Palo
by Jedox

plan analyse report

Commercial Open Source
Business Intelligence für
Performance Management

Jedox WHITEPAPER

Jedox WHITEPAPER

Jedox WHITEPAPER

INTRODUCTION

Over the years many terms have been used that attempt to summarize the need to collect, align, plan, analyse, consolidate, report and distribute corporate data throughout the organization. Some of the better known terms are "MIS", "EIS", OLAP, "Business Intelligence", "Corporate Performance Management", "Enterprise Performance Management" and "Business Performance Optimisation". The term "Corporate Performance Management"(CPM) is currently popular and seems to encapsulate the most complete vision. CPM is built on the foundation of Business Intelligence, but ties it to the planning and control cycle of the enterprise – embracing enterprise planning, consolidation and modelling capabilities. As a simplification, we will still use the term Business Intelligence (BI), but we refer to the extended meaning of Corporate Performance Management. Jedox' vision of the future of BI is simple fast and user-driven technology instead of the centrally controlled products of the past.

Business Intelligence does not have to be complicated. Many vendors and key influencers position BI as an expensive and highly sophisticated solution, but it basically means managing numbers and performing basic arithmetic operations on the numbers, and even division is a bit exotic, primarily being used for extracting the so called 'Key Performance Indicators'. So BI is actually based on generic and relatively easy mathematical operations. The actual calculations are generated by combining these generic operations in the "right way", which then leads to a sophisticated and possibly more complex solution.

From experience we also know that the "right way" is very different from organization to organization. This explains why there isn't a defined standard solution for Business Intelligence. This makes BI quite different from ERP or CRM. Even competing organizations in the same industry build very different BI solutions suited to their specific needs. Indeed in some cases, such as the retail industry, details of the BI system design are regarded as proprietary information that gives the company a competitive advantage. It seems that in the same way as successful organizations each have their unique way of delivering value to the customer, they also have to model their individual BI solution to get the most value out of their various corporate data sources.

As there are different approaches to a business and different business models, there are different needs to the Business Intelligence in companies. Therefore Jedox has designed a flexible BI solution, which is easy to adapt to the business needs.

Open source is more than just free software. It defines a new relationship between the user and the software provider based on openness, fairness and transparency. Palo does not come with hidden costs or arbitrary license fees. Instead the users only pay for the services they require, which are priced in a straightforward way.

Jedox believes in open source, as represents transparency and trust between the software vendor and the user.

PUTTING THE USER FIRST

The expertise needed to find the right BI solution for a given company comes from the business users. Business intelligence is an information technology, but IT departments usually lack the expert knowledge needed for modelling and design. Instead, business departments such as finance and controlling, sales, marketing, and manufacturing have the expert knowledge if not the technical skills for data modelling. Requirements for BI solutions change continuously. If the IT department is responsible for modelling and implementation is often delayed by the feedback cycles between business users and developers.

As a result, business users need to take a leading role in building and managing successful BI solutions. IT departments do come into play when it comes to technical issues such as data acquisition from ERP, CRM or other IT managed data sources or providing IT-infrastructure for the BI solution such as server hardware, networks, installation, and so on. Finding the right balance between IT and business users is the key to successful BI initiatives. Software that requires a minimum of IT involvement makes it easier to find this balance.

Finally, today's BI solutions are available to only a small portion of their potential users. Many companies have such solutions but in most cases they are only made available to a small number of employees. The relatively high license costs of today's CPM and BI solutions are one important reason for the limited spread of CPM solutions. Significant training and installation costs for new software as well as poor data integration to Microsoft Office applications such as Excel also explain why today's BI solutions are still unable to leverage their full potential.

Big BI solutions are often only available to a small group of people, due to license and installation costs. This often produces an "island" within the company. Palo is available for many users at the same time, allowing a full integration into the Business process.

JEDOX PALO SUITE OVERVIEW

Jedox offers an integrated and easy-to-use BI solution to collect, plan, analyse, consolidate, report and distribute corporate data based on a centralized multi-dimensional data model. The entire suite is open source and can be freely downloaded for immediate use. Jedox technology provides simple generic BI functions that can be combined in a variety of ways and scale to the level of complexity that is comfortable for the user. Modelling and administration are simple enough to be carried out by the business departments. Both the modelling/ administration features and the end-user features can be accessed with the Web-based BI spreadsheet or with Microsoft Excel. Palo works well as a desktop solution but its complete Web support also positions it as a platform for software as a service (SaaS). The underlying open source license allows its free and widespread use at any workplace in the organization.

Palo is built to avoid the hurdles companies encounter when rolling out a BI solution. Palo can freely spread within the corporate user base and be administered in a spreadsheet environment with minimal involvement by the IT department. As a result Jedox has made a major leap towards idea of "democratizing business intelligence" within the organizations. Palo's liberates BI from complexity and overhead, enabling business users to build BI solutions quickly.

