

Unleash the Power of Your Data Using Open Source



Presenter:
Christopher Lavigne, Partner, Breadboard BI.

Agenda

- BBBI Introduction, Analytic Modules
- Good Data – Breadboard BI Cloud Prototype Overview
- YachtWorld.com - Boats.com Case Study
- Questions & Answers

Introduction

- Breadboard BI (www.breadboardbi.com) - business intelligence (BI) consulting services.
- Cooperate with partners in Australia, Belgium, Chile, Mexico, and Portugal.
- MySQL partner since 2007.
- Offer a suite of **FREE** customizable BI modules based on MySQL 5.1 and Pentaho that form a modular and scalable enterprise solution. Run local or in the cloud.

BBBI Analytic Module List

Customer 360°

- + Case Management
- + Clickstream
- + Contact Center
- + Marketing
- + Returns
- + Sales Leads
- + Sales Opportunities
- + Sales Orders*

Finance

- + Accounts Payable
- + Accounts Receivable
- + Billing
- + Budget & Forecast
- + Ledger*

Supply Chain

- + Inventory*
- + Purchase Orders
- + Requisitions
- + Shipments
- + Shipping Notices
- + Warehouse Receipts
- + Work Orders

Workforce

- + Compensation
- + Recruitment
- + Snapshot

More information at <http://www.breadboardbi.com>

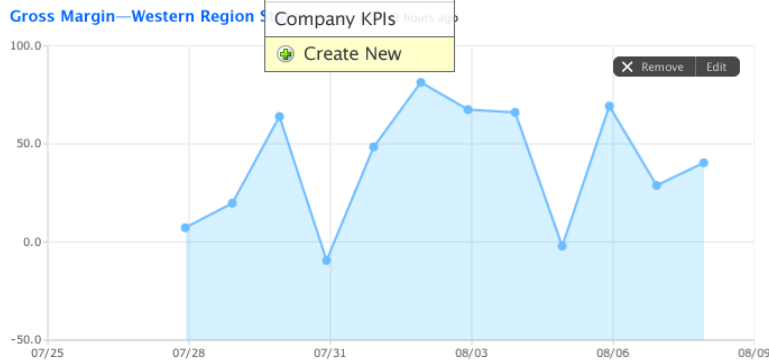
* Available in Spanish.

BBBI Enterprise Data Model

- An enterprise dimensional model behind the 23 analytic modules running over MySQL 5.1.
- Conformed dimension, fact, bridge, tree tables.
- Consistent naming conventions across objects.
- Metadata – table and column descriptions, relationships, primary and alternate keys, indexes, etc.

Good Data – Breadboard BI Cloud Prototype Overview





Add Report

Application Pulse

All Others Me

- Announcement**
Good Data Operations are just moving your project to faster node. See the [Documentation](#) to figure out what does it mean.
 - Yesterday**
Peter Stormare added a comment to report Low Fat Products Profit. "What is the profit per unit?"
 - Last Seven Days**
Peter Stormare added report Store Sales Share. "It's not perfect, I think management is going to be angry!"
You added a comment to report Store Sales Share. "These data seems to be nonsense..."
 - Fourteen Days and More Ago**
Austin Dwell added a comment to report Store Costs. "I've added the averages."
- [See older updates in Feed Archive](#)

Get More People

You've just started, get more people here so you can work together.

Invite Colegues

Favorite Reports

- Overpaid Time
 - Low Fat Products
 - Top 10 Employees
 - Store Costs
 - Best 5 Months
 - Store Sales by Customer Edu...
 - Average Salary by Time
 - Top 10 Product Brands
- [Go to Favorites](#)

Tag Cloud

sales decline 2007 low_fat cost profitable percentage GameDay-Special supply top balance revenues positions type salaries time states quarter brand wages share growth ratio ledger store sheet profit salary promotion average customer margin 2006 prod-uct year

[Show as List](#)

Collaborative Analytics

Good Data provides a simple and easy-to-use service to view, analyze and report on the data that drives your business.

All in a secure online space without the typical costs and headaches.

