Lezione 20

- Mondrian: Un OLAP Server Open Source

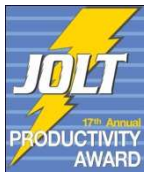
Creating Interactive OLAP Applications with MySQL Enterprise and Mondrian

Agenda

- Pentaho Introduction
- Mondrian features and architecture
- Schemas and queries
- olap4j
- Roadmap
- Case Studies
- Business Intelligence suite
- Q & A

Pentaho Introduction

- World's most popular enterprise open source BI Suite
  - 2 million lifetime downloads, averaging 100K / month
  - Founded in 2004: Pioneer in professional open source BI
- Management - proven BI and open source veterans
  - from Business Objects, Cognos, Hyperion, JBoss, Oracle, Red Hat, SAS
- Board of Directors – deep expertise and proven success in open source
  - Larry Augustin - founder, VA Software, helped coin the phrase "open source"
  - New Enterprise Associates – investors in SugarCRM, Xensource, others
  - Index Ventures – investors in MySQL, Zend, others
- Widely recognized as the leader in open source BI
  - Distributed worldwide by Red Hat via the Red Hat Exchange
  - Embedded in next release of OpenOffice (40 million users worldwide)



pentaho™
open source business intelligence™

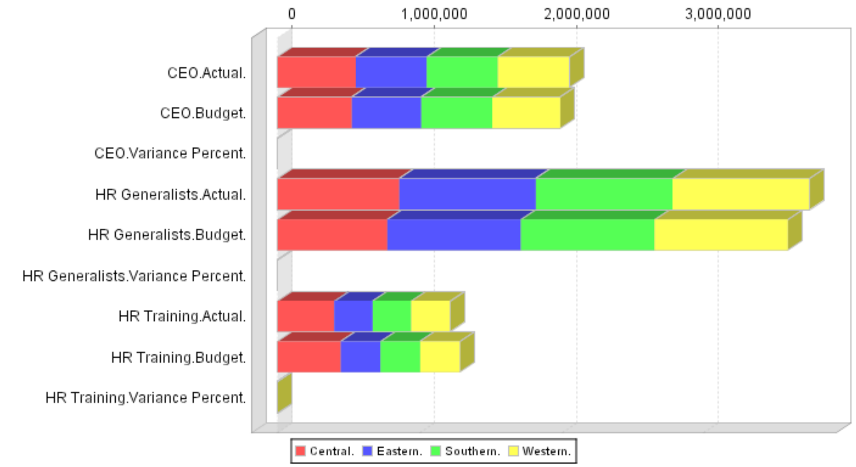
What is OLAP?

- View data "dimensionally"
  - i.e. Sales by region, by channel, by time period
- Navigate and explore
  - Ad Hoc analysis
  - "Drill-down" from year to quarter
  - Pivot
  - Select specific members for analysis
- Interact with high performance
  - Technology optimized for rapid interactive response

		Region			
Positions	Measures	Central	Eastern	Southern	Western
CEO	Actual	549,625.00	500,000.00	500,000.00	500,000.00
	Budget	522,250.00	488,750.00	498,750.00	478,750.00
	Variance Percent	-5.24%	-2.30%	-.25%	-4.44%
HR Generalists	Actual	856,190.00	961,000.00	961,000.00	961,000.00
	Budget	771,225.00	940,158.00	940,158.00	938,158.00
	Variance Percent	-11.02%	-2.22%	-2.22%	-2.43%
HR Training	Actual	397,473.00	271,200.00	271,200.00	271,200.00
	Budget	443,570.00	279,674.00	279,674.00	277,674.00
	Variance Percent	10.39%	3.03%	3.03%	2.33%

Slicer: [(All)=All Departments]

Slicer: (All)=All Departments



Mondrian features and architecture

Key Features

- On-Line Analytical Processing (OLAP) cubes
  - automated aggregation
  - speed-of-thought response times
- Open Architecture
  - 100% Java
  - J2EE
  - Supports any JDBC data source
  - MDX and XML/A
- Analysis Viewers
  - Enables ad-hoc, interactive data exploration
  - Ability to slice-and-dice, drill-down, and pivot
  - Provides insights into problems or successes

