

Open Source
for the Enterprise
Day 2

Seth Grimes

Alta Plana Corporation

+1 301-270-0795 -- <http://altaplana.com>

Technology Transfer

May 7-8, 2007

Alta Plana

Agenda – Day 1

The Business Case for Open Source

History, developments, and trends

Licensing and support

Open Source operating systems

Any questions?

Alta Plana

Agenda – Day 2

Software frameworks: Java EE, .Net, and LAMP

Programming tools and development environments

Database servers

Office and Enterprise Applications

Business Intelligence and Data Warehousing

Best Practices, implementation strategy, and resources

Software frameworks

A software framework is a reusable design for a software system (or subsystem).

en.wikipedia.org/wiki/Software_framework

Software frameworks provide –

reusable software components.

a means of plugging in your own components.

a mechanism for component interoperability.

Software frameworks are the foundation of modern enterprise application development.

Java – Java EE

Java –

originated as a programming language and run-time environment for device-embedded software.

got attention as a tool for developing rich Internet client applications that run in the Web browser.

Plain-old HTML and JavaScript won that role however, supplemented by Flash and some Microsoft ASP.

became a tool of choice for server-side programming.

has only recently been released as OS by Sun Microsystems.

Java – Java EE

Java specification development is –

governed by the Java Community Process (JCP).

See jcp.org/en/home/index

determined by JSRs, Java Specification Requests.

Java application development may be performed within a number of OS frameworks.

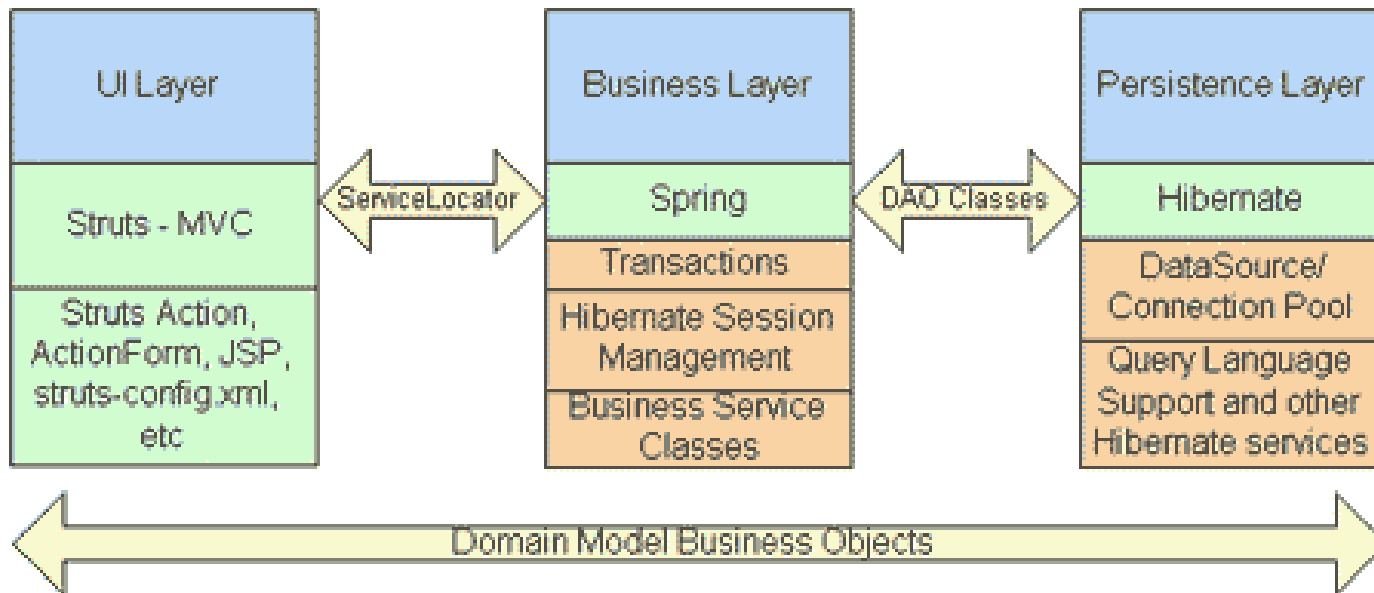
MVC oriented: Struts-Spring-Hibernate.

Component oriented: JavaServer Faces, Tapestry.

Eclipse.

Java – Java EE

One can combine frameworks, for instance for Web development –



www.onjava.com/pub/a/onjava/2004/04/07/wiringwebapps.html

Java – Java EE

Java EE is the Java Platform, Enterprise Edition.

Java EE was formerly known as Java 2 Enterprise Edition (J2EE).

Java EE 5 is the current version and includes:

Enterprise JavaBeans 3.0 with Plain Old Java Objects (POJOs).

[Object] Persistence API.

Enhanced web services and support for Service Oriented Architectures (SOAs).

JavaServer Faces (JSF); Java Server Pages (JSP), and the JSP Standard Tag Library (JSTL).

See java.sun.com/javae/technologies/

Spago

J2EE framework.

GNU LGPL license.

Hosted by ObjectWeb Consortium.

Sponsored by Engineering Ingengnerla
Informatica in Rome.

spago.eng.it/

Spago World community.

spagoworld.com/

SpagoBI business intelligence project.

spagobi-info.eng.it/SpagoBISiteENG/target/docs/index.html

.Net

.Net is Microsoft's application development framework. .Net is –

Currently out in version 3.0.

“Microsoft software for connecting information, people, systems, and devices. .NET provides XML-based interoperability and is being incorporated across Microsoft clients, servers, services, and tools. For example, products like Microsoft Windows and Microsoft Office will use .NET to connect with other systems and applications. For developers, .NET is manifested in the programming model delivered in the Microsoft .NET Framework.”

See *www.microsoft.com/net/default.aspx*

The open source Mono project –

“provides the necessary software to develop and run .NET client and server applications on Linux, Solaris, Mac OS X, Windows, and Unix.”

is commercially supported by Novell.

See *www.mono-project.com/Main_Page*

LAMP

LAMP is an alternative, open-source stack consisting of –

1. Linux operating system.
2. Apache Web server.
3. MySQL database.
4. Perl/Python/PHP programming.

The programming layer includes application frameworks.

LAMP

LAMP essentially provides an alternative Web development/delivery platform.

PostgreSQL easily substitutes for MySQL.

Ruby substitutes for the programming-layer Ps.

Each of the programming-layer alternatives may be packaged within one or more software frameworks.

Zend is one of many frameworks for PHP.

Rails is a noted framework for Ruby development.

Django is one of many frameworks for Python.

Eclipse

Eclipse—

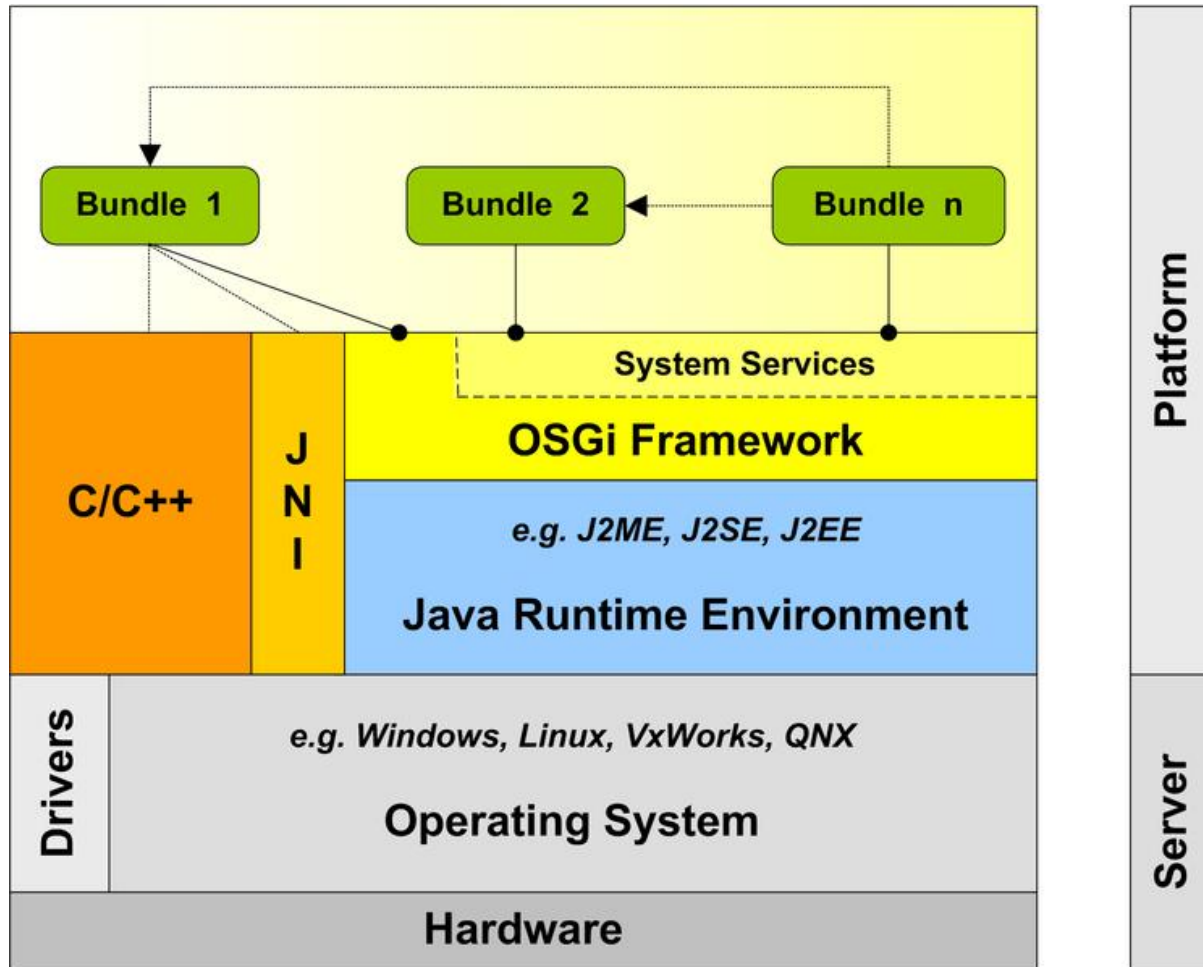
is a “universal tool platform.”

offers an integrated development environment (IDE) for software development.

provides a framework, a “rich-client platform” (RCP), for applications.

implements the Open Services Gateway initiative (OSGi) framework for service/component interoperations.

Framework integration



en.wikipedia.org/wiki/Image:Osgi_layer.png

Application servers

An application server is/runs middleware serving as a component container for –

Clients/presentation.

Business logic.

Data access.

Open source examples include –

Apache Tomcat.

ObjectWeb JOnAS.

Sun GlassFish.

Zope (Python).

Apache

Started as a HTTP (Web) server.

Became a nexus for Apache-licensed projects that adhere to “The Apache Way”:

- collaborative software development.

- commercial-friendly standard license.

- consistently high quality software.

- respectful, honest, technical-based interaction.

- faithful implementation of standards.

- security as a mandatory feature.

Apache Tomcat & Geronimo

Tomcat is a container for –

Java servlets.

Java Server Pages.



Geronimo supports the entire Java EE stack –

Servlet container.

Enterprise Java Beans (EJB) container.

Messaging via Java Message Service (JMS) API.

Java Connector Architecture (JCA) container.

IBM WebSphere Application Server Community

Edition = Geronimo.

JBoss

The leading commercial-OS app server; LGPL.

JEMS = JBoss Enterprise Middleware.

Provides applications services within an SOA.

Supports many projects including –

Enterprise Service Bus.

Portal.

jBPM business process management engine.

Seam application framework for POJOs/EJBs.

Architecture trends

Service Oriented Architecture (SOA) –

Prizes interoperability and therefore standards.

Typically includes process orchestration.

Offers an Enterprise Service Bus (ESB) as one interconnection option.

Interoperation –

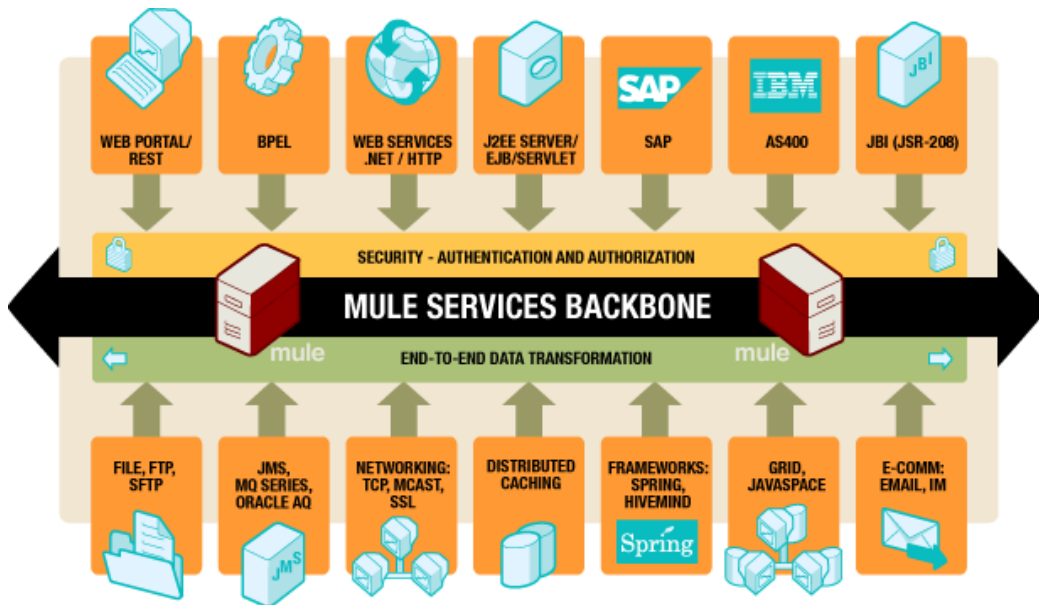
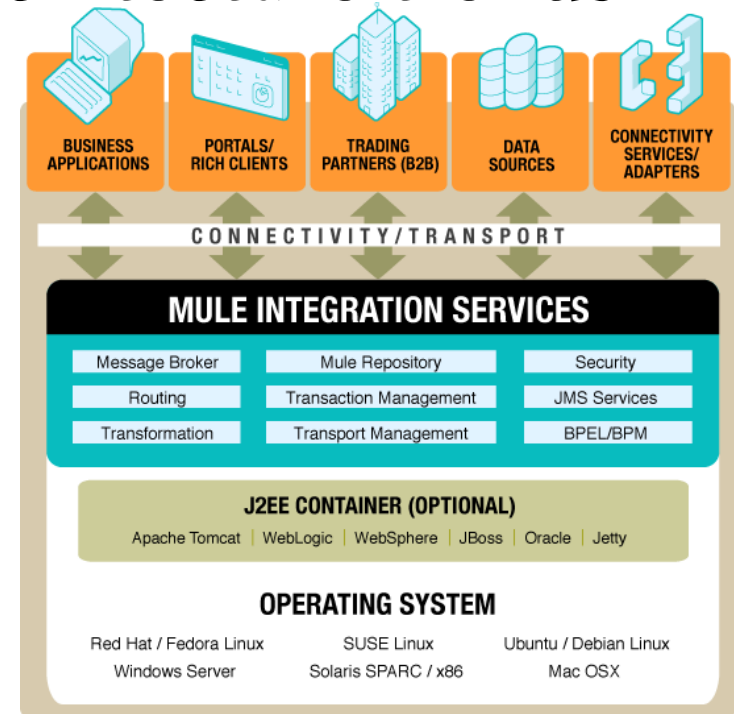
A software development framework supports interoperation within an application.

