

UNIVERSIDADE DE BRASÍLIA

Faculdade do Gama

Sistemas de Banco de Dados 2

Tecnologias de Banco de Dados (TI-BD)

Bancos de Dados Ativos

Gabriel Moretti de Souza - 200018205

Brasília, DF

2023

Definição da Tecnologia Pesquisada:

Os bancos de dados ativos são sistemas de bancos de dados (BD) estendidos com um sistema de regras (ou gatilhos) definidas, incluídas no modelo de execução de um BD juntamente ao seu sistema tradicional. É uma tecnologia que integra as funcionalidades convencionais dos bancos de dados com um monitoramento que pode assegurar um modelo comportamental ao banco.

Este mesmo sistema é capaz de reconhecer determinados eventos e executar ações correspondentes a quaisquer regras que sejam ativadas por tais. Essas mesmas regras geralmente são baseadas em três componentes, seguindo a sigla do paradigma E-C-A, que são:

- Evento (quando avaliar): Indicador da ocorrência de uma determinada situação externa ou interna ao banco de dados. Podem ser classificados em três tipos, sendo eles temporais, normalmente com horário e frequência de repetições pré-definida; definidos pelo usuário, tendo como exemplo um login de usuário; e as próprias operações do banco de dados: insert, delete, update, select. Irão iniciar o processo de avaliação das regras.

- Condição (o que avaliar): Predicado referente a um estado do banco de dados, geralmente implementadas por consultas, define uma validação para a execução de ações;

- Ação (como responder): Conjunto de operações que serão executadas quando um determinado evento ocorrer e sua respectiva condição for validada. Tipicamente as ações realizadas são de modificação ou consulta no banco de dados;

Desse modo, podemos visualizar a seguir, de forma esquemática, o paradigma E-C-A:



Juntamente ao paradigma E-C-A, os bancos de dados ativos possuem três componentes básicos para a obtenção de seus sistemas ativos, tais

componentes são responsáveis pelo tratamento ativo de informações dentro do sistema. São eles:

- Monitoramento de eventos: Módulo encarregado de detectar os eventos que ocorram e ativar as regras que dependam desse evento. Para que essa detecção de eventos possa ocorrer em situações reais e ser monitorada de acordo, devem ser empregadas linguagens de definição de eventos, que permitem a modelagem de situações complexas relacionadas aos mesmos. Alguns exemplos dessas linguagens são: Ode, Samos, Compose e Snoop. Essas linguagens, por sua vez, utilizam-se de operadores comuns que estão incorporados a elas. Tais operadores se resumem em:

- Sequência: O evento composto acontece quando todos os eventos que constituem uma sequência ocorram em sua ordem pré-determinada;

- Conjunção: O evento composto acontece quando todos os eventos que formam a conjunção ocorrem, comparável ao operador lógico “E”. Independe de ordem.

- Disjunção: O evento composto acontece quando pelo menos um dos eventos que formam a disjunção ocorrem, comparável ao operador lógico “OU”. Independe de ordem.

- Negação: Tem como base o tempo decorrido. O evento composto ocorre caso os eventos descritos em sua expressão de negação não ocorram em um determinado intervalo de tempo.

Desse modo, esses operadores são utilizados geralmente em conjunto acoplados ao banco de dados ativo, para assim obter o monitoramento dos eventos que ocorram no mesmo.

- Avaliação de Condição: Seguindo a detecção do evento, a avaliação da condição é responsável pela validação e avaliação eficiente das condições definidas para tal evento e regra. Desse modo, todas as regras cujas condições respectivas sejam verdadeiras serão assim passadas para o próximo componente, o Executor de Ações.

Essas condições são expressas por fórmulas, em que são atribuídas um valor booleano quando executadas (verdadeiro ou falso), e tais fórmulas são

geralmente escritas em linguagens de consulta, onde seu resultado, quando avaliada, ser vazio ou não determinará a veracidade da condição.

- Executor de Ações: Módulo responsável por coordenar e sincronizar a detecção de eventos e a execução de suas ações. Desse modo, existem duas formas de execução dessas ações, sendo divididas entre ações imediatas, que ocorrem antes do final da transação que disparou a respectiva regra; e ações tardias, ocorrendo em uma transação independente da transação atual.