pentaho™
open source business intelligence

How Mondrian Extends MySQL for OLAP Applications

MySQL Provides

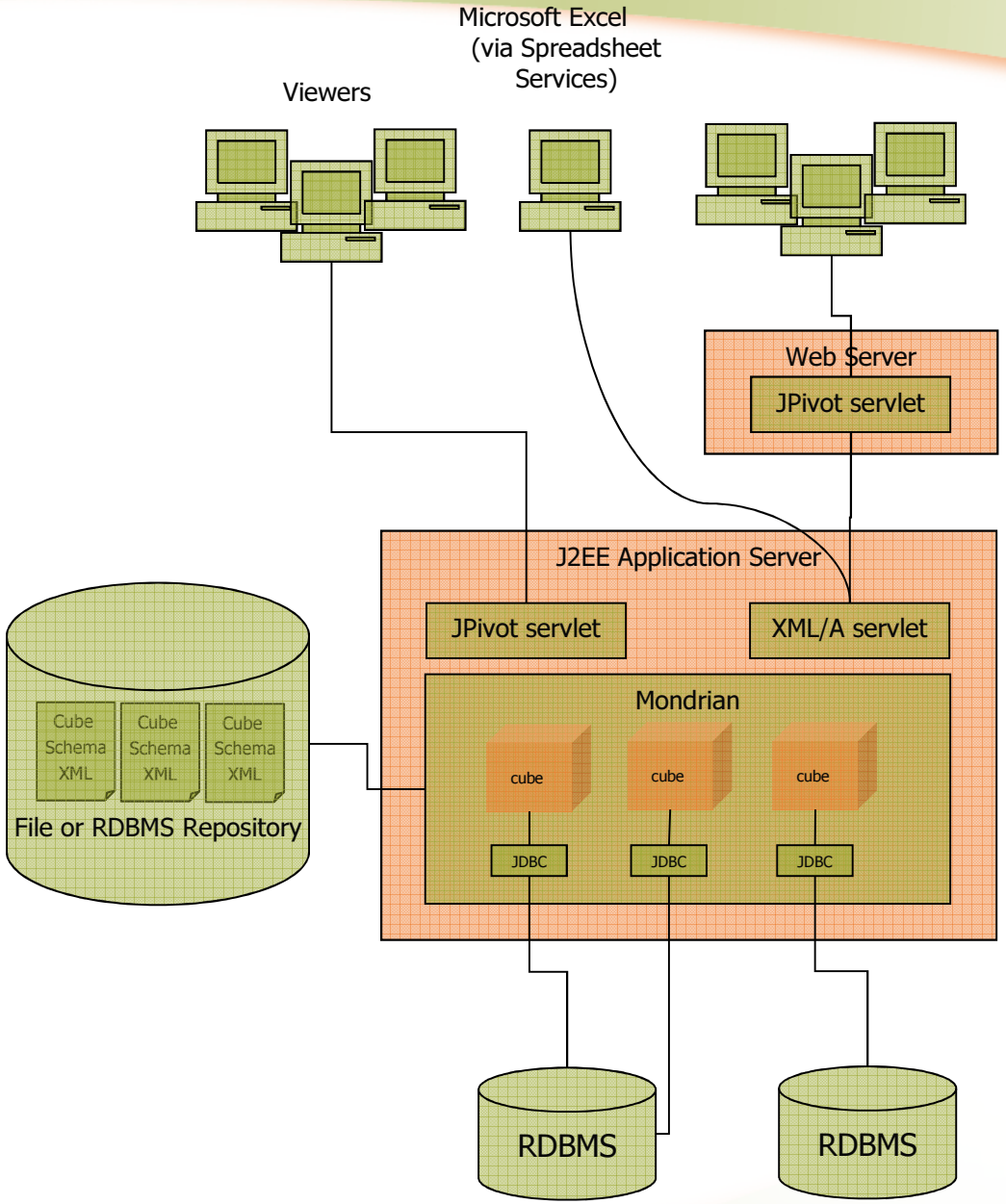
- Data storage
- SQL query execution
- Heavy-duty sorting, correlation, aggregation
- Integration point for all BI tools

Mondrian Provides

- Dimensional view of data
- MDX parsing
- SQL generation
- Caching
- Higher-level calculations
- Aggregate awareness

Open Architecture

- Open Standards (Java, XML, MDX, XML/A, SQL)
- Cross Platform (Windows & Unix/Linux)
- J2EE Architecture
  - Server Clustering
  - Fault Tolerance
- Data Sources
  - JDBC
  - JNDI



pentaho™
open source business intelligence

<mondrian/jpivot demonstration>

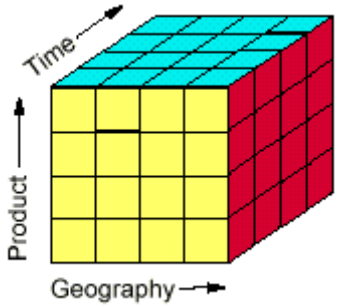
pentaho™
open source business intelligence

Schemas and queries

A Mondrian schema consists of…

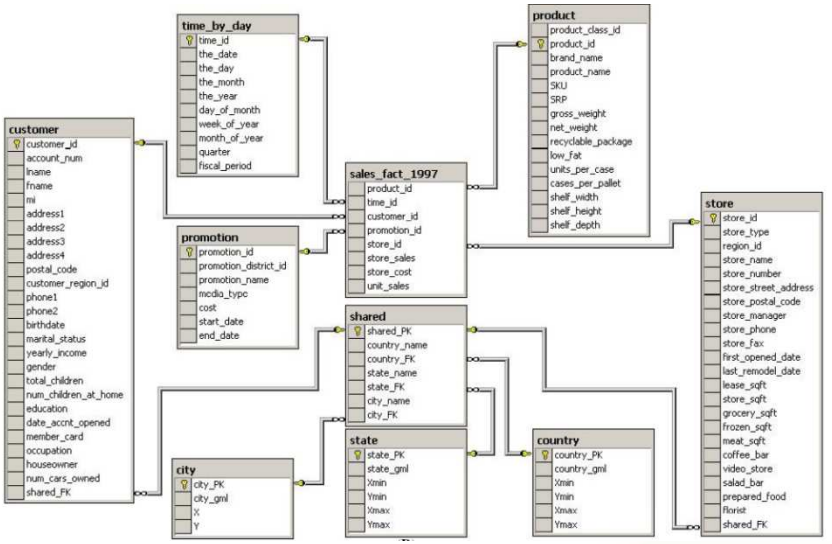
A dimensional model (logical)

