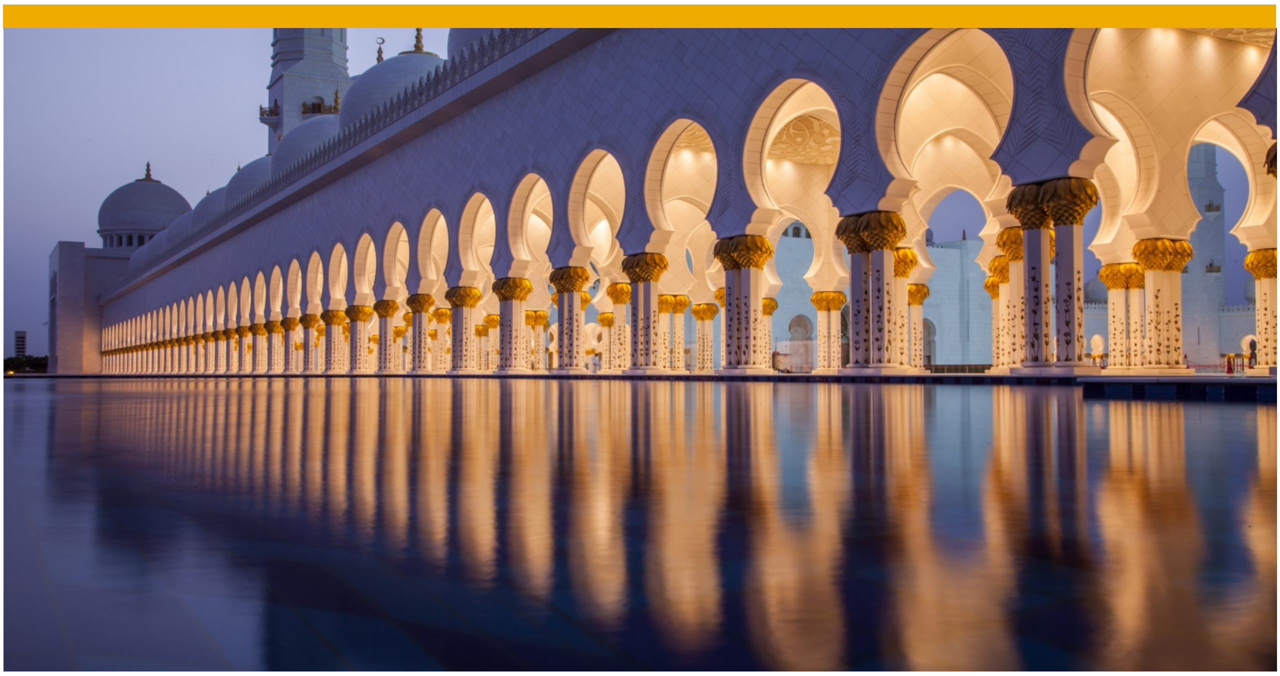

Analytics powered by SAP HANA and SAP NetWeaver BW

SAP's Next Generation Decision Platform

Ulrich Christ
Christian Dressler
Jürgen Hagedorn
Gunther Piller

TABLE OF CONTENTS

Being able to process huge amounts of data in a fraction of a second using in-memory technology opens up genuine new possibilities for business intelligence systems. Long-standing restrictions of traditional analytic systems can now be overcome, and decision processes can be supported with new, powerful analytic capabilities.

What is the architecture of SAP's analytic platform, designed to maximally leverage the in-memory capabilities of SAP HANA? What are its core components and features? Which classical and new use cases or analytical applications will it support? These are central questions in the following discussion.

To start with, we recall characteristic properties of traditional business intelligence systems. We outline typical concepts and sketch their benefits and drawbacks. In addition, we list current market trends and requirements. We then introduce the basic architecture of an in-memory based analytic platform built on SAP HANA and SAP NetWeaver BW. Leveraging HANA's capabilities allows the innovation of SAP's analytic offering – making many traditional restrictions irrelevant and laying the foundation for next generation analysis scenarios. Typical use cases illustrate how operative, tactical and strategic decision making can substantially change and improve, while the classical separation of IT systems supporting these different kinds of decisions becomes less distinctive.

1 TRADITIONAL LIMITATIONS AND MARKET TRENDS

Business intelligence (BI) includes applications, infrastructure, tools, and best practices that enable access to and analysis of information to improve and optimize decisions and performance [Gar13]. Companies commonly support diverse BI usage scenarios with different tools and architectures. Three typical setups are illustrated in Figure 1, and a summary of the corresponding pros and cons is shown in Table 1.

Data analysis in operational systems: Reports, directly generated from transactional data in operational systems, are often used to support decisions in operative day-to-day business activities. An accounting clerk might regularly view a list of customers with payments overdue in an ERP system for example and trigger appropriate dunning activities. As has been known since the early 1990s however, the direct analysis of information in traditional operational systems has at least two basic shortcomings: If applications use conventional databases, scanning larger amounts of data for analytical purposes puts strain on systems, to the detriment of operational processing. As a consequence, response times for transactions and analytics in traditional operational systems can become unacceptably long. A second shortcoming of conventional operational reporting is that information from different systems needs to be explored in a consistent way for many investigations. But the corresponding data integration, harmonization and consolidation is beyond the scope of operational systems. Analytics in transactional systems also has advantages: Reports typically include data from the latest transactions. Furthermore – at least in principle – fine granular data at document level is available for analyses.

