



**PÓS-GRADUAÇÃO LATO SENSU**

**Curso: Banco de Dados**

*Disciplina: Laboratório de Data Warehouse e Business Intelligence*  
*Professor: Fernando Zaidan*

Unidade 2  
2012




Crédito dos Slides: Clever Junior



**História e Apresentação**



**PENTAHO**

Atualmente é a plataforma de inteligência de negócios (BI) que utiliza código aberto mais popular e poderosa no mundo.

A sede da empresa está localizada em Orlando na Flórida, com escritórios em São Francisco, California e por toda a Europa.



**História**

De acordo com o site oficial, o Pentaho nasceu do desejo de se alcançar uma mudança positiva no mercado de análise de negócios dominada por megavendedores burocráticos, que ofereciam produtos "pesados" baseados em plataformas antigas, direcionadas as suas aplicações já existentes e às custas de suas capacidades de inovação na área de BI.

Sob esta perspectiva, cinco veteranos da indústria (por isto o nome Penta) fundaram o Pentaho em 2004.



**Origem do nome Pentaho**

Além do significado da palavra penta, alega-se que o nome surgiu do índios Pentaho que viveram na Flórida nos séculos XVIII e XIX e supostamente foram responsáveis pela domesticação do peixe-boi das Índias Ocidentais, até então animais selvagens e agressivos, que de acordo com alguns historiadores chegaram a tornar o comércio e comunicação ao longo das águas da Flórida praticamente impossível.



## Soluções

- Geração de relatórios empresariais;
- ETL (Extração, Transformação e Carga);
- Análise de informações (OLAP);
- Painéis para controle (Dashboards);
- Mineração de Dados (Data Mining).

7

## Vantagens

- Licença não tem custo no Pentaho;
- Indicadores com acesso pela web não importando o SO (Windows, Mac, Linux, iPhone);
- Customização dos Indicadores de Negócio (Dashboards, Cubos e Relatórios).

8

## Desvantagens

- Documentação insuficiente;
- Ainda sem suporte;
- Referências (poucas) na Internet.

9

## Licença

- Licença Pública da Pentaho (PPL);
- É distribuída como código aberto;
- Esta licença é similar à Licença Pública do Mozilla V1.1.

10

## Custos de versões

- Possui dois tipos:
- => Pentaho CE (Community Edition) = Gratuita;  
 => Pentaho EE (Enterprise Edition) = Paga.

11

## Versões Free

- CE: Relatórios Básicos, Análise Simples, Integração de dados, Conectividade com SGBD, Instalar, Implantar, Administração, Segurança, DashBorads, Interactive Reporting, Análise interativa, Data Discovery e Visualization.
- Paga: EE – na casa de dezena de milhares de dólares – inclui suporte – atualizações – todos os recursos (teste por 30 dias).

12

### Quem utiliza no Brasil?



13

### Quem utiliza no mundo?



14

### Conex o com Banco de Dados



15

### Ferramentas do Pentaho

- Pentaho Data Integration - Kettle (ETL);
- Cubos OLAP;
- Analysis View - Mondrian;
- Pentaho Reporting;
- JPIVOT.

• A partir da vers o 3.5 vem no bi-server o Analysis View e o Reporting. As outras ferramentas precisam do download individual.

16

### Kettle (ETL)

- Tamb m conhecido como PDI (Pentaho Data Integration).
- L  e escreve em mais de 10 formatos de SGBD, como Oracle, PostgreSQL, SQL Server, etc.
- Importa arquivos de texto, planilhas de Excel e base de dados.



17

### Cubos OLAP

OLAP: (On-Line Analytical Processing) M todo de Processamento Analtico On-Line.

Cubos OLAP: o modelo do BD dimensional n o   suficiente para o Pentaho identificar os cubos. Ent o,   necess rio criar arquivos XML e definir quais os cubos, dimens es, conex es de dados e tabelas associadas.

Al m disto, a partir destes arquivos XML, gerar um XACTION – arquivo identificado pelo Pentaho.

18

### Ferramentas para Cubos

**Schema Workbench (tem integração com o Pentaho) :** É um designer de interface que permite criar e testar esquemas cubo Mondrian OLAP visualmente.

**Cube Designer :** Ferramenta open source para criação de cubos e bancada para edição, que pode ser baixada no site do Pentaho.