- 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



pentaho™
open source business intelligence

Writing a Mondrian Schema

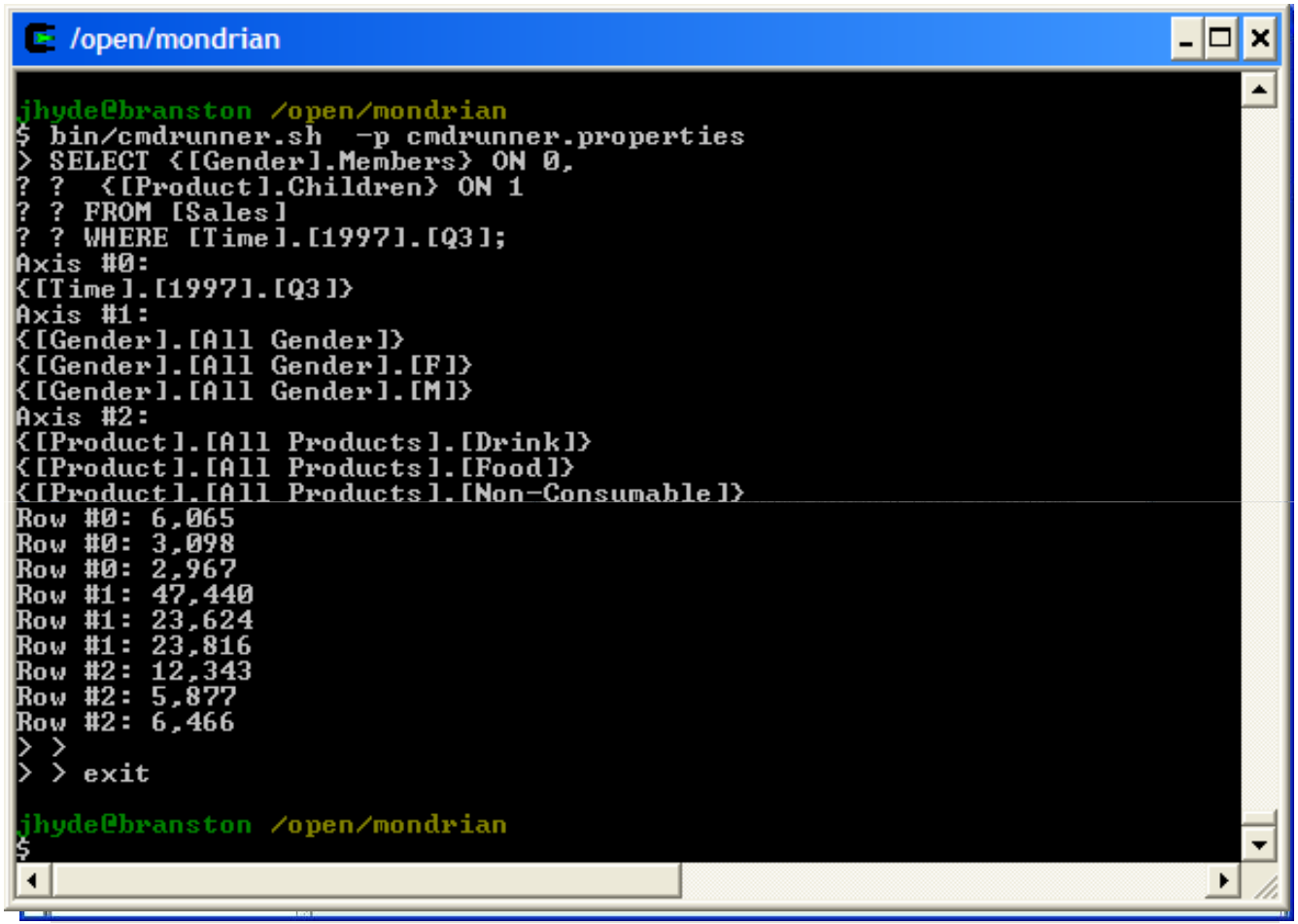
- Regular cubes, dimensions, hierarchies
- Shared dimensions
- Virtual cubes
- Parent-child hierarchies
- Custom readers
- Access-control

```
<!-- Shared dimensions -->
  <Dimension name="Region">
    <Hierarchy hasAll="true“
        allMemberName="All Regions">
      <Table name="QUADRANT_ACTUALS"/>
      <Level name="Region" column="REGION“
          uniqueMembers="true"/>
    </Hierarchy>
  </Dimension>
  <Dimension name="Department">
    <Hierarchy hasAll="true“
        allMemberName="All Departments">
      <Table name="QUADRANT_ACTUALS"/>
      <Level name="Department“
        column="DEPARTMENT“
        uniqueMembers="true"/>
    </Hierarchy>
  </Dimension>
```

(Refer to http://mondrian.pentaho.org/documentation/schema.php)

Tools

- Schema Workbench
- Pentaho cube designer
- cmdrunner



pentaho™
open source business intelligence

MDX – Multi-Dimensional Expressions

- A language for multidimensional queries
- Plays the same role in Mondrian's API as SQL does in JDBC
- SQL-like syntax

```
SELECT {[Measures].[Unit Sales]} ON COLUMNS,
    {[Store].[USA], [Store].[USA].[CA]} ON ROWS
FROM [Sales]
WHERE [Time].[1997].[Q1]
```

- … but un-SQL-like semantics

(Refer to http://mondrian.pentaho.org/documentation/mdx.php)