Palo is designed in such a simple manner that it can be modelled and administered by the end user.

Jedox believes in democratizing BI, which means that Palo is easy and quick to install, and therefore reduces time and investment costs significantly, in comparison to proprietary software.

It provides results without having to write large complex specifications and without having to set up monster projects that take months or years before delivering any visible results. Palo makes Business Intelligence easy, both in a sense of simply doing it and of doing it simply.
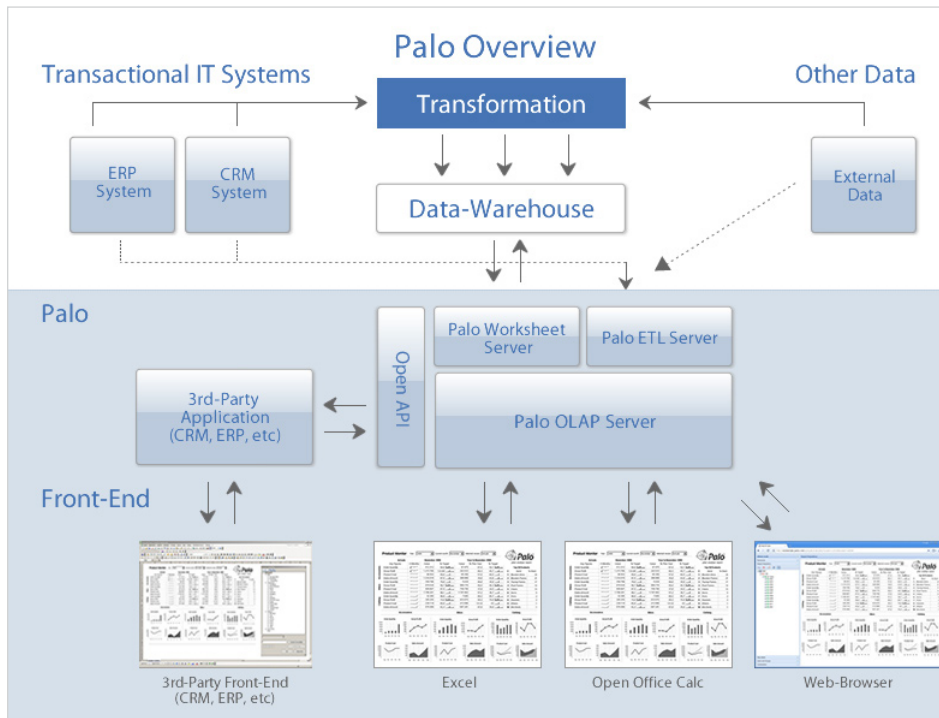
DESIGN PRINCIPLES

Jedox' primary design principle is to minimize the amount of overhead, time, cost and IT involvement in the creation and maintenance of BI solutions. This is achieved by pursuing the following key development guidelines.

- Optimize the UI for the business user. The user interface is built with the business user, not the IT technical person.

- Use technical concepts business users can understand. Business users may be expected to understand cells, lists, hierarchies, formulas and cell coordinates. They are less likely to understand aggregation tables, star or snowflake schema designs, parent-child relationships or SQL commands.

- Support Microsoft Office. With Microsoft Office being the standard software environment for the business user, Palo should be compatible to Microsoft Office, especially Microsoft Excel and the spreadsheet user interface.

The user interface is designed with regard to technical knowledge level of the business user.

With Microsoft Office being the leading equipment of business users, Palo is fully compatible with theses programs.

- Make the product responsive. Response times for screen actions should be very fast to get the same user experience as in products like Microsoft Excel. This should apply to recalculation and to interactive exploration of different calculations schemes and scenarios. With spreadsheets using in-memory technologies to provide real-time calculations, the same should apply for BI functionalities residing on the back-end server side.

- Get technology out of the user's way. This applies to infrastructure, technical pre-requisites, specific operating system versions, services packs, and required databases.
Palo's user management is also much simpler than most database applications, while still providing powerful data security features.

- Provide platform independence. The technology should not be bound to one OS only (such as Windows). The software should be truly platform independent (both client and server).

Jedox pursues similar goals in its licensing policies. The license model should not impose unnecessary risks on the owner of the BI systems. High initial costs such as license fees should be avoided. More favourable are open source or SaaS models, as they either do not impose any license fees or are bound only to a monthly payment that can easily be adjusted to the growing or shrinking needs of the customers.

Palo is designed to provide response times, that are comparable to those in MS Excel, letting the business user experience real time calculations.

Even though provided with many security features, Palo reduces the administrational effort i.e. updates or downloads to a minimum.

COMPONENTS OF THE PALO BI SUITE

The Palo BI Suite is made up of three tightly integrated components:

- The Palo OLAP Server, a multidimensional in-memory OLAP Server.

