

OLAP

OLAP é uma ferramenta de Business Inteligente utilizada para apoiar as empresas na análise de suas informações, visando obter novos conhecimentos que são empregados na tomada de decisão. O termo OLAP refere-se a um conjunto de ferramentas voltadas para acesso e análise ad hoc de dados, com o objetivo final de transformar dados em informações capazes de dar suporte as decisões gerenciais de forma amigável e flexível ao usuário e em tempo hábil. É um software cuja tecnologia de construção permite aos analistas de negócios, gerentes e executivos analisar e visualizar dados corporativos de forma rápida, consistente e principalmente interativa.

A funcionalidade OLAP é inicialmente caracterizada pela análise dinâmica e multidimensional dos dados consolidados de uma organização permitindo que as atividades do usuário final sejam tanto analíticas quanto navegacionais.

As ferramentas OLAP (do inglês, Online Analytical Processing) são geralmente desenvolvidas para trabalhar com banco de dados desnormalizados. Essas ferramentas são capazes de navegar pelos dados de um Data Warehouse, possuindo uma estrutura adequada tanto para a realização de pesquisas como para a apresentação de informações.

Características Das Ferramentas OLAP

Uma das características que devem estar presentes em ferramentas OLAP é a capacidade de efetuar algumas operações, como:

- **Drill Across:** ocorre quando o usuário pula um nível intermediário dentro de uma mesma dimensão. Por exemplo, a dimensão tempo é composta por ano, semestre, trimestre, mês e dia. A operação Drill Across é executada quando o usuário passa de ano direto para trimestre ou mês;
- **Drill Down:** ocorre quando o usuário aumenta o nível de detalhe da informação, diminuindo a granularidade (A granularidade determina quais os tipos de consultas podem ser feitas no DW. Ela influencia diretamente na velocidade do acesso às informações e no volume de dados armazenados).
- **Drill Up:** é o contrário do Drill Down, ocorre quando o usuário aumenta a granularidade, diminuindo o nível de detalhamento da informação;
- **Drill Throught:** ocorre quando o usuário passa de uma informação contida em uma dimensão para uma outra. Por exemplo: Inicia na dimensão do tempo e no próximo passo analisa a informação por região;
- **Dlice and Dice:** é uma das principais características de uma ferramenta OLAP. Como a ferramenta OLAP recupera o microcubo (No OLAP, as informações são armazenadas em cubos multidimensionais, que gravam valores quantitativos e medidas, permitindo visualização através de diversos ângulos. Estas medidas são organizadas em categorias descritivas, chamadas de dimensões e formam, assim, a estrutura do cubo), surgiu a necessidade de criar um módulo, que se convencionou de Slice and Dice, para ficar responsável por trabalhar esta informação. Ele serve para modificar a posição de uma informação, trocar linhas por colunas de maneira a facilitar a compreensão dos usuários e girar o cubo sempre que tiver necessidade.

Arquitetura de ferramentas OLAP

Apesar de obedecer a uma estrutura cliente/servidor multiusuário, as ferramentas OLAP podem ser implementadas de diversas formas, classificadas em cinco tipos a seguir:

- **MOLAP** (Multidimensional On Line Analytical processing);
- **ROLAP** (Relational On Line Processing);
- **HOLAP** (Hybrid On Line Analytical Processing);
- **DOLAP** (Desktop On Line Analytical Processing);
- **WOLAP** (Web On Line Analytical Processing).

O local de armazenamento pode ser um banco de dados multidimensional, como nos casos de MOLAP e ROLAP; um banco de dados relacional, como em HOLAP e DOLAP, ou arquivos em base de dados locais ou mesmo em memória de estações clientes, como em DOLAP e WOLAP. Na arquitetura MOLAP os dados ficam armazenados em um banco de dados multidimensional, onde o servidor MOLAP atua e o usuário trabalha, monta e manipula os dados diferentes no servidor.

Os dados de um banco multidimensional são armazenados em um espaço menor que o utilizado para armazenar os mesmos dados em um banco de dados relacional. No banco multidimensional, os dados são mantidos em estruturas de dados do tipo array de maneira a prover um melhor desempenho ao acessá-los. Além de ser uma arquitetura rápida uma outra vantagem é o rico e complexo conjunto de funções de análises presentes nos bancos multidimensionais.

Uma de suas limitações é a possibilidade dos dados serem esparsos (nem todo cruzamento das dimensões contém dados), ocorrendo a chamada explosão de armazenamento de dados, ou seja, um imenso banco de dados multidimensional contendo poucos dados armazenados. Outras limitações dessa ferramenta estão relacionadas ao fato dos bancos multidimensionais serem sistemas proprietários que não seguem padrões, ou seja, cada desenvolvedor cria a sua própria estrutura para o banco e as próprias ferramentas de suporte (CARVALHO, 2004).