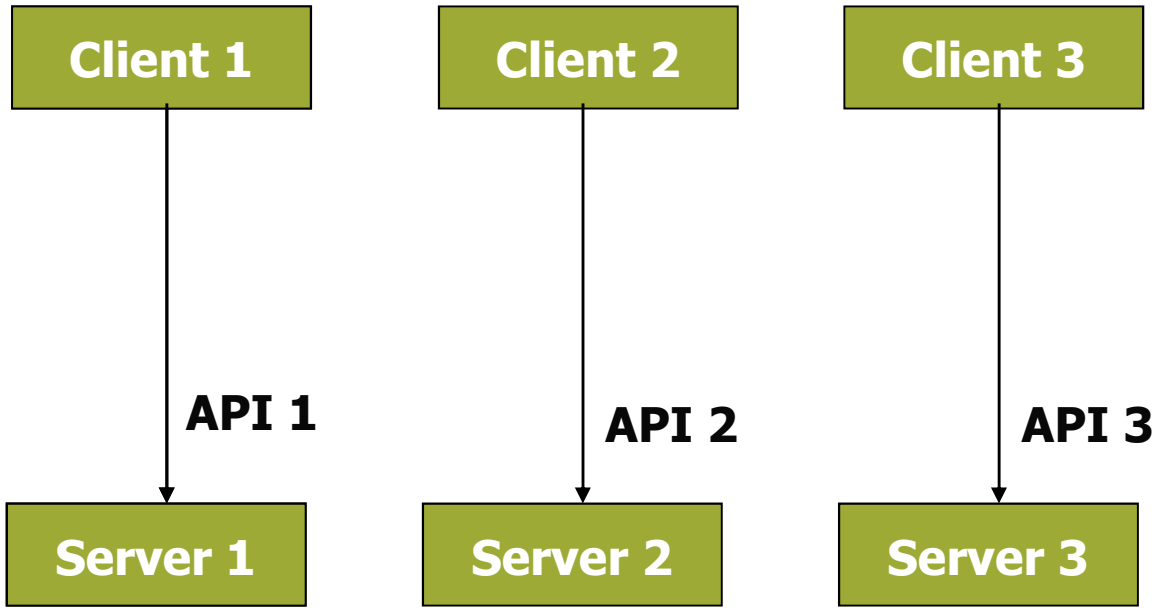
olap4j

OLAP APIs

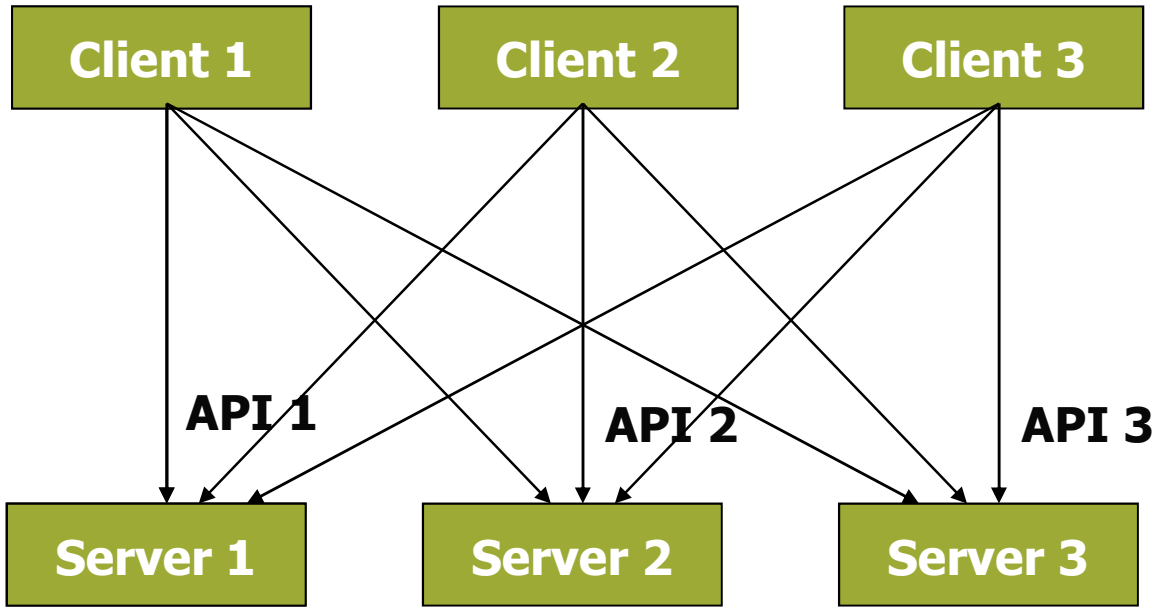
- OLAP APIs that failed:
  - OLAP Council's MDAPI and OLAPI
  - Sun's JOLAP
- OLAP APIs that succeeded:
  - Microsoft's OLE DB for OLAP, ADOMD, XMLA

- Mondrian has an API for creating running MDX queries:
  - Powerful and intuitive
  - Features the MDX language
  - Used by Mondrian's XMLA provider, JPivot, other clients
  - But it's Mondrian-only