An app server supports application interoperation.

SOA adds an abstraction layer.

Architecture trends

SOA/ESB architecture illustration from the open-source Mule platform.



Alta Plana

Copyright © 2007 Alta Plana Corporation

The Mule license is based on Mozilla's. The company sells support subscriptions.

www.mulesource.com/products/,
www.mulesource.com/solutions/

Technology Transfer

Questions?

Discussion?

Next: Programming tools and development environments

Agenda – Day 2

Software frameworks: Java EE, .Net, and LAMP

Programming tools and development environments

Database servers

Office and Enterprise Applications

Business Intelligence and Data Warehousing

Best Practices, implementation strategy, and resources

Software development

To create software you need –

Team management and development approach.

Programming tools.

Development environment.

Repository with version and build control.

Test tools.

Software development

Traditional software development focuses on –

Quality assurance via attention to –

Processes and procedures.

Knowledge management and learning.

Management practices.

Quality control via –

Verification and validation.

Structured testing processes: unit, integration-system, acceptance.

OS development deviates from the traditional model given the nature of teams, coordination.

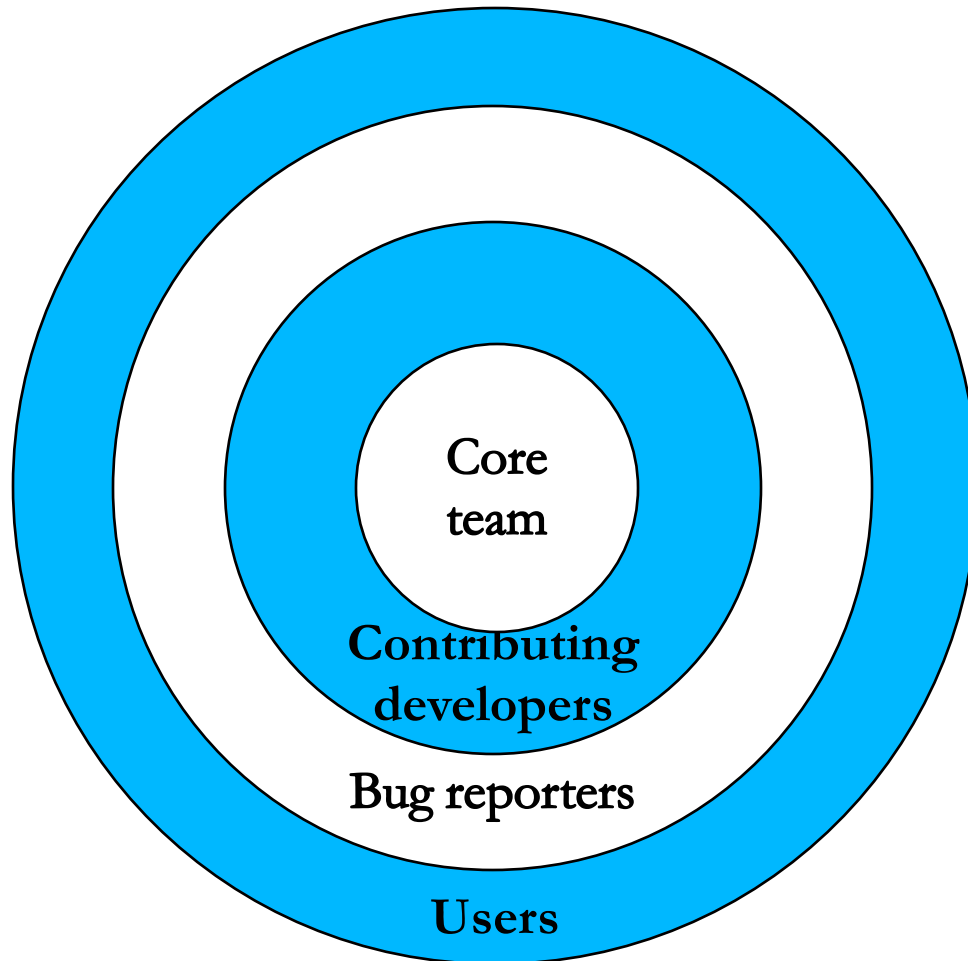
Software development

Closed source	Open source
Well defined development methodology	Development methodology often not defined or documented
Extensive project documentation	Little project documentation
Formal, structured testing and quality assurance methodology	Unstructured and informal testing and quality assurance methodology
Analysts define requirements	Programmers define requirements
Formal risk assessment produces – monitored and management throughout project	No formal risk assessment process
Measurable goals used throughout project	Few measurable goals
Defect discovery from black-box testing as early as possible	Defect discovery from black-box testing late in the process
Empirical evidence regarding quality used routinely to aid decision making	Empirical evidence regarding quality isn't collected
Team members are assigned work	Team members choose work
Formal design phase is carried out and signed off before programming starts	Projects often go straight to programming
Much effort put into project planning and scheduling	Little project planning and scheduling

Mark Aberdour, "Achieving Quality in Open Source Software," *IEEE Software*, vol. 24, no. 1, pp. 58-64, Jan/Feb, 2007.

Software development

“The onion model of a sustainable software development community.”



Mark Aberdour, "Achieving Quality in Open Source Software," *IEEE Software*, vol. 24, no. 1, pp. 58-64, Jan/Feb, 2007.

Programming languages

GCC (GNU Compiler Collection)

Includes C++, Fortran, Java.

Perl

Python

Ruby

Popular Rails Web-development framework.

Java

Tcl (Tool Control Language)

Python

Scripting language –

Created by Guido Van Rossum in 1991.

Managed by the Python Software Foundation.

BSD-style Python Software Foundation License
(PSFL).

Object oriented.

Standard library plus many modules.

Dynamic (duck), strong typing create flexibility.

Automatic memory management.

www.python.org

Eclipse development environment

There are many development environments.

Our primary interest will be in Eclipse.

Originated as an IBM integrated development environment (IDE) designed to replace the Visual Age toolset.

Released to open source in 2001.

Managed by the Eclipse Foundation.

www.eclipse.org/

Eclipse development environment

Eclipse projects include –

Business Intelligence and Reporting Tools (BIRT)

Data Tools Platform

Device Software Development Platform

Eclipse Modeling Project

SOA Tools

... and technology, programming and Web tools,
and more.

Subversion

Globally distributed software development teams need to version and share source code.

Tigris.org –

“A mid-sized open source community focused on building better tools for collaborative software development,” hosted by CollabNet.

Subversion –

An open source system for version control.

CollabNet corporate sponsorship, hosted at *tigris.org*.

www.collab.net/products/open_source_subversion/

GForge

Collaborative development environment.

Based on the original *SourceForge.net* system, which was closed by VA Linux in 2001.

Italian translation available.

Free but not pure open source.

Source code is available to customers who purchase a source code license and sign the source code NDA.

Free license available by application for public websites, charities, and open-source projects.

GForge

Features include:

Message forums; mailing lists.

Tools to create and control access to Source Code Management repositories.

Additional Tools: Manage File Releases

Document Management.

News announcements.

Surveys for users and admins.

Issue tracking with "unlimited" numbers of categories, text fields, etc.

Task management.

GForge



(c) 2004, GForge Group, L.L.C.

GForge

Used by:

SugarCRM.

Pentaho.

PGfoundry.

Splunk.

SpikeSource.

Compare to:

Bugzilla – from Mozilla.

Jira – free for OS projects.

Questions?

Discussion?

Next: Database servers

Agenda – Day 2

Software frameworks: Java EE, .Net, and LAMP

Programming tools and development environments

Database servers

Office and Enterprise Applications

Business Intelligence and Data Warehousing

Best Practices, implementation strategy, and resources

Database systems

A DBMS is...

Four data-management modes:

Content publishing.

Transactional.

Analytical.

Operational (embedded).

Database systems

“Traditional”:

Ingres



PostgreSQL



MySQL



Firebird



Java:

Apache Derby



HSQldb



Embedded + Java:

Oracle Berkeley DB (Sleepycat)

Alta Plana

Copyright © 2007 Alta Plana Corporation

ORACLE

Technology Transfer

Apache DB Project

Subprojects:

DdlUtils

Component for working with Database Definition (DDL) files.

XML files that contain the definition of a database schema, e.g. tables and columns.

Derby

Java Data Objects (JDO)

Plain old Java object (POJO) database persistence.

OBJ (ObjectRelationalBridge), Torque

Object-DB mapping.

Apache Derby

Embedded relational database.

Implemented entirely in Java.

Apache License, Version 2.0.

Embedded JDBC driver that lets you embed Derby in any Java-based solution.

Small footprint -- about 2 megabytes for the base engine and JDBC driver.

Also supports the more familiar client/server mode with the Derby Network Client JDBC driver and Derby Network Server.

db.apache.org/derby/

HSQLDB

Another relational database written in Java.

2001 fork from the Hypersonic SQL Project.

Database engine in OpenOffice.org 2.0.

Modified BSD license.

In memory & disk based; embedded & server modes.

Used as a database and persistence engine in many OS projects: one applet version has a 100k footprint.

hsqldb.org/

HSQLDB

Hyper-
XtremeSQL
is an
application-
compatible
commercial
equivalent.

www.hxsq.com - Mozilla Firefox

File Edit View History Bookmarks Tools Help deljcio.us

http://www.hxsq.com/

In the test runs below, the operating system provides high-speed caching of data via a memory-mapped file for both HSQLDB and HXSQ. COUNT, SELECT, UPDATE and DELETE queries run at several times the speed of HSQLDB. The only exception in all Disk Engine tests is shutdown speed after extensive changes. This is due to the fact that HXSQ manages the memory cache more effectively and keeps modified rows much longer than HSQLDB, resulting in a larger number of rows to be saved at shutdown.

OPERATION	rows	HSQLDB timings		HXSQ timings		speedup
		time (ms)	rows / s	time	rows / s	
insert	256,000	15,953	16,046	11,422	22,410	1.40
shutdown		1,859		1,734		1.07
reopen		63		47		1.34
count (index on id)	256,000	6,484	39,475	547	467,153	11.85
select random id	256,000	54,641	4,685	4,703	54,421	11.62
select random zip (zip table)	256,000	1,062	240,827	922	277,356	1.15
update with random id	256,000	78,766	3,250	19,719	12,981	3.99
count (index on id)	256,000	6,609	38,729	422	605,200	15.66
delete with random id	32,000	8,266	3,870	1,187	26,936	6.96
count (index on id)	224,000	5,812	44,039	360	709,141	16.14
shutdown		2,328		3,046		
total test time		181,969		44,171		4.12

In the second set of test runs with the same number of rows, the memory-mapped file is not used and there is less reliance on the operating system caching mechanisms. The speedup is generally higher than the previous test, with both SELECT and DELETE queries running at about 12 times the speed.

OPERATION	rows	HSQLDB timings		HXSQ timings		speedup
		time (ms)	rows / s	time	rows / s	
insert	256,000	19,360	13,222	11,375	22,503	1.70
shutdown		1,968		1,328		1.48
reopen		63		47		1.34
count (index on id)	256,000	15,672	16,333	594	430,252	26.38
select random id	256,000	131,640	1,944	10,625	24,091	12.39
select random zip (zip table)	256,000	1,032	247,821	907	281,938	1.14
update with random id	256,000	190,484	1,343	69,187	3,700	2.75
count (index on id)	256,000	16,031	15,968	469	544,680	34.18
delete with random id	32,000	22,625	1,414	1,906	16,780	11.87
count (index on id)	224,000	14,188	18,042	406	628,992	34.95
shutdown		1,281		3,485		

Find: Next Previous Highlight all

Oracle Berkeley DB

A family of open source embeddable databases:

Berkeley DB –

A popular embeddable database engine.

Berkeley DB Java Edition –

A pure Java version of Berkeley DB optimized for the Java environment.

Berkeley DB XML –

An embeddable XML database with XQuery-based access to documents stored in containers and indexed based on their content.

Oracle Berkeley DB

History:

Developed in the early '90s and distributed with BSD Unix.

Commercialized by Sleepycat Software in 1996.

XML and Java engines added in 2003, 2004.

Acquired by Oracle in February, 2006.

Licensing:

Dual license scheme –

OS for internal use and open-source distribution.

Commercial for embedding in a proprietary application.