- The Palo ETL-Server, a data acquisition tool for Palo OLAP Server, which also supports drill through back into the source data.

- The Palo Worksheet-Server is a web based client for modelling, analysis, planning, reporting connected to the Palo OLAP Server and additionally relational data sources.

The Palo BI Suite in the Community version (100% Open Source) features the following:

• Palo OLAP Server
• Palo ETL Server
• The Palo Excel Add-in
• Palo Worksheet-Server

The Palo OLAP Server also supports the Palo Supervision Add-In for workflow capabilities, event handling and scripting. In addition Jedox provides powerful add-ins for Microsoft Excel and OpenOffice.org Calc which both act as powerful end user clients for Palo OLAP Server.
The Palo OLAP Server is the core component of Jedox BI offering. In combination with the free-of-charge Excel Add-in it is a complete BI solution by itself, providing features for data collection, reporting, analysis, planning and consolidation in the enterprise.
The Palo BI Suite with Palo Worksheet-Server and Palo ETL-Server add significantly more value to Jedox BI offering, but Palo-Server and the Excel Add-in are a highly competitive departmental tool in their own right. Usually Palo OLAP Server and the Excel Add-In are the first components to be employed.

As the amount of data and the number of users increase with time the Palo based BI applications become more mission-critical and the remaining components of the Palo BI suite move in focus.
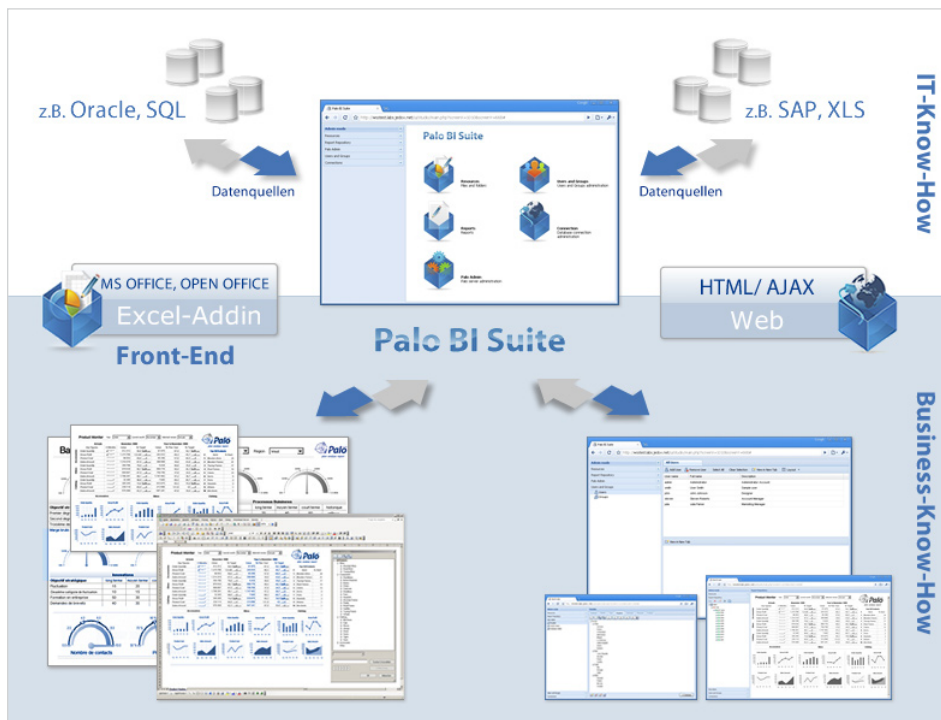
PALO OLAP SERVER

Palo OLAP Server is a multidimensional in-memory OLAP Server. All the data it holds is organized in cubes, dimensions, elements and element attributes. A similar model has been used over and over again in various forms in a wide variety of products going back to the beginnings of computing. It is so durable because it accurately reflects the way users think of their business. Palo offers a model which is ideal for departmental solutions and self-service BI because it is particularly easy to use.

The most rewarding aspect of the in-memory technology is the speed advantage. Compared to an old-style ROLAP system in-memory technologies have the potential to be as much as 100 times faster than an OLAP system with a disk based relational database. Speed is especially important for Palo, as it supports multidimensional formulas and write-back of data.

When the data in the cube changes, all aggregated numbers change as well as all calculation results are instantly available real-time. The cubes do not have to be reprocessed. They are ready to use as soon as the users defines them. Palo's real-time aggregation and calculation capability is particularly important, since Jedox view of business intelligence includes the planning – not just the analysis and reporting features of traditional OLAP.

The Palo Olap server works in memory and provides real time calculation, as well as cross data source calculations.

The server supports multi-core and multi processor architectures, and offers a completely open API that can easily be accessed from any third party application.