Sign up for free at <http://www.gooddata.com>

Good Data Service

- Good Data provides the underlying infrastructure to host multi-dimensional data, the tools to analyze that data in a collaborative environment, and the means to share the results with others. All as an on demand service.
- Free trial at <http://www.gooddata.com>

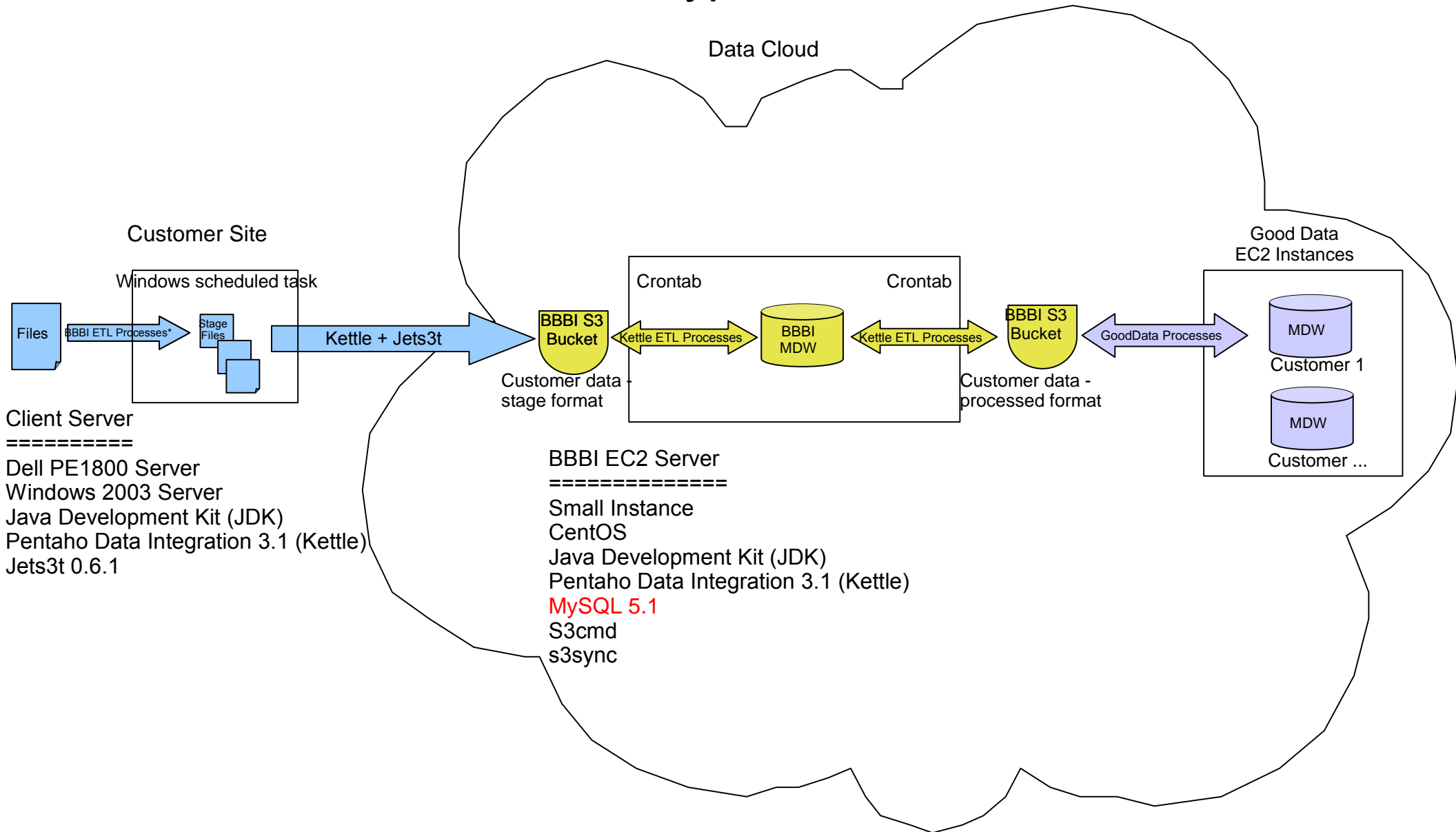
Company Overview

- Founded by Roman Stanek (founder of NetBeans, Systinet)
- Headquartered in San Francisco, engineering in the Czech Republic
- Company DNA: high-scale analytics processing, service-oriented architecture, user experience

Good Data – Breadboard BI Cloud Prototype Goals

- Test framework to deploy BBBI analytic modules on MySQL 5.1 in the cloud.
- Ability to support SaaS business intelligence companies like Good Data, or customers with their own presentation server (e.g., Pentaho) in the cloud.