Oracle Berkeley DB

Usage:

Estimated 200 million deployments.

Motorola cell phones.

Microsoft/Groove's collaboration suite

Sun's Java Enterprise System.

Ericsson routers.

EMC storage systems.

www.oracle.com/technology/oramag/oracle/07-jan/o17opensource.html

Firebird

Open source continuation of the Interbase RDBMS, www.firebirdsql.org/.

(1984) Groton Database Systems → (1986-91) Ashton Tate → (1991) Borland → (1998) Inprise / Borland → (2000) code fork, open source.

Innovative multi-generational architecture.

InterBase Public Licence, a variant of Mozilla Public Licence (MPL).

Supported by the Firebird Foundation.

IBPhoenix development portal.

PostgreSQL

BSD license.

Use, modify and distribute PostgreSQL in any form, open or closed source.

Press kit in Italian:

www.postgresql.org/about/press/presskit82.html.it

PostgreSQL foundry, *pgfoundry.org/*

The PostgreSQL Development Group's site for developing and publishing PostgreSQL-related software that is not part of the core product.

Other tools development at *gborg.postgresql.org/*.

PostgreSQL

Features –

Runs stored procedures in more than a dozen programming languages, including Java, Perl, Python, Ruby, Tcl, C/C++, and its own PL/pgSQL, which is similar to Oracle's PL/SQL.

Interfaces for Java (JDBC), ODBC, Perl, Python, Ruby, C, C++, PHP, Lisp, Scheme, and Qt.

Triggers and stored procedures can be written in C and loaded into the database as a library.

Framework for custom data types along with supporting functions and operators.

MySQL

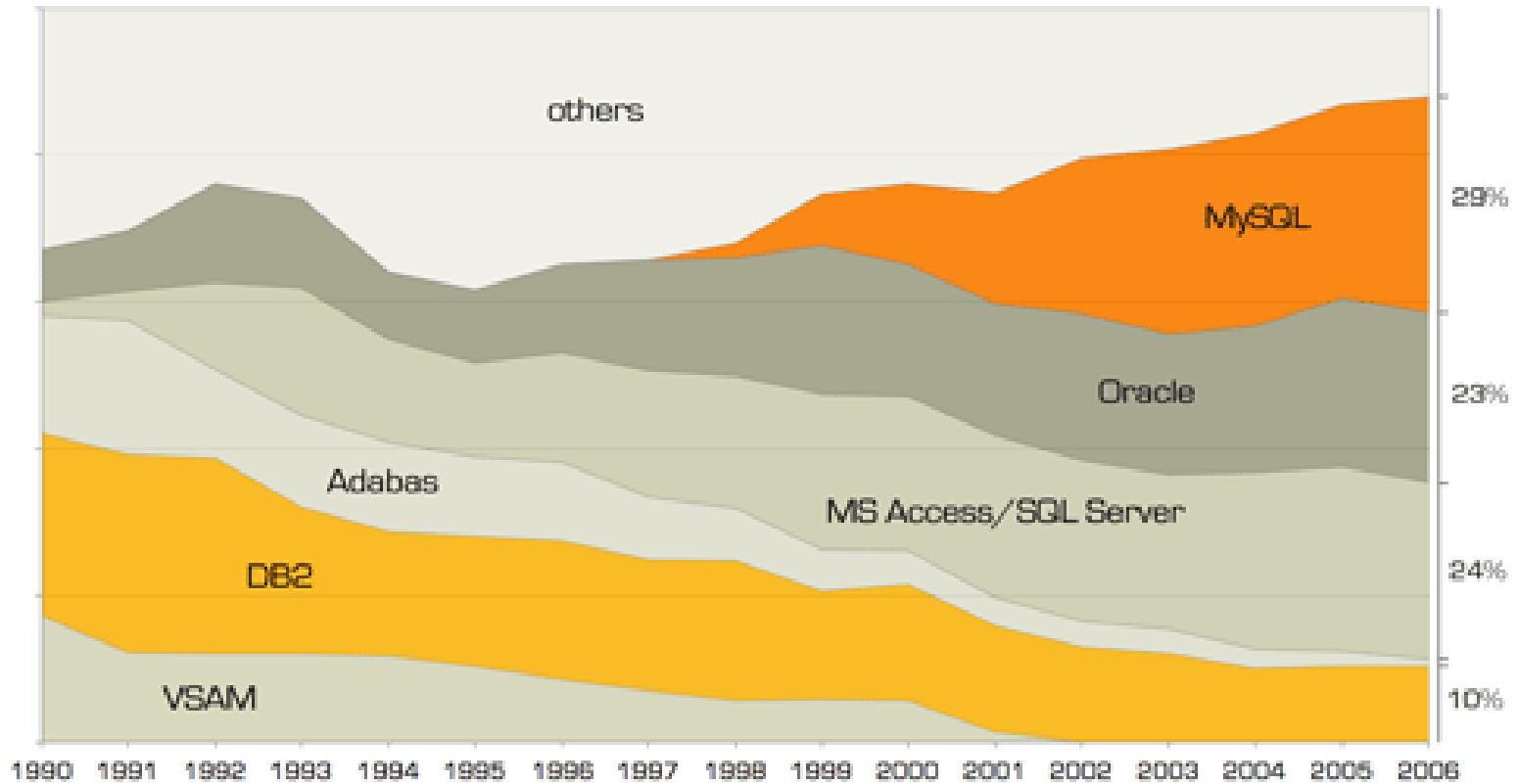
MySQL is the most popular open-source DBMS.

Early success for Web content as part of the LAMP stack.

Later success for transactional, analytical, and operational (embedded) systems.

Targets developers, ISVs, VARs, hardware vendors, and network appliance.

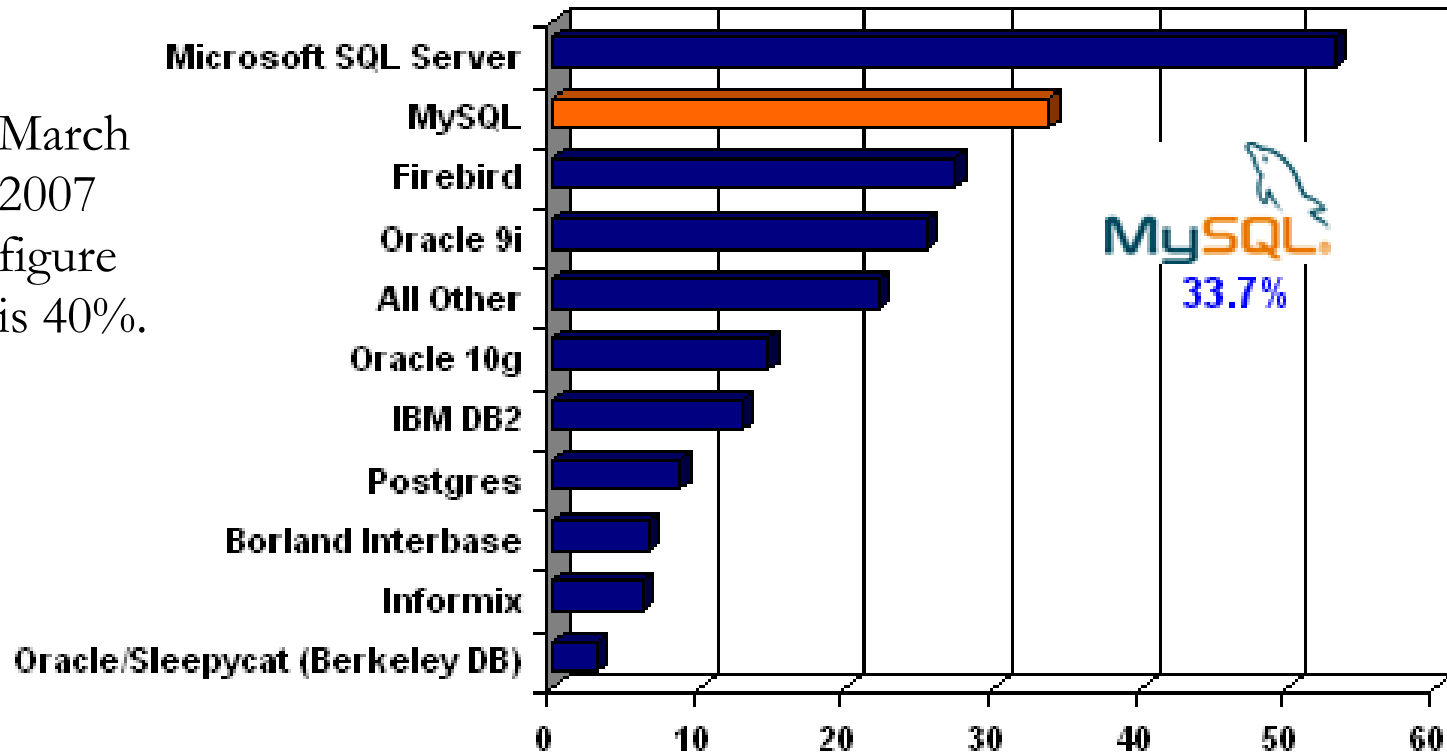
MySQL market share



Source: JoinVision E-Services GmbH, July 2006, survey of the **JoinVision community**, www.mysql.com/why-mysql/marketshare/ and www.joinvision.com/jv/ext/infow/itfacts/200606/itfacts200606.pdf

MySQL market share

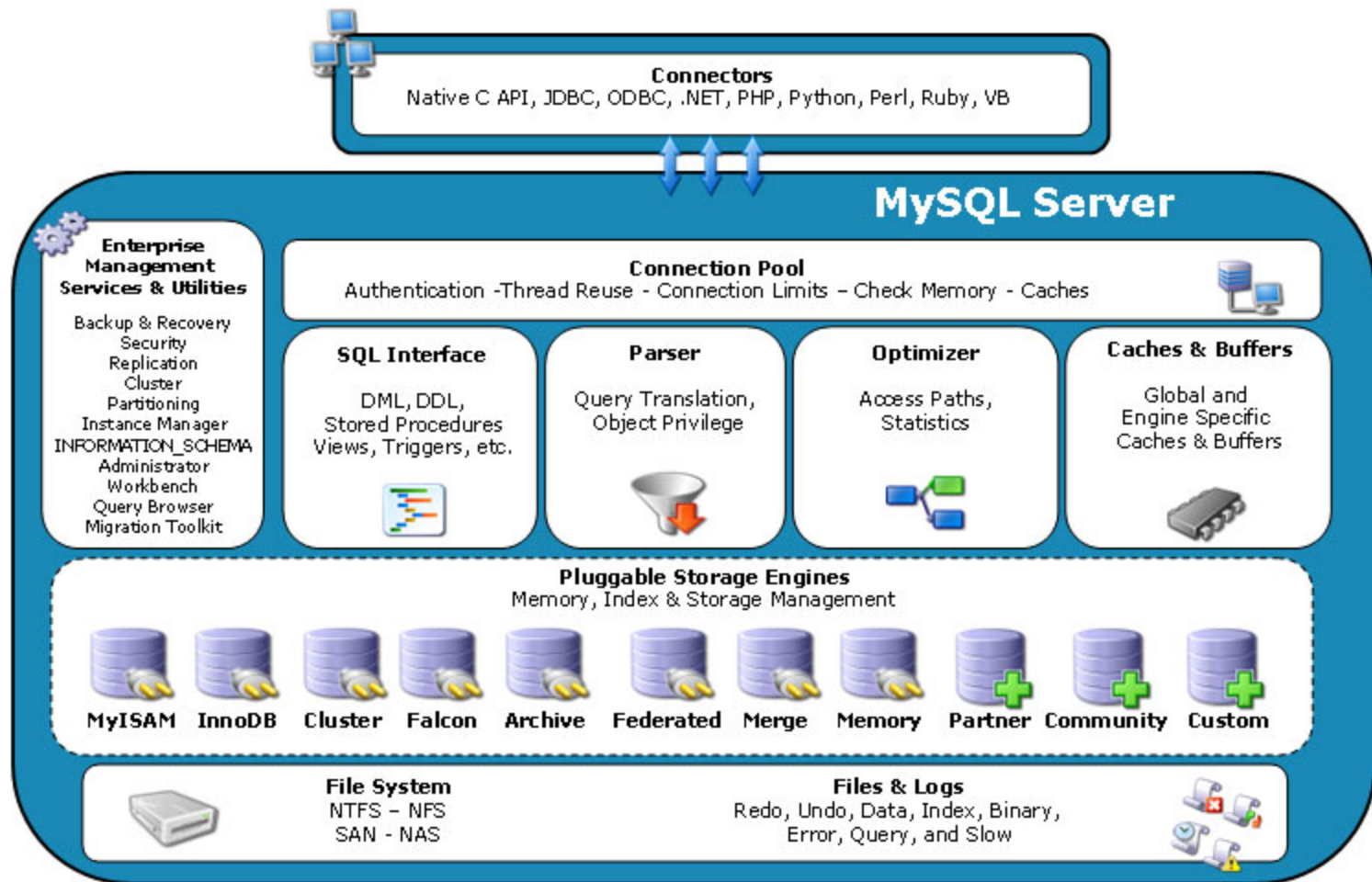
March 2007 figure is 40%.



Source: Evans Data Corp., October 2005

Evans Data Database Development Survey, October, 2005, examining the usage patterns and attitudes of **database developers**, www.evansdata.com/survey_database_topical.shtml

MySQL architecture



MySQL resources

Italian site:

www-it.mysql.com/

Italian white papers at:

www-it.mysql.com/why-mysql/white-papers/

My white paper, “MySQL 5.0 - Ready for Prime Time Business Intelligence with Pentaho” is at:

dev.mysql.com/tech-resources/articles/mysql_5.0_pentaho.html

PostgreSQL enhancements

ExtenDB:

Layered on PostgreSQL and DW optimized & parallelized but NOT open source.

EnterpriseDB:

Layered on PostgreSQL and Oracle compatible but NOT open source.