If SAP HANA is used as the database for operational systems, such as SAP ERP, the performance problems mentioned above disappear. The benefits of data analysis in operational systems can then be fully realized, also assuring high performance of both transactional and analytical processes (see e.g. [PZ12]).

Data warehouses based analytics: To overcome the shortcomings mentioned above, enterprises have introduced BI systems which are separated from operational ones. The ability to integrate data from different sources into one dedicated environment for analytics is a major value proposition of data warehouse-based BI.

By means of ETL processing, enterprise data warehouses (EDW) offer subject-oriented data with common structure and semantics, regardless of their origin. EDWs also store data permanently and serve as a source of non-volatile and time-variant information [Inm05]. This allows analysis tasks to leverage consistent historic data as a basis for reporting and more advanced analytics, like planning and simulation. Separation from operational systems also enables consistent and stable analysis results which do not vary in an uncontrolled way. In the case of strategic decision processes in particular, it is important that reports do not suddenly change and momentarily seem to be inconsistent, for example when documents from the most current transactions have not yet been completely posted and included, or real-time updates in departmental reports have not yet been rolled up into divisional accounts. From a performance perspective, suitable data design and processing and multidimensional analysis methods normally ensure acceptably fast response times in EDW systems.

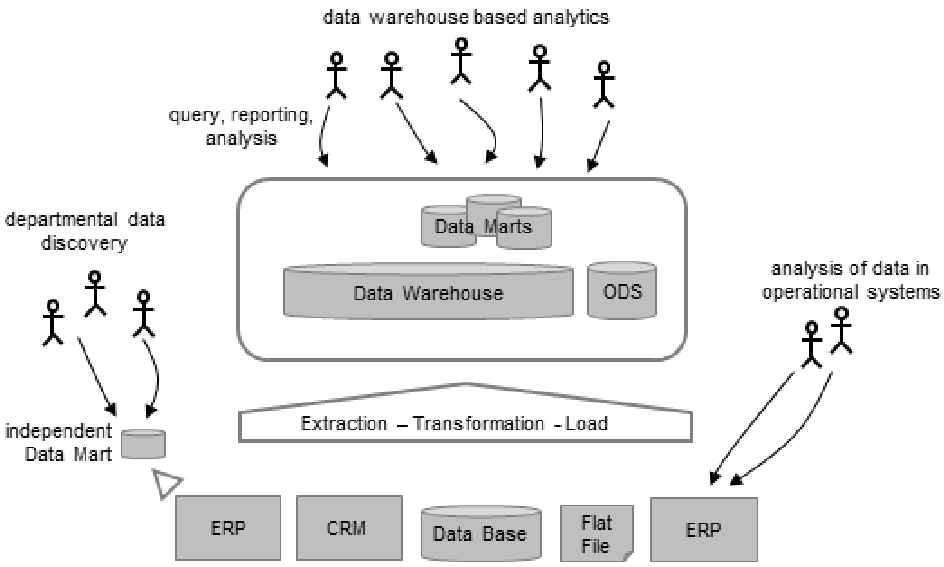

Figure 1: Typical BI usage scenarios.

Classical EDW systems also have their shortcomings however: Latency due to ETL processing often leads to outdated reports and forecasts. Loading data into an EDW by means of nightly batch jobs for example results in reports that do not include transactions from the current business day. Processing data through the different layers of an EDW also requires time consuming and sometimes costly quality assurance measures. For high volume analyses, data is sometimes aggregated and persisted with respect to predefined structures in order to optimize response times. Flexible root-cause investigations at document level are then no longer possible.

Low agility and high maintenance are further burdens of EDW systems: Analyses are typically pre-defined and based on standardized semantic models. These often do not meet business users' needs to rapidly and iteratively combine and explore data, for example with respect to new characteristics or key figures. Changing an existing data model in an EDW usually requires time-consuming modifications in all layers of the data warehouse. Loading and processing data into changed structures requires additional time and resources.

For processing of more current information, data warehouses are often enhanced with Operational Data Stores (ODS). Here, data is typically of high granularity and is updated frequently [Inm99]. Traditional ODS structures mostly integrate data from a few source systems only and do not keep historic snapshots.

Departmental data discovery: The wish to interactively investigate and combine data without depending on IT resources sometimes leads business users to invest in departmental solutions. Known as data discovery platforms (for example [HSR12]), these enable users to explore local data with high flexibility – without being limited by standardized enterprise data models and tools. New analyses are easily defined and can be performed rapidly. They are not slowed down by the lengthy deployment processes often prevalent in IT departments. Typical examples of investigations of this type are exploration of sales data for individual sales areas or analysis of marketing data for specific regions or product categories.

Data discovery solutions usually store and explore data in locally owned, independent repositories however, resulting in fragmented silos of data and definitions. This often renders a common understanding of results and their interpretation in a broader enterprise context for root cause analyses impossible.

BI Approach	Benefits	Shortcomings
Data analysis in operational systems	➢ Inclusion of latest transactional data ➢ Fine granular data on document level	➢ Low performance of analyses *) ➢ Slow down of operational data processing *) ➢ Coverage of data from single system only
Data warehouse based analytics	➢ Integration, harmonization and consolidation of data from diverse systems ➢ Non-volatile and time-variant data ➢ Consistent and stable analysis results ➢ Fast response times – as compared to data analysis in operational systems	➢ Data latency, partly outdated analyses ➢ Costly data processing and quality assurance ➢ Low agility ➢ High maintenance
Departmental data discovery	➢ High flexibility ➢ Rapid deployment ➢ Independence from IT resources	➢ Fragmented silos of data and definitions ➢ Heterogeneous BI solutions, "Shadow-IT" with hidden costs

Table 1: Pros and cons of traditional BI setups. Shortcomings which become obsolete if SAP HANA is used as database for operational systems are indicated by *).

