

Performance evaluation of SPARQL queries in triplestore management systems

Felipe Luzzardi da Rosa
Universidade Federal de Pelotas
Rua Gomes Carneiro, 1
fldrosa@inf.ufpel.edu.br

Adenauer Corrêa Yamin
Universidade Federal de Pelotas
Rua Gomes Carneiro, 1
adenauer@inf.ufpel.edu.br

Roger da Silva Machado
Universidade Federal de Pelotas
Rua Gomes Carneiro, 1
rdsmachado@inf.ufpel.edu.br

Ana Marilza Pernas
Universidade Federal de Pelotas
Rua Gomes Carneiro, 1
marilza@inf.ufpel.edu.br

ABSTRACT

The number of research areas that model information through ontologies is increasing, but there is an absence of publications referring to the persistence of these data models. Taking into account that the use of suitable managers for its storage could be a key factor to the scalability of applications, it is perceivable that an analysis of said managers is necessary. Therefore, this work presents a comparison based on queries execution time in three triplestore database managers. AllegroGraph, Stardog, and Virtuoso were analyzed using the WatDiv benchmark. In the performed tests, Virtuoso achieved the best performance in the vast majority of the queries.

KEYWORDS

Ontology, Triple Store, Queries Evaluation

1 INTRODUÇÃO

É crescente o número de esforços de pesquisa ligados a aplicações que utilizam informações vindas de ontologias, principalmente em áreas como Big Data, Internet das Coisas e Web Semântica. Embora diversas propostas apresentadas neste âmbito façam uso de tais modelos de dados, poucas versam sobre a maneira de garantir a persistência dos mesmos, ou então, quando fazem, utilizam repositórios relacionais para tal fim [4, 6]. Uma forma mais eficiente de persistir vocabulários ontológicos é utilizando um modelo de armazenamento em triplas.

O modelo de armazenamento em triplas pode ser dividido em três grandes categorias, com base na arquitetura de sua implementação: em memória; nativo; e externo. No armazenamento em memória o modelo de triplas é mantido na memória principal do dispositivo, porém, tal método é ineficiente para o armazenamento de grandes volumes de dados [5]. O modelo de armazenamento nativo de triplas fornece a persistência com a implementação de uma base de dados própria da ferramenta utilizada, sendo possível recuperar os dados posteriormente por meio de consultas utilizando uma linguagem apropriada, como a SPARQL. Na categoria de armazenamento externo, as triplas são configuradas para executarem em

bancos de dados de terceiros, utilizando modelos relacionais de armazenamento, por exemplo.

Nota-se que o armazenamento nativo de triplas está ganhando força e popularidade, em grande parte devido a seus tempos de carregamentos inferiores [3]. Sendo assim, é importante conhecer as diversas ferramentas disponíveis para armazenamento nativo de triplas, analisando suas características e desempenho, de forma a se escolher a mais eficiente. Deste modo, este artigo propõe-se a fazer uma comparação de desempenho baseada no tempo de execução de consultas entre as três das ferramentas de armazenamento de triplas mais populares atualmente: AllegroGraph [1]; Stardog [7]; e Virtuoso [8]. Estas três ferramentas se propõem a gerenciar o armazenamento e a consulta de dados em triplas, variando principalmente com relação aos formatos de dados aceitos e a forma de interação para as consultas.

Para os testes de desempenho foi utilizado o *benchmark* Waterloo SPARQL Diversity Test Suite (WatDiv), desenvolvido pela Universidade de Waterloo [2]. O WatDiv fornece um gerador de *datasets* com um fator de escala variável, permitindo assim a geração de *datasets* de diversos tamanhos. Além disso, é fornecido um gerador de consultas SPARQL, permitindo que sejam geradas consultas distintas, separadas em quatro grupos: “Lineares” (L), mais simples e diretas; “Estrela” (S) consultas que referenciam diversos nodos em um formato de estrela; “Floco de Neve”(F), consultas que referenciam vários nodos que por sua vez referenciam outros nodos; e “Complexas” (C), consultas que utilizam uma mistura dos formatos anteriores.