Essa relação entre execução de ações e a regra é denominada de modo de acoplamento.

As linguagens utilizadas para a especificação de ações são divididas entre três grupos, sendo eles:

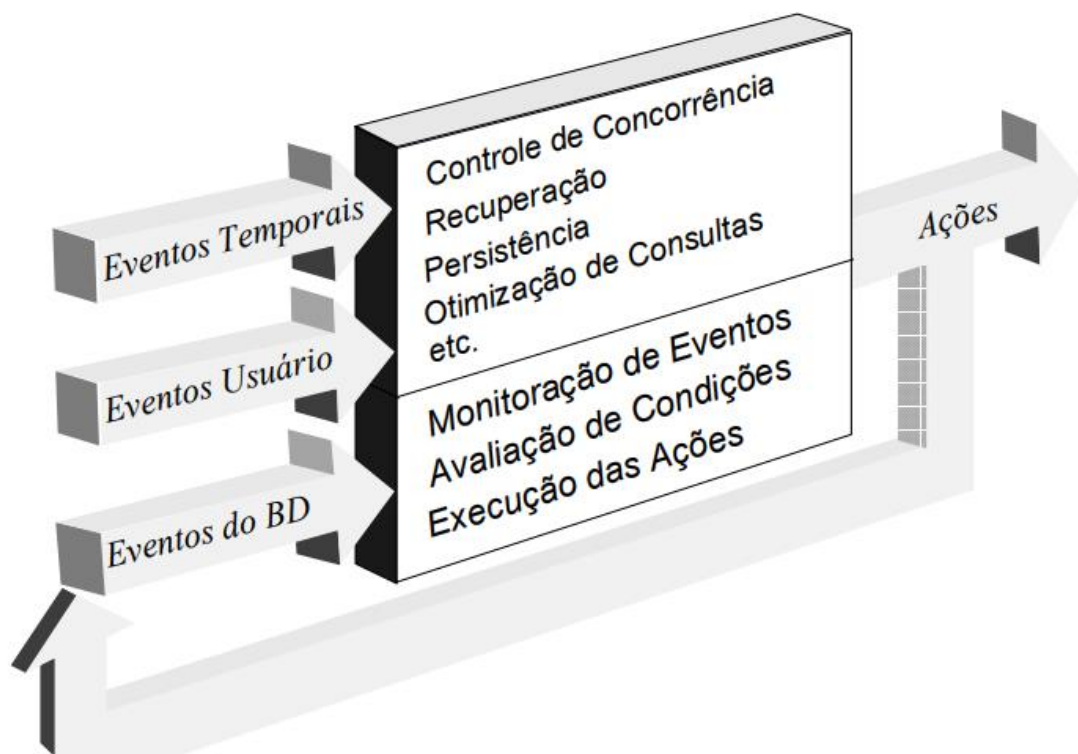
- Linguagens de consulta: Nessas linguagens, as ações serão especificadas com a linguagem de consulta padrão do banco de dados em questão, estando sobre o controle do próprio sistema.

- Linguagens de consulta estendidas: Para que seja possível que as regras interajam com o ambiente externo, é necessário que a linguagem de consulta seja estendida, ou seja, ela pode basear-se em duas partes: a linguagem de consulta e a linguagem de comunicação, sendo a última baseada em chamadas que terão conexão com o ambiente externo, como por exemplo “send” ou “receive”.

- Acesso direto ao banco de dados: Linguagem algorítmica, utilizando-se de um número pequeno de elementos se comparado às linguagens citadas anteriormente, e assim possuindo maior expressividade em suas ações.

Com uma organização e conexão definida entre esses três módulos do banco de dados ativos é possível definir e acoplar as regras previamente citadas, automatizando as respostas do banco para quaisquer eventos e condições atendidas.

Segue, portanto, uma imagem representando a integração desses três módulos com o sistema tradicional de um banco de dados (não ativo):



Objetivo(s) principal(is) da Tecnologia Pesquisada:

Com essa definição, é possível entender os objetivos gerais dos bancos de dados ativos. Como alguns dos seus principais objetivos como tecnologia podemos citar: automatizar o tratamento de eventos e processos que ocorram no banco de forma constante e periódica; Para casos de grande porte, o monitoramento de uma grande quantidade de dados pode se tornar inviável para tratamento humano, desse modo, a utilização dos bancos de dados ativos visa implementar um modo de tornar esse tratamento possível e eficiente, procurando diminuir também a quantidade de erros que possam vir de precedência humana nessas ações; Do mesmo modo, essa tecnologia visa também facilitar na tomada de decisões, visto que as ações podem ocorrer em tempo real; Juntamente a todos esses objetivos, também é possível citar a facilidade de integração desse banco de dados com outros sistemas, com seu tratamento rápido e automático de eventos.