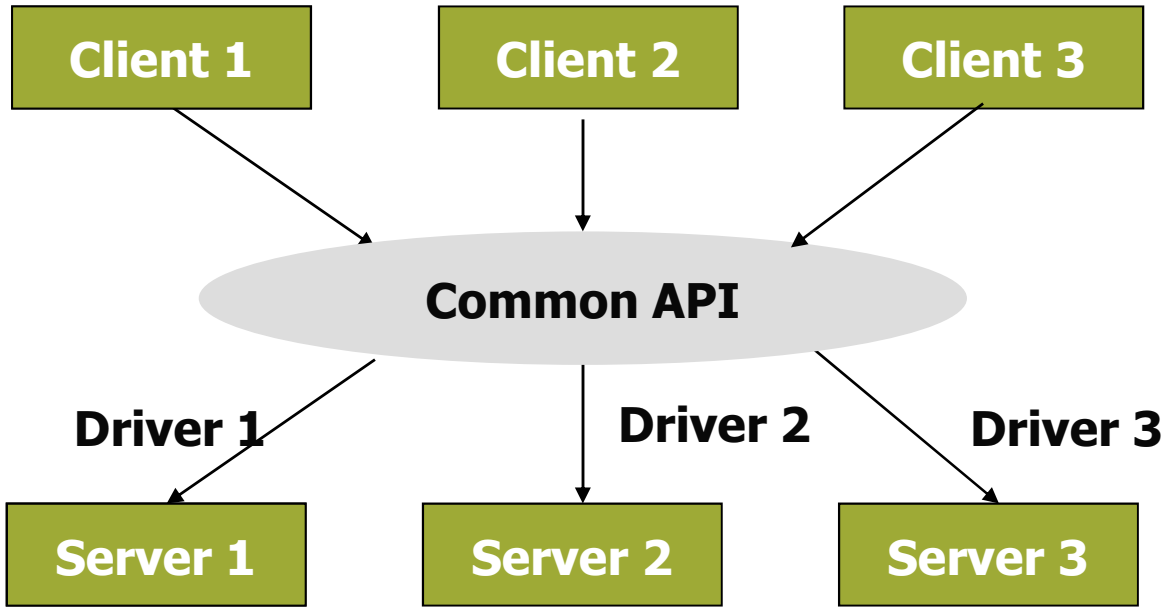
The problem with APIs



The problem with APIs #2



The problem with APIs – the solution



olap4j

- olap4j aims to be the "JDBC for OLAP"
  - An extension to JDBC
  - Also inspired by ADOMD.NET
  - Implementations for many OLAP servers
  - Enable one client to work against many servers
  - Break the 'lock-in'
    - Encourage more businesses to 'take a chance' on open-source

- Backed by:
  - Companies: Jedox, JasperSoft, Loyalty Matrix, LucidEra, Pentaho, Tensegrity, Tonbeller AG
  - Projects: Halogen, JPivot, JRubik, Mondrian, OpenI, PALO
  - Community at SourceForge.net

olap4j connecting to mondrian in-process

```
import org.olap4j.*;

Class.forName("mondrian.olap4j.MondrianOlap4jDriver");
OlapConnection connection =
  DriverManager.createConnection(
    "jdbc:mondrian:Jdbc=jdbc:mysql://localhost/foodmart;" +
    "JdbcUser=foodmart;JdbcPassword=foodmart;" +
    "Catalog=/WEB-INF/queries/FoodMart.xml;" +
    "Role='California manager'");
OlapConnection olapConnection =
  connection.unwrap(OlapConnection.class);
OlapStatement statement =
  olapConnection.createOlapStatement();

OlapResult result =
  statement.execute(
    "SELECT {[Measures].[Unit Sales]} ON COLUMNS,\n" +
    "  {[Product].Members} ON ROWS\n" +
    "FROM [Sales]");
```

olap4j connecting to SQL Server Analysis Services via XMLA

```
import org.olap4j.*;

Class.forName("org.olap4j.driver.xmla.XmlaOlap4jDriver");
OlapConnection connection =
  DriverManager.createConnection(
    "jdbc:xmla:Server=http://marmalade/xmla/msxisapi.dll;" +
    "Catalog=FoodMart;" +
    "Role='California manager'");

OlapConnection olapConnection =
  connection.unwrap(OlapConnection.class);
OlapStatement statement =
  olapConnection.createOlapStatement();

OlapResult result =
  statement.execute(
    "SELECT {[Measures].[Unit Sales]} ON COLUMNS,\n" +
    "  {[Product].Members} ON ROWS\n" +
    "FROM [Sales]");
```