2 TESTES E RESULTADOS

Foram realizados testes focados no tempo de execução de cada uma das vinte consultas SPARQL fornecidas pelo WatDiv, utilizando *datasets* com 1, 5 e 10 milhões de triplas. Os testes com 10 milhões de triplas não foram realizados na ferramenta Allegrograph devido a limitações de sua licença gratuita. Cada teste foi realizado cinco vezes, sendo registradas as médias do tempo de execução em segundos para cada consulta. Os resultados são apresentados nas Tabelas 1, 2, 3 e 4.

Observando os tempos, é possível concluir que a ferramenta Virtuoso apresentou os melhores resultados, tendo sido a mais rápida na grande maioria das consultas. Também pode ser notado que o

In: Sessao de Pôsteres do WebMedia'2017, Gramado, Brasil. Anais do XXIII Simpósio Brasileiro de Sistemas Multimídia e Web: Workshops e Pôsteres. Porto Alegre: Sociedade Brasileira de Computação, 2017.

© 2017 SBC – Sociedade Brasileira de Computação.
ISBN 978-85-7669-380-2.

Tabela 1: Tempo de execução, em segundos, das consultas do tipo S

Ferramentas	S1		S2		S3		S4		S5		S6		S7	
Allegro	1	0,0668774	1	0,06412	1	0,0613696	1	0,0613696	1	0,0543772	1	0,0548694	1	0,057317
	5	0,0963936	5	0,0793994	5	0,0750856	5	0,0863822	5	0,085407	5	0,0563396	5	0,0554092
	10	-	10	-	10	-	10	-	10	-	10	-	10	-
Stardog	1	0,0947322	1	0,1018032	1	0,0962996	1	0,0971264	1	0,095974	1	0,0978992	1	0,0976978
	5	0,1044178	5	0,1024508	5	0,1031698	5	0,1041022	5	0,100918	5	0,0988374	5	0,0976624
	10	0,1009872	10	0,0952024	10	0,097306	10	0,094586	10	0,1024596	10	0,0993094	10	0,0996264
Virtuoso	1	0,3289344	1	0,0609126	1	0,058814	1	0,0604578	1	0,0597912	1	0,0593226	1	0,0587208
	5	0,348566	5	0,059431	5	0,0615134	5	0,0616746	5	0,060899	5	0,0582094	5	0,0600702
	10	0,326622	10	0,0636436	10	0,0627762	10	0,064229	10	0,063713	10	0,0659842	10	0,0637428

Tabela 2: Tempo de execução, em segundos, das consultas do tipos C

Ferramentas	C1		C2		C3	
Allegro	1	0,0728248	1	0,0770824	1	0,4375378
	5	0,1296854	5	0,0967634	5	1,958557
	10	-	10	-	10	-
Stardog	1	0,0986196	1	0,0967634	1	0,0996898
	5	0,1008178	5	0,1020082	5	0,0997854
	10	0,1006606	10	0,0994658	10	0,0935464
Virtuoso	1	0,064128	1	0,066897	1	0,0629626
	5	0,0643752	5	0,0737042	5	0,0637796
	10	0,068031	10	0,0739352	10	0,0634726

Tabela 3: Tempo de execução, em segundos, das consultas do tipo L

Ferramentas	L1		L2		L3		L4		L5	
Allegro	1	0,0643506	1	0,0568794	1	0,0616546	1	0,063656	1	0,0678878
	5	0,0616158	5	0,0571544	5	0,0627976	5	0,0590924	5	0,0614466
	10	-	10	-	10	-	10	-	10	-
Stardog	1	0,0947406	1	0,094898	1	0,0998588	1	0,0956404	1	0,097169
	5	0,1030084	5	0,099584	5	0,0984724	5	0,1004204	5	0,1025072
	10	0,0980908	10	0,095526	10	0,0916584	10	0,098767	10	0,0938232
Virtuoso	1	0,059044	1	0,058974	1	0,057987	1	0,0576816	1	0,0580234
	5	0,0577476	5	0,0598486	5	0,0623434	5	0,0568444	5	0,0582824
	10	0,0609034	10	0,0603972	10	0,0625004	10	0,0634382	10	0,0632302