Já a arquitetura ROLAP é uma simulação da tecnologia OLAP feita em banco de dados relacionais que, por utilizar a estrutura relacional, possui a vantagem de não restringir o volume de armazenamento de dados. Essa ferramenta não utiliza cubos pré-calculados como a MOLAP. À medida que o usuário monta sua consulta em uma interface gráfica, a ferramenta acessa os metadados ou quaisquer outros recursos que possua, para gerar uma consulta SQL (Structured Query Language).

A sua principal característica é a possibilidade de fazer qualquer consulta, atendendo melhor os usuários que não têm um escopo de análise bem definido. Essa ferramenta tem a vantagem de utilizar tecnologia estabelecida, de arquitetura aberta e padronizada, beneficiando-se da diversidade de plataformas, escalabilidade e paralelismo de hardware. Sua desvantagem é o conjunto pobre de funções para análises dimensionais e o baixo desempenho da linguagem SQL na execução de consultas pesadas.

A arquitetura HOLAP, ou processamento híbrido, está se tornando a mais popular para os produtos atuais, porque consegue combinar a capacidade e a escalabilidade das ferramentas ROLAP com o desempenho superior dos bancos de dados multidimensionais. Por exemplo, suponha uma base de 50.000 clientes distribuídos em 500 cidades, 23 estados, 5 regiões e um total geral. Até o nível de cidades o armazenamento multidimensional resolveria as consultas para levantar totais de vendas. Porém, se fosse necessário consultar o total de vendas de um determinado cliente, o banco de dados relacional responderia com muito mais rapidez à solicitação.

Essa situação é típica para a indicação da arquitetura HOLAP (OLIVEIRA, 2002). Uma outra arquitetura é a DOLAP que é uma arquitetura desktop do OLAP, ou seja, é uma ferramenta para usuários que possuam uma cópia da base multidimensional ou de um subconjunto dela ou ainda, que queiram acessar um repositório de dados central localmente.

O usuário ao acessar este repositório, dispara uma instrução SQL e acessa os cubos já existentes no banco de dados multidimensional residente no servidor OLAP e obtém de volta um para ser analisado em sua estação de trabalho.

A vantagem dessa arquitetura é a redução da sobrecarga no servidor de banco de dados uma vez que todo o processamento OLAP acontece na máquina cliente e a desvantagem é o tamanho do microcubo que não pode ser muito grande, caso contrário, a análise pode ser demorada e a máquina do cliente não suportar. Finalmente, a arquitetura WOLAP é a utilização de uma ferramenta OLAP a partir de um browser.

Essa arquitetura tem duas tecnologias que estão em constante evolução, a primeira é a Web e a segunda são as ferramentas OLAP. A diferença desta ferramenta para as outras é que ela utiliza a Web, facilitando assim, a distribuição da ferramenta, o acesso remoto dos dados a serem analisados e a utilização da aplicação independente de plataforma.

1. Browser é o software usado para a navegação na Internet (navegador).

2. Atualmente o uso de Web browsers para acesso OLAP está sendo muito divulgado, entretanto, poucos são os sites que utilizam esta ferramenta. O WOLAP, segundo pesquisas realizadas, será a chave para aplicações na Internet e deverá oferecer um caminho simples e barato no acesso a dados do DW.

A tecnologia OLAP permite que o usuário trabalhe com cálculos complexos, através de consultas com maior flexibilidade e funcionalidade. Seu uso proporciona um bom andamento e um excelente gerenciamento de uma empresa. A essência das ferramentas OLAP é a rapidez, a sumarização e a análise flexível dos dados.

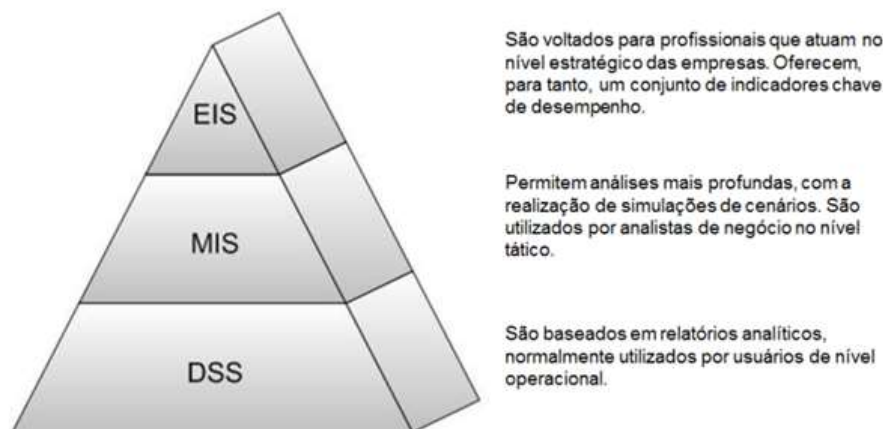
As aplicações bem sucedidas de OLAP tendem a aumentar a produtividade de gerentes, desenvolvedores e até da organização como um todo. As ferramentas OLAP aliadas à filosofia de Data Warehouse revolucionaram o modo de se fazer negócios, garantindo agilidade e flexibilidade nas investigações e análise de dados para os usuários finais, isto é, aqueles que realmente fazem parte do processo de tomada de decisões.