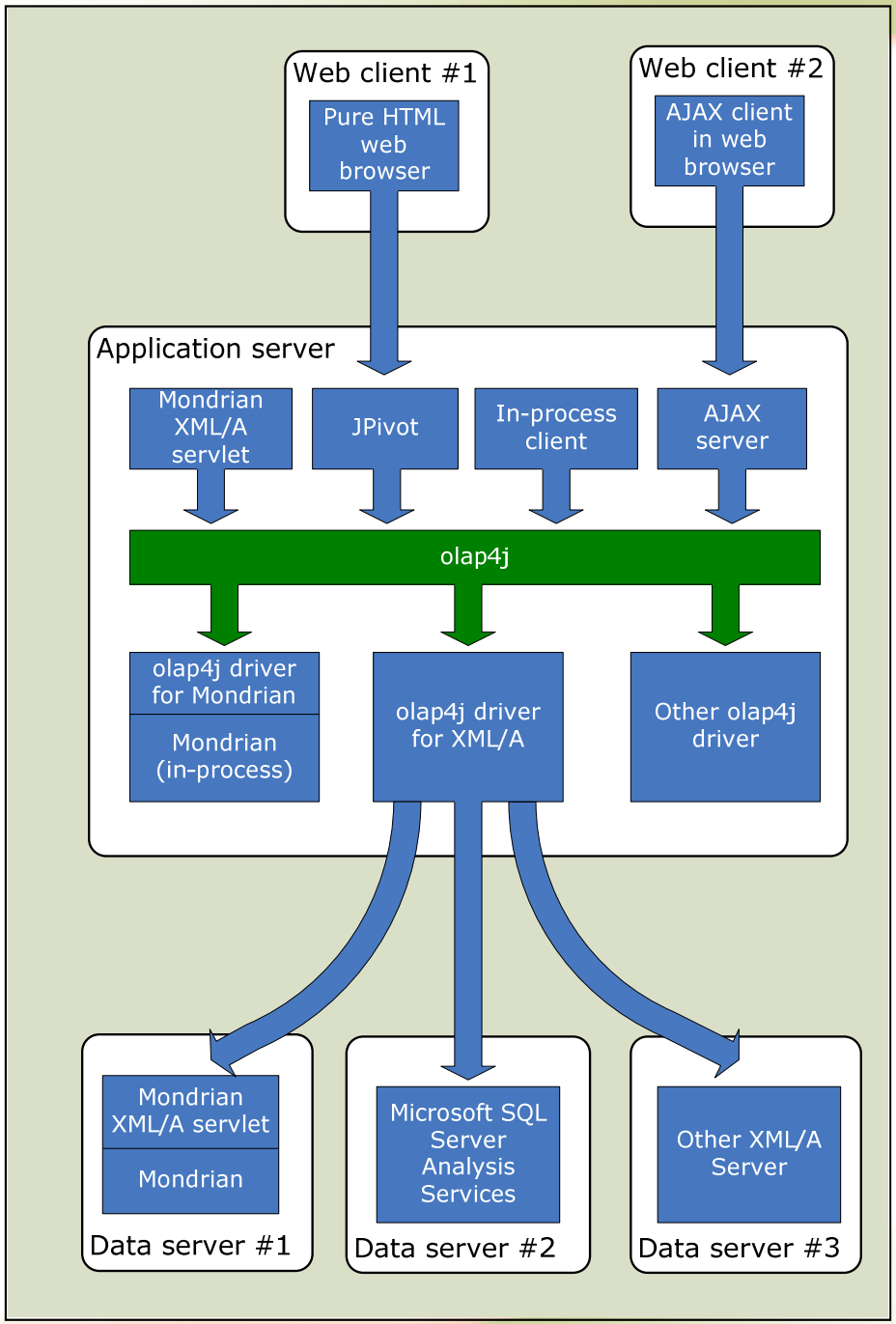
olap4j architecture

Client:

- In-process
- XMLA
- HTML
- AJAX

Server:

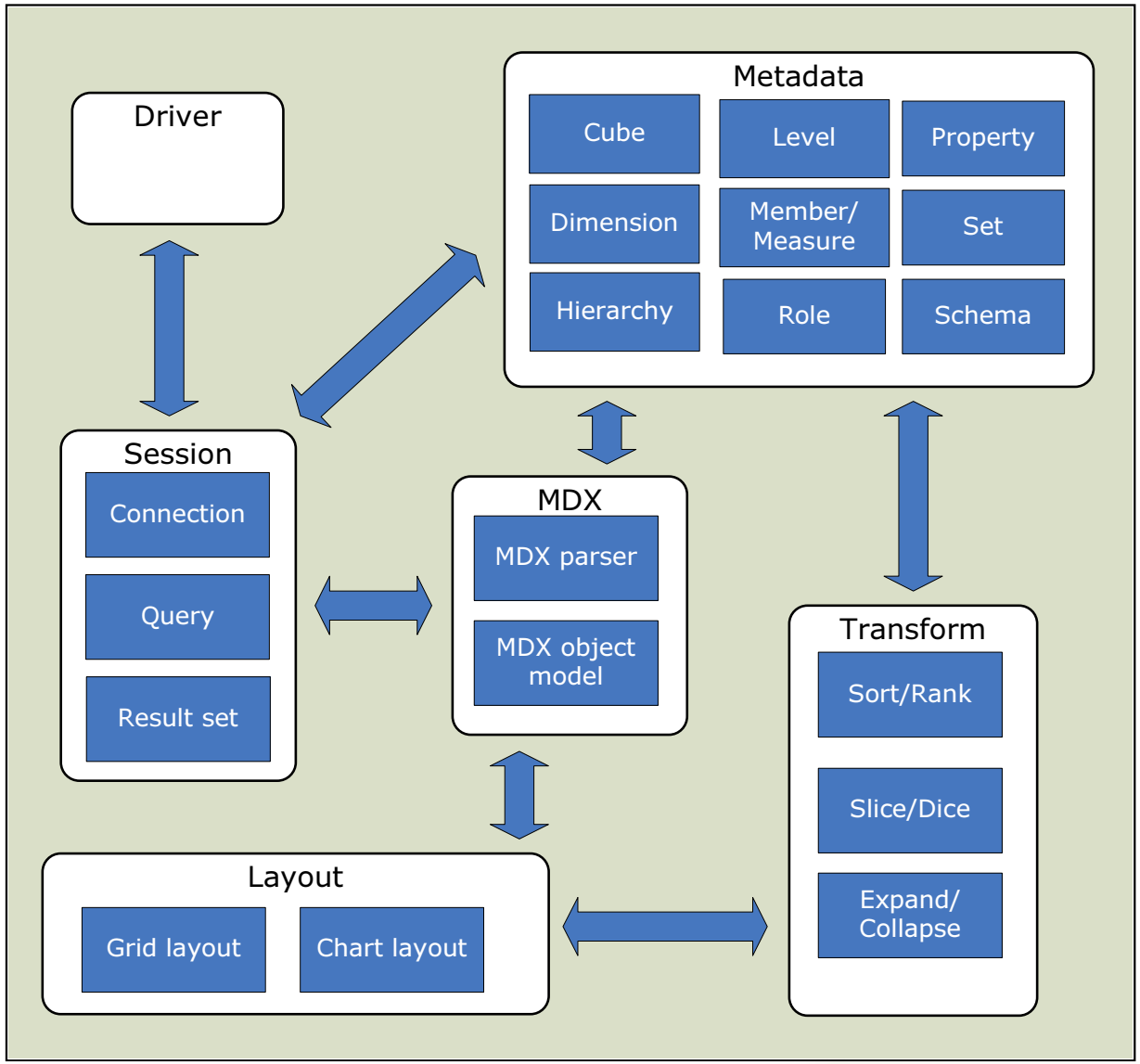
- Mondrian in-process
- Mondrian via XMLA
- Microsoft SSAS via XMLA
- Other OLAP server via XMLA
- Other OLAP server via dedicated olap4j driver