Palo OLAP Server also supports real-time modelling. This means users can modify hierarchies and even create new cubes on the fly in the familiar Excel environment. The database returns new aggregated values immediately. Real time modelling is a powerful simulation feature, that lets planners play out what-if scenarios involving changes in their corporate structures. This is another reason why the Jedox designers chose in-memory technology when they started to implement Palo OLAP Server.

The Palo OLAP Server has an advanced design that is far more powerful than a traditional OLAP technology. Like most OLAP products Palo OLAP Server has cubes, dimensions, elements and element attributes. However in addition to the traditional aggregation functionality of OLAP cubes Palo also offers a sophisticated calculation engine that supports cross-dimensional business rules. Some planning tools provide time intelligence, currency conversion workflow and inter-company eliminations as expensive add-ons for complex analysis and planning scenarios. But Palo OLAP users can model these features themselves, and the server can carry out the calculations in real-time.

Palo provides easy-to-use features to modify and cleanse data before it is imported into the cube. In real-life projects the source data is often plagued by quality problems that make analysis difficult. Misspellings and empty database fields are the most common problems. Palo ETL-Server offers wide ranging features for filtering and modifying data – the suite is specifically designed to cope with poor data quality.

A single Palo cube can contain data from multiple data sources. This makes it straightforward to analyze across data sources – for example to compare actuals and plan data. Traditional OLAP systems build their cubes on the structure of some source database. As a result, the cubes and the cube structure can only be as good as the underlying database. In contrast, Palo users can import the data they need incrementally from any data source Palo can access and store it a structure they define themselves. It doesn't matter what the source data looked like. Palo's data architecture is also simple to understand and doesn't bring along unnecessary baggage from its data sources such as complex ERP structures. Instead, the user can model the data the ways he needs it.

PALO EXCEL-ADD-IN

Microsoft Excel is the most widely used BI-Tool, but it is often used inappropriately as a store of data instead of a presentation tool with ad hoc calculation abilities. But this changes dramatically if Excel is connected to a multidimensional and centralized database like Palo OLAP Server. Excel then becomes a highly sophisticated BI front end without the well known "Excel hell" issues which arise when a spreadsheet is misused as a database. Most important, with a database in the background Excel puts a huge amount of power in the hands of the people who need it most – the business users.

Palo Excel Add-in provides a cell based system based on Excel formulas that directly access the database. This allows users to create a wide variety of reports without

The Excel add-in is one option to access the data in the Palo OLAP server.

The Palo Excel add-in turns Excel into a powerful BI front end, with a multi dimensional database calculation the data in the background.
The Palo Excel add-in makes it easy for a group of people to enter their data into a spreadsheet, as it is centralized on the server for everyone to see.

the usual limitations of a single view on the data that nearly all other BI systems have. Because each cell accesses the database separately, it is no problem for a single report to access several data sources and combine the results. It is even possible to use Excel formulas that combine values from two separate databases in a single cell. In addition the system provides a dynamic ad hoc analysis system which gives business users immediate access to stored data. Because Excel is so flexible the Palo Excel Add-in can be used for all scenarios from simple reports to be shared among a few colleagues up to complete applications for large audiences of end users.

Palo Excel-Add-in is also the environment for creating data models. End users put off by the idea of creating their own data models will be pleasantly surprised by the simplicity of Palo modelling tools. A cube can be created in a few minutes by a business user. It is a simple as defining a few word lists as dimensions and putting them together in a cube. As soon as the cube is defined it can be filled with data – either by typing it in by hand or by using spreadsheet formulas to fill it. Palo also provide spreadsheet formulas to put long lists in Excel into dimensions. As a result, with a little practice a business user can convert an existing Excel based reporting system into a database system in a matter of hours.

Palo Excel Add-in offers build-in functionality for importing data from files or relational databases into a Palo OLAP Server. One particularly powerful method is the internal loop function of the data import wizard, which allows non-technical users to apply the power of spreadsheet calculations to the problem of data import. This spreadsheet approach is ideal for business users who need ad-hoc imports because it is simple and leverages the user's knowledge of spreadsheet functions. In fact Palo is one of the very few products on the market to address this real world issue – most BI vendors would question the very idea of an ad hoc data import, even though it is a common requirement in departmental solutions.

Palo's Excel-based import capability puts a lot of power into the hands of the business user, but it is not intended to cover all the bases. For high volume, fully automatic data imports which are not limited by Windows, Jedox also provides the Palo ETL-Server.

PALO WORKSHEET-SERVER

The Palo Worksheet-Server offers a complete Web based reporting and analysis system that can be maintained by business users. Reports are created and published in the web, and flexible secure content management is provided. The product can be used as a basis for an intranet solution or for the exciting market for BI software as a service (SaaS).

