tipos de

ferramentas disponíveis no mercado, as quais propõem minimizar problemas de desempenho de execução de consultas, conforme apresentado na Tabela 1.

Tabela 1. Linguagens ou Ferramentas Utilizadas para Consultas OLAP.

N.	Paper	Linguagem de Consultas OLAP
1	[2]	WCF + SOA protocols (webservice)
2	[4]	OMML
3	[6]	CloudMdsQL
4	[18]	MDX
5	[19]	Própria baseada em SQL
6	[20]	Amazon EC2 (Amazon Web Service)
7	[21]	Amazon Web Services
8	[22]	TPC-H query
9	[23]	MDX-XML (Xquery)
10	[24]	BigTable
11	[25]	Hive-QL
12	[26]	Amazon EC2 (Amazon Web Service)
13	[27]	Hive-QL
14	[28]	Amazon EC2 (Amazon Web Service)

Conforme ilustrado na Figura 4, vinte e três por cento dos trabalhos utilizaram a Amazon EC2, uma ferramenta sem código aberto e cujo plano gratuito não está disponível em todas as regiões. Essa ferramenta faz uso do AWS (Amazon Web Service) para envio de consultas. Desta forma, a existência de uma padronização seria relevante porque as ferramentas OLAP não estariam atreladas a ferramentas proprietárias e pagas, podendo ser implementadas uniformemente em código aberto e gratuito. Por outro lado, A MDX e a Hive-QL, da Microsoft e Treasure Data respectivamente, são padrões abertos disponíveis sob licença gratuita as quais são usadas em 30% dos trabalhos analisados (15% cada uma).

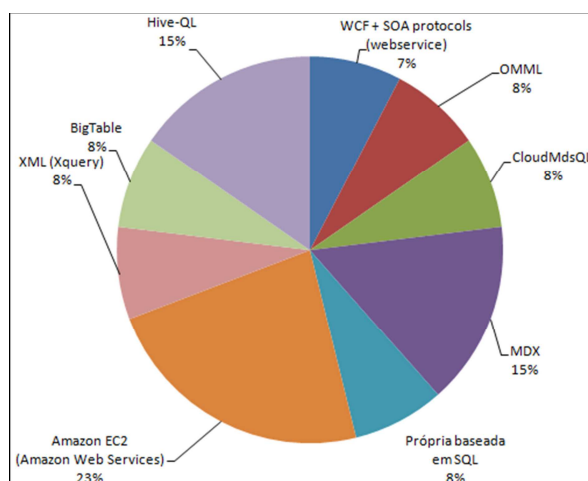


Figura 4. Ferramentas e linguagens utilizadas em consultas OLAP baseada nas nuvens

É importante frisar que a MDX é um padrão de consulta largamente utilizado nas ferramentas OLAP em bancos de dados relacionais e em ambientes desktop e web. Alguns trabalhos abordaram o uso de MDX em bancos de dados XML [32]. Contudo, sua utilização em ambientes em nuvens tem tido adesão moderada.

A Hive-QL surgiu no contexto de bancos de dados (e data warehouses)

armazenados nas nuvens, no projeto Hive [29]. Para realizar a consulta o Hive-QL, a ferramenta Hadoop é utilizada para mapear os bancos de dados distribuídos nas nuvens [31], a qual trata um *framework* que permite o processamento distribuído de um grande volume de dados configurado em *clusters* de computadores [30].

4 CONCLUSÃO E TRABALHOS FUTUROS

Os trabalhos encontrados, que tratam sobre consultas OLAP nas nuvens, possuem foco na resolução ou mitigação de problemas de desempenho, i.e. com relação ao tempo de execução em ambientes distribuídos e em nuvens. As linguagens de consultas utilizadas não são projetadas a partir de qualquer norma de padronização internacional.

As consultas OLAP são feitas por meio de mecanismos e ferramentas construídos para atender o uso específico de cada ferramenta – proprietária ou não - sem foco na padronização. Mesmo havendo um forte estímulo de órgãos internacionais (e.g. ISO) para a uniformização dos sistemas de informação, as linguagens de consultas OLAP baseada em nuvens não possuem um padrão que norteia sua implementação. Observa-se, ainda, a adaptação no ambiente OLAP de mecanismos de consultas que não foram especificados para ambientes multidimensionais (e.g. Web Service, SQL).