pentaho™
open source business intelligence™

olap4 components

- Driver management
- Session
- Metadata
- MDX
- Transform & layout



olap4j/mondrian roadmap

2006 September olap4j 0.5 (draft)

2007 October – olap4j 0.9 (beta)

2008 February – olap4j driver for XMLA

2008 July – olap4j 1.0

2007 August – mondrian 2.4

2008 March – mondrian 3.0 featuring olap4j API

2008 August – mondrian 3.1 featuring olap4j 1.0 API

mondrian 3.0 features

- olap4j API
- Rollup policy
- Aggregate roles
- Allow distinct-count measures to be aggregated
- Improve dimension sharing
- JNDI in connect string
- Over 90 new MDX functions

Halogen

- Pentaho incubator project
- Slice/dice OLAP client
- Built using GWT → AJAX
- Built on olap4j → portable across OLAP servers

<halogen demonstration>

pentaho™
open source business intelligence

Case studies

Case Study: Frontier Airlines

"The competition is intense in the airline industry and Frontier is committed to staying ahead of the curve by leveraging technology that will help us offer the best prices and the best flight experience.... [the application] fits right in with our philosophy of providing world-class performance at a low price."

FRONTIER AIRLINES

Frontier Airlines

Key Challenges

- Understanding and optimizing fares to ensure
  - Maximum occupancy (no empty seats)
  - Maximum profitability (revenue per seat)

Pentaho Solution

- Pentaho Analysis (Mondrian)
- Chose Open Source RDBMS and Mondrian over Oracle
- 500 GB of data, 6 server cluster

Results

- Comprehensive, integrated analysis to set strategic pricing
- Improved per-seat profitability (amount not disclosed)

Why Pentaho

- Rich analytical and MDX functionality
- Cost of ownership

pentaho™ open source business intelligence

Pentaho at Loma Linda University Health Care

"Pentaho Customer Support has been exceptional. This is a strategic application at LLUHC, and working with Pentaho has accelerated our deployment and improved our overall application delivery."

Leading Healthcare Provider

Key Challenges

- Providing analytics for billing and operations supporting 500,000 patients and 600 doctors

Pentaho Solution

- Pentaho Analysis Subscription
- Selected over Business Objects and Cognos
- Microsoft Windows Server with SQL Server
- Integrated with LDAP and Active Directory

Results

- Comprehensive analysis of time periods, services provided, billing groups, physicians
- Centralized, secured, consistent information delivery (versus prior Excel-based system)
- Ability to drill and analyze down to the individual patient level

Why Pentaho

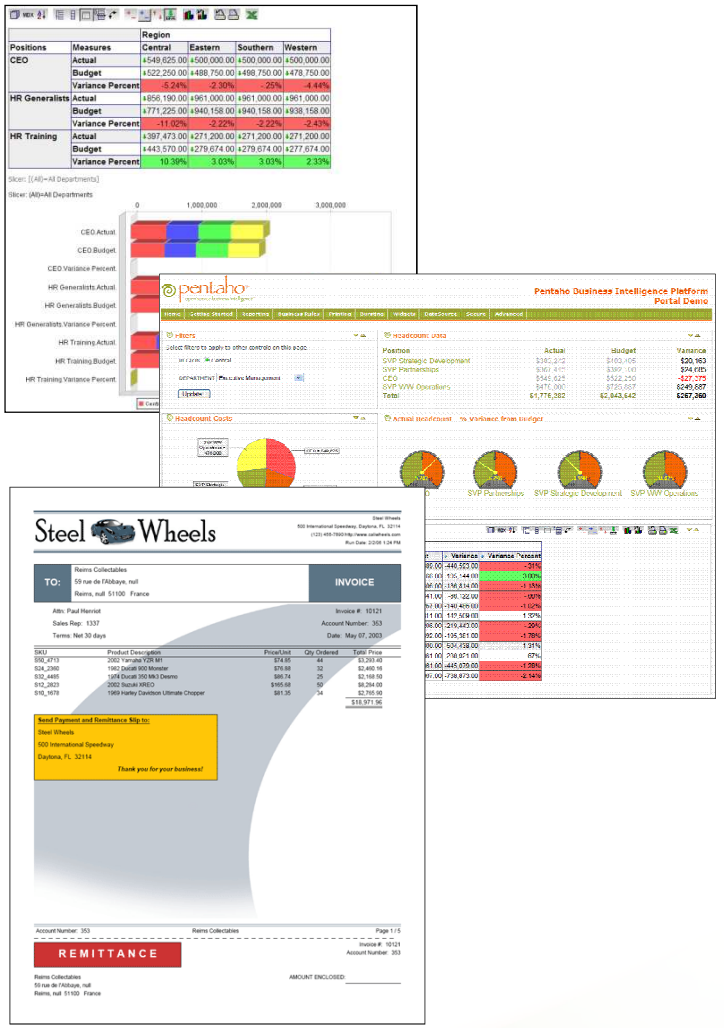
- Open standards support and ease of integration
- Cost of ownership