Palo Worksheet-Server has a user-friendly report creation system which succeeds by being a logical extension of the spreadsheet environment so familiar to hundreds of millions of business users. It provides a unique method of combining cell based spreadsheet formulas and formatting with database storage and flexibility. Palo Worksheet Server is a clear improvement over most spreadsheet based products, which fall into the following main classes:

With the Palo Worksheet-Server, Jedox presents a powerful web front end, making it easy for example to create and publish reports in the intranet or the web.

The Worksheet-Server provides the user with the familiar look and feel of Excel, but also features innovations, such as the Dyna Ranges (a dynamic formatting technology).

- Dynamic solutions that use Excel as a simple grid, and do not really deliver on the promise to extend Excel. These products have their own formatting and formulas, and provide little or no ability to mix Excel functions and formatting with the data from the database.

- Static solutions, offering cell based access to the database, which is better than storing data in Excel. These systems leverage the strength of Excel, but the reports are static and require manual updating when they change.

The Palo Worksheet-Server's DynaRanges offer a natural logical method of combining spreadsheet calculations and formatting with dynamic database features. Because it is so similar to Excel, business users become productive after a few hours training, creating reports the same way they would in Excel, but the resulting spreadsheet are automatically updated when the database is modified – for example when new elements are added to a dimension. Furthermore filtering and sorting features based on data in the server are also provided, and Palo OLAP Server's comprehensive security system also applies to the reports, guaranteeing that each user only sees the data he is permitted to see. It also supports single sign-on with Windows security.

PALO ETL-SERVER

The ETL-Server provides industrial strength extraction of mass data from a wide range of data sources. It transforms and loads both master data and transactional data into a wide range of target databases. The ETL-Server not limited to loading the Palo OLAP Server, but it has been tailored to the specific requirements of mass data imports and exports to and from Palo models.

With the help of Palo ETL-Server, data can be imported simply, flexibly and automatically. The ETL-Server can build an entire Palo model from scratch or add to an existing database incrementally on a regularly scheduled basis. As a pure Java application, Palo ETL-Server is platform independent. It is even possible to have a test environment under Windows OS and a productive environment under Linux.

The Palo BI Suite Enterprise Edition extends the connectivity features of the ETL-Server by offering complete support for SAP BW. The Enterprise Edition can access any multidimensional ODBO data source using the industry standard MDX query language.

Palo ETL-Server can be used in two different modes: In stand-alone mode as a simple desktop application or in Client-Server mode. In the latter case, an ETL service is installed on the server machine which manages all ETL projects centrally and executes them on demand. The ETL-functionality is the same in both cases.

The Palo ETL-Server extracts mass data of various data sources into a wide range of target databases, an of course into the Palo OLAP Server.

Tailored to the specific needs of the Palo PALO server, the ETL-Server communicates the data automatically.

Palo ETL Server : Architecture

Palo ETL Server, 08/2008

Jedox
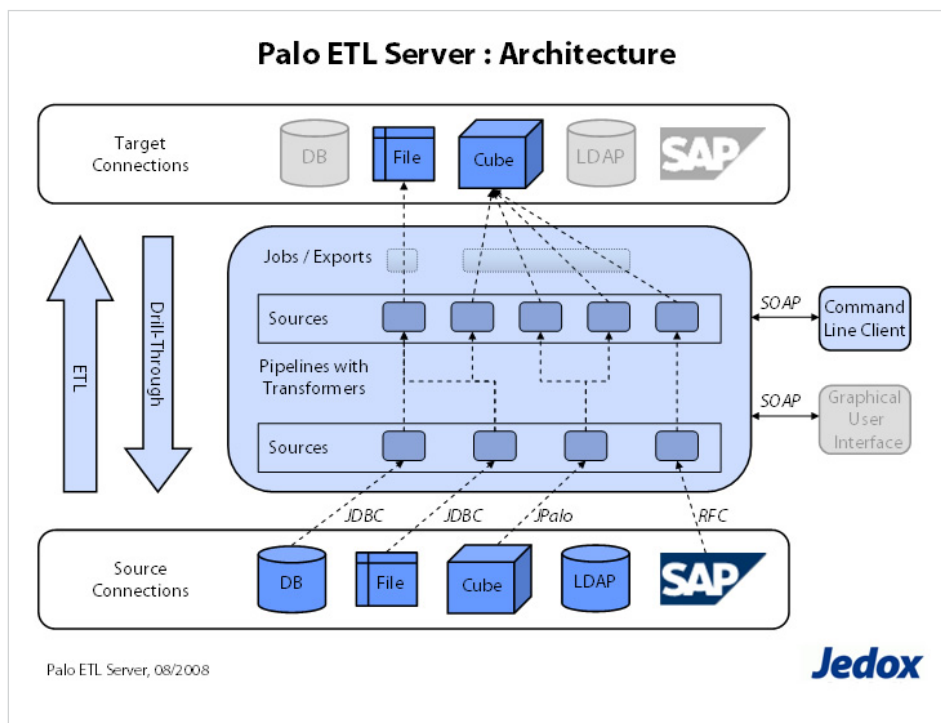
The Palo ETL-Server transfers one or several sources into a target by mapping, aggregating, filtering and joining the data. The joined data can then be drilled down, to show the level below the aggregated figures within a doubleklick, right back down to the raw data.