A maior adesão de Amazon Web Service (AWS), entre os trabalhos identificados, não configura o surgimento de um padrão. A maior propagação da Amazon EC2, no mercado e academia, justifica o uso de sua sintaxe (i.e. AWS) na maioria dos trabalhos pesquisados. Desta forma, os trabalhos apenas fazem uso de ferramentas específicas disponíveis no mercado para testar a arquitetura e validar o desempenho de execução, não importando qual mecanismo de consulta (e linguagem) é utilizado.

Com relação a questão de pesquisa, conclui-se que o cenário atual das linguagens de consultas OLAP na computação em nuvens:

- Não é padronizado ou especificado;
- Faz uso de recursos tecnológicos advindos de outras áreas da ciência da computação, os quais não foram concebidos para atender dados multidimensionais (e.g. web service, consultas transacionais SQL);
- Não há evidência de interesse na especificação de um padrão, uma vez que os recursos tecnológicos utilizados tem atendido ao proposto pelas ferramentas que dominam o mercado.

Fomenta-se como investigação futura - considerando a inexistência de trabalhos que tratam sobre a normalização das linguagens de consultas OLAP nas nuvens - a construção de um padrão tecnológico que considere a multidimensionalidade dos bancos de dados que suportam sistemas OLAP e o modelo distribuído proposto pela computação em nuvens. Esta normalização tornaria o desenvolvimento das soluções OLAP uniformes por fabricantes de software, harmonizando as especificações técnicas de produtos e serviços no contexto OLAP.

AGRADECIMENTOS

Este trabalho foi apoiado pela Fundação CAPES, por meio de bolsa de estudo de doutorado fornecida ao primeiro autor.

REFERÊNCIAS

- [1] International Organization for Standardization (2017). Disponível em: <https://www.iso.org/>. Acessado em 10 de julho de 2017.
- [2] Park, NH, Joo, KH (2014) Query Processing on OLAP System with Cloud Computing Environment. International Journal of Multimedia and Ubiquitous Engineering. <Http://dx.doi.org/10.14257/ijmue.2014.9.5.16>.
- [3] Russell, S., Yoon, V., Forgionne, G. (2010) "Cloud-based Decision Support Systems and Availability Context: The Probability of Successful Decision Outcomes". Inf Syst E-Bus Manage (2010) 8: 189. doi:10.1007/s10257-010-0126-4.
- [4] P. Brezany, Y. Zhang, I. Janciak, P. Chen and S. Ye. (2011) "An Elastic OLAP Cloud Platform," 2011 IEEE Ninth International Conference on Dependable, Autonomic and Secure Computing, Sydney, NSW, 2011, pp. 356-363. Doi: 10.1109/DASC.2011.76
- [5] Strímbei, C. (2012) "OLAP Services on Cloud Architecture". Journal of Software and Systems Development. Article ID 840273 (16 pages). Doi: 10.5171/2012.840273.
- [6] Boyan Kolev, Patrick Valduriez, Carlyna Bondiombouy, Ricardo Jimenez-Peris, Raquel Pau, Jose Pereira (2016) CloudMdsQL: Querying Heterogeneous Cloud Data Stores with a Common Language. Distrib Parallel Databases 34: 463. doi:10.1007/s10619-015-7185-y.
- [7] Hassan, Qusay (2011). "Demystifying Cloud Computing". The Journal of Defense Software Engineering. CrossTalk Jan/Feb 2011. Acessado em 07 julho 2017.
- [8] Chaudhuri, S. Dayal, U. (1997). An overview of data warehousing and OLAP technology. ACM SIGMOD Record Homepage archive. Volume 26 Issue 1, March 1997. Pages 65-74.
- [9] B. Kitchenham (2007) Guidelines for performing Systematic Literature Reviews in Software Engineering. EBSE Technical Report. EBSE-2007-01.
- [10] C Wohli (2014) Guidelines for snowballing in systematic literature studies and a replication in software engineering. EASE '14 Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering. Article No. 38 ISBN: 978-1-4503-2476-2 doi>10.1145/2601248.2601268.
- [11] IEEE (2017) Advanced Search Options. Disponível em: http://ieeexplore.ieee.org/search/adv_search.jsp. Acessado em 12/07/2017.
- [12] ACM Digital Library (2017) Advanced Search. Disponível em: <http://dl.acm.org/advsearch.cfm>. Acessado em 12/07/2017.
- [13] IADIS Digital Library (2017). Disponível em: <http://www.iadisportal.org/digital-library/showsearch>. Acessado em 12/07/2017.