pentaho™
open source business intelligence

The big picture

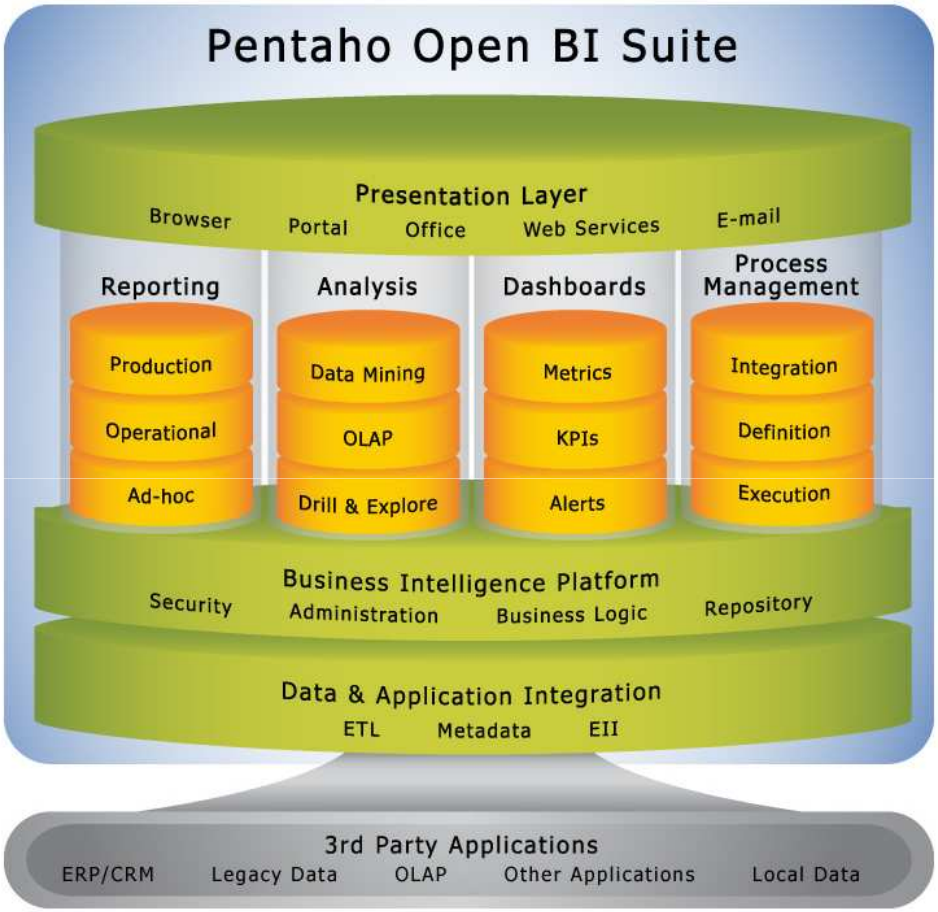
Business Intelligence Suite

- Mondrian OLAP
- Analysis tools:
  - Pivot table
  - Charting
  - Dashboards
- ETL (extract/transform/load)
- Integration with operational reporting
- Integration with data mining
- Actions on operational data
- Design/tuning tools



pentaho™
open source business intelligence

Pentaho Open Source BI Offerings



All available in a Free Open Source license

pentaho™
open source business intelligence

A Sample of Joint MySQL-Pentaho Users

"Pentaho provided a robust, open source platform for our sales reporting application, and the ongoing support we needed. The experts at OpenBI provided outstanding services and training, and allowed us to deploy and start generating results very quickly."

"We selected Pentaho for its ease-of-use. Pentaho addressed many of our requirements -- from reporting and analysis to dashboards, OLAP and ETL, and offered our business users the Excel-based access that they wanted."

iStockphoto

MySQL

DivX

ZIP REALTY
Your home is where our heart is.

UNIONFIDI

Expedia.com

pentaho
open source business intelligence

Next Steps and Resources

- More information http://www.pentaho.org and http://mondrian.pentaho.org
- Pentaho Community Forum http://community.pentaho.org
  - Go to Developer Zone
  - Discussions
- Pentaho BI Platform including Mondrian http://www.pentaho.org/download/latest
- Mondrian OLAP Library only http://sourceforge.net/project/showfiles.php?group_id=35302

Thank you for attending!

pentaho™
open source business intelligence