Palo ETL-Server itself features a SOAP based ETL-API interface, which allows open access to all modelling, monitoring and execution functionalities.

As source systems, all common relational databases such as Oracle, Microsoft SQL Server or MySQL are supported. The connection is established via a JDBC driver (Java Database Connectivity), a widely used interface to access relational databases. Palo OLAP Server itself can be used as a source system; transformation processes between two Palo models and export to flat files are possible. Data from flat files and LDAP Servers, e.g. Microsoft Active Directory Service, can be imported as well.

Data import and Modelling

More advanced projects require a complete ETL model, made up of individual ETL tasks. The transformation of data is performed in pipelines. They transfer one or several sources into a target by mapping, aggregating, filtering and joining the data. The field-conversions are described in transformers: There are pre-defined transformers used for modifications such as mapping transactional data to the right master data, data modifications such as adding prefixes to product names and date handling. To cover project-specific transformation requirements, custom transformers can be developed. The programming language is standard Java.

The ETL-Server provides advanced modelling features for hierarchies. It has detailed support for defining consolidations, attributes and weights. This is an area which most ETL tools have trouble with, but it plays a crucial role in the business logic of many multidimensional solutions. Various formats of hierarchies are supported including parent-child hierarchies, full hierarchies, and N-C-format hierarchies. The product also offers advances features such as transformations and joins on hierarchies.

ENTERPRISE VERSION

The Extension Packs of the Enterprise Version cover several areas.

- Extended Control: The Enterprise Version offers extended automation features and control over the application with process control, alerts and workflow with the Palo Supervision Server.

- Extended Scalability: The Enterprise Version supports multicore shared memory architeture for multi-core servers and extremely high-end GPU-based dedicated hardware solutions.

- Extended Connectivity: With the SAP Connector the Enterprise Version provides semantic support for SAP BW and R/3. The Enterprise Version can also access any multidimensional ODBO data source using the industry standard MDX query language. Palo BI Suite Enterprise provides a SAP interface for the ETL-Server.

- Extended Support: Jedox provides a range of services and extensions to the Palo BI Suite Enterprise. In particular Jedox has a complete support program with telephone support and timely hot fixes for individual projects.

- Extended insurance: Jedox' Software Assurance program provides companies a guarantee that the product will perform as expected.

The enterprise edition includes useful security and performance features, such as the SAP connector, the supervision Server, Software Support and Software control.

PALO SUPERVISION SERVER

Palo Supervision Server is an additional component of the Palo BI Suite designed to simplify and standardize event handling and process control. This is an area that can be complex and is often addressed by BI vendors large and small in a piecemeal fashion, with additional servers and components added over the years as the need arises. The results tend to be unsatisfactory, and vendors rarely discuss them when presenting the product to perspective customers.

Jedox's design principle is to reduce technical overheads and simplify project implementation. The Palo Supervision Server is a single component that deals with process control, workflow, events and alerts. It also offers features for integrating third party applications. The server applies the same logic to the business user oriented planning processes and business calculations such as line item consolidations as it does to technically oriented processes such as data imports. The same logic can also be applied to creating business alerts. The Palo Supervision Server also provides version controlling for reports and tracks changes in the database itself, making it a key element in internal control system as mandated by regulatory requirements such as Sarbanes Oxley.

The Palo Supervision-Server is able to log and control data processing and other Palo activities in order to extend process control.

Problem solving with Palo

There are many problems, the business user encounters quite frequently that the Palo BI Suite can solve. In the following you will find some examples of different kind of troubles our customers have solved with the Palo BI Suite and its components.

1) Standard life Budget planning process with excel and Palo

When planning the budget with Excel, Standard Life encountered the typical problem when working with Excel and having to report to and receive figures ie. for a report from a larger group of people.

Before Using Palo, a master sheet was created, which was then adapted to the needs of the recipients. The spreadsheets were rolled out via mail, where every recipient got his or her own spreadsheet to be filled out with the corresponding data.