Prototype Model



Prototype Process at the Customer Site

On a scheduled basis:

1. Incrementally extract data from source systems, load pre-defined stage files.
2. Compress files.
3. Transmit files to S3.

Required Components:

1. Pentaho ETL Server
2. Breadboard BI pre-defined ETL objects
3. Zip
4. Jets3t 0.6.1

Comments:

1. Simple, low-maintenance ETL tool at customer site. No database or presentation servers required. ETL tool interchangeable, or use scripts. (Only target file structure is important.)
2. All tools are licence-free, can run on any operating system.

Prototype Process in BBBI Cloud

On a scheduled basis:

1. Get files from client-specific S3 bucket.
2. Decompress files.
3. Extract file data, transform data into dimensions and facts, load MySQL 5.1 database.

Required Components:

1. Pentaho ETL Server
2. Breadboard BI pre-defined ETL objects
3. Zip
4. S3cmd, s3sync, and/or Jets3t 0.6.1
5. MySQL 5.1 (partitioned by customer).
6. Breadboard BI pre-defined enterprise data model.

Comments:

1. Simple, lower-maintenance architecture – single, partitioned database for all customers.

Prototype Process BBBI Hand-Off to Good Data

On a scheduled basis:

1. Create customer-specific MySQL data files.
2. Compress files.
3. Push files to S3 area accessible by Good Data.
4. Set ACLs on files.

Required Components:

1. Pentaho ETL Server
2. Breadboard BI pre-defined ETL objects
3. Gzip
4. S3cmd, s3sync, and/or Jets3t 0.6.1
5. MySQL 5.1
6. Breadboard BI pre-defined enterprise data model.

Comments:

1. Data stored in BBBI & Good Data – redundant, but provides another backup. BBBI serves as a dimensional ODS.

Prototype Lessons Learned

- All prototype goals were met.
- BBBI components are well-suited to the data cloud. Only minor changes were required.
- BI in the data cloud offers businesses a low-cost, highly-scalable alternative.

YachtWorld.com - Boats.com Case Study



YachtWorld.com - Boats.com Business

- YachtWorld.com is the premier online sales channel for yacht brokers around the world. Boats.com provides marketing and web services to new boat dealers and builders, and offers a "For Sale By Owner" classified service.
- Over 140,000 new and used boat listings worldwide offered by over 4,000 brokers, dealers and builders in 115 countries. Visited by over 4 million boating consumers every month who click through over 95 million page views.
- Headquartered in Seattle, Washington, European headquarters in the United Kingdom, sales offices in Germany, Italy and Russia, and sales representation in Dubai, Australia and China. Both companies are business units of Dominion Enterprises, based in Norfolk, Va.

YachtWorld.com – Boats.com Solution Challenges

- Organize oceans of data from diverse operational systems into a comprehensive market intelligence solution.

Millions of daily page views from each of their popular web sites in Apache web server logs;