Em resumo, a tecnologia objetiva trazer maior agilidade, eficiência e automação às aplicações que possam utilizá-la, trazendo benefícios significativos para qualquer sistema que dependa ou não do monitoramento de

informações críticas.

Vantagens da Tecnologia Pesquisada:

Uma das principais vantagens da tecnologia em questão é a utilização de tais regras para automatizar e, portanto, evitar perda de tempo com trabalho manual dos mantenedores do banco de dados acerca de determinadas condições. Algo que gastaria constantemente o tempo de um desenvolvedor com a execução das ações em um banco relacional tradicional, que podem ser diárias, pode ser substituído pelo sistema E-C-A da tecnologia.

Por outro lado, também é possível citar como vantagem da utilização de bancos de dados ativos a possibilidade de evitar os erros de precedência humana que podem ocorrer em atividades constantes e periódicas de um banco relacional comum, pois as ações determinadas pelo sistema E-C-A são definidas e rígidas e buscam evitar tais erros.

Como outra vantagem, é possível citar o potencial de implementações dentro de uma rede de sistemas de bancos de dados ativos. Com a automação das condições e ações de um banco, é possível tornar o sistema persistente, atemporal e de possível evolução, podendo ser integrado, melhorado ou utilizado em outros projetos ou sistemas, facilitando algo que seria também comum para um sistema simples de bancos de dados relacional.

Desvantagens da Tecnologia Pesquisada:

Do mesmo modo que a tecnologia possui diversas vantagens, também é possível citar algumas desvantagens sobre a mesma. Como primeiro exemplo, temos a possível sobrecarga de processamento em um sistema de bancos de dados ativos, visto que o monitoramento crescente de ações e eventos pode ocupar um espaço indesejado na carga de processamento do mesmo, levando a uma possível queda no desempenho do sistema, onde em um sistema comum de um banco de dados poderia ter controle sobre.

Outra desvantagem seriam os possíveis conflitos que podem ocorrer entre as regras do banco de dados ativo, principalmente em um sistema que possua uma grande quantidade de regras ativas, onde algumas dessas podem, por exemplo, designar ações opostas ou inconsistentes, possibilitando a ocorrência de erros que dificilmente ocorreriam abaixo da vigilância de um

sistema operado manualmente de um banco de dados relacional comum.

Apesar disso, também pode se tornar difícil para se manter constante a manutenção de um sistema complexo de regras que esse sistema pode resultar, assim tornando, por exemplo, uma atualização, algo que seria comum em um banco de dados relacional, delicada e de difícil implementação.

Exemplo(s) de uso interessante(s) em empresas, organizações, projetos ou instituições dessa tecnologia de Banco de Dados pesquisada:

Consequência do seu potencial de utilização, a tecnologia de bancos de dados ativos possui diversas aplicações em organizações, empresas ou projetos do mercado, assim como implementações em diversos bancos de dados amplamente utilizados, que são as bases para essas aplicações.

Dentre esses bancos, é possível citar três bancos de dados que possuem o suporte necessário para a implementação de princípios dos bancos de dados ativos, sendo eles:

- PostgreSQL: sistema de gerenciamentos de bancos de dados relacionais que suporta o sistema de gatilhos (triggers) e procedimentos armazenados, assim suportando as funcionalidades de bancos de dados ativos. Por ser de natureza livre (sem custo de licença), é amplamente utilizado em aplicativos da web e conhecido por sua escalabilidade, confiabilidade e recursos avançados de segurança.

Como exemplo de sucesso desse banco de dados no mercado, podemos citar a NASA, onde o seu sistema de monitoramento de satélites foi implementado utilizando os gatilhos do banco de dados ativos para detectar quaisquer condições anormais nos dados dos satélites, acionando assim alertas para notificar os operadores dos mesmos.