Once the data was filled in, the recipients would send back their spreadsheets to the initiator, who would then check the returned spreadsheets manually for typos, formula corrections, empty data, etc. Once having all the information necessary, the spreadsheets were then manually consolidated into one spreadsheet that was used for the presentation to the board.
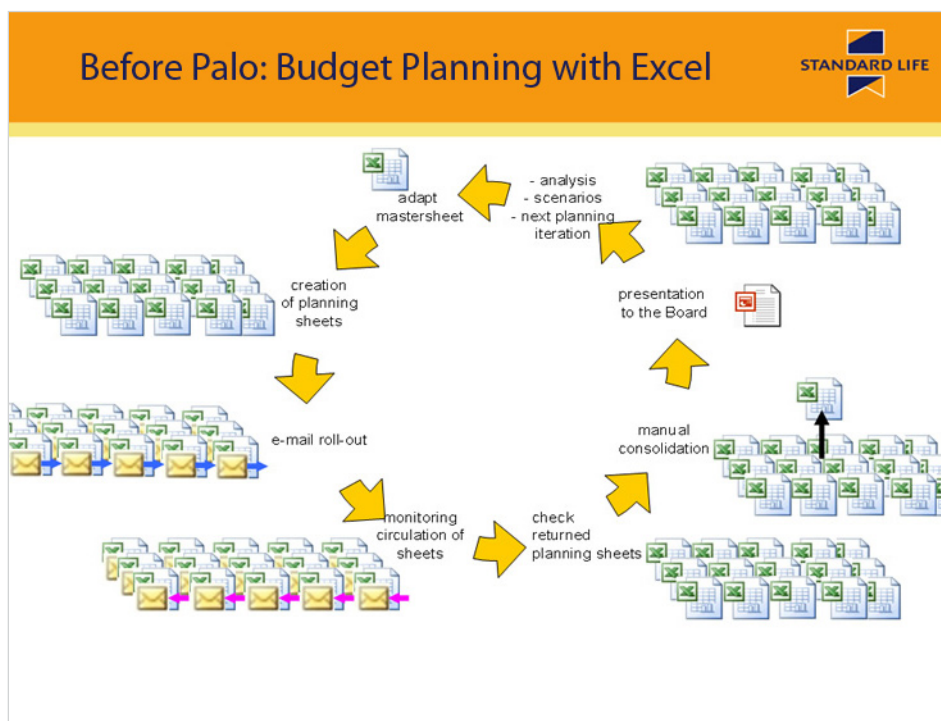
Standard life Budget planning process with Palo

For the process, a cube or configuration file had been created. An email was written to the recipients, with the link where the cube/ configuration file was to be found.
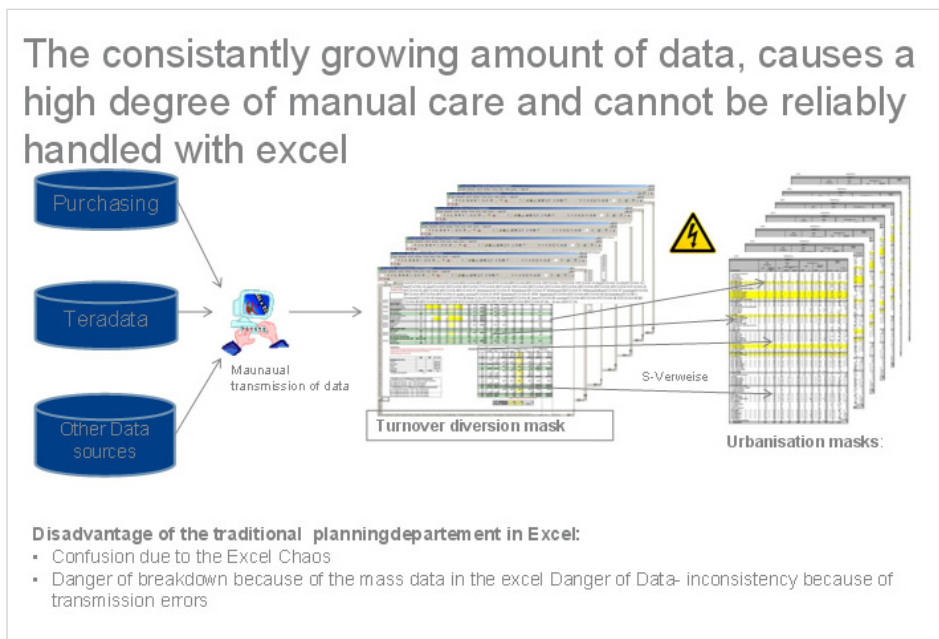Every recipient could enter the corresponding data into the same centralized spreadsheet, which then did not need checking or manual consolidation.
The Data is consolidated automatically, and could directly be presented to the board, having the cycle lose 4 steps and simplifying the process daramtically.
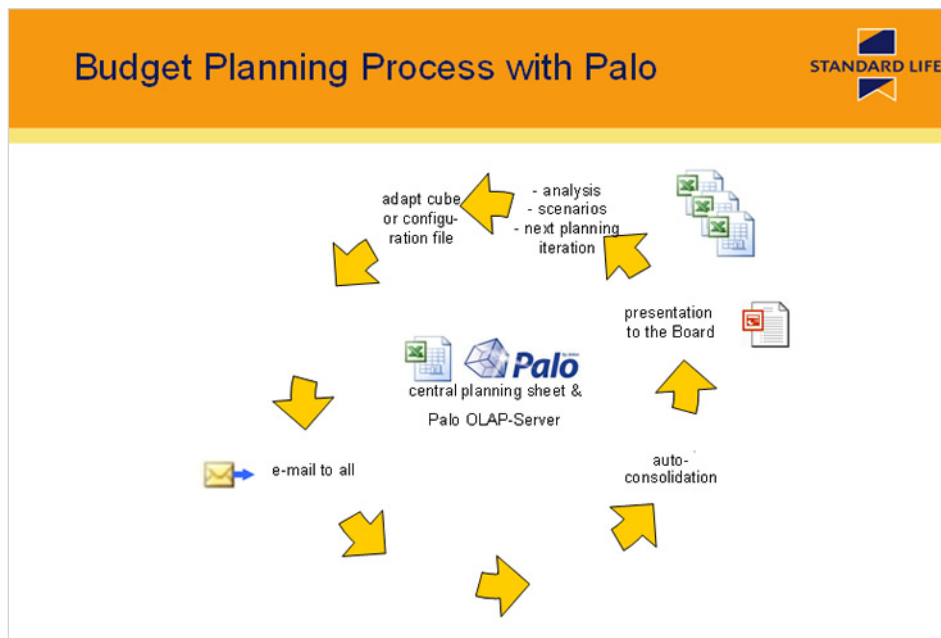
2) Single Point of Truth at Otto Purchasing Dept.