Bizgres (Greenplum)

Open source; Bizgres MPP is parallelized, designed for data warehousing, but not OS.

DBMS tools

MySQL GPL tools include –

Administrator, Query browser, Migration utility.

Free but not open source:

Toad for MySQL.

www.quest.com/toad-for-mysql/

Not free, not open source:

MySQL Maestro

www.sqlmaestro.com/products/mysql/maestro/download/

PostgreSQL Maestro.

www.sqlmaestro.com/products/postgresql/maestro/

Content management

A content management system is a document/text equivalent to a DBMS. There are many packages. Consider –
Alfresco.

Java; GPL with FLOSS exception to allow interoperation.

Drupal.

PHP; GPL; www.drupalitalia.org/



Joomla.

GPL; a 2005 fork of Mambo over corporate control.

Open Source Matters provides an administrative umbrella for legal and financial support for the Joomla! Project.

Questions?

Discussion?

Next: Office and enterprise applications

Agenda – Day 2

Software frameworks: Java EE, .Net, and LAMP

Programming tools and development environments

Database servers

Office and Enterprise Applications

Business Intelligence and Data Warehousing

Best Practices, implementation strategy, and resources

Office productivity tools

Office productivity suites consist of desktop:

Word processing.

Spreadsheet.

Presentation.

E-mail, directory, calendar.

(Database.)

Poll: What software does your company use?
Why?

Office productivity tools

The market:

Microsoft achieved and maintains a dominant position by –

- Beating Lotus & WordPerfect from DOS to Windows.

- Strong Windows and inter-tool integration.

- Effective (and predatory) relationship with PC makers.

Competitors (WordPerfect, etc.) hang on by offering compatibility and low prices.

OpenOffice.org from Sun is open source.

Google docs & spreadsheets are free, hosted but not OS.

Hosted office tools

Hosted versions provide an alternative delivery model but do not redefine the concept.

Also known as Software as a Service (SaaS) and On Demand software.

They foster collaboration, location independence.

An emerging business model is to host OS software for a fee.

Reason for the Affero General Public License.

Google spreadsheet:
hosted,
not open source.

The screenshot shows a Google Docs & Spreadsheets interface in a Mozilla Firefox browser. The spreadsheet is titled "ft1dv10se22ma221mi221100yr2000" and is autosaved at Mar 20, 2007 2:52:41 PM EDT. The spreadsheet contains financial data for "Income Tax Returns of Active Corporations with Accounting Periods Ended July 2000 Through June 20001 --". The data is organized into columns for asset categories: Total, Zero Assets, 1 to under 100, 250 to under 500, 500 to under 1,000, and 1,000 to under 5,000. The rows list various financial metrics such as Total assets, Cash, Notes and accounts receivable, and Total liabilities.

	Total	Zero Assets	1 to under 100	250 to under 500	500 to under 1,000	1,000 to under 5,000
1 Number of returns	885	9	*247	*200	*121	*108
2 Total assets	519,234,712	0	*5,393	*66,721	*76,834	*264,151
3 Cash	9,451,713	0	*2,530	*56,581	*14,053	*14,804
4 Notes and accounts receivable	42,609,963	0	0	86	*21,963	*6,330
5 Less: Allowance for bad debts	697,575	0	0	0	0	0
6 Inventories	4,493,505	0	0	0	*6,395	0
7 U.S. govt. obligations, total	1,426,285	0	0	0	0	0
8 Tax-exempt securities	24,777	0	0	0	0	0
9 Other current assets	42,892,704	0	0	*5,699	*1,768	*2,729
10 Loans to shareholders	239,587	0	0	*1,780	*2,542	*6,663
11 Mortgage and real estate loans	1,419,012	0	0	0	0	0
12 Other investments	87,033,454	0	0	377	0	*59,222
13 Depreciable assets	392,914,921	0	*17,944	*29,519	*45,831	*235,553
14 Less: Accumulated depreciation	156,647,479	0	*15,081	*27,384	*17,509	*80,205
15 Depletable assets	1,281,715	0	0	0	0	0
16 Less: Accumulated depletion	666,393	0	0	0	0	0
17 Land	2,048,645	0	0	0	0	*6,814
18 Intangible assets (Amortizable)	15,265,890	0	0	0	*1,114	*17,103
19 Less: Accumulated amortization	1,717,630	0	0	0	*279	*4,871
20 Other assets	77,861,617	0	0	64	*955	9
21 Total liabilities	519,234,712	0	*5,393	*66,721	*76,834	*264,151
22 Accounts payable	29,119,949	0	0	927	*4,918	*20,222
23 Mort. notes, and bonds under 1 yr	34,523,586	0	7,407	0	0	*190

Extended office tools

Hosted and server versions boost collaboration.

The spreadsheet – VisiCalc, then Lotus 1-2-3, then Excel – were “killer apps” instrumental in the PC devolution: personal power at the expense of best practices.

Office tools capture “knowledge” on desktop system.

Not secure.

Poor documentation and standards compliance.

Little sharing or collaboration.

No testing.

Enterprise applications

Enterprise applications include:

Enterprise Resource Planning (ERP).

Customer Relationship management (CRM).

Supply Chain Management (SCM).

Human Resources Management (HRM).

Financial and Planning applications.

Sales Force Automation (SFA).

Marketing Automation, e.g., campaign management.

Systems for manufacturing, logistics, and other functions.

Enterprise applications

Open source doesn't significantly compete with Oracle, SAP, Salesforce.com, or even Microsoft Dynamics.

Target small-medium enterprises (SMEs).

SugarCRM – FOSS and commercial versions.

Uses PHP and MySQL.

CentricCRM – all OS.

Java.

Compiere ERM and CRM –

Enterprise applications

Others:

ERP5.

Apache OFBiz.

Opentaps.

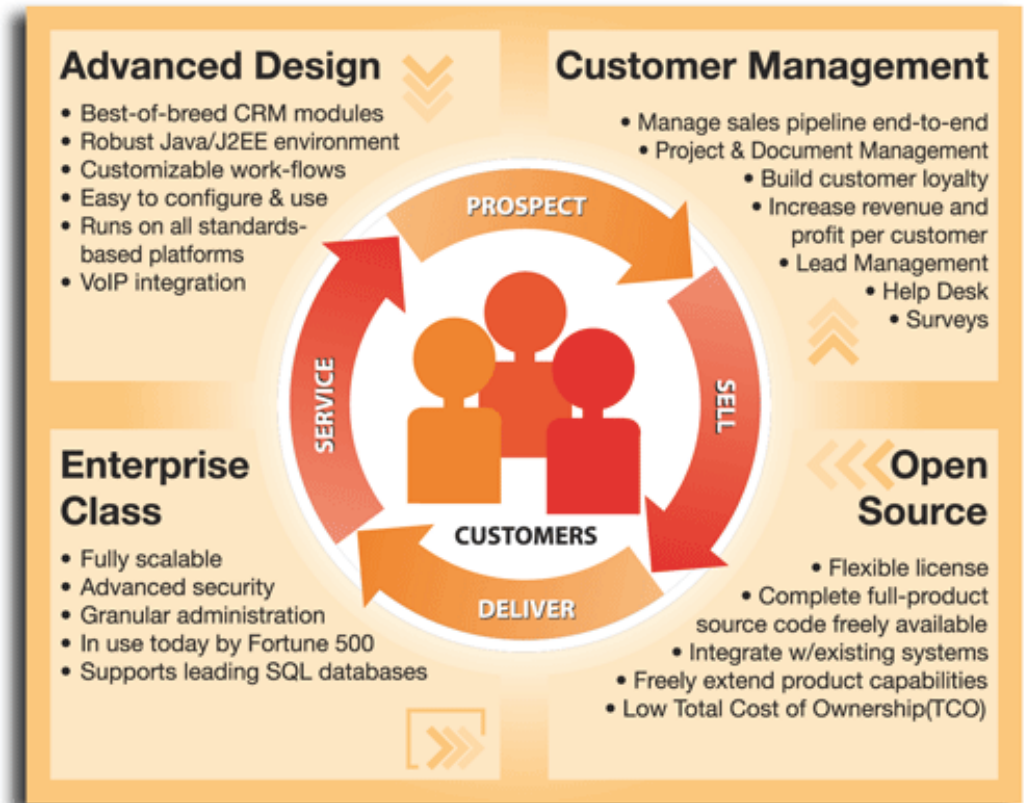
Does Apache inclusion guarantee quality, capabilities, community?

Centric CRM

J2EE, 100% open source.

Centric Public License assigns ownership of derivative work to developer.

Company sells maintenance and support.



SugarCRM

License fee for enterprise edition.

FOSS version has 50%-75% of full capabilities.

Uses PHP and MySQL.

SugarCRM Public License, an amended version of the Mozilla Public License (MPL) v1.1.

Derivations belong to SugarCRM.

Full Italian translations available as language packs.

Compiere

ERP, CRM, POS, SCM applications; integrated workflow.

Over 1.2 million Sourceforge.net downloads.

What does this include?

Now GPL licensed; alternative license offered.

Adempiere was a September 2006 fork due to dissatisfaction of project control.

More community led: “The Bazaar takes over and makes it happen.”

Enterprise applications

Adaptive Planning.

Budgeting, forecasting, and reporting.

Express Edition is a free (GPL) for entry-level budgeting and forecasting – **very limited**.

Corporate Edition and Enterprise Edition are commercially-licensed versions that provide enhanced performance-management features as well as expanded support and indemnification.

Accessed as an on-demand, web-based service, or deployed on-premises in a company's datacenter.

www.adaptiveplanning.com/home.shtml

Enterprise applications

Sample of other enterprise tools available as OS:

Business process management (BPM):

Intalio, www.intalio.com/.

Identity management:

Based on Liberty Federation and WS stds., OpenLiberty.org.

Search:

Apache Lucene, lucene.apache.org/.

Networking/telecommunications:

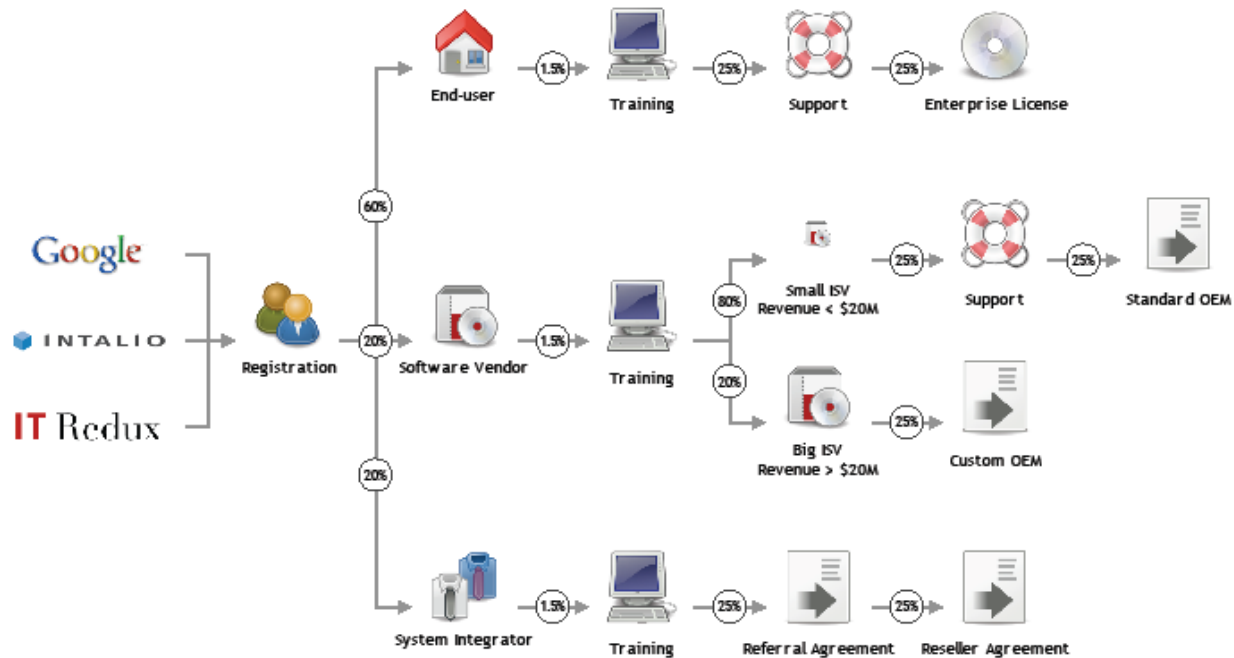
Asterix PBX for VoIP Internet telephony.