After our sketch of the benefits and drawbacks of typical BI solutions, we briefly summarize current trends and requirements. According to recent analyst reports (see [HSR12] for example) typical aspects of future developments and investments in BI platforms are:

- **Analysis of large, volatile and diverse data ("Big Data")**: Companies want to analyze huge and steadily increasing amounts of structured and unstructured data.
- **Real-time analytics**: To react immediately to changes, many analyses need to include the most recent transactional data. Explorative and interactive investigations also often require immediate response from analytic tools.
- **Integration of flexible data exploration and enterprise BI**: To avoid the proliferation of independent data silos and to allow cross-departmental insights, it is important to provide capabilities which unify departmental analysis results with enterprise data models.
- **Advanced analytics at the point of decision**: More complex analytic capabilities - like simulation and forecasting - need to be embedded in business processes and should be directly actionable.
- **Easy, mobile consumption**: A large and growing percentage of BI functionality will be consumed via handheld devices. Business users will then expect a similar experience from BI tools as they do with personal devices like smartphones and tablets.

In the following we will show how SAP's analytic platform based on SAP HANA and SAP NetWeaver BW can combine the benefits of different BI approaches. We also explain how HANA's capabilities make it possible to lessen or even to overcome the shortcomings of past approaches, and how they help to address current market trends.

2 ARCHITECTURE OF SAP'S NEW ANALYTIC PLATFORM

The general structure of SAP's analytic platform is illustrated in Figure 2. SAP NetWeaver BW powered by HANA enables the design and implementation of a HANA-Optimized Classical EDW. It offers the typical advantages of data warehouse based analytics. However, based on HANA's high-performance capabilities, a broad spectrum of additional benefits can be achieved: First of all, the speed of analyses can be strongly increased. The amount of persistent layers of a typical EDW can also be significantly reduced. This makes it possible to reduce maintenance and costs for elaborate data staging. The introduction of virtual structures instead of persistent ones also increases agility when implementing new analysis requirements.

Stable analyses based on EDW data can be combined with HANA real-time extensions, i.e. real-time data obtained instantly from source systems using appropriate HANA tables and views. This allows one to reduce data latency for many time-critical analyses to near zero. Furthermore, real-time transactional data from individual source systems can immediately be explored with respect to consolidated and harmonized information, such as master data residing in the HANA-Optimized Classical EDW. In many cases, this makes it possible to overcome the fragmentation of data and definitions resulting from departmental data discovery (as discussed in Section 1).[1] These use cases are examples of what are known as "mixed scenarios", where data from native HANA can be provided to BW for consumption, and vice versa. Details and further use cases are illustrated in Sections 2.2 and 2.3.

To understand, how these improvements are achieved, we first sketch different parts of the architecture displayed in Figure 2. We then introduce scenarios, characterized by typical analysis requirements and involved source systems. In general, this discussion is based on the recommended architecture for SAP NetWeaver BW powered by SAP HANA, called LSA++, a layered scalable architecture tailored to HANA's in-memory capabilities. A detailed discussion of LSA++ can be found in [Hau12].[2]

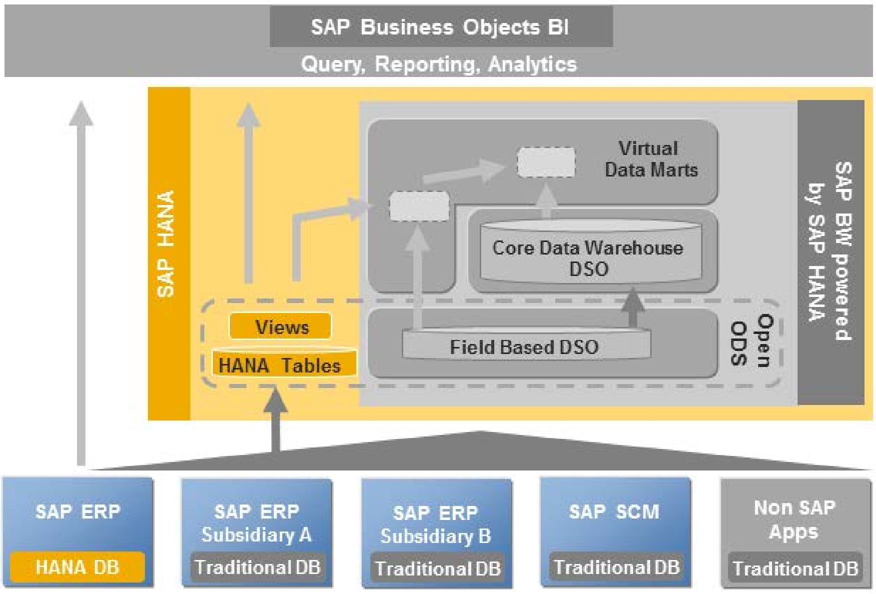

Figure 2: SAP's analytic platform based on SAP HANA and BW. Typical data flow for load processes are illustrated by dark arrows, data flow for analysis and querying by grey ones.

