ADMINISTRAÇÃO DE BANCO DE DADOS

Márcio Motta

SAGAH SOLUÇÕES EDUCACIONAIS INTEGRADAS

OBJETIVOS DE APRENDIZAGEM

Ao final deste texto, você deve apresentar os seguintes aprendizados:

- Identificar os desafios da administração de banco de dados.
- Conhecer comandos que auxiliam na administração do banco de dados.
- Reconhecer as técnicas de administração de banco de dados.

INTRODUÇÃO

O banco de dados é um dos elementos da administração de um ambiente de informação. Para um usuário final a organização do banco de dados é invisível, apenas aparecendo os seus resultados quando este é bem formatado e executado. No caso de problemas, o usuário final identificará dados inconsistentes ou problemas com a sua performance, mas não visualizando o banco de dados em si.

Três áreas continuam a ser os maiores desafios de gerenciamento para os administradores de banco de dados. São elas.

TECNOLOGIAS

Nos últimos anos, o surgimento de forças, como a TI híbrida, virtualização, computação em nuvem, convergência continuada de infraestrutura e BYOD (traga seu próprio dispositivo) ofereceram aos profissionais de tecnologia novas formas de trabalhar, revolucionando o modelo tradicional até então praticado na área. A expectativa é de que as pessoas que trabalham nessa área consigam colocar em prática conceitos e tendências tecnológicas, como bancos de dados embutidos, Internet das Coisas (IoT) e Cloud Computing, mas sem deixar de lado as habilidades de gerenciar tecnologias e infraestruturas tradicionais.

Cabe ao Administrador de Banco de Dados e sua equipe considerarem as possibilidades para então definir o uso de SGBD físico ou a utilização de um serviço na nuvem, sobre essa decisão pesarão diversos fatores que deverão ser analisados antes de tomar uma decisão final. O mesmo vale para o tipo de SGBD escolhido de acordo com as suas características, funções, desempenho e custos.

Com sua promessa de economia de custos e maior flexibilidade e agilidade, a Nuvem está atraindo muitas organizações como uma alternativa para a implantação de novos aplicativos, incluindo aqueles com altos requisitos de desempenho de banco de dados. Entretanto, essa transição cria novas complexidades e desafios para os DBAs, especialmente porque os DBAs ainda são, em última análise, os responsáveis tanto pelo desempenho dos bancos de dados quanto pela segurança dos dados, independentemente de eles estarem nas próprias instalações ou na nuvem.

São algumas considerações para o gerenciamento dos dados na nuvem que os DBAs devem ponderar antes de sua escolha:

- Ao considerar quais bancos de dados são migrados para a nuvem, levar em conta o processo de transferência de dados e a latência, além de como manter os bancos de dados em sincronia, se necessário, especialmente se for preciso integrar os aplicativos com outros que não residam na mesma implantação na nuvem.
- Um banco de dados com desempenho insatisfatório nas instalações também apresentará um mau desempenho na nuvem. Mudar o problema de lugar não é solução. O escalonamento na nuvem para compensar um mau desempenho pode sair caro rapidamente, além de ser a abordagem incorreta.
- Entender os serviços e as capacidades do provedor de serviços, avaliar a arquitetura recomendada e estar atento à manutenção programada.
- Refletir, planejar e gerencar o backup e recuperação para garantir que dados importantes não sejam perdidos caso ocorra uma falha ou interrupção no fornecedor.
- Mantenha-se à frente da segurança, percebendo que a criptografia é apenas a ponta do iceberg – considere quem monitorará o acesso ao banco de dados impedindo o acesso mal-intencionado ou não autorizado, prepare-se para o pior e tenha um plano de ação documentado para o caso de violação da segurança ou perda de dados.
- Se é importante monitorar e otimizar as implantações nas instalações, isso é ainda mais importante na nuvem, dada sua natureza dinâmica, sendo ideal usar um conjunto de ferramentas consistente em ambos os lados dos ambientes de TI híbrida.

