

Conceção e modelação de um esquema híbrido de base de dados: transacional e analítico

Design and modeling of a hybrid database schema: transactional and analytical

André Pinheiro Costa, Universidade do Minho, Portugal, a69195@alunos.uminho.pt

Jorge Oliveira e Sá, Universidade do Minho, Centro Algoritmi, Portugal, jos@dsi.uminho.pt

Resumo

Os sistemas de bases de dados em contexto organizacional estão segmentados em duas categorias: bases de dados operacionais e bases de dados analíticas. Como consequência da constante evolução tecnológica, vários esforços foram levados a cabo com o objetivo de reavaliar o paradigma atual. Assim, a conceção de bases de dados híbridas capazes de suportar o processamento transacional e analítico é vista como possível e viável. No entanto, é clara a presença de uma lacuna no que diz respeito à existência de abordagens que demonstrem como modelar esquemas conceptuais para este tipo de bases de dados. Neste póster é apresentada uma conceção e modelação de um esquema lógico de uma base de dados híbrida.

Palavras-chave: OLAP, OLTP, HTAP, Esquema conceptual de dados

Abstract

Database systems in organizational context are segmented in two categories: operational databases and analytical databases. Several efforts have been undertaken to re-evaluate the current paradigm, based on the constant technological evolution. Thus, the design of hybrid databases, capable of supporting transactional and analytical processing is possible and feasible. Even so, it is clear a gap about the existence of approaches regarding the modelling of conceptual schema for this type of databases. In this poster, is presented a logical schema for a hybrid database.

Keywords: OLAP, OLTP, HTAP, Conceptual data modelling

1. DESCRIÇÃO DO TRABALHO (PÓSTER)

A tecnologia de base de dados está no centro de vários tipos de sistemas de informação organizacionais. No contexto atual, é clara a segmentação destes sistemas em On-Line Transaction Processing (OLTP) e On-Line Analytical Processing (OLAP).

Durante muito tempo considerou-se boa prática manter os sistemas transacionais separados dos sistemas analíticos, com o intuito de prevenir que o processamento analítico condicionasse o desempenho do processamento operacional.

Recentemente, e como consequência da evolução da tecnologia, a validade deste paradigma foi alvo de uma reavaliação por parte de investigadores e profissionais. A possibilidade de coexistência do ambiente transacional e do ambiente analítico numa única base de dados é vista, nos dias de hoje, como uma possibilidade real e viável (Bog, 2014; Bog, Plattner & Zeier, 2011). Este novo paradigma é denominado de Hybrid Transactional Analytical Processing (HTAP) (Pezzini, Feinberg, Rayner & Edjlali, 2016).

Este novo paradigma traz como benefícios a: (i) diminuição da complexidade técnica e simplificação da arquitetura; (ii) diminuição da latência e; (iii) eliminação, ou pelos menos diminuição, da duplicação de dados (Pezzini, et al., 2016).

São vários os trabalhos desenvolvidos que tratam desta problemática, contudo, todos eles têm como foco a conceção/implementação física (Kemper & Neumann, 2011). Posto isto, e apesar dos esforços levados a cabo por diversos autores para providenciar uma base científica para este novo paradigma, é clara a existência de uma lacuna relativa à conceção e modelação do modelo lógico de uma base de dados híbrida.

Procurando constituir um primeiro passo para a diminuição desta lacuna, irá ser realizado um projeto, como prova de conceito, para a criação de um repositório HTAP de informação de alunos do 3º ciclo da Escola de Engenharia da Universidade do Minho. Este repositório HTAP deverá ser capaz de suportar o processamento operacional diário da organização, bem como disponibilizar um conjunto de ferramentas para uma análise detalhada, a partir de diferentes perspetivas, dos mais diversos aspetos relativos tanto a alunos como a cursos.

Espera-se com a realização deste projeto desenvolver uma abordagem para a criação de esquemas lógicos para bases de dados híbridas HTAP. Este esquema de dados HTAP deverá ser capaz de suportar o processamento de transações (componente operacional – OLTP) e, simultaneamente, a realização de análises sobre os dados (componente analítica - OLAP).

A abordagem Hybrid Transactional Analytical Schema (HTAS) toma como base o modelo relacional e o modelo dimensional. A modelação relacional é comumente encontrada em implementações de bases de dados operacionais nas organizações segundo a 3ª Norma Formal. Já a modelação dimensional está presente nas bases de dados que suportam os sistemas analíticos organizacionais.

Esta abordagem é constituída por um conjunto de regras que, tendo como *input* o modelo relacional e modelo dimensional, orienta e regula a conceção do modelo híbrido capaz de suportar ambos os ambientes. As regras possuem especificidade própria, sendo aplicadas ao modelo relacional, ao modelo dimensional ou a ambos. Um exemplo é a regra relativa à dimensão tempo. Esta regra é específica do modelo dimensional e obriga a inclusão desta dimensão no modelo híbrido com o objetivo de garantir a característica de historicidade do sistema analítico.

Tendo em consideração o progresso do projeto, neste documento ainda não é possível apresentar o conjunto final de regras que integram a abordagem HTAS. Esse conjunto estará presente na versão final do póster.

REFERÊNCIAS

- Bog, A. (2014). Benchmarks for Transaction and Analytical Processing Systems. In *Benchmarking Transaction and Analytical Processing Systems* (pp. 45–62). Springer.
- Bog, A., Plattner, H., & Zeier, A. (2011). A mixed transaction processing and operational reporting benchmark. *Information Systems Frontiers*, 13(3), 321–335.
- Kemper, A., & Neumann, T. (2011). HyPer: A hybrid OLTP&OLAP main memory database system based on virtual memory snapshots. In *Data Engineering (ICDE), 2011 IEEE 27th International Conference on* (pp. 195–206). IEEE.
- Pezzini, M., Feinberg, D., Rayner, N., & Edjlali, R. (2016). *Real-time Insights and Decision Making using Hybrid Streaming, In-Memory Computing Analytics and Transaction Processing*.