As far as accessing data from source systems is concerned, various situations are possible: If transactional data from just one single application running on a HANA database is explored, data processing through an EDW is often not necessary. In this case, BI analysis tools can directly access source data. For the SAP Business Suite in particular, the Virtual Data Model (VDM)[3] can be used as a broad consumable basis, optimized for analysis with respect to existing Business Suite data models. This case is not the focus of our discussion.

Conventional operative reporting for single SAP ERP systems using traditional databases can be optimized using real-time replication of data into a HANA database. Fast and flexible analyses are then possible using

[1] Note that further functionality for data discovery is provided by BW Workspaces which are not detailed here. More information on their design and capabilities can be found in e.g. [BN11].

[2] It should be mentioned that our discussion partly includes also functionality planned for releases in the near future, in particular SAP NetWeaver BW 7.40.

[3] VDM is a core part of the SAP HANA Analytic Foundation (SHAF), which provides views for the analysis of data of SAP Business Suite applications. Focus of SHAF is operational reporting.

dedicated views and reports. For many customers, this use case – often referred to as HANA Data Mart scenario – serves as an attractive, non-disruptive entry point to SAP's in-memory technology.

If more than one source system is involved, data can be extracted by request or replicated in real-time into a HANA-optimized BW. Of course, if a SAP Business Suite application uses HANA as its database, its data can also be directly read and combined with data warehouse information, avoiding replication and additional storage.

For data processing through the EDW, the following layers are essential:

- **Open Operational Data Store Layer (Open ODS Layer)**: Here, data is provided at field level, still reflecting the data models of its origin. As a core service, this layer supports the acquisition of data from different source systems and allows immediate querying without staging into further EDW layers. For data acquisition, scheduled extraction or real-time data replication from source system databases with SAP System Landscape Transformation (SLT) is available.[4] As far as data storage is concerned, there are two options: Field Based DSOs modeled and controlled in BW, or native HANA tables accessible through HANA views.

- **Core Data Warehouse Layer (Core DW Layer):** In this layer, cleansed, harmonized and consolidated data from different source systems is provided. Source data can be transformed, enhanced and merged, to support broad usage and easy consumption for a wide-range of analysis purposes. Data is not yet aggregated, but stored at line-item level in Core Data Warehouse DSOs (Core DW DSO). Core DW DSOs thus serve as a reusable consistent data foundation used by Data Marts in the next layer.

- **Virtual Data Mart Layer**: Based on application-specific business requirements, data from lower layers or source systems themselves are combined and mapped to Data Marts. These structures then serve as query targets for reporting and analytics. Due to HANA, the structures in this layer can be virtual only – unlike in classical data warehouse architectures. This means that persistent Data Marts and InfoProviders – InfoCubes for example – that limit flexibility and agility, often become obsolete.

It should be emphasized that immediate querying on all layers of the HANA and BW based analytic platform is possible: consolidated data from the Core DW DSO, combined data from HANA views and the Core DW DSO, or data from the Open ODS Layer. If a source system uses HANA as its database, its data can also be analyzed directly in BW using virtual InfoProviders, with or without including data stored in the data warehouse itself.

In the next section, we show how these layers and components work together for different types of analyses and outline corresponding architectures.

2.1 HANA-Optimized Classical EDW

Here we focus on use cases which are typically addressed through analytics built on data warehouses. If SAP BW is using HANA as its database however, classical EDW architectures can be simplified a lot, as the number of persistent layers can be significantly reduced. The benefits of a HANA-Optimized Classical EDW that go beyond the advantages of traditional data warehouses were already mentioned above: Query performance and transformation speed can be dramatically improved. Data latency and maintenance can be reduced, as data staging is simpler. As pre-defined data aggregation in InfoCubes becomes obsolete, analyses can use highly granular line item information, which provides maximal flexibility during data exploration. Data models in virtual structures can also be enhanced and changed more easily, resulting in greater agility.

2.1.1 Characteristics and Data Flow

The use cases addressed by this analysis type are identical to the ones for traditional EDWs. They are typically characterized by the following criteria:

- **Source systems**: Analysis of data from heterogeneous sources.
- **Harmonization and consolidation**: Analyses with respect to enterprise wide structures and semantics.

[4] Traditional services of the Persistent Staging Area (PSA) are contained in the Open ODS Layer.

- **Actuality**: Right-time, i.e. data is updated by means of scheduled data flow execution requests.
- **Stability**: Well-defined analyses, with clearly specified data snapshots.
- **Consistency**: Analyzed data is consistent across processes and organizations.
- **Historization**: Availability of non-volatile, time-variant data as a basis for reporting and advanced analytics.