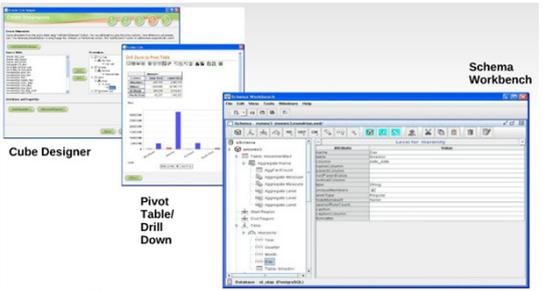
**Mondrian :** É um Servidor OLAP adquirido pelo Pentaho, escrito em Java e open source.

**Pivot Table :** Uma tabela dinâmica de sumarização de dados extraídos.



19

### Ferramentas para Cubo

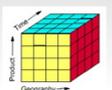



20

### Ferramentas para Cubo - Mondrian

A dimensional model (logical)

- Cubes & virtual cubes
- Shared & private dimensions
- Calculated measures in cube and in query language
- Parent-child hierarchies



... mapped onto a star/snowflake schema (physical)

- Fact table
- Dimension tables
- Joined by foreign key relationships

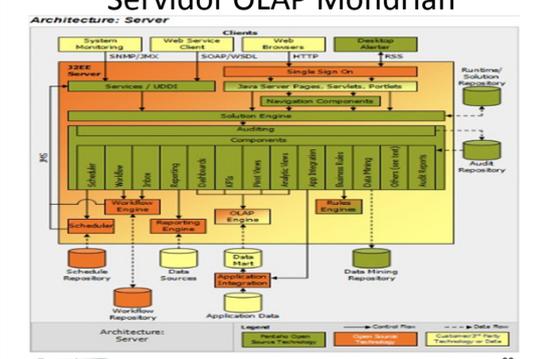


Fonte: Apresentação Julian Hyde



21

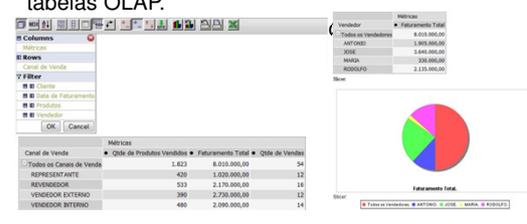
### Servidor OLAP Mondrian




22

### JPivot

Definição : É uma biblioteca de identificadores (tags), responsável pelo desenvolvimento de tabelas OLAP.




23

### Relatórios

Definição:

- Também conhecido como PRD (Pentaho Report Designer);
- Auxilia na criação de relatórios.

Ferramentas:

- Report Designer (integrado ao Pentaho).
- Jasper Reorts
- BIRT



24

### Relatórios

The image shows several Pentaho reporting tools. At the top left is Eclipse Birt, a web-based reporting tool. Below it is Ad-Hoc Reporting. To the right is Report Designer, a tool for creating reports. At the bottom right is Jfreereport, a Java-based reporting tool. The background shows a dashboard with various charts and tables.

25

### Data Mining Weka

- Criada em 1993, "University of Waikato", Nova Zelândia.
- Desenvolvido em Java.
- Comprada pela Pentaho em 2006.

Definição:

- Conjunto de ferramentas para aprendizado de máquinas de mineração de dados.
- Ajuda a melhorar desempenhos futuros através de análise preditivas.

26

### Data Mining

The image shows the Weka Knowledge Flow Environment. It features a central workspace with a Knowledge Flow diagram. The diagram includes nodes for DatafileLoader, AttributeSummarizer, ScatterPlotMatrix, CrossValidationFoldMaker, AttributeSelection, and Classifier. Arrows indicate the flow of data between these components. On the right, there is a visualization window showing a bar chart. The top menu includes Evaluation, Visualization, Datafiles, Filters, Classifiers, and DataSources.

27

### Dashboards

Definição:

- É uma das formas de apresentação dos dados (gráficos e relatórios).
- Usado muito em empresas para apresentação.
- Fácil acessibilidade.
- Apresentação em Navegadores.

- A versão Free possui o CDF (Community Dashboard Framework), porém requer conhecimento em HTML, JavaScript, JQuery e CSS (para facilitar, foi criado o Dashboard Framework), com drag and drop.