In big corporations it is hard to find "the single point of truth" when it comes to reporting your KPI's and planning the next fiscal year. Bayer Healthcare had made an effort to consolidate the data from multiple data Sources into one simple to handle spreadsheet. Reason for this decision was the lack of transparency, due to the highly complicated formulas which linked across multiple data sheets, the risk of producing false date with the manual import of data and the fear of a breakdown which would erase the large amount of data in the multiple excel spreadsheets.

The consistantly growing amount of data, causes a high degree of manual care and cannot be reliably handled with excel

Disadvantage of the traditional planningdepartement in Excel:
- Confusion due to the Excel Chaos
- Danger of breakdown because of the mass data in the excel Danger of Data- inconsistency because of transmission errors

With the Palo Olap Technology and the Palo ETL server the data was automatically aggregated and centralized leaving no room for mistakes in the process of transmitting the data from one system into another.

With the ETL Server being able to write in both directions, into the cube and back into the data origin, the planning data is also saved in the data sources connected to the ETL, making it easy to compare plan and actual figures to each dataset at any point of time.

Because the Palo technology processes data real time, the danger of data loss is reduced to a minimum, making the BI solution simple and easy.

PALO is by OTTO for the planning the Seasonal Purchase
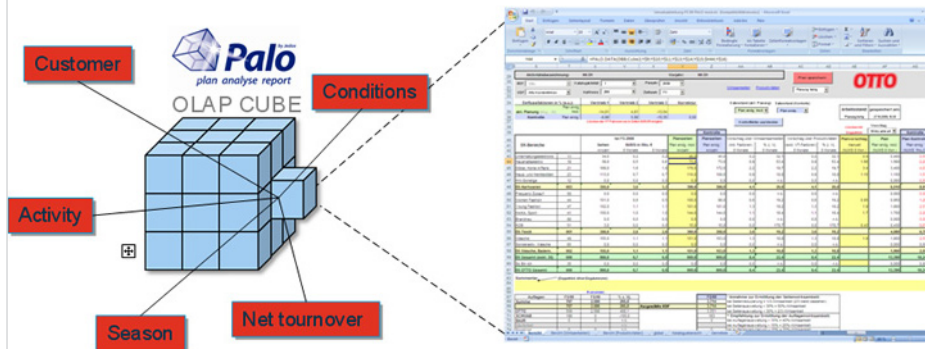
Significant Improvements in process with Palo OLAP:
- Use of a data bank technology with Excel Frontend – user friendly
- One standard data source
- Updating possibilities and automatic data-aggregation and –consolidation
- Flexibility with the model adaption
- New functions such as splashing

FUTURE DEVELOPMENTS

Palo BI Suite by 2010 and beyond:

- Palo on GPU processors (based on Cuda/OpenCL) to increase the performance by factor 20 (or more). This project is worked on in cooperation with the Freiburg University / DFG and Amitava Datta from the University of Perth.

- Integration of additional data providers for DynaRanges

- ROLAP: MS Analysis Server, SAP BW, Mondrian,

- MOLAP: Applix TM1, Infor PM10

- RDBMS: ODBC, JDBC

- Advanced functionality:

- More chart types and chart format options

- Offline planning & analysis

15