In Figure 3, the typical architecture for this scenario is illustrated: Data is extracted from source systems into Field Based DSOs. Besides providing data acquisition services, Field Based DSOs can also be queried directly. Cleansed, harmonized and transformed data is stored in a Core DW DSO, based on common, enterprise-wide information models. Further processing and provisioning of data according to specific analysis requirements is performed using virtual Data Marts and virtual InfoProviders. Additional storage of data in persistent Data Marts is avoided wherever possible. [5]

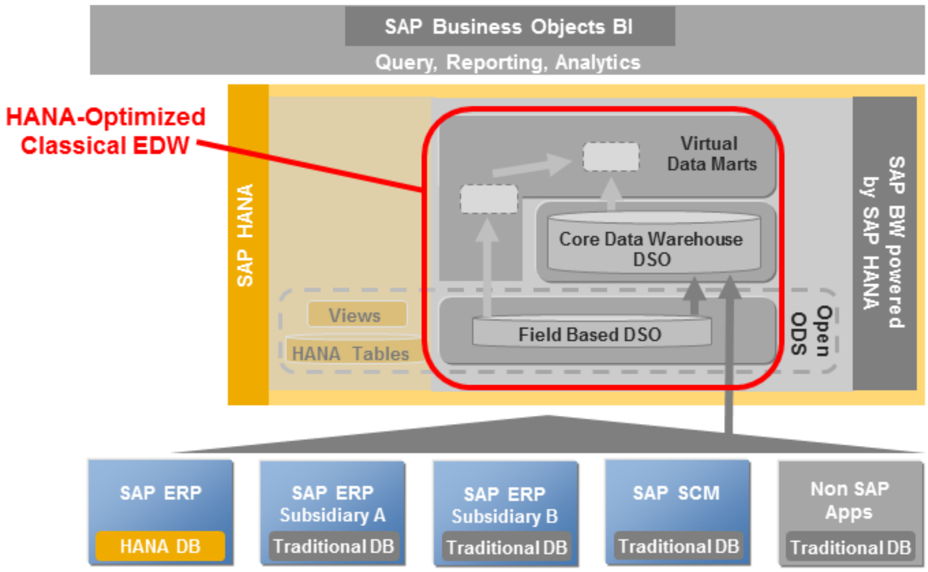

Figure 3: HANA-Optimized Classical EDW.

2.1.2 Advantages from combined HANA and BW capabilities

So far, we have been quite general in pointing out the benefits of HANA for enterprise data warehousing. To shed more light on the potential BW powered by HANA provides for classical EDW capabilities, we highlight several new possibilities. By way of example, we show how analytics can profit from the combination of HANA and BW during query execution at runtime, data staging, data modeling at design-time and data life-cycle management.

- **Moving query operations to HANA**: Since DSOs are based on in-memory, columnar data storage, read access in query operations is much faster than in traditional databases. An even higher boost to query performance can be achieved by moving OLAP operations and other query related computations from BW to the HANA calculation engine. Hierarchy processing, the calculation of restricted key figures, exception aggregation and currency conversion are typical examples.

- **Pushing transformation logic to HANA**: Data transformation during data processing to persistent or virtual DSOs and Data Marts can also be optimized by moving transformation logic to HANA. Simple transformations, like one-to-one mappings or formulas, will automatically be carried out on HANA. More complex transformation logic containing ABAP code can still be executed on the BW application server. In this case too however, the BW transformation framework also allows a replacement of these

[5] At a detailed level, data can also be extracted from source systems and directly be stored in Core DW DSOs after appropriate transformations, while Field Based DSOs also serve as single persistent memory for source data at field level.

transformations with code that can be directly executed in HANA. The resulting improvements in performance can be significant and should be analyzed on a case-by-case basis.

- **HANA Analysis Processes**: BW powered by HANA allows the use of high-performance native HANA operations in BW analyses. In this context, a rich library of HANA capabilities for data mining, cleansing and so on is made available to BW through "HANA Analysis Processes". This library contains statistical functions via R-script as well as algorithms for forecasting, clustering or associations, for example for customer segmentation, ABC analysis, customer recommendations in web shops, address cleansing or geocoding.

- **Virtualization of persistent Data Marts**: Most InfoCubes that were created for performance reasons become obsolete and can be replaced with virtual Data Marts. There are circumstances however, where InfoCubes are currently still relevant, for example for non-cumulative key-figures or in cases with complex business logic joining data from several DSOs [Nag12]. In these cases, HANA-optimized InfoCubes can be used. These are based on a simplified star schema which enables easier modeling and enhancements, as well as three to five times faster loading. The conversion of standard InfoCubes to HANA-optimized InfoCubes is simple and can be performed using standard routines without disruption.

- **Incremental bottom-up modeling**: As high performance reduces the need for pre-defined optimized data structures, new approaches to data modeling are possible: Classical InfoObject-based data models offer a lot of features and services, but they require relatively high upfront design and implementation efforts. With the Open ODS layer a converse approach is possible: InfoProviders can be defined through source fields without the need for corresponding InfoObjects. To begin with, analyses and queries can then already be executed. After that, existing InfoObjects can be linked to the initial InfoProviders, adding master data attributes, texts or hierarchies. This facilitates step-by-step development and enhancement of data models.

- **Optimized Data Storage**: BW powered by HANA provides dedicated services to optimize the storage location of data, leading to appropriate hardware sizing and lower TCO. BW's Non-active Data Concept allows the automatic identification and removal of infrequently accessed "warm" data from main memory to disk. If non-active data needs to be accessed, through a query for example, the system only loads the smallest possible amount – the columns of the relevant partition – to the main memory. Sporadically used "cold" data can also be archived through a near-(on)line storage solution (NLS). Here, data is usually stored in highly compressed form, with fewer backups, thus making it more cost efficient.

2.1.3 Application Examples

Let us illustrate the use of a HANA-Optimized Classical EDW with a few examples: Strategic decisions center on long term plans and have a large sphere of influence, often the entire company. Corresponding analyses need stable, harmonized and consolidated enterprise data. Traditionally – without in-memory technology – performance considerations often make it necessary to pre-aggregate data in addition to batch processing.