Routers: Quagga, Vyatta, Zebra.

www.opensourcenetworks.org/

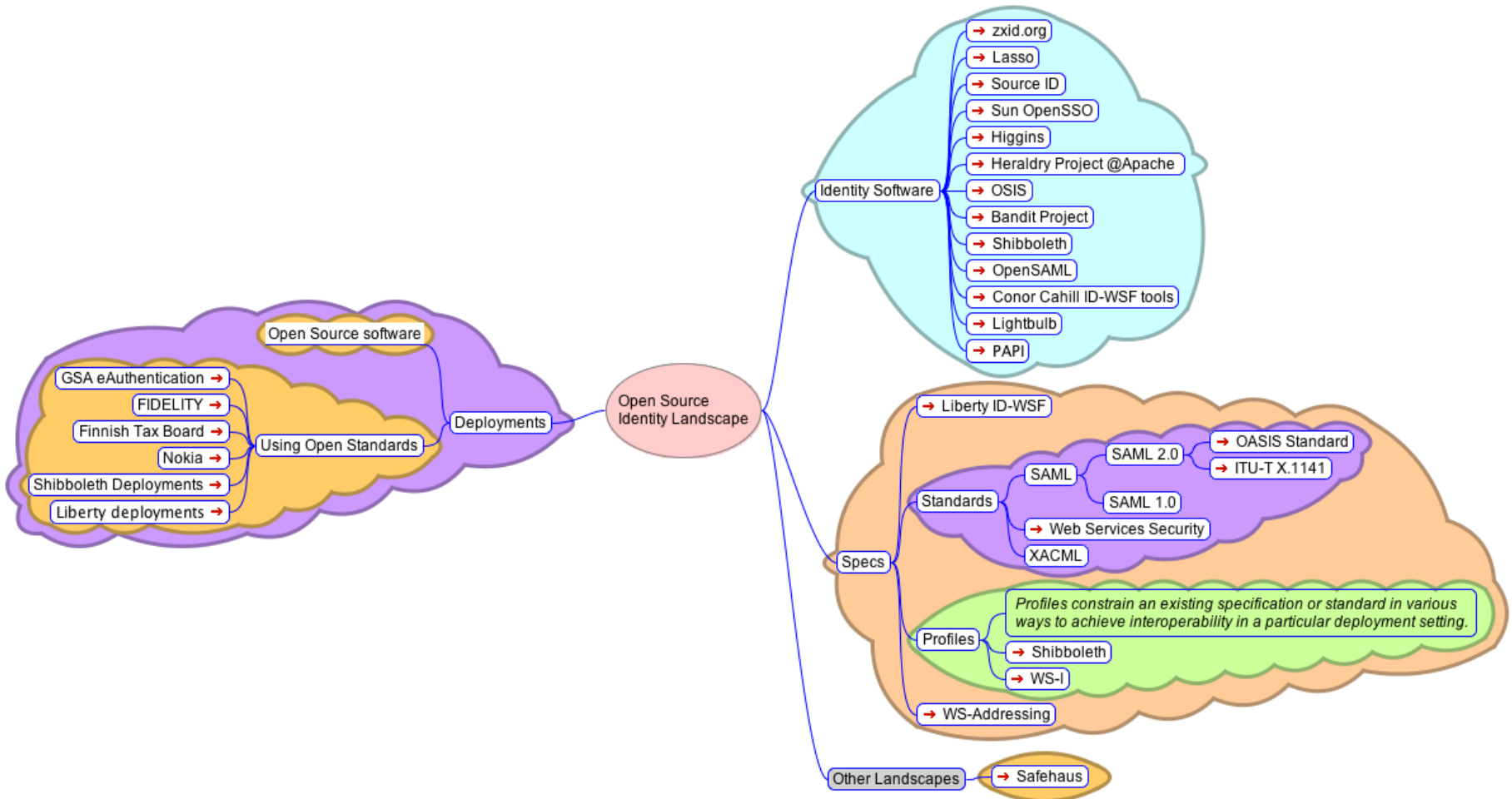
Intalio BPM

INTALIO | Sales Process



Intalio, Leader in Open Source BPM

OpenLiberty.org



openliberty.org/wiki/uploads//0/06/OSLandscapeBig.png

Questions?

Discussion?

Next: Business Intelligence and Data
Warehousing

Agenda – Day 2

Software frameworks: Java EE, .Net, and LAMP

Programming tools and development environments

Database servers

Office and Enterprise Applications

Business Intelligence and Data Warehousing

Best Practices, implementation strategy, and resources

Business Intelligence

What's BI?

A technologist will answer “software.”

Big-picture BI encompasses:

Process: event>data>analysis>decision.

Software.

Information: a highly contextual business driver.

For that matter, what's open source?

Analogously:

Process: problem>collaboration>solution.

Software.

Culture: community, framework.

Breaking Down BI

At its simplest, business intelligence analyzes data derived from the business itself (as opposed to such external data as market information); that analysis arrives in the form of answers to questions, either canned or ad hoc. Within that broad range you'll find these subcategories of solutions from over 300 companies.

Business intelligence platform

Business intelligence software

Retrospective BI

- Querying, reporting, and analysis tools

Operational BI

- In-process and transaction analysis, alerting, and reporting (including dashboards and KPIs)

Semistructured and unstructured analytics

- Search
- Dynamic summarization

Predictive analytics

- Modeling and analysis tools

Specialty tools & add-ons

- Geospatial analytics and presentation
- Natural language query
- Content analysis of speech and freeform text

Data management platform

Data sources

- Data warehouse, data marts, and other data repositories
- Data sources
- Data models

Data management

- ETL
- Data cleansing
- Data integration mechanisms

www.infoworld.com/article/07/04/02/14F_Ebizintel_1.html?source=NLC-SR&cgd=2007-04-03

Business Intelligence

BI software consists of:

Reporting; dashboards; ad-hoc query.

Analysis, especially OLAP.

Advanced analytics, e.g., statistics and data mining.

Office/applications integration including EAI.

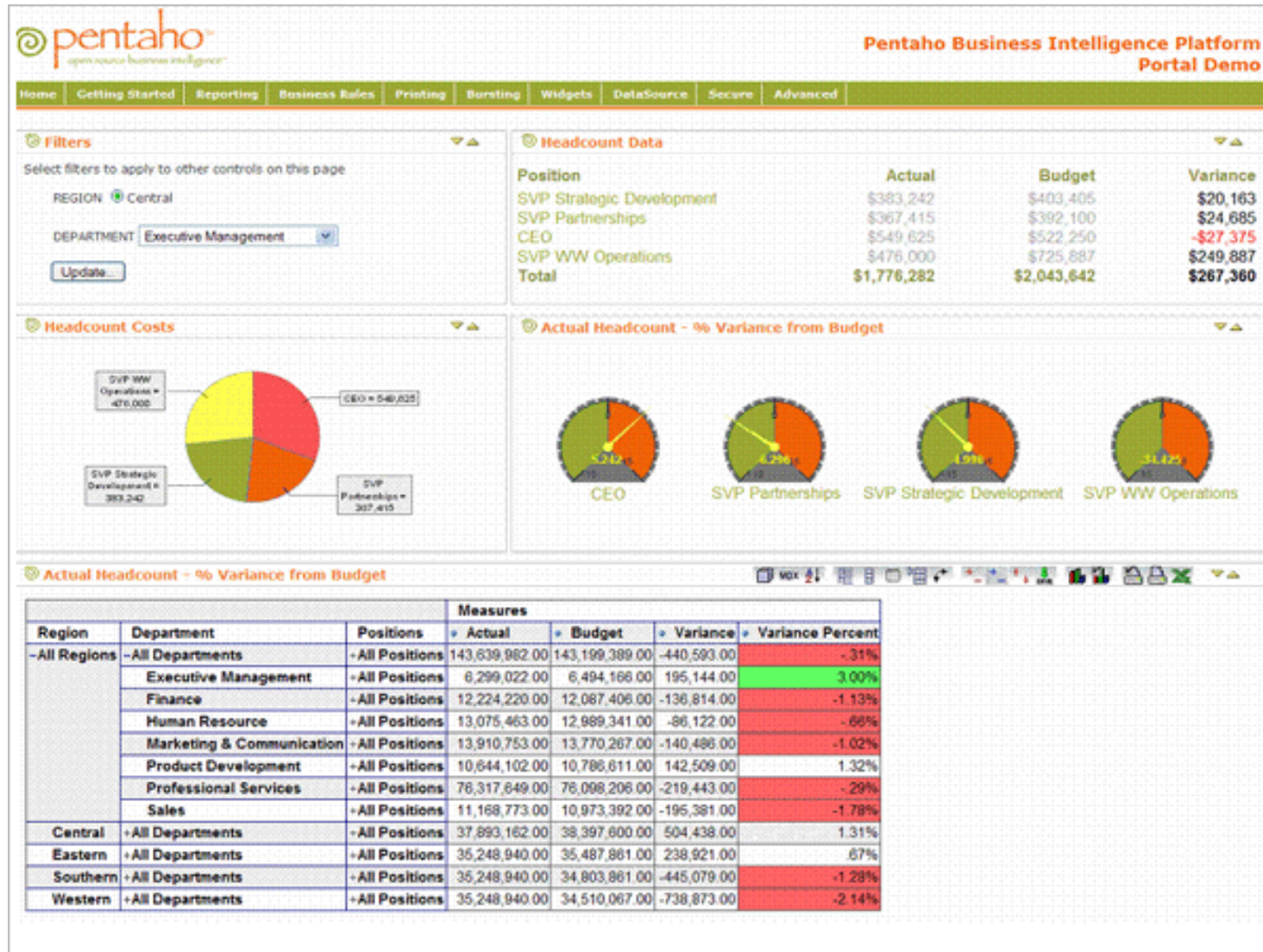
BI relies on:

Information movement & integration, e.g., ETL.

Data warehousing; metadata management.

Visualization.

Search.



<http://www.pentaho.com/products/dashboards/>

Data Warehousing

What's a data warehouse?

A reference database structured for analysis.

Non-transactional, ACID not required.

Contents are cleansed, harmonized, and comprehensive.

Partitioning, bitmap indices, star joins, materialized views,
& cluster/grid/SMP support help.

... with plenty of room for controversy:

Kimball versus Inmon/Imhof versus Teradata.

Normalized versus “dimensional” models.

DW vs. data mart vs. operational data store (ODS).

Real-time and “unstructured” data needs.

The Data Warehousing Scene

There's plenty of DW going on, but:

Teradata is the only *notable* DW pure-play...

blazing a trail for other DW appliance vendors, e.g.,

DATAllegro (Ingres), Netezza, and **Sun-Greenplum**.

Every major DBMS vendor supports data warehousing.

Analytical tools will generally work with any DBMS that supports *standard* APIs/access methods.

What does this mean?

DW techniques are portable to any DBMS platform with the necessary capabilities and tool support.

The Business Intelligence Scene

There are many BI vendors:

Pure-/almost-pure-plays: Business Objects, Cognos, Hyperion, Information Builders, Microstrategy.

The (would-be) dominators: IBM, Microsoft, Oracle
... and their toadies such as Panorama.

Visualization, performance management, reporting, dashboard specialists: Actuate, Applix, arcplan, Pilot, Spotfire, Tableau.

Analytics heavyweights: SAS, SPSS.

Data mining: Angoss, Fair Isaac, KXEN, Megaputer, Salford Systems.

The Business Intelligence Scene

... and then there's the Excel problem, an artifact of the PC devolution.

What does this crowded-segmented field mean?

Vendor lock-in.

When it comes to end-user BI, **open source is nowhere to be seen.**

But let's look at mainstream perceptions...

The BI World According to Gartner



<http://mediaproducts.gartner.com/reprints/cognos/vol3/article2/article2.html>

As of January 2007

What Do the Analysts Think?

Nigel Pendse is author of the OLAP Report –

Actually, I've been quite surprised at how little impact open source BI solutions seem to be having. I was expecting much more.

I guess there are two parallel universes: customers in OSW (open-source world) have decided for idealistic, economic or technical reasons that they must have an open-source solution, and don't even consider any proprietary options, while most other people ignore open-source solutions....

Current OS OLAP solutions are quite weak (at least a decade behind the current proprietary products), whereas the reporting solutions may be better...

The proprietary BI software vendors seem to be genuinely unconcerned by open-source BI. They never mention it to me, and they seem quite surprised if I ask them about it. A few have looked at briefly products like Pentaho, and seem totally unimpressed/unconcerned. I guess they don't sell into OSW anyway, and therefore aren't losing any business to OS BI that they are aware of.

Category Error

My guru friends have made a “category error.”

Open source does not succeed (best) by replicating commercial, proprietary, closed source software and processes.

The most successful open source projects are not imitative, they are innovative.

Think about Internet, server, and desktop computing in this light.

OSBI has NOT aimed to replace closed-source, commercial solutions.

Database Management Systems

There are really two OS-DBMS players in the BI & DW world:

MySQL.

PostgreSQL.

Ingres is possibly the most enterprise worthy, but it enjoys little mindshare.

“We switched from Postgres to Ingres a few years ago after determining that Ingres was much more suited for data warehousing.” -- Stuart Frost, CEO of DATAlegro.

Database Management Systems

MySQL, popular but limited DW capabilities.

Multiengine architecture. We're interested in –

MyISAM.

Merge.

Big strides with MySQL 5, out in late 2005.

Native functions, user defined functions, stored procedures.

Views.

5.1 will add true partitioning.

Nice query, admin & migration utilities. Toad for MySQL is free.

Database Management Systems

PostgreSQL is a more robust enterprise & DW platform.

Greenplum's Bizgres (which is open source) & Bizgres MPP, which is parallelized, are designed for data warehousing.

ExtenDB is layered on PostgreSQL and is DW optimized & parallelized but NOT open source.

BI Components

For reporting –

JasperReports.

Eclipse Business Intelligence and Reporting Toolkit
(BIRT) from Actuate.

JFreeReports.

For data mining –

R is an open source implementation of AT&T's S
statistical programming language.

R-Python links let you extend Postgres!

Weka is a machine learning and data mining tool.