- MySQL: sistema de gerenciamento de banco de dados relacional de código aberto que também suporta recursos de banco de dados ativos, como gatilhos e procedimentos armazenados. Também é de uso livre, e amplamente utilizado para aplicativos empresariais e da web, sendo conhecido por sua velocidade e facilidade de uso.

- Manhattan: é um banco de dados ativo desenvolvido pelo Twitter, e

projetado para armazenar e processar grandes quantidades de dados em tempo real. Ele é usado pelo Twitter para gerenciar as informações de usuários, tweets, tendências e muito mais, permitindo que a plataforma suporte milhões de usuários e bilhões de tweets.

O Manhattan foi desenvolvido para ser altamente escalável e tolerante a falhas, usando um modelo distribuído de armazenamento e processamento de dados para atender seus objetivos.

Desse modo, as tecnologias de bancos de dados ativos, juntamente com suas implementações em sistemas amplamente utilizados, demonstram a importância e utilidade desses sistemas em organizações de diversos segmentos e características.

O PostgreSQL, o MySQL e o Manhattan são exemplos de como os bancos de dados ativos podem ser utilizados para processar grandes quantidades de dados em tempo real, além de também suportarem funcionalidades avançadas de segurança, automação e escalabilidade que podem possuir implementações notáveis no mercado.

Com isso, é possível perceber que os bancos de dados ativos têm um papel fundamental no suporte à tomada de decisões e no gerenciamento de informações críticas em diversas organizações modernas.

Bibliografias Pesquisadas:

CERI, S.; WIDOM, J. Active Database Systems: Triggers and Rules for Advanced Database Processing. 1. ed. Morgan Kaufmann Publishers, 1995.

DEV MEDIA. Bancos de Dados Ativos. Revista SQL Magazine, nº 94, p. 44-53, jan. 2013. Disponível em: [https://www.devmedia.com.br/bancos-de-dados-ativos-revista-sql-magazine-](https://www.devmedia.com.br/bancos-de-dados-ativos-revista-sql-magazine-94/23025#:~:text=O%20Banco%20de%20Dados%20Ativo,modelo%20comportamental%2C%20denominadas%20regras%20ativas)

94/23025#:~:text=O%20Banco%20de%20Dados%20Ativo,modelo%20comportamental%2C%20denominadas%20regras%20ativas. Acesso em: 17 abr. 2023.

CILIA, Mariano. Bancos de Dados Ativos. Dissertação de mestrado - Universidad Nacional del Centro de la Provincia de Buenos Aires, Facultad de Ciencias Exactas, Instituto de Sistemas Tandil, 2005.

Universidade Estadual do Oeste do Paraná. GRANDO, L. S.; MAGALHÃES, H. J. Extensões de Modelos de Dados para Aplicações Avançadas. Paraná: CASCAVEL, 2009. Capítulo 2, p. 2-8. Disponível em: <https://www.inf.unioeste.br/~olguin/4458-semin/G8-monografia.pdf>. Acesso em: 17 abr. 2023.

RAMAKRISHNAN, R.; VENKATARAMAN, S. Active Databases: Not Just a Matter of Triggers. IEEE Computer, [S.l.], v. 31, n. 6, p. 38-44, jun. 1998. ISSN 0018-9162. Disponível em: <https://ieeexplore.ieee.org/document/658844>. Acesso em: 17 abr. 2023.

MEU NEGÓCIO UOL. MySQL x SQL Server x MongoDB x Postgres: Qual banco de dados escolher? Disponível em: <https://meunegocio.uol.com.br/blog/mysql-x-sql-server-x-mongodb-x-postgres-qual-banco-de-dados-escolher/>. Acesso em: 17 abr. 2023.

RANGANATHAN, Karthik; KELLOGG, Chris. Manhattan: A real-time multi-tenant distributed database for Twitter. In: Proceedings of the Eighth ACM International Conference on Web Search and Data Mining. ACM, 2015.

TWITTER. Manhattan: our real-time, multi-tenant distributed database for Twitter scale. Twitter Engineering Blog, 2014. Disponível em: https://blog.twitter.com/engineering/en_us/a/2014/manhattan-our-real-time-multi-tenant-distributed-database-for-twitter-scale. Acesso em: 17 abr. 2023.