

CONIC SEMESP

15º Congresso Nacional de Iniciação Científica

TÍTULO: ANÁLISE CIENTÍFICA DE DESEMPENHO DE BANCO DE DADOS NOSQL

CATEGORIA: EM ANDAMENTO

ÁREA: CIÊNCIAS EXATAS E DA TERRA

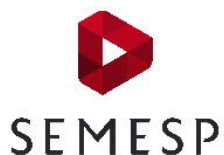
SUBÁREA: COMPUTAÇÃO E INFORMÁTICA

INSTITUIÇÃO: FACULDADE DE TECNOLOGIA DE SÃO JOSÉ DO RIO PRETO

AUTOR(ES): JOSIANE RODRIGUES DE SOUZA NAKAGAWA

ORIENTADOR(ES): SERGIO RICARDO BORGES JUNIOR

Realização:



Apoio:



1. Resumo

O presente projeto busca elaborar um comparativo entre alguns SGBDs NoSQL que se diferenciam entre si pela forma de armazenamento de dados. Nesta perspectiva, selecionou-se para comparação os SGBDs NoSQL: Berkeley DB [1], Cassandra [2], Mongo DB [3] e Neo4J [4], que tratam e manipulam os dados utilizando estruturas do tipo chave/valor, colunas, grafos e documentos, respectivamente. Um plano de teste funcional foi elaborado com os seguintes critérios: instalação, transação, concorrência, consistência, integridade e recursos de segurança. Os resultados preliminares demonstram que a escolha do SGBD NoSQL depende do uso e do tipo de aplicação.

2. Introdução

No ambiente corporativo as condições mudam rapidamente e para acompanhar a evolução da tecnologia e a rapidez das informações, grandes empresas investem em novos mecanismos e sistemas de armazenamento de dados. Esses investimentos buscam principalmente atender a demanda dos clientes, aumentar vantagens competitivas e torna-se líder de mercado.

Entretanto, o grande desafio é encontrar soluções para os problemas de escalabilidade, desempenho e custo. Devido a esses problemas e necessidades, recentemente, surgiram os sistemas gerenciadores de banco de dados (SGBDs) *NoSQL (Not Only SQL)*, cujo termo é usado para descrever soluções de dados distribuídos não relacionais. Nessa perspectiva, um dos principais fatores que os SGBDs *NoSQL* oferecem é agilidade em armazenamento de grande volume de dados, simplificando a replicação e oferecendo escalabilidade.

3. Objetivos

O presente projeto tem por objetivo elaborar um estudo comparativo entre alguns dos sistemas gerenciadores de banco de dados *NOSQL*, destacando-se, dentre eles: Berkeley DB, Cassandra, Neo4j e Mongo DB, com base em critérios que buscam destacar suas vantagens e desvantagens.

4. Metodologia

O presente trabalho teve caráter exploratório e aplicado, pois busca entender o funcionamento dos sistemas gerenciadores de banco de dados NoSQL e, em seguida, elaborar um estudo comparativos com base nos critérios: interface gráfica,

desempenho em operação básicas (inserção, exclusão, alteração e consulta), usabilidade, recursos de segurança, transações/concorrências, consistência e integridade. Para isso, a pesquisa bibliográfica foi concentrada em livros, artigos científicos, manuais online, entre outros.

O plano de trabalho a ser executado envolve as seguintes atividades:

Atividade 1: Levantamento bibliográfico;

Atividade 2: Estudo e investigação dos SGBDs NoSQL selecionados;

Atividade 3: Elaboração de um Plano de Teste;

Atividade 4: Realização dos testes conforme definidos no Plano de Teste;

Atividade 5: Análise e interpretação dos resultados obtidos nos testes;

Atividade 6: Elaboração de artigos para demonstrar o estudo comparativo realizado dos SGBDs NoSQL.

O acompanhamento da pesquisa foi realizado pelo professor orientador por meio de reuniões semanais com objetivo de verificar a evolução dos estudos e também para traçar diretrizes e direcionamentos.

5. Desenvolvimento

Inicialmente, a instalação e configuração padrão de cada SGBD NOSQL foi realizada. Nesta etapa, não houve muitos desafios a serem superados. No entanto, um dos entraves foi entender como cada SGBD armazena as informações, ou seja, como as informações são estruturadas. Neste sentido, para os SGBDs NoSQL tem-se: Berkeley DB (chave/valor), Cassandra (colunas), Mongo DB (documentos) e Neo4J (grafos). Note que se buscou comparar SGBDs que tratam o armazenamento dos dados de forma diferente entre si, justificando-se a escolha dos SGBDs a serem comparados. Em seguida, buscou-se a execução de comandos básicos para inserção, alteração, exclusão e consulta aos dados. Assim, optou-se primeiro em compreender o funcionamento de cada SGBD para, em um segundo momento, avaliar os critérios do estudo comparativo.

Após as duas etapas já citadas, elaborou-se um plano de teste. Optou-se pelo teste funcional, no qual o sistema é considerado uma caixa preta e, para testá-lo, são fornecidas entradas e as saídas são avaliadas. Neste sentido, selecionou-se as seguintes funções de teste: instalação, transação, concorrência, consistência, integridade e recursos de segurança.

6. Resultados Preliminares:

Os resultados preliminares apontam que a utilização de um SGBD NoSQL depende das características e da necessidade da aplicação, ou seja, não há um SGBD NoSQL que atenda todos os tipos de aplicações e suas finalidades. A Tabela 1 apresenta um breve comparativo dos SGBDs comparados.

Tabela 1 – Breve comparativo dos SGBDs NoSQL.

SGBD's	Pontos Fortes	Pontos Fracos
Cassandra	Boa distribuição de dados e rapidez nas consultas.	Em consultas complexas pode ocorrer perda da consistência dos dados.
Mongo DB	Apresenta melhor solução Web.	Manutenção e operação difícil.
Neo4j	Conexões rápidas.	Os recursos de segurança deixam a desejar quando a memória é limitada.
Berkeley DB	Mecanismos de armazenamento confiável.	Baixo desempenho em transações e concorrências.

De acordo com os testes iniciais pode-se constatar que o SGBD Cassandra oferece agilidade em armazenamento de grande volume. Já em relação à configuração, o Mongo DB apresenta melhor solução Web, mas necessita de especialistas para utilizar a ferramenta. O Neo4j destaca-se por oferecer excelente qualidade em conexões. Por fim, o SGBD Berkeley DB apresenta mecanismo de armazenamento confiável, mas apresenta baixo desempenho em transações e concorrências.

7. Fontes Consultadas:

[1] Oracle Berkeley DB 12c Release 1. Disponível em:

http://docs.oracle.com/cd/E17076_04/html/index.html >. Acesso em: 28 mai. 2015.

[2] The Project Cassandra Apache. Disponível em: <<http://cassandra.apache.org>>. Acesso em 28 jun. 2015.

[3] MongoDB. Disponível em: <<https://www.mongodb.org/>>. Acesso em: 28 abr. 2015.

[4] Neo4j, the World's Leading Graph Database. Disponível em: <<http://neo4j.com>>. Acesso em: 28 jun. 2015.