A Figura 1 a seguir procura relatar o ranking dos SGBDs mais utilizados e traçar uma relação entre as suas tecnologias:

Figura nº 1 – Ranking dos SGBDs

Rank Jan 2017	Rank Dec 2016	Rank Jan 2016	DBMS	Database Model	Score Jan 2017	Score Dec 2016	Score Jan 2016
1.	1.	1.	Oracle	Relational DBMS	1416.72	+12.32	-79.36
2.	2.	2.	MySQL	Relational DBMS	1366.29	-8.12	+67.03
3.	3.	3.	Microsoft SQL Server	Relational DBMS	1220.95	-5.70	+76.89
4.	↑ 5.	4.	MongoDB	Document store	331.90	+3.22	+25.88
5.	↓ 4.	5.	PostgreSQL	Relational DBMS	330.37	+0.35	+47.97
6.	6.	6.	DB2	Relational DBMS	182.49	-1.85	-13.88
7.	7.	↑ 8.	Cassandra	Wide column store	136.44	+2.16	+5.49
8.	8.	↓ 7.	Microsoft Access	Relational DBMS	127.45	+2.75	-6.59
9.	9.	↑ 10.	Redis	Key-value store	118.70	-1.20	+17.54
10.	10.	↓ 9.	SQLite	Relational DBMS	112.38	+1.54	+8.64
11.	11.	↑ 12.	Elasticsearch	Search engine	106.17	+2.90	+28.96
12.	12.	↑ 14.	Teradata	Relational DBMS	74.17	+0.79	-0.78
13.	13.	↓ 11.	SAP Adaptive Server	Relational DBMS	69.10	-1.32	-14.08
14.	14.	↓ 13.	Solr	Search engine	68.08	-0.92	-7.32
15.	15.	↑ 16.	HBase	Wide column store	59.14	+0.51	+5.77

Fonte: Austrian IT Consulting, disponível em: http://db-engines.com/en/
Acesso em: 28/05/2017

Nas 4 primeiras posições desse ranking são vistos 2 modelos de SGBDs de licença proprietária (Oracle e SQL Server) e 2 de licença Open Source (Livre, de código aberto, MySQL e MondoDB). O MondoDB vem em forte crescimento no mercado mundial, ocupando já a quarta posição superando o também bastante utilizado PostreSQL. Um dos motivos dessa ascensão é o fato do MongoDB não trabalhar com o modo Relacional de tabelas como os outros SGBDs, mas com um sistema de Índices também conhecido como NO-SQL.

PERFORMANCE

A performance do banco de dados é um fator de grandes desafios para o BDA. O conjunto que envolve Hardware+Sistema+SGBD deve estar equacionado para uma performance satisfatória. Porém, conforme o crescimento e a demanda das informações armazenadas essa tarefa torna-se cadê vez mais complexa. Muitas vezes a escolha inicial de infraestrutura pode não ser mais suficiente, não atendendo as demandas necessárias necessitando de migração ou atualização, ou a sua manutenção e configuração não

estão ocorrendo de forma organizada e coerente com as necessidades da plataforma, causando transtornos e inconsistências.

