

Reestruturação de Estruturas Multidimensionais de Dados em Tempo Útil

Jorge Loureiro

jloureiro@di.estv.ipv.pt

Departamento de Informática
Instituto Superior Politécnico de Viseu
Escola Superior de Tecnologia de Viseu
Campus Politécnico de Repeses
3500 Viseu
PORTUGAL

Orlando Belo

obelo@di.uminho.pt

Departamento de Informática
Escola de Engenharia
Universidade do Minho
Campus de Gualtar
4710-057 Braga
PORTUGAL

Resumo. O sucesso ou insucesso de uma organização empresarial está ligado intimamente às suas actividades de decisão que, para serem sustentadas de forma credível, dependem da análise em tempo útil da informação que têm à sua disposição, independentemente do local onde esta está armazenada. A globalização das organizações tem provocado uma grande mobilidade dos seus recursos humanos e, obviamente, a distribuição geográfica das suas competências decisórias. Ao longo do tempo, a conjugação destes factores fez evoluir os sistemas de processamento analítico para estruturas distribuídas. A distribuição do processamento e dos sistemas de dados por vários nós, surge assim como uma solução interessante, porque permite um crescimento sustentado da capacidade de processamento dos equipamentos, a custos controlados, possibilita a diminuição de custos de comunicação, ao se aproximar as fontes da sua real utilização, e um aumento significativo na autonomia do próprio sistema, já que elimina a dependência de uma única fonte, de um único sistema. Todavia, esta solução também tem os seus custos, particularmente em aspectos relacionados com questões de administração dos sistemas instalados e na gestão dos dados. A administração da distribuição dos dados assume nestes sistemas um papel preponderante. Esta deve ser realizada de forma a ir ao encontro do objectivo primordial de todo o sistema: a minimização do tempo de consulta percebido pelos agentes de decisão. Este objectivo deve ser alcançado com o menor custo possível de processamento, comunicação e manutenção, tendo em conta eventuais constrangimentos relevados pelo espaço de armazenamento disponível. A esta problemática, que poderemos denominar de problema de selecção de cubos distribuídos, pode ser dada uma resposta sob a forma da minimização de uma função de custos que permita propor uma distribuição mais eficiente para uma base de dados multidimensional, pelos diversos nós que integrem um sistema de processamento analítico de forma mais atempada e automática. Esta função terá como efeito imediato a aproximação das estruturas multidimensionais dos locais onde estão os seus utilizadores, conhecendo-se as suas necessidades em termos analíticos. Este artigo apresenta uma proposta para um sistema analítico que, utilizando a história das interrogações analíticas colocadas pelos utilizadores num sistema, terá a capacidade de propor uma distribuição espacial dinâmica das suas estruturas multidimensionais, tendo como base os diversos nós de processamento disponíveis – servidores OLAP - e utilizando uma abordagem evolucionária, atendendo às suas dependências características no que se refere às hierarquias dimensionais e agregações inter-dimensionais representadas pelo lattice do cubo OLAP, aproximar essas estruturas de quem realmente as utilize e reestruturá-las, em tempo útil, conforme as necessidades dos seus utilizadores. Desta forma, pretende-se não só rentabilizar o espaço ocupado no sistema analítico pelos diferentes cubos materializados, como também melhorar, de forma significativa, o tempo de acesso à informação neles contida.

Palavras-Chave: Sistemas de Processamento Analítico, Estruturas Multidimensionais de Dados, Optimização de Interrogações em Sistemas de Processamento Analítico, Alocação e Distribuição de Estruturas Multidimensionais, Avaliação de Perfis de Interrogação de Dados e Computação Co-Evolucionária.