Sample: R

The screenshot displays the R IDE interface with several windows open:

- R Console:** Shows the following code:


```

rgl.sr> ylen <- ylim[2] - ylim[1] + 1
rgl.sr> colorlut <- terrain.colors(ylen)
rgl.sr> col <- colorlut[y - ylim[1] + 1]
rgl.sr> rgl.clear()
rgl.sr> rgl.surface(x, z, y, color = col)
            
```
- R Data Editor:** Displays a table with two columns: height and weight.

height	weight
58	115
59	117
60	120
61	123
62	126
63	129
64	132
65	135
66	139
67	142
68	146
69	150
70	154
71	159
72	164
- Quartz (2) - Active:** Shows a 3D plot of a landscape with a grid overlay. The x-axis is labeled 'long' and the y-axis is labeled 'lat'.
- R Workspace Browser:** Lists objects in the workspace:

Object	Type	Structure
dati	data.frame	dim: 20 4
g	factor	levels: 10
l	numeric	length: 12
n	numeric	length: 1
opar	list	length: 2
pie.sales	numeric	length: 6
pin	numeric	length: 2
scale	numeric	length: 1
usr	numeric	length: 4
women	data.frame	dim: 15 2
height	numeric	length: 15
weight	numeric	length: 15
x	numeric	length: 87
- R Package Manager:** Shows installed and available packages:

status	Package	Description
<input checked="" type="checkbox"/> loaded	graphics	The R Graphics Package
<input type="checkbox"/> not loaded	grid	The Grid Graphics Package
<input type="checkbox"/> not loaded	lattice	Lattice Graphics
<input checked="" type="checkbox"/> loaded	methods	Formal Methods and Classes
<input type="checkbox"/> not loaded	mgcv	GAMs with GCV smoothness estimation
- RGL device 1 (active):** Shows a 3D plot of a landscape with a grid overlay, similar to the Quartz window.
- Code Editor:** Shows a function definition:


```

BoxDens=function(data, npts = 200., x = c(0.,
      add = TRUE, col = 11., border=FALSE, collin
      {
      dens <- density(data, n = npts)
      dx <- dens$x
      dy <- dens$y
      if(add == FALSE)
      plot(0., 0., axes = F, main = "", xlim = x, ylim = y,
          ylab = "")
      if(orientation == "paysage") {
      dx2 <- (dx - min(dx))/(max(dx) - min(dx)) * (x[2.] - x
          x[1.]
      dy2 <- (dy - min(dy))/(max(dy) - min(dy)) * (y[2.] - y
          y[1.]
      seqbelow <- rep(y[1.], length(dx))
      if(Fill == T)
      confshade(dx2, seqbelow, dy2, col = col)
      if (border==TRUE) points(dx2, dy2, type = "l", col = c
      }
      else {
      dy2 <- (dx - min(dx))/(max(dx) - min(dx)) * (y[2.] - y
          y[1.]
            
```

Components

Here are a few more –

JPivot JSP (Java Server Page) tag library.

Mondrian Relational OLAP Server.

Palo Multidimensional OLAP Server.

Enhydra Octopus, Kettle ETL, Kinetic Networks
Extract Transform and Load (KETL), Talend.

See www.manageability.org/blog/stuff/open-source-etl/view .

Related open source packages –

AT&T Graphviz.

Gate, Lucene, UIMA for search/text analytics.

Packages

Pentaho –

JPivot, Mondrian, JFree, Kettle, Weka, Excel services with portal tools and workflow management in a comprehensive framework.

JasperIntelligence is a recent entry (as a suite) combining JasperReports, JasperServer, and Mondrian.

OpenI and SpagoBI provide other frameworks for Mondrian and Jpivot.

Sample: Palo.net – Tensegrity - Eclipse

The screenshot displays the Palo Eclipse Client interface. On the left is the 'Palo Database Explorer' showing a tree view of the database structure, including Dimensions (Datatypes, Months, Measures, Products, Regions, Elements, Subsets), Cubes (Sales), and Views (Sales West, Sales East, Sales North). The main window shows 'Cube 'Sales'' with dimensions 'Measures' (Units) and 'Regions' (France). A data table is displayed for the year 2002, showing variance, budget, and actual values for various product categories across quarters and months.

		2002					
		Year	Qtr. 1			Qtr. 2	
			Jan	Feb	Mar	Apr	
All Products	Variance	-6,618.00	322.00	-523.00	6,286.00	-5,441.00	-3,110.00
	Budget	-782.00	263.00	873.00	840.00	-1,450.00	-3,033.00
	Actual	-1,118.00	-1,699.00	-853.00	0.00	-846.00	715.00
Desktop L	Budget	6,136.00	1,699.00	853.00	0.00	846.00	1,424.00
	Actual	5,018.00	0.00	0.00	0.00	0.00	2,139.00
	Variance	3,510.00	2,255.00	1,527.00	728.00	0.00	-734.00
Desktop Pro	Budget	5,650.00	976.00	727.00	249.00	0.00	2,329.00
	Actual	9,160.00	3,231.00	2,254.00	977.00	0.00	1,595.00
	Variance	798.00	656.00	132.00	576.00	-52.00	-1,255.00
Desktop Pro XL	Budget	5,468.00	681.00	348.00	0.00	333.00	1,955.00
	Actual	6,266.00	1,337.00	480.00	576.00	281.00	700.00
	Variance	441.00	79.00	-158.00	318.00	-81.00	-346.00
Desktop High XQ	Variance	-1,262.00	44.00	-262.00	159.00	147.00	-369.00
	Budget	1,380.00	-199.00	-286.00	416.00	-329.00	66.00
	Actual	-1,356.00	52.00	725.00	-587.00	-86.00	-421.00
Server Power TT	Variance	-242.00	142.00	462.00	-535.00	215.00	-472.00
	Budget	-2,372.00	-920.00	35.00	-381.00	-574.00	-458.00
	Actual	-561.00	-147.00	-449.00	146.00	156.00	241.00

Palo Eclipse Client - Technical Preview III (June 2006)

Applications / Deployments

A few open-source, end-user applications use these component –

Jrubik is one.

As do some hosted applications –

LucidEra is one.

A number of bigger-name organizations have deployed open-source DW & BI.

Ask the vendors about them.

What's Missing?

What are the gaps in the OS stack for BI & DW?

I don't know of robust –

- Master Data Management.

- Data Cleansing.

- Data Profiling.

- Applications for verticals.

– in addition to the lack of end-user BI applications.

Other short-comings?

- Tool integration.

Market Reaction

How have vendors of proprietary, closed source, commercial software reacted?

By porting to Linux, providing limited MySQL support, and exploring Eclipse.

I interpret these steps as mostly positioning for now.

By moving up the applications stack into –

Business Performance Management.

Planning & Budgeting, Compliance.

Industry verticals.

By attempting platform lock-in –

Microsoft Sharepoint.

Market Reaction

But the established vendors shifted tactics before OSBI emerged. What pushed them?

Competition.

Commoditization: Microsoft SQL Server OLAP, Analysis Services.

Opportunity (i.e., \$\$) generated by the enterprise-applications space: SAP, Siebel, Oracle.

Market Analysis

Is OS BI-DW a threat to established vendors?

Not while OS projects/vendors are providing tools but few solutions.

Not until it establishes an end-user presence.

Not until there are more, credible user stories showing robustness, scalability, reliability.

Not while alliances break out of the open-source/small-shop world.

Market Analysis

The answer to the “category error”?

OS BI-DW is doing quite nicely providing developer tools for end-user and embedded applications.

Their route to enterprise acceptance is:

- by leveraging the OS stack.

- by appealing to in-house developers.

- by supporting development shops.

Will OSBI provide the tools (and cost model) to enable the much-talked-about democratization of BI?

Questions?

Discussion?

Next: Best Practices, implementation strategy,
and resources.

Agenda – Day 2

Software frameworks: Java EE, .Net, and LAMP

Programming tools and development environments

Database servers

Office and Enterprise Applications

Business Intelligence and Data Warehousing

Best Practices, implementation strategy, and resources

Best Practices

There are many thousand OS projects.

How do you find the tools you need?

How do you assess tool:

Capabilities?

Total Cost of Ownership (TCO)?

Viability?

How do you evaluate candidate tools?

How can you contribute back?

... and why should you?

Finding #1

Starting point #1: development and distribution sites.

Sourceforge.net.

ObjectWeb.org.

Eclipse.org.

Microsoft's CodePlex, *codeplex.com/*.

GNU, *www.gnu.org/software/*.

There are also tools/vendor specific “forges”:

JasperForge.org, JasperIntelligence BI.

SugarForge.org, SugarCRM.

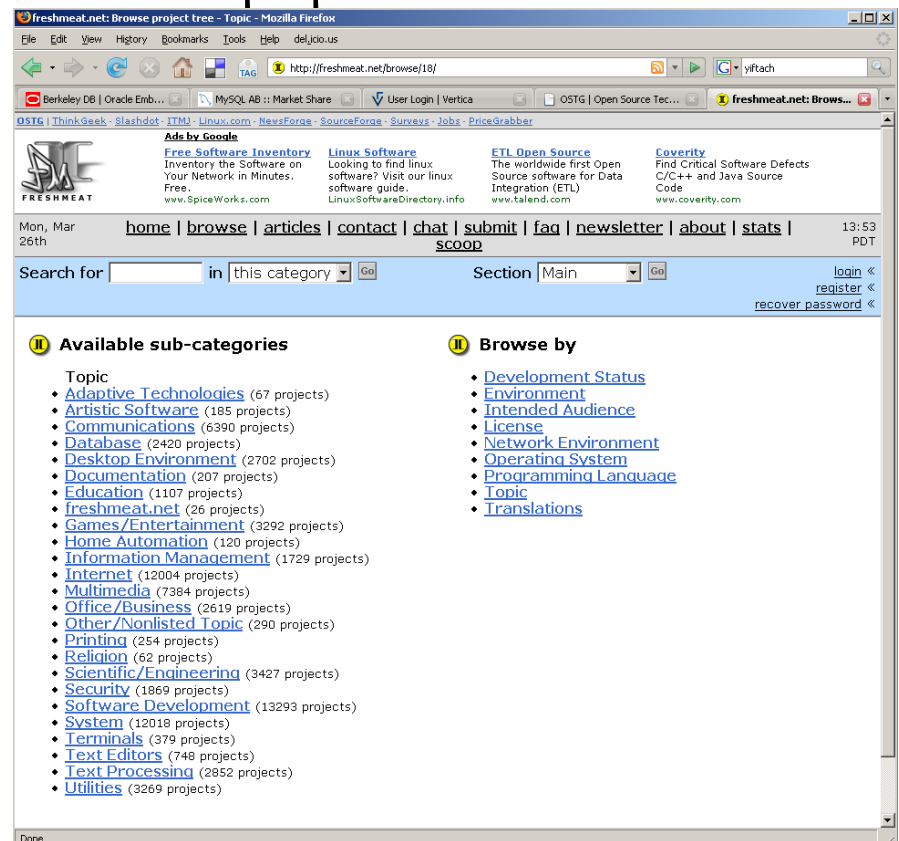
Finding #2

Starting point #2: news and information sites.

Newsforge.com, “the online newspaper for Linux and open source.”

Freshmeat.net

try the browse page –



Freshmeat

“freshmeat maintains the Web's largest index of Unix and cross-platform software, themes and related ‘eye-candy,’ and Palm OS software. Thousands of applications, which are preferably released under an open source license, are meticulously cataloged in the freshmeat database, and links to new applications are added daily. Each entry provides a description of the software, links to download it and to obtain more information, and a history of the project's releases, so readers can keep up-to-date on the latest developments.”

freshmeat.net/

Finding #3

Starting point #3: key open-source platform / stack / solution providers.

Apache.

GNU.

Red Hat/JBoss.

Novell.

MySQL.

A good place to start if you already run software that is a part of a stack.

Finding #3: Red Hat

Red Hat:

Planning an online exchange of open source applications certified to work with the new system.

Would offer user feedback, discussion forums, and download counts.

Finding #4

Starting point #4: Attend project / vendor / open-source / business conventions.

... or just review their Web sites.

Speakers and panels.

Exhibitors.

Finding #5

Another possibility: Look into industry alliances

—

Open Solutions Alliance.

Interop Vendor Alliance.

Established by Microsoft; open source and not.

Includes Red Hat, Sun, SugarCRM, Novell.

interopvendoralliance.org/Directory.aspx

This possibility may be most useful if you already use one of the alliance solutions.

And use the course Resources, which follow.

Assessing

Once you've identified candidate software:

Look for published surveys and assessments.

Check out ohloh – more in a moment – and look for blog entries and articles at sites such as SlashDot.

Look for implementation case studies.

Prominent users and users whose IT environments and/or business needs are like yours.

There's a presumption that market presence (“mindshare”) + and funding = worthiness.

Is this true?

Assessing: Ohloh

Ohloh –

“Mapping the open source world by collecting objective information on open source projects”

“Ohloh collects software metrics from a variety of sources including the project’s source code and the software development infrastructure used by the project’s development team.”

“So far we’ve indexed over 3,000 projects and 220 million lines of source code.”

www.ohloh.net/

Sugar Open Source Analysis Report - Mozilla Firefox

File Edit View History Bookmarks Tools Help del.icio.us

http://www.ohloh.net/projects/77

ohloh
yo!yo

Home > Projects > Sugar Open Source

Sugar Open Source *popular!*

<http://www.sugarcrm.com/>

Report Code Contributors Edit

3 stacks
Stack It!

Sugar Open Source enables sales, marketing, and support organizations to manage their customer interactions more efficiently and profitably and is intended for both large and small companies worldwide. Platform agnostic, Sugar Open Source is the least complex, most portable, and most cost-efficient customer relationship management (CRM) solution. Natively built on the LAMP (Linux, Apache, MySQL and PHP) stack, Sugar Open Source can be deployed in virtually any environment.

business cms crm linux management mysql php sugarcrm web | [Edit Tags](#)

Ohloh Summary

- Mostly written in PHP
- Only a single active developer
- PHP License may conflict with GPL
- Decreasing year-over-year development activity

Updated 25 Sep 2006 21:46 UTC

Average Rating
★★★★☆
(4.0/5.0)
Based on 2 user ratings.

Your Rating
☆☆☆☆☆
Click to rate this project.

Project Cost

This calculator estimates how much it would cost to hire a team to write this project from scratch.

Include	Markup And Code
Codebase	258,228 LOC
Effort (est.)	67 Person Years
Avg. Salary	\$ 55000 /year
\$3,689,746	

Embed this badge on your website.

Sugar Open So...

Find: Next Previous Highlight all

Done

The screenshot shows a Mozilla Firefox browser window displaying the Ohloh website for the 'Sugar Open Source' project. The browser's address bar shows the URL 'http://www.ohloh.net/projects/77/analyses/latest'. The page header includes the Ohloh logo and navigation links for 'Login', 'Sign up', and 'Forgot password?'. Below the header, there are buttons for 'Report', 'Code', 'Contributors', and 'Edit'. The main content area is titled 'Code' and includes a 'Statistics' link. A paragraph explains that the project is over 2 years old and is listed in a sugarforge.org repository. The 'Licenses' section shows a table of licenses used in the project's source code. The 'Codebase History' section features a stacked area chart showing the growth of code, comments, and blanks from January 2005 to January 2007. The 'Languages' section includes a pie chart and a list of programming languages and their percentages.