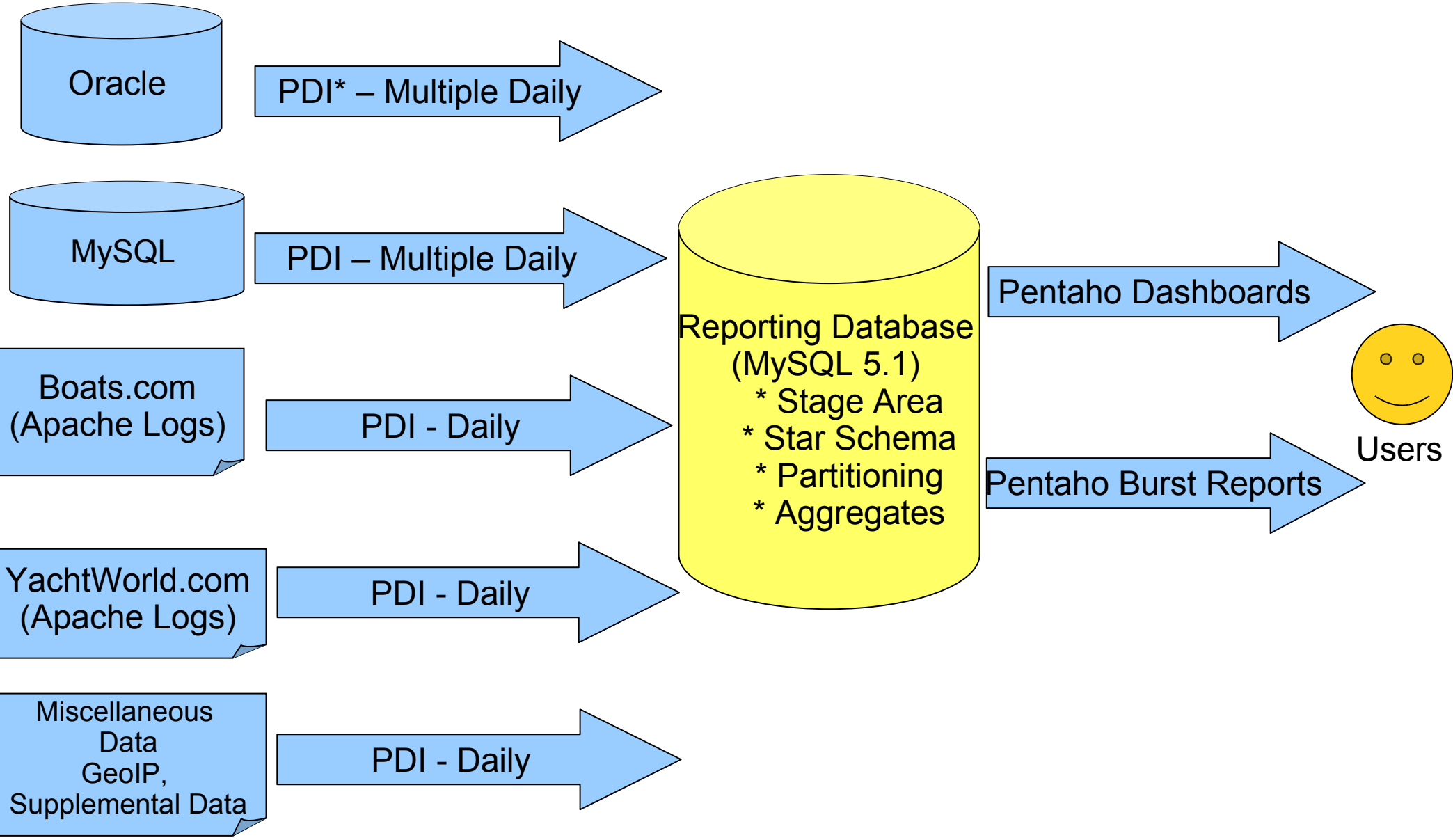
Boat listing inventory (1.7 million+) from Oracle and MySQL databases;

Sales lead emails and toll free calls to their affiliated brokers, dealers, and builders in Oracle and MySQL databases; and

Supporting data in various flat files and spreadsheets.

- Build the solution within a reasonable budget.

YachtWorld.com – Boats.com Environment Overview



* Pentaho Data Integration (PDI).

YachtWorld.com – Boats.com MySQL Database Layer

- MySQL 5.1
 - MyISAM and Memory engines.
 - Table partitioning (key and list).
 - Aggregation, denormalization, indexing.
- Star schema design with many aggregate tables.
 - ~20 fact tables (including aggregate facts), 30+ dimension tables.

Why MySQL 5.1 for YachtWorld.com - Boats.com?

- YachtWorld.com – Boats.com already had MySQL in-house.
- Table partitioning (key and list) availability.

Very large stage tables utilize list partitions for instantaneous deletes.

Each stage table maintains data for multiple business units. List partitioning via business unit supports fast delete for a subset of a table's data.

Fact tables initially partitioned by key partition to allow for very large tables (overcome file size limitations). Client may transition to range (partitioning column is already smart format - YYYYMM).

YachtWorld.com – Boats.com PDI Layer

- Pentaho Data Integration 3.1 used for all data movement.
Source Database/File → Stage.
Stage → Star.
- Modular – One Job calls it all (nested jobs) or child jobs can be run individually.
- Great integration with MySQL, data files, etc.

YachtWorld.com – Boats.com Pentaho Reporting and Dashboards Layer

- Pentaho Reporting 1.6 (Design Studio & Report Designer).
Complex email burst .pdf report using subreports.
- Pentaho Dashboards 1.6.
Utilizes dashboard widgets and drill reports.

Questions & Answers

Chris Lavigne - chris_lavigne@breadboardbi.com

Web Site with Link to Demo Server -
<http://www.breadboardbi.com/>