- [14] Scopus (2017). Disponível em: <https://www.scopus.com/>. Acessado em 12/07/2017.
- [15] Google Scholar (2017). Disponível em: <https://scholar.google.com.br/>. Acessado em 12/07/2017.
- [16] UFPE (2017). Disponível em: <http://www2.cin.ufpe.br/site/index.php>. Acessado em 12/07/2017.
- [17] Fundação CAPES (2017). Disponível em: <http://www.capes.gov.br/>. Acessado em 12/07/2017.
- [18] A. D. Patil and N. D. Gangadhar, "OLaaS: OLAP as a Service," 2016 IEEE International Conference on Cloud Computing in Emerging Markets (CCEM), Bangalore, 2016, pp. 119-124. Doi: 10.1109/CCEM.2016.029
- [19] Chen, G.; Chen, K.; Jiang, D.; Ooi, B.C.; Shi, L.; Vo, H.T.; Wu, S. (2012) E3: an Elastic Execution Engine for Scalable Data Processing. Electronic Preprint for Journal of Information Processing. Vol. 20. No. 1. Information Processing Society of Japan.
- [20] Cătălin Strîmbei (2012) OLAP Services on Cloud Architecture. IBIMA Publishing Journal of Software & Systems Development. Vol. 2012 (2012), Article ID 840273, 16 pages DOI: 10.5171/2012.840273.
- [21] F. Dehne, D. Robillard, A. Rau-Chaplin and N. Burke, "VOLAP: A Scalable Distributed System for Real-Time OLAP with High Velocity Data," 2016 IEEE International Conference on Cluster Computing (CLUSTER), Taipei, 2016, pp. 354-363. Doi: 10.1109/CLUSTER.2016.29.
- [22] T. Shioi and K. Hatano, "Rule- and Cost-Based Optimization of OLAP Workloads on Distributed RDBMS with Column-Oriented Storage Function," 2016 IEEE 4th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW), Vienna, 2016, pp. 165-170. doi: 10.1109/W-FiCloud.2016.44.
- [23] R. Dkaich, I. El Azami and A. Mouloudi, "XML OLAP cube in the cloud: Towards the DWaaS," 2016 International Conference on Engineering & MIS (ICEMIS), Agadir, 2016, pp. 1-5. doi: 10.1109/ICEMIS.2016.7745331.
- [24] Google Inc. (2012) An Inside Look at Google BigQuery - Google Cloud Platform. Disponível em: <https://cloud.google.com/files/BigQueryTechnicalWP.pdf>. Acessado em 05 de julho de 2017.
- [25] B. Arres, N. Kabbachi and O. Boussaid, "Building OLAP cubes on a Cloud Computing environment with MapReduce," 2013 ACS International Conference on Computer Systems and Applications (AICCSA), Ifrane, 2013, pp. 1-5. doi: 10.1109/AICCSA.2013.6616498.
- [26] F. Dehne, Q. Kong, A. Rau-Chaplin, H. Zaboli and R. Zhou, "A distributed tree data structure for real-time OLAP on cloud architectures," 2013 IEEE International Conference on Big Data, Silicon Valley, CA, 2013, pp. 499-505. Doi: 10.1109/BigData.2013.6691613.
- [27] A. Ettaoufik and M. Ouzzif, "Query's optimization in data warehouse on the cloud using fragmentation," 2014 International Conference on Next Generation Networks and Services (NGNS), Casablanca, 2014, pp. 145-148. Doi: 10.1109/NGNS.2014.6990243.

- [28] Hyuck Han, Young Choon Lee, Seungmi Choi, Heon Y. Yeom, Albert Y. Zomaya (2013) Cloud-Aware Processing of MapReduce-based OLAP Applications. AusPDC '13 Proceedings of the Eleventh Australasian Symposium on Parallel and Distributed Computing - Volume 140. Pages 31-38.
- [29] Apache Hive TM (2017). Disponível em: <https://hive.apache.org>. Acessado em 10 de julho de 2017.
- [30] Apache Hadoop (2017). Disponível em: <http://hadoop.apache.org/>. Acessado em 08 de julho de 2017.
- [31] Treasure Data (2017). Hive (SQL-style) Query Language | Treasure Data. Disponível em: <https://docs.treasuredata.com/articles/hive>. Acessado em 12 de julho de 2017.
- [32] Silva, P.C., Times, V.C., Ciferri, R.R. and Ciferri, C.D., 2012. Analytical Processing Over XML and XLink. International Journal of Data Warehousing and Mining (IJDWM). Volume 8, Issue 1. IGI Global.