The capabilities of a HANA-Optimized Classical EDW can improve this situation significantly: Analyses, simulations and forecasts can be performed in real-time, always using a fine granular data basis – even for huge data volumes. This allows interactive exploration of past performance as well as various future business scenarios, possibly during management meetings.

The analysis of contribution margins for the validation and planning of product strategies can serve as a good example here: Working without data aggregates allows the investigation of contribution margins down to customer-product-combinations, as well as a detailed evaluation of deviations related to changes in quantity, price, costs, discounts or structural variations. In-memory data analysis also offers great potential for the efficient implementation of statistical analysis methods, as well as data mining algorithms. Partially automated top-down-navigation in the context of a gross margin flow analysis is a good example. The aim here is to acquire a detailed understanding of main causes for aggregated deviations from planned, target or historic figures. High performance enables flexible on-the-spot drill downs – if required also on mobile devices like tablets.

Investigations at a tactical level can also profit from fast analyses of non-aggregated, enterprise-wide harmonized and consolidated data, as provided by a HANA-Optimized Classical EDW. Sales analyses can serve as a typical example. Here, sales managers regularly monitor key performance indicators to ensure that marketing activities generate the number, quality and flow of deals necessary to meet sales targets. To flexibly analyze the influence of product and customer characteristics, as well as marketing or sales activities, it is important to use fine granular data without prior aggregation. Further insights can be obtained by comparing sales reports and corresponding plans from different channels and regions. For this purpose, harmonization and consolidation of data from different sources is often needed.

2.2 Real-Time Extended EDW

This analysis type covers use cases where information from a HANA-Optimized Classical EDW is enhanced with real-time data from individual source systems. Here, real-time means that data is available for analyses as soon as it is entered in an application.[6] If source systems work with traditional databases, data required for analysis can be replicated in real-time into tables in the HANA instance used by the EDW. Both components – HANA and BW – then stand side-by-side, facilitating a combination of complementary analysis capabilities. On the other hand, if a source system such as SAP ERP is using HANA as its database, the situation becomes simpler: The source system itself can directly provide its data for analysis purposes through HANA views – predominantly the VDM. Both options are illustrated in Fig.4.

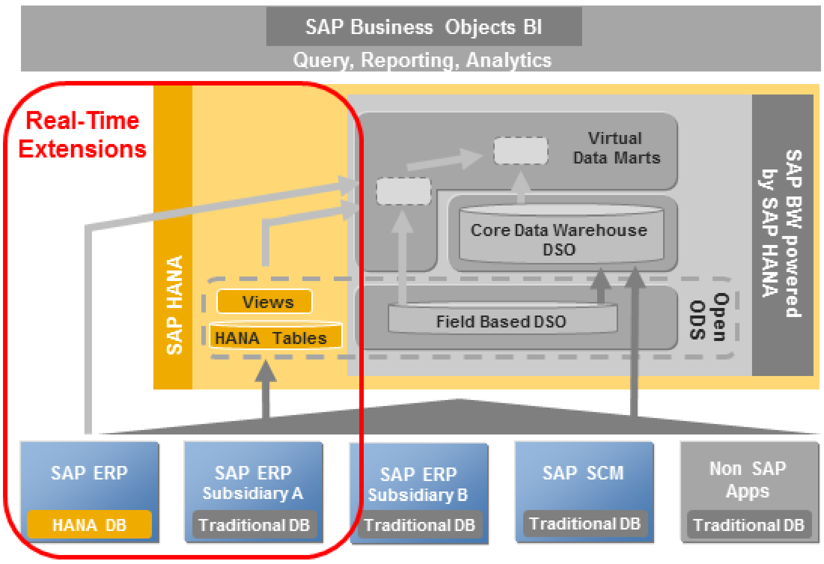

Figure 4: Real-Time Extended EDW.

2.2.1 Characteristics and Data Flow

The typical criteria for this analysis type add to the ones from the previous case:

- **Source systems**: Heterogeneous sources.
- **Stability and actuality**: Enhancement of well-defined, EDW-based analyses with real-time data from the most current transactions.
- **Data latency**: Instant analysis of real-time data without overhead from data processing through EDW layers.
- **Flexibility**: Exploration of real-time data without being limited by pre-defined EDW models.
- **Harmonization and consolidation**: Direct querying of real-time transactional data at departmental level with respect to company-wide consolidated master data.

[6] This goes beyond BW's Real-Time Data Acquisition capabilities (RDA), where source data can be made available to BW within short time intervals.

An important basis for analyses of this type is provided by harmonized and consolidated data, such as master data for customers or materials. This often originates from different source systems. In this case – as outlined in the previous section – the functionalities of a HANA-Optimized Classical EDW can be used. They make it possible to provide cleansed and transformed data, appropriate for standardized enterprise data models that are stored centrally in Core DW DSOs.

On the other hand, real-time extensions – source system-specific transactional data required immediately for querying – can be taken directly from a SAP Business Suite application for example, without the latency of an EDW caused by extraction and transformation. If a source system uses a traditional database, transactional data can be replicated in real-time into native HANA tables within the Open ODS Layer using SLT. Data structures for analytics are then provided by corresponding HANA views. Finally, EDW information and real-time data from HANA is joined using dedicated InfoProviders.