Os Fundamentos de Sistemas Analíticos

Nos últimos anos o termo Business Intelligence (BI) tem sido largamente utilizado no mercado como sinônimo de sistemas analíticos, OLAP, cubos, entre outros. Embora essas denominações possam estar associadas entre si, são conceitualmente distintas.

A rigor, Business Intelligence pode ser obtido por qualquer artefato, seja tecnológico ou não, que permita a extração de conhecimento a partir de análises do negócio. Por razões óbvias, a efetividade destas análises será maior se os dados estiverem disponíveis de modo consistente e, preferencialmente, consolidado. Este é um dos objetivos dos Data Warehouses.

Soluções informatizadas de BI geralmente contém sistemas analíticos, que podem ser de diversos tipos, dependendo do objetivo das análises e do perfil do usuário, conforme ilustrado na Figura:



OLTP vs OLAP

Sistemas transacionais, também conhecidos como sintéticos ou ainda OLTP – Online Transactional Processing – são aqueles que, como o nome sugere, baseiam-se em transações. Alguns exemplos deste tipo de sistemas são:

- Sistemas Contábeis;
- Aplicações de Cadastro;
- Sistemas de Compra, Estoque, Inventário;
- ERPs, CRMs.

Os sistemas transacionais se caracterizam pela alta taxa de atualização, grande volumes de dados e acessos pontuais, ou seja, pesquisas cujo resultado seja de pequeno volume (até milhares de linhas, mas preferencialmente menos).

Já os sistemas analíticos, ou OLAP – Online Analytical Processing – se caracterizam por fornecer subsídio para tomadas de decisão, a partir de análises realizadas sobre bases de dados históricas, por vezes com milhões de registros a serem totalizados.

A tabela abaixo sintetiza as principais diferenças entre sistemas transacionais e analíticos.

Característica	Sistemas Transacionais (OLTP)	Sistemas Analíticos (OLAP)
Atualizações	Mais freqüentes	Menos freqüentes
Tipo de Informação	Detalhes	Agrupamento
Quantidade de Dados	Poucos	Muitos
Precisão	Dados atuais	Dados históricos
Complexidade	Baixa	Alta
Consistência	Microscópica	Global
Exemplos	CRM, ERP, Supply Chain	MIS, DSS, EIS
Terminologia	Linhas e Colunas	Dimensões, Medidas e Fatos

Conforme ilustrado na tabela, o fato dos sistemas transacionais refletirem a situação atual de um determinado tipo de dado conduz todas as demais características, como:

- A realização de atualizações freqüentemente, mantendo os dados atuais;
- Informação detalhada com a maior granularidade possível (consistência microscópica);
- Pesquisas pontuais, portanto, de baixa complexidade, no tocante ao negócio (do ponto de vista técnico, a pesquisa pode ser bem elaborada).

Do mesmo modo, o fato das análises serem realizadas sobre dados históricos leva às seguintes características:

- Uma vez que os dados são históricos, as atualizações não precisam ser tão freqüentes. Por exemplo, numa comparação entre a produtividade de três filiais de uma empresa para um determinado produto nos últimos quatro meses, por mês, o dia de hoje ou mesmo ontem não é, em geral, de grande representatividade;
- As análises geralmente agrupam informações, sendo tais agrupamentos mais importantes neste contexto do que os dados detalhados. No exemplo do item anterior, o importante é a produção conjunta mensal, e não a produção de uma unidade particular do produto analisado.

Os diferentes tipos de sistemas também sugerem diferentes abordagens técnicas, seja na forma de armazenamento ou de busca.

Arquiteturas OLAP

ROLAP (Relational On Line Analytical Processing)

A consulta é enviada ao servidor de banco de dados relacional e processada no mesmo, mantendo o cubo no Servidor.

MOLAP (Multidimensional On Line Analytical Processing) processamento realizado em um servidor multidimensional.

HOLAP (Hybrid On Line Analytical Processing)

É uma mistura de tecnologias onde há uma combinação entre ROLAP e MOLAP.

Alicerce Relacional