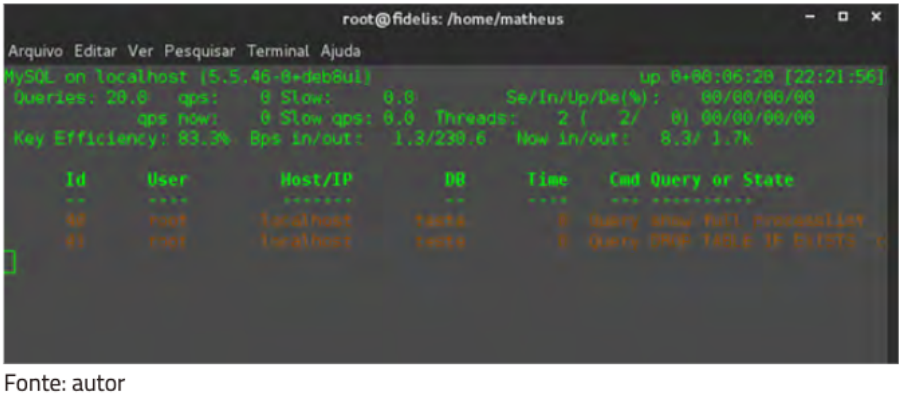
A modelagem dos dados e a construção de consultas otimizadas são extremamente importantes no desempenho do banco de dados. Campos mal definidos, com tamanhos de dados equivocados tanto para textos quanto para números são determinantes para a performance, toda a definição dos dados (DDL) deve ser dimensionada para o tipo de dado que cada registro receberá. O mesmo vale para a organização das consultas. Índices e *constraints* (restrições) já devem ser criados na etapa de definição para que possam ser utilizadas de forma ágil durante a manipulação dos dados(DML) com consultas simplificadas e de resultados diretos. Muitas vezes uma base possui milhares ou até milhões de registros, se uma consulta não for restrita ao que se deseja, o tempo de espera poderá ser bastante demorado, prejudicando todo o sistema.

Uma das formas de monitorar o desempenho de um SGBD (MySQL neste exemplo) é através da ferramenta MYTOP (Linux/Open Source). Para usá-la é necessário executar o comando:

mytop -u <usuario> -p <senha> -h <host>

Ao executá-lo, serão mostradas todas as informações dos bancos de dados armazenados assim como os tempos de acesso e possíveis problemas, como são apresentados na Figura 2 abaixo:

Figura nº 2 – Tela do MyTOP no Linux



Fonte: autor

Um erro bastante comum é se preocupar com segurança e desempenho apenas diante da reclamação dos usuários de um dado alterado indevidamente ou lentidão na utilização do sistema, pois um monitoramento pontual não terá fundamento para mostrar que o banco de dados não é o culpado. É necessário monitoramento e relatórios constantes que auxiliem o DBA a identificar gargalos, tomando as ações necessárias preventivamente, ou mesmo comprovando que o problema não está no banco de dados, podendo direcionar o tratamento do problema para o local correto.

SEGURANÇA

A preocupação com a criação e manutenção de ambientes seguros se tornou a ocupação principal de administradores de redes, de sistemas operacionais e de bancos de dados. Pesquisas mostram que a maioria dos ataques, roubos de informações e acessos não- autorizados são feitos por pessoas que pertencentes à organização alvo. De modo geral, os mecanismos de segurança referem-se às regras impostas pelo subsistema de segurança do SGBD, que verifica todas as solicitações de acesso, comparando-as com as restrições de segurança armazenadas no catálogo do sistema. Entretanto existem brechas no sistema e ameaças externas que podem resultar em um servidor de banco de dados comprometido ou na possibilidade de destruição ou no roubo de dados confidenciais.

As ameaças aos bancos de dados podem resultar na perda ou degradação de alguns ou de todos os objetivos de segurança aceitos, são eles: integridade, disponibilidade, confidencialidade. A integridade do banco de dados se refere ao requisito de que a informação seja protegida contra modificação imprópria. Os bancos de dados SQL implementam mecanismos que restringem ou permitem acessos aos dados de acordo com papeis ou roles fornecidos pelo administrador. O comando GRANT concede privilégios específicos para um objeto (tabela, visão, seqüência, banco de dados, função, linguagem procedural, esquema ou espaço de tabelas) para um ou mais usuários ou grupos de usuários.
São tipos de controles que podem ser implementados para garantir maior segurança nos bancos de dados:

Controle de Acesso

É todo controle feito quanto ao acesso ao BD, impondo regras de restrição, através das contas dos usuários. O DBA é o responsável superior por declarar as regras dentro do SGBD. Ele é o responsável por conceder ou remover privilégios, criar ou excluir usuários, e atribuição de um nível de segurança aos usuários do sistema, de acordo com a política da empresa.