License	Files
GNU Library or "Lesser" GPL	191 files
PHP License	7 files
GNU General Public License	1 files

Language	Percentage
PHP	69%
JavaScript	20%
HTML	10%
Other	
Perl	<1%
Python	<1%
CSS	<1%
XML	<1%

Assessing: Ohloh

What are Ohloh's "factoids"?

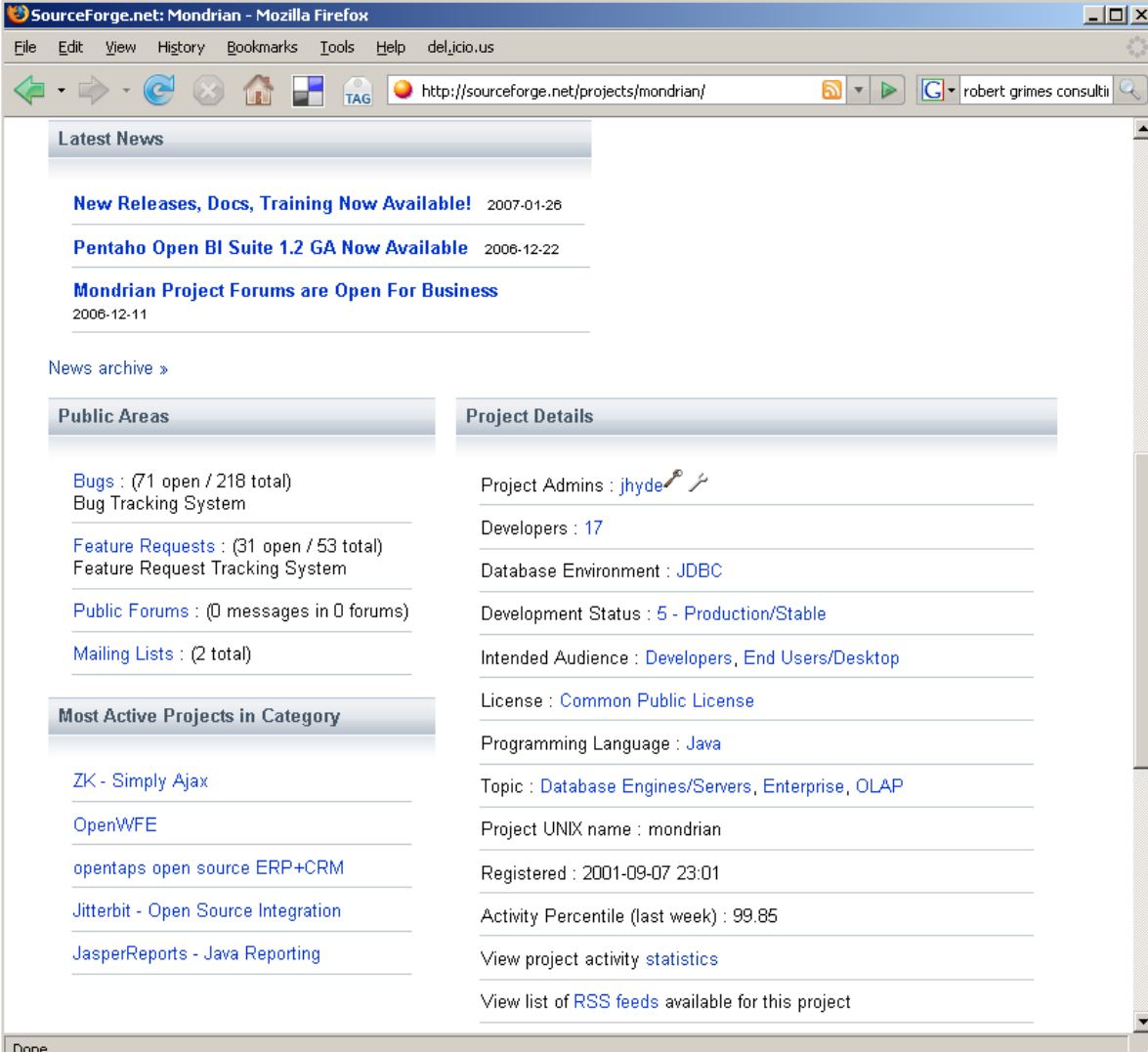
“Mapping the open source world by collecting objective information on open source projects”

“Ohloh collects software metrics from a variety of sources including the project's source code and the software development infrastructure used by the project's development team.”

www.ohloh.net/

Assessing: *SourceForge.net*

Look for project information and for maturity and progress indicators.

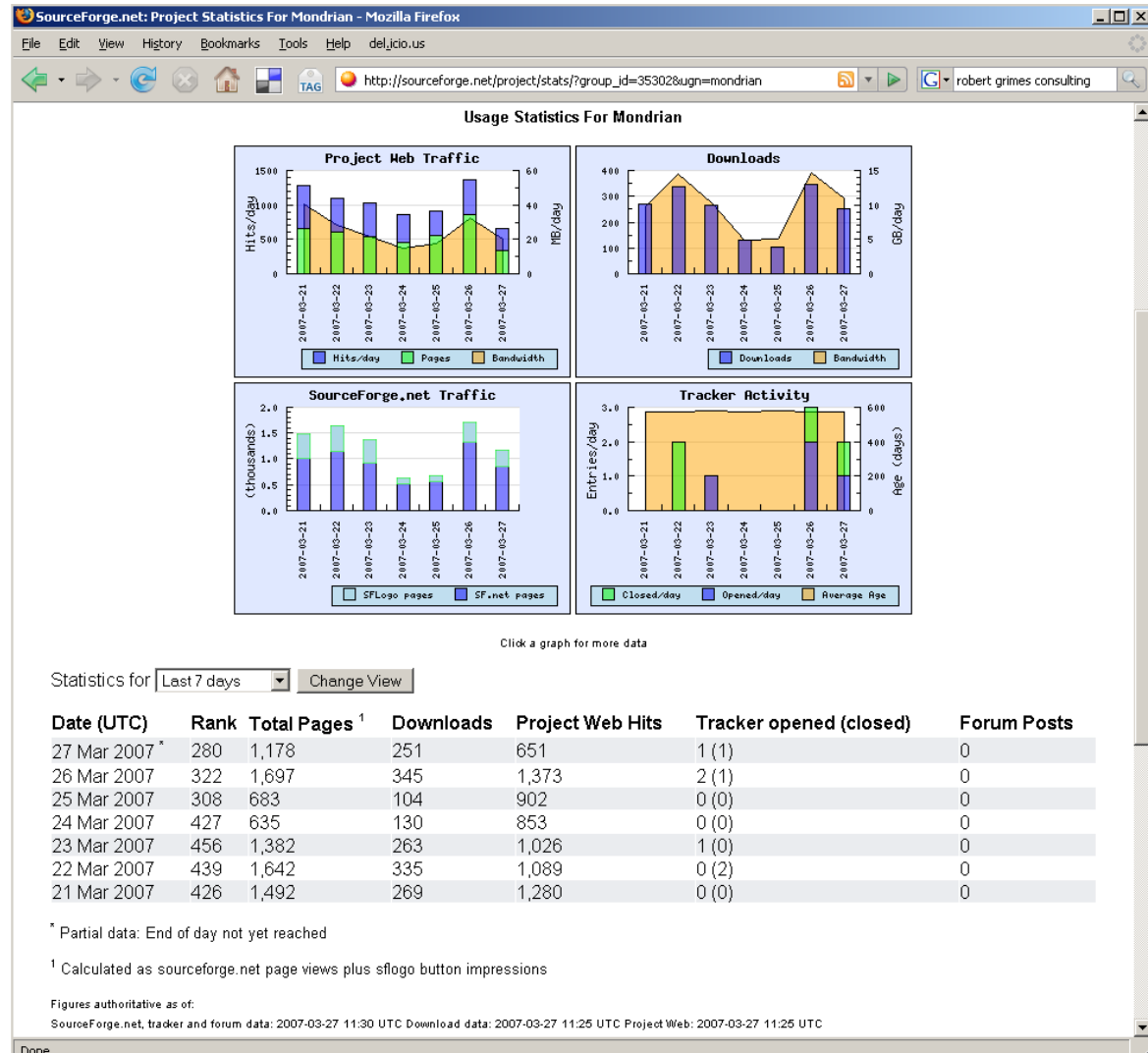


The screenshot shows a Mozilla Firefox browser window displaying the SourceForge.net project page for Mondrian. The browser's address bar shows the URL <http://sourceforge.net/projects/mondrian/>. The page content is organized into several sections:

- Latest News:** A list of recent news items, including "New Releases, Docs, Training Now Available!" (dated 2007-01-26), "Pentaho Open BI Suite 1.2 GA Now Available" (dated 2006-12-22), and "Mondrian Project Forums are Open For Business" (dated 2006-12-11).
- Public Areas:** A section listing various project areas with their respective counts: Bugs (71 open / 218 total), Feature Requests (31 open / 53 total), Public Forums (0 messages in 0 forums), and Mailing Lists (2 total).
- Most Active Projects in Category:** A list of related projects including ZK - Simply Ajax, OpenWFE, opentaps open source ERP+CRM, Jitterbit - Open Source Integration, and JasperReports - Java Reporting.
- Project Details:** A detailed overview of the Mondrian project, including Project Admins (jhyde), Developers (17), Database Environment (JDBC), Development Status (5 - Production/Stable), Intended Audience (Developers, End Users/Desktop), License (Common Public License), Programming Language (Java), Topic (Database Engines/Servers, Enterprise, OLAP), Project UNIX name (mondrian), Registered date (2001-09-07 23:01), Activity Percentile (last week) (99.85), and links to view project activity statistics and RSS feeds.

Assessing: *SourceForge.net*

Look at project statistics.



Assessing: positioning

Look at download numbers...

... which may be distorted. What's being counted?

Review forum entries –

What's the nature of bug reports, feature requests?

What is the competitive positioning...

Relative to competing open- and closed source?

Related to your business domain, users, and IT environment?

Assessing: resources

What resources are required to support implementation and maintenance?

Assessing: licensing

What are the license terms?

Are terms compatible with those of other software packages and hardware you're using?

Some hardware manufacturers' warranties are invalidated by open source (Linux) installation!

You must be extra support from Microsoft to run Windows in certain virtual environments!

Will you create a derivative product for free or commercial distribution?

Assessing: licensing

What are the license terms?

Does the free license scale to your needs?

Capability/functions.

Number of users allowed.

Number of installed nodes.

Number of objects (e.g., database tables).

Number of data records.

Assessing: project

How is the project run?

Vision and roadmap.

Past forks.

Participation

Number of deciders, committers, and users.

Variety of backgrounds, skills, and interests.

Sponsors and backers; partnerships and alliances.

Management approach –

Use of mature management methods and tools.

Documentation and supporting material.

Assessing: project

What support is available?

Community.

Vendor.

Integrator.

Consultant.

What in-house capabilities have you?

Assessing

Thesis:

Many or most business-related assessment questions can be avoided by going with a distribution of a supported stack.

Thesis:

Most technical evaluation points are the same or at least analogous for open source and closed solutions.

Assessing: OSMM

Open Source Maturity Model

Developed by Capgemini in 2003.

Factors grouped in categories.

Scoring model that prioritizes factors.

www.seriouslyopen.org/nuke/html/modules/Downloads/osmm/GB_Expert_Letter_Open_Source_Maturity_Model_1.5.3.pdf

Assessing: OSMM

Usability – The intended user audience, the experience of that group.

Interfacing – Required connectivity, which standards are applicable. How does this fit into the organisation?

Performance – The expected load and processing capability. The performance demands that must be met.

Reliability – What level of availability should the product deliver?

Security – What security measures are required, what restrictions are imposed onto the product.

Proven technology – Does the product use technology that has proven itself in daily production?

Vendor independence – What level of commitment between supplier and user does the product demand?

Platform independence – Is the product available for particular ICT environments only, or does the product allow a wide range of platforms.

Support – What level of support is required.

Reporting – What reporting facilities are required.

Administration – Does the product allow the use of existing maintenance tools, the demands for operational management.

Advice – Does the client require validation / recommendation by independent parties, if so, what is required.

Training – Required training and facilities.

Staffing – Is product expertise bought, taught or hired.

Implementation – Which implementation scenario is preferred?

Assessing: Business Readiness Rating

BRR was a 2005-6 attempt at “a new standard model for rating open source software.”

Invitation-only community sponsored by Carnegie Mellon West Center for Open Source Investigation, O'Reilly CodeZoo, SpikeSource, Intel.