28

### Dashboards

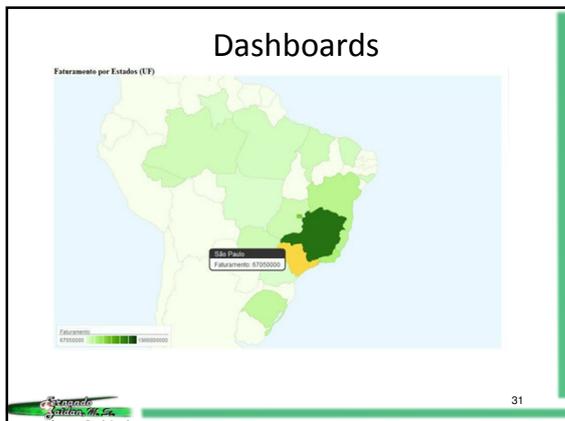
The image shows a Pentaho dashboard with multiple widgets. On the left, there is a map of Brazil. In the center, there is a pie chart titled 'Headcount Spending by Region'. To the right, there is a bar chart showing data for different regions. The dashboard also includes a table with data and a search bar at the top.

29

### Dashboards

The image shows a dashboard with a map on the left and a line chart on the right. The map displays a geographical area with a red location marker. The line chart shows data points over time, with a legend indicating 'Quantidade Vendida' and 'Previsão Vendida'. The chart has a grid and a y-axis labeled 'Médias em milhões de toneladas'.

30



### Perfil e Comunidades

A comunidade Pentaho é formada principalmente por empresas públicas, empresas privadas, universidades, alunos de graduação, alunos de pós-graduação e professores.

Lista de Discussões

- Pentaho Brasil (<http://pentahobrasil.org/>)
- Pentaho Community (<http://community.pentaho.com/>)

32

### Pentaho Staff

#### MANAGEMENT TEAM

 <b>Quentin Gallivan</b> Chairman and Chief Executive Officer <a href="#">Bio</a>	 <b>Richard Daley</b> Founder and Chief Strategy Officer <a href="#">Bio</a>	 <b>Doug Johnson</b> Executive VP and Chief Operating Officer <a href="#">Bio</a>
 <b>Rosanne Saccone</b> Chief Marketing Officer <a href="#">Bio</a>	 <b>Rod Squires</b> Executive Vice President Worldwide Sales <a href="#">Bio</a>	 <b>Jake Cornelius</b> Senior Vice President of Products <a href="#">Bio</a>

33

### Pentaho Staff

#### TECHNICAL LEADS

 <b>James Dixon</b> Founder and Chief Geek / CTO <a href="#">Bio</a>	 <b>Doug Moran</b> Founder and Big Data Product Manager <a href="#">Bio</a>	 <b>Marc Batchelor</b> Founder and Chief Engineer <a href="#">Bio</a>
 <b>Will Gorman</b> Vice President, Engineering <a href="#">Bio</a>	 <b>Matt Casters</b> Chief Architect, Product Project Founder <a href="#">Bio</a>	 <b>Dave Henry</b> Senior Vice President, Enterprise Solutions <a href="#">Bio</a>
 <b>Mark Hall</b> Senior Pentaho Data Mining Consultant / Engineer <a href="#">Bio</a>	 <b>Julian Hyde</b> Chief Architect, Pentaho, Analytics (OLAP) Microsoft Program Project Founder <a href="#">Bio</a>	 <b>Thomas Morgner</b> Chief Architect, Pentaho Reporting <a href="#">Bio</a>

34

### Referências

Pentaho História. Disponível em <<http://www.multilinguairchive.com/ma/enwiki/pt/Pentaho>> . Acesso em: 09 fev. 2012.

Pentaho. Disponível em <<http://www.pentaho.com>> . Acesso em: 19 fev. 2012.

Blog Professor Coruja. Disponível em <<http://blog.professorcoruja.com>>. Acesso em: 19 fev. 2012.

Pentaho Community. Disponível em <<http://community.pentaho.com>> . Acesso em: 19 fev. 2012.

Pentaho BI Suite – Enterprise Edition Intro. Disponível em <<http://www.youtube.com/watch?v=Vy2coaxJfCo>> . Acesso em: 20 fev. 2012.

35