Controle de Inferência

É um mecanismo de segurança para banco de dados estatísticos que atua protegendo informações estatísticas de um individuo ou de um grupo. Bancos de dados estatísticos são usados principalmente para produzir estatísticas sobre várias populações. O banco de dados pode conter informações confidenciais sobre indivíduos. Os usuários têm permissão apenas para recuperar informações estatísticas sobre populações e não para recuperar dados individuais, como, por exemplo, a renda de uma pessoa específica.

Controle de Fluxo

É um mecanismo que previne que as informações fluam por canais secretos e violem a política de segurança ao alcançarem usuários não autorizados. Ele regula a distribuição ou fluxo de informação entre objetos acessíveis. Um fluxo entre o objeto A e o objeto B ocorre quando um programa lê valores de A e escreve valores em B. Os controles de fluxo têm a finalidade de verificar se informações contidas em alguns objetos não fluem explicita ou implicitamente para objetos de menor proteção. Dessa maneira, um usuário não pode obter indiretamente em B aquilo que ele ou ela não puder obter diretamente de A.

Criptografia de Dados

Você pode ler aqui um pouco mais sobre criptografia. É uma medida de controle final, utilizada para proteger dados sigilosos que são transmitidos por meio de algum tipo de rede de comunicação. Ela também pode ser usada para oferecer proteção adicional para que partes confidenciais de um banco de dados não sejam acessadas por usuários não autorizados. Para isso, os dados são codificados através da utilização de algum algoritmo de codificação. Assim, um usuário não autorizado terá dificuldade para decifrá-los, mas os usuários autorizados receberão chaves para decifrar esses dados. A criptografia permite o disfarceda mensagem para que, mesmo com o desvio da transmissão, a mensagem não seja revelada.

Privilégios

Os privilégios são permissões únicas dadas a cada usuário ou grupo. Eles definem permissões para tipos de autorização. Pelos privilégios é possível autorizar o usuário a modificar ou alcançar determinado recurso do Banco de Dados.

O SGBD deve oferecer acesso seletivo a cada relação do banco de dados baseando-se em contas específicas. As operações também podem ser controladas; assim, possuir uma conta não necessariamente habilita o possuidor a todas as funcionalidades oferecidas pelo SGBD. Informalmente existem dois níveis para a atribuição de privilégios para o uso do sistema de banco de dados:

- Nível de conta: Nesse nível, o DBA estabelece os privilégios específicos que cada conta tem, independente das relações no banco de dados.
- Nível de relação (ou tabela): Nesse nível, o DBA pode controlar o privilégio para acessar cada relação ou visão individual no banco de dados.

Em alguns casos, interessa conceder um privilégio temporário a um usuário. Por exemplo, o proprietário de uma relação pode querer conceder o privilégio SELECT a um usuário para uma tarefa específica e depois revogar aquele privilégio quando a tarefa estiver completada. Por isso, é necessário um mecanismo para a revogação de privilégios. Em SQL, um comando REVOKE é introduzido com o intento de cancelar privilégios.

REFERÊNCIAS

Embarcadero - **Database Trends Survey Report.** Disponível em: http://www.embarcadero.com/images/dm/technical-papers/database-survey-report.pdf. Acessado em: 28 de maio de 2017.

PFLEEGER, Charles P.; PFLEEGER, Shari L. - **Security in computing**. 4ª ed. Massachusetts: Editora Prentice Hall, 2012.

SHASHA, DENNIS, BONNET, PHILIPPE - **Database Tuning**, Morgan Kaufmann Publishers, 2003.

Conteúdo:

SAGAH
SOLUÇÕES EDUCACIONAIS INTEGRADAS