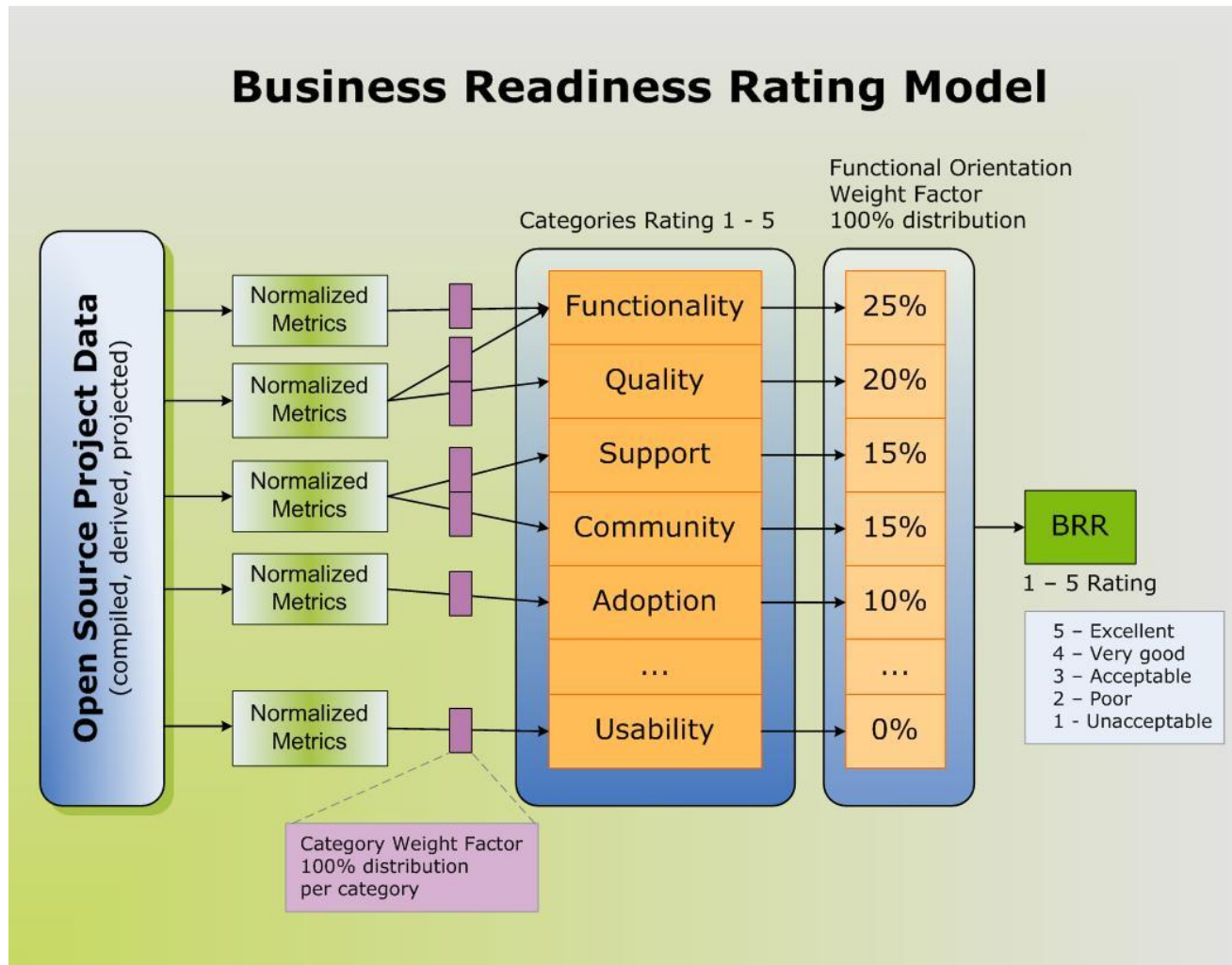
Currently moribund, but let's look at their material...

www.openbrr.org/

The project's 2005 white paper is worth a look:

www.openbrr.org/wiki/images/d/da/BRR_whitepaper_2005RFC1.pdf

Assessing: Business Readiness Rating



Assessing: Navica OSMM

Navica is a systems integrator that has published an Open Source Maturity Model.

Structured process to objectively perform assessments along critical dimensions:

Software.

Support.

Documentation.

Training.

Integration.

Service availability.

www.navicasoft.com/pages/osmm.htm

Assessing: Navica OSMM

	Phase 1: Assess Element Maturity				Phase 2	Phase 3
	Define Requirements	Locate Resources	Assess Element Maturity	Assign Element Score	Assign Weighting Factor	Calculate Product Maturity Score
Product Software						
Support						
Documentation						
Training						
Product Integrations						
Professional Services						

© Navica 2004
 Licensed under the Academic Free License version 2.1
 Open Source Maturity Model is a Service Mark of Navica

**Open Source
Maturity Model**
Product Requirements Template
OSMM Product Requirements Checklist
 © Navica 2004-2005
 Licensed under the Academic Free License version 2.1
 Open Source Maturity Model is a Service Mark of Navica

Product Requirements	
<i>Method</i>	<i>Notes</i>
Create Requirements Task Force	Consider: User Organization, Development/Engineering, Operations, Legal, Customer Service
Poll User Community for Requirements	
Group 1 Requirements	Purpose for Requirement
Group 2 Requirements	Purpose for Requirement
Review Applicable Standards	List formal standards and current status (e.g., draft, comment, etc.); also analyze future standards under development
	List informal or industry standards as appropriate
Standard 1 (Name)	Applicable functionality
	Current Status
	Future Status
Standard 2 (Name)	Applicable functionality
	Current Status
	Future Status
Review Commercial Vendor Product Materials	
Vendor 1 (Name)	Whitepapers
	Technical Specification Papers
	Functionality Checklist
Vendor 2 (Name)	Whitepapers
	Technical Specification Papers
	Functionality Checklist
Review Applicable Analyst Firm Materials	
Firm 1 (Name)	Functionality Checklist
	Vendor Recommendations
Firm 1 (Name)	Functionality Checklist
	Vendor Recommendations

Technical evaluation

Not much different from evaluation of a commercial product.

Ascertain needs.

Prioritize requirements.

Create an assessment approach:

- Function points to test.

- Evaluation criteria and scoring.

- Decision criteria.

Conduct evaluation.

Perform a Gap Analysis: What's missing?

Technical evaluation

Evaluate:

Function points.

Security.

Integration.

Interoperation.

Non-functional criteria:

Scalability.

Usability.

Extensibility.

Performance and throughput.

Technical evaluation

Process:

Design and conduct a functional evaluation.

Create a proof-of-concept prototype.

Vertical slice: All functions for a subset of data or users.

Horizontal slice: Selected functions for all data.

Tests capacity and scalability.

Create a nominal implementation plan.

Nominal = high-level.

Cost evaluation

You must account for:

License fees.

(If a “professional” or “enterprise” or indemnified version.)

Support costs.

Infrastructure: hardware and computing environment.

Typically less for open source given platform scalability, deployment flexibility (e.g., add as needed).

Required mediation, that is, gap bridging.

Participation cost, that is, to “give back.”

Migration strategy

Rationalize your computing architecture.

Tiers:

User interface (UI).

Middleware, business logic.

Back-end, e.g., database, content management, and legacy services.

Hardware and network.

Understand:

APIs (application programming interfaces).

Business processes.

Model your computing environment.

Migration strategy

Rebaseline your requirements.

What functions are no longer used?

What new functions are needed?

What “non-functional” characteristics need improvement?

Assess your budget and staff.

Plan!

Migration strategy

“Big bang” system replacement.

Faster.

Cheaper if it works.

Riskier.

“Divide and conquer” component replacement.

You can start with a “face lift” replacing UIs.

You can possibly replace the back-end systems or the middleware while supporting current UIs.

You could possibly move (some of) the current software to an OS operating system.

Migration strategy

Infrastructure/encapsulation.

Replace APIs, interoperation framework, for instance with SOA.

Platform.

Move existing applications to a new operating system and possibly hardware.

Not always feasible because of application dependencies.

Migration example

Windows desktop to Linux Desktop.

Choose your distribution. Test on your hardware and network with your devices.

Select a virtualization tool.

On Windows:

Migrate Microsoft Office to OpenOffice; selectively convert files; train users.

Migrate desktop databases to MySQL.

Migrate directory/identity management & authentication.

Install Linux and relocate applications.

Decommission Windows.

Open Office migration

One take, from
Network Computing
magazine...

9 | STEPS TO A SUCCESSFUL OPEN- OFFICE MIGRATION

1. Prepare fiscal analysis, compare with other opportunities
2. Enlist upper-management support
3. Conduct "lunch-and-learn" demonstrations
4. Communicate "what's in it for me?" message
5. Decide on document format strategy
6. Execute end-user, train-the-trainer and helpdesk/support training
7. Deploy software 80 percent-20 percent if you must
8. Do perception surveys and address any concerns
9. Create long-term strategy for 100 percent standardization

www.networkcomputing.com/gallery/2007/0219/0219f2b.jhtml

Open Office migration

Do a cost comparison...

COST SAVINGS OF A MICROSOFT DIVORCE AND AN OPENOFFICE MARRIAGE

	YEAR 0	YEAR 1	YEAR 2	ASSUMPTIONS/CONSTANTS
Savings on Office licenses	\$60,000	\$60,000	\$60,000	Cost of OpenOffice and Evolution licenses 0
Training ("train the trainers")	-\$10,000			Number of office productivity users 1200
Productivity loss during training	-\$20,000			Cost of OEM office license 200
External support, hourly contract	-\$8,000	-\$5,000	-\$5,000	OEM license turnover/year 300
Total	\$22,000	\$55,000	\$55,000	Cost of capital for organization 6%
NPV* of switching	\$122,837			

*For a primer on discounted cash flow, an analysis that produces a project's NPV (net present value), see nwc.com/showArticle.jhtml?articleID=171000416.

Using the principles of discounted cash flow and applying opportunity cost to the cost of capital projects, this number represents the present value (net of inflows and outflows) of future cash expenditures and savings.

www.networkcomputing.com/gallery/2007/0219/0219f2e.jhtml

Policy

“Is it now possible to build a complete open source enterprise software stack? I put together the following table that suggests you could get pretty close.”

-- Matthew Aslett, February 2006

www.businessreviewonline.com/os/archives/2006/02/a_complete_open.html

Anti-spam - SpamAssassin/SendMail/Thunderbird
Antivirus - ClamAV/Open Antivirus
Instant messaging - Jabber
Email client - Evolution/Thunderbird
Browser - Firefox
Office productivity - OpenOffice.org/Koffice
Mobile infrastructure - Funambol
Content management - Alfresco/Plone
Collaboration software - Openexchange/Sendmail/Zimbra
BI applications - Pentaho/JasperSoft/GreenPlum
CRM applications - SugarCRM/Compiere/Daffodil
ERP applications - Compiere/ERP5/OFBiz
Programming - PHP/Perl/Python
IDE/dev tools - Eclipse/Mono/NetBeans

Web server - Apache
ESB - Celtix/Mule
Middleware tools - JBoss/Apache/ObjectWeb
App server - JBoss/Geronimo
Database - MySQL/PostgreSQL/Ingres
File/print services - Samba
Virtualisation - Xen/OpenVZ
Operating system - Linux/BSD/OpenSolaris
Systems management - openQRM/Groundwork
Network management - OpenNMS/Groundwork
Backup/archive - Amanda
Application security - AppArmor/SE Linux
Storage management - Aperi
Security - OpenSSH/OpenVPN/OpenLDAP

Policy

Consider using a (commercial) solutions packager. Red Hat is one. Others include – SpikeSource. SourceLabs.

“SourceLabs simplifies and mitigates the risk for large companies adopting open source software.”

SASH consists of 50+ open source projects and runs on all leading J2EE application servers.

SourceLabs tests, applies patches, packages software into a single distribution, and provides support, maintenance, and upgrades.

sourcelabs.com/

Policy

SpikeSource's package is called SpikeIgnited.

Integration: Ensures that all open source components ...work together seamlessly.

Configuration: Easy installation and configuration of open source components.

Testing and Security: Extensive testing, code reviews and packaging by SpikeSource ...plus immediate notification of security issues and delivery of patches via SpikeNet.

Maintenance: [Keep] up with open source infrastructure versions and updates.

Support: 24x7 unlimited support from SpikeSource.

www.spikesource.com/

Resources

Given that open source revolves around community and collaboration, there are many:

Organizations.

Portals.

Conferences.

Blogs.

Projects.

Publications.

I will provide a sampling...

Organizations

Free Software Foundation

www.fsf.org/

www.fsf-europe.org/index.it.html



The Linux Foundation

Merger of the Open Source Development Lab (OSDL) and the Free Standards Group (FSG)

www.linux-foundation.org/en/Main_Page

Open Source Initiative (OSI)

www.opensource.org/index.php

Open Solutions Alliance

osa.org/

Associazione per il Software Libero...

Organizations

L'Associazione per il Software Libero...

... è un'associazione senza scopo di lucro che ha come obiettivi principali la diffusione del software libero in Italia ed una corretta informazione sull'argomento.

Site at *www.softwarelibero.it/* contains a glossary, material on licenses, e-mail lists, etc.

GNUvox

“Le voci di FSFE e delle comunità italiane del Software Libero: uno strumento per fare informazione, mettere in comune strumenti e cultura, sostenere iniziative condivise a sostegno della libertà di pensiero.”

www.gnuvox.info/

Portals

Microsoft's Port 25.

port25.technet.com/

OSDir.com.

News, software tours.

osdir.com/

OSTG, the Open Source Technology Group.

A network of technology sites that includes:

Freshmeat.net.

Newsforge.com.

Slashdot.org.

Sourceforge.net.

www.ostg.com/

Conferences

O'Reilly's Open Source Convention

conferences.oreillynet.com/os2007/

conferences.oreillynet.com/euos2006/

Will there be a 2007?

LinuxWorld

www.linuxworld.com/

ApacheCon Europe

www.eu.apachecon.com/

Open Source Business conference

www.osbc.com

Open Source Think Tank

thinktank.olliancegroup.com/



Blogs

Matt Asay and Dave Rosenberg, Open Sources, in InfoWorld

weblog.infoworld.com/openresource

451 CAOS [Commercial Adoption of OS] Theory

blogs.the451group.com/opensource/

Alex Fletcher's Open Source Unleashed

alexfletcher.typepad.com/all_bets_off/

O'Reilly Open Source

radar.oreilly.com/open_source/

Dana Blankenhorn

blogs.zdnet.com/open-source/

Blogs

Roberto Gallopini, Commercial Open Source Software

robertogaloppini.net/

Eben Moglen's Freedom Now

emoglen.law.columbia.edu/blog

Projects

FLOSSMetrics

Sponsored by the European Commission.

Stands for Free/Libre Open Source Software Metrics.

Aim is to “construct, publish and analyse a large scale database with information and metrics about libre software development.”

flossmetrics.org/



Publications

Open Source Forge

News, tutorials, articles, press releases.

www.osforge.com/

Enterprise Open Source Journal

Suspended publication, but back issues are useful.

www.eosj.com/

The 451 CAOS Community

www.the451group.com/caos/caos_community.php

Questions?
Discussion?

Thank you!