Tabela 4: Tempo de execução, em segundos, das consultas do tipo F

Ferramentas	F1		F2		F3		F4		F5	
Allegro	1	0,0641824	1	0,0578822	1	0,0669106	1	0,067333	1	0,0914706
	5	0,115556	5	0,0673398	5	0,0629318	5	0,0857004	5	0,108796
	10	-	10	-	10	-	10	-	10	-
Stardog	1	0,095435	1	0,0950576	1	0,0956896	1	0,0997422	1	0,0990498
	5	0,1007864	5	0,1013582	5	0,1061386	5	0,1019098	5	0,0987932
	10	0,0950134	10	0,0924756	10	0,1020956	10	0,099365	10	0,0950358
Virtuoso	1	0,063648	1	0,0801034	1	0,0642348	1	0,0792464	1	0,0635378
	5	0,0642144	5	0,0796122	5	0,0656168	5	0,0792714	5	0,0612986
	10	0,0656276	10	0,073142	10	0,0642534	10	0,085988	10	0,067115

Stardog obteve resultados muito similares em todos os testes realizados, tendo sido a ferramenta com menor alteração de desempenho entre tamanhos diferentes de base de dados.

Para se identificar a relevância dos resultados obtidos, foram desenvolvidos teste t de Student¹ com intervalo de confiança de 95%. Nos testes, foi detectado que os resultados das consultas S2, S3, S4, S7, L2, F1 e F3 para o tamanho de 1M assim como os das consultas L3, F3 e S6 para o tamanho de 5M entre as ferramentas AllegroGraph e Virtuoso não apresentam diferenças estatísticas significativas.

3 CONSIDERAÇÕES FINAIS

Este trabalho realizou uma comparação de desempenho relacionada ao tempo de execução de consultas SPARQL entre três ferramentas de armazenamento em formato de triplas: AllegroGraph, Stardog e Virtuoso. Os testes consideraram o tempo de execução, em segundos, de vinte consultas diferentes em três tamanhos de bases de dados, contendo 1, 5 e 10 milhões de triplas cada uma. Ao final dos testes gerados foi possível concluir que a ferramenta Virtuoso apresenta o melhor desempenho geral, apresentando o tempo de execução mais baixo na grande maioria das consultas.

Como trabalhos futuros destaca-se a realização de testes com outras ferramentas de armazenamento em triplas, assim como testes com bases de dados maiores e com métodos estatísticos mais sofisticados para aferição do tempo médio de execução.

REFERÊNCIAS

- [1] Allegrograph. 2017. acesso em agosto 2017. (2017). Disponível em: <franz.com/agraph/allegrograph/>.
- [2] Gunes Aluç. 2014. *Diversified Stress Testing of RDF Data Management Systems*. Master's thesis. David R. Cheriton School of Computer Science, Waterloo, ON, Canada.
- [3] BioOntology. 2011. Comparison of Triple Stores. Disponível online em: <https://www.bioontology.org/wiki/images/6/6a/Triple_Stores.pdf>. (2011). acesso em novembro 2016.
- [4] K. Bitencourt, F. Durão, and M. Mendonça. 2015. EmergencyFire: An Ontology for Fire Emergency Situations. In *WebMedia - Brazilian Symposium on Multimedia and the Web*.
- [5] Shridevika Maharajan. 2012. *Performance of native SPARQL query processors*. Masters Theses. Department of Information Technology, Uppsala University.
- [6] M. V. O. Sena and R. F. Bulcão Neto. 2016. A Solution to Discard Context Information using Metrics, Ontology and Fuzzy Logic. In *WebMedia - Brazilian Symposium on Multimedia and the Web*.
- [7] Stardog. 2017. acesso em agosto 2017. (2017). Disponível em: <stardog.com/>.
- [8] Virtuoso. 2017. acesso em agosto 2017. (2017). Disponível em: <http://virtuoso.openlinksw.com/>.

¹Teste estatístico que tem por objetivo detectar se duas médias podem ser consideradas estatisticamente diferentes.