For source systems running on the same HANA database as BW, virtual InfoProviders can directly access source data, using VDM views for example. As an entry point to in-memory data management, companies sometimes choose to replicate data from a SAP Business Suite system using a traditional database into a separate HANA system. The situation is then quite similar to the one just discussed, except that side-by-side HANA components use a different database instance to BW itself.[7] For complex landscapes with different systems and databases, direct data access and a side-by-side approach can be combined.

2.2.2 Typical Challenges

While the combination of real-time data from source systems with information from a HANA-Optimized Classical EDW has benefits, it does of course also pose certain challenges. Integrating real-time data from just one source system is often easy to handle. The level of complexity will normally increase with the number of real-time data sources to be combined. Case-by-case investigations are then required in order to decide whether real-time extensions are still the best solution, or whether a pure HANA-Optimized Classical EDW would be better. Typical points of discussions and corresponding approaches are:

- **Consistent separation of data**: Real-time and EDW data used jointly in an analysis have to be separated consistently. One possibility is to filter data within queries and key figure calculations, depending on the values for selected characteristics, such as time or status information. For example, if data from transactional systems is loaded into BW on a daily basis, appropriate filtering can ensure that data from past days is read from Core DW DSOs, while transactions from the current day are taken from HANA. Redundancies between data in BW DSOs and native HANA tables can be avoided through appropriate routines which remove aged data from HANA.

- **Semantic fit of data models**: To combine real-time, source system-dependent data with EDW information, the involved data models need to fit together. One typical use case is the analysis of real-time transactional data from different sources with respect to consolidated master data. Here, transactional data is processed through source-dependent HANA views and virtual InfoProviders. Finally, they are mapped to consolidated InfoObjects. These use key-mapping information to provide local, system-specific reports as well as consolidated views with respect to EDW master data.

- **Efficient data storage**: If data from native HANA views and BW is joined through virtual InfoProviders, for querying for example, a combined solution for data storage with respect to the frequency of its use is needed. With the Non-active Data Concept, BW provides dedicated services for the optimal storage of “hot”, “warm” and “cold” data from Field Based DSOs and Core DW DSOs, for example also leveraging NLS capabilities. These services have to be combined with appropriate storage mechanisms for the involved data from native HANA.

2.2.3 Application Examples

Operative decisions deal with single transactions which are part of day-to-day activities in a company. An analysis of the most current and highly granular data is therefore crucial. To continue an operative process seamlessly, results also need to be available without delays. Operative decisions also often benefit from analyses which include data from more than one transactional system. A combination of real-time and consolidated data is then important. Here, the Real-Time Extended EDW can deploy its unique strength.

[7] It is assumed that in future virtual InfoProviders can also integrate data from different HANA instances.

To decide on the approval of a major customer order or shipment for example, it is important to know the current level of outstanding payments and recent changes in payment patterns. For larger companies with several subsidiaries, this often requires an investigation of consolidated information from several accounting systems. In this case, the Real-Time Extended EDW allows analysis of actual data from individual transactional systems on latest invoices or payments with respect to consolidated and harmonized master data for debitors.

The point-of-sales analysis in retail is a further example of the combination of EDW data with real-time transactional information. Here, consistent and stable inventory data from BW is joined with real-time point-of-sales information from native HANA tables. Based on these, new key figures are calculated which allow immediate identification of out-of-stock risks.

An example from the domain of tactical decisions is the consolidation of up-to-date sales figures, like expected or closed deals. The goal is to decide on actions needed to meet a sales unit's quarterly objectives. In many industries, closing deals is more prominent towards the end of the quarter - sometimes driven by budget cycles. As a consequence, many transactions are created very late, just before quarter closing. To allow for short-term interventions of sales management teams, it is necessary to provide results with the most current data as quickly as possible. Comparing actual results from sales units at corporate level using consolidated master data helps to identify reasons for deviations from plans, on account of the region or product for example. For this purpose the capabilities of the Real-Time Extended EDW can be leveraged.

2.3 Variations and Extensions

Through the tight integration with BW, HANA can add many more capabilities to typical analysis scenarios. Examples beyond the HANA-Optimized or Real-Time Extended EDW, which we illustrated in the previous sections, are:

- **Complex calculations in HANA**: The powerful calculation capabilities of HANA can be used to serve BW with results from complex computations carried out on native HANA. These calculations can either use data from native HANA tables or can operate on BW data. In the latter case, data stored in Core DW DSOs can be accessed directly in HANA. In calculation views, a broad spectrum of highly flexible computation capabilities, such as SQL Script, can then be utilized. Afterwards, the corresponding results are available in BW by means of virtual InfoProviders for further analyses. In practice, this approach has already been implemented in a solution for the gross margin flow analysis in controlling. Here, profitability analysis data from BW is further processed using complex calculation schemas in native HANA. The corresponding results are consumed in BW, where they are made available for querying and reporting.

- **BW data models within HANA views**: Data models of SAP BW powered by HANA can be published as views for consumption using native HANA. Analysis of BW data in HANA can then profit from common BW services, for authorization or security for example. An example of a use case for this scenario is the integration of a company's central EDW based on BW with a custom-built data warehouse from a subsidiary. Running both of these on a common HANA database instance using separate schemas enables easy and secure consumption of EDW data in analyses built upon the subsidiary's data warehouse.

- **Native HANA tables and views as data source for BW DSOs**: Using BW, data transfer and transformation services make it possible to transfer data from native HANA tables and views in the Open ODS layer to Field Based DSOs or Core DW DSOs. The appropriate EDW layer depends on the quality or harmonization level of data, as well as the corresponding consolidation requirements. Taking up the example from the previous point, this function allows integration of data from a custom-built data warehouse of a subsidiary into a central EDW. A typical challenge for this integration use case is the availability of valid delta criteria for native HANA data.

- **SLT data replication into BW**: Besides extraction, data can also be replicated into BW in real-time. A potential use case is the replacement of master data extraction and load from single source systems by SLT replication into BW powered by HANA. In situations where delta extraction is restricted, uploads of mostly unchanged records for entities with slowly changing dimensions can thus be avoided. However, as the extraction of some core entities, such as material, involves joins between several source tables, and consistency between transactional and master data needs to be ensured, using SLT like this must always be carefully considered.

3 SUMMARY

In this paper, we have introduced the basic concepts of an innovative analytical platform which combines the capabilities of SAP HANA and BW. As we have illustrated, this platform offers new and valuable opportunities for analyses, enabling superior operational, tactical and strategic decision making. We have outlined how using virtual structures can drive down maintenance costs and increase agility with respect to new analysis requirements. In particular, we have described how the different components of HANA and BW work together for major analysis types, introducing the HANA-Optimized EDW and the Real-Time Extended EDW.

Coming back to our initial discussion of traditional limitations and market trends, it becomes apparent that an analytic platform based on HANA and BW is able to combine most benefits of the different BI setups summarized in Section 1. On the other hand, it is possible to overcome shortcomings or at least to take the sting out of them. Table 2 provides a summary. Analytics with SAP NetWeaver BW powered by HANA and data analysis in operational systems using HANA as the database are two approaches that nicely complement each other.

BI Approach		Benefits	Shortcomings
Data analysis in operational systems	SAP ERP on HANA	➢ Inclusion of latest transactional data ✓ ➢ Fine granular data on document level ✓	➢ Low performance of analyses ✗ ➢ Slow down of operational data processing ✗
	SAP HANA and BW based analytic platform		➢ Coverage of data from single system only ✗
Data warehouse based analytics	SAP HANA and BW based analytic platform	➢ Integration, harmonization and consolidation of data from diverse systems ✓ ➢ Non-volatile and time-variant data ✓ ➢ Consistent and stable analysis results ✓ ➢ Fast response times – as compared to data analysis in operational systems ✓	➢ Data latency, partly outdated analyses ✗ ➢ Costly data processing and quality assurance ↘ ➢ Low agility ↘ ➢ High maintenance ↘
Departmental data discovery	SAP HANA and BW based analytic platform	➢ High flexibility ✓ ➢ Rapid deployment [-> BW Workspace] ➢ Indep. from IT resources [-> BW Workspace]	➢ Fragmented silos of data and definitions ↘ ➢ Heterogeneous BI solutions, "Shadow-IT with hidden costs ↘

Table 2: Revisiting Table 1. Checkmarks indicate which benefits can be addressed with a HANA and BW based analytic platform. Shortcomings becoming obsolete are crossed out. Reduced ones are indicated by arrows. Characteristics of data analysis in operational systems based on HANA are also displayed.

The HANA and BW based analytic platform also makes it possible to address the trends listed in Section 1: Analysis of large data volumes ("Big Data") at real-time is a core capability of HANA, which applies both to structured and unstructured data. The Real-Time Extended EDW makes it possible to combine departmental data with enterprise models, contributing to the integration of flexible data exploration and enterprise BI. Integration of advanced calculation capabilities of HANA into BW (see HANA Analysis Processes, Section 2.1.2.) enables the implementation of sophisticated analytic scenarios with short response times. Finally, the high performance of a HANA-based BW is the basis for successful provisioning of analytics through mobile applications.

4 REFERENCES

[BN11] S. Bratz, K. Nagel: What is a BW Workspace? SAP AG, Walldorf, Germany, 2012. http://www.sdn.sap.com/irj/scn/go/portal/prtroot/docs/library/uuid/0022d123-13ed-2e10-9f99-ac0a3acd35af?QuickLink=index&overridelayout=true&52849572578976. Retrieved February, 8th, 2013.

[Gar13] Gartner: IT Glossary, Business Intelligence. http://www.gartner.com/it-glossary/business-intelligence-bi. Retrieved February, 8th, 2013.

[Hau12] J. Haupt: LSA++ (Layered Scalable Architecture) for SAP NetWeaver BW on SAP HANA. SAP TechEd session EIM203, Madrid, 2012. http://www.saphana.com/servlet/JiveServlet/previewBody/2814-102-2-5167/EIM203_updated_Madrid.pdf. Retrieved February, 8th, 2013.

[HSR12] J. Hagerty, R.L. Sallam, J. Richardson: Magic Quadrant for Business Intelligence Platforms. Gartner, Feb. 2012.

[Inm05] W. H. Inmon: Building the Data Warehouse. John Wiley & Sons, 4th edition, 2005.

[Inm99] W. H. Inmon: Building the Operational Data Store. John Wiley & Sons,1999.

[Nag12] K. Nagel: Aspects of the HANA-optimized InfoCube. SAP AG, Walldorf, Germany, 2012. http://www.saphana.com/servlet/JiveServlet/previewBody/1363-102-2-1810/SDN_HANA_opt_InfoCube%20FINAL.PDF. Retrieved February, 8th, 2013.

[PZ12] H. Platter, A. Zeier: In-Memory Data Management: Technology and Applications. Springer, 2nd edition, 2012.

www.sap.com