openSAP
SAP Business Warehouse powered by SAP HANA

WEEK 1, UNIT 1

00:00:14 Hello and welcome to the openSAP “SAP Business Warehouse powered by SAP HANA” course.

00:00:21 My name is Marc Hartz, I'm member of the product management of SAP HANA, and I'm focusing there on SAP BW as a product and the whole topic data warehousing itself.

00:00:31 My name is Ulrich Christ, I'm a colleague of Marc, and we are very happy to guide you through this four-week course about SAP BW on HANA.

00:00:40 The course will have six units each week. After each unit, we'll have a self-test.

00:00:45 After each week, we'll have a weekly assignment for you. And at the end of the course, we'll have a final exam for you.

00:00:50 Right. So the course is, I think we have many interesting topics in there. It's basically built in way that

00:00:58 everybody should be able to follow the course. A pre-BW knowledge is definitely a plus, but we will also guide you through the main concepts, the main ideas

00:01:07 of BW also as a common starting point, which was by the way also for us interesting again to go over the very basics and foundations.

00:01:18 OK, so let's get started. I think we can get started with week 1. And we start – so no worries – we will go deeper

00:01:27 after this unit, but we will definitely start with the “What is BW again?” and “What is the product doing?”. And

00:01:35 the answer is SAP BW is basically our data warehousing application, so it's the SAP application for data warehousing.

00:01:43 Means we are adding in BW services on top of a database. And services to build a whole data warehouse.

00:01:51 And especially for this course, we will show you these services on the example of SAP HANA.

00:01:59 Maybe let's have a quick look at the services which Marc was mentioning. One very central thing is modeling, so you're modeling a lot of different objects in a data warehouse environment. You're modeling

00:02:10 analytic reports, you're modeling the data foundation for analytic reports – in BW that's typically master data and transaction data.

00:02:18 Very important is also the ETL part, that's the extract, load, and transform from source systems, but also within the data warehouse to prepare the data for the analytic reports.

00:02:29 Then a very important thing is, of course, the OLAP functionality, which allows you to run the reports on top of the data which you have prepared.

00:02:36 So this basically everything what deals with the analytic view on top of the data. So the data warehouse, and like you mentioned, we have the services there,

00:02:46 we bring up the data from several sources – like indicated in our graphic here – and then these services are hamonizing the data.

00:02:54 This analytical functionality comes on top, and then we do have a possibility to create, for instance, corporate reporting.

00:03:03 One thing which also very important as a service here is that obviously you want to automate all of this stuff.

00:03:10 So you want to automate and schedule the loading processes, all these kinds of things.

00:03:14 So you need scheduling and monitoring in order to detect errors and correct errors.

00:03:20 That's a very important aspect. Lifecycle management is also a very important aspect in two ways: It's about data lifecycle management but it's also about lifecycle management of your data models.

00:03:29 So whenever you change your data models, you have to make sure that you bring things from your development environment into the production environment in a correct way.

00:03:36 All this is handled by BW in an integrated way, so that each of these services is aware of all the relevant services around it.

00:03:43 And changes in one of the areas kind of affect the other areas.

00:03:49 These services will also guide us through this week 1, so we will get back to them later.

00:03:57 But for now, let's just recap also a bit about BW, and especially powered by SAP HANA.

00:04:02 So BW itself is a very long-standing product in the market, so we have a very strong installed base.

00:04:10 We have 14–15,000 customers running their data warehouse with BW on a database, and like we said,

00:04:20 BW was, I think three years ago, one of the first applications which were optimized, which were powered by SAP HANA.

00:04:28 And this is exactly our driver here. Since we have the HANA underneath, the SAP BW, we have a lot of customer demands and a lot of questions, so it's an amazing time since HANA is there.

00:04:41 And therefore, let's have a look what is different compared to a traditional environment.

00:04:48 And it's basically the main guiding principle during this course will also be

00:04:53 that many of the services where we have a certain functionality behind will be pushed down natively to SAP HANA, to run there in memory-optimized and in a special way.

00:05:06 Maybe it's important to mention why we did not do this before. We did not have the chance to optimize in such a way, because we were supporting and we are still supporting many other databases, which are always...

00:05:16 Each of them is a little bit different in certain areas. There's a common standard, of course, on which BW resides.

00:05:24 But we can't really optimize on specific platforms in a highly integrated way. There are some optimizations obviously,

00:05:31 but with SAP HANA we can go much further. That's the whole point about this picture.

00:05:36 And this is how we then achieve this excellent performance in reporting and data loading.

00:05:43 But I think you all are aware of that HANA delivers a certain performance.

00:05:48 And that's...we will show this in the course, but our main driver is that HANA was also introducing a lot of simplification in BW.

00:05:59 So you will see ways and concepts in this course how we, in an agile way, build data models with a reduced amount of layers.

00:06:08 This was in the past very, also painful for a lot of customers. And this is exactly where we would like to go deeper into in this course.

00:06:19 The course will basically consist of the mentioned four weeks already, we're right now on the introduction.

00:06:26 And then we are basically focusing on every important task with BW, and especially the optimization and the key aspects with SAP HANA.

00:06:35 So in the next week, we will take a look about how can you get data into BW, what are the new ways. For example,

00:06:43 one really interesting part is the real-time replication into SAP BW.

00:06:48 In the next chapter, in the third week, we will then focus on how you can build data models in BW with SAP HANA as database.

00:06:58 Also interesting to see that, of course, we will also show you how to combine BW modeling with HANA native modeling. That's very important aspect.

00:07:06 That's one key aspect, which then goes along in reporting, and we will do an end-to-end process and end up with the planning process.

00:07:16 This whole process is supported by a demonstration model. So we will use the E-Procurement Management Model, which comes with every NetWeaver installation.

00:07:29 This model consists out of several tables and is representing a fictive company.

00:07:39 Fictitious sales company which sells hardware in this example. You should think of this company as a global player, which has subsidiaries in many countries.

00:07:48 So you probably have a lot of different sources, different source systems which you want to integrate into one consolidated and harmonized data view, data warehouse.

00:08:00 And I think in the next slide we will already see some details of the data model which we're interested in.

00:08:06 The data model itself is actually much, much bigger and has much more details. We're going to focus on the sales part.

00:08:13 So here you see, for example, the sales order header tables and the sales order item tables. And you see the relevant master data, which is the business partner –

00:08:19 basically the customer or the provider of the hardware which we are selling.

00:08:27 We also have the product, and we have invoice data as well. And if you look at these tables, as Marc mentioned earlier, this is delivered with every NetWeaver instance.

00:08:36 So if you get, if you have one or if you get one on the web, you will find these tables in your NetWeaver.

00:08:44 And the interesting part is, and this is what we are leveraging there as well, there is a huge and a really good and mature infrastructure around these tables and models.

00:08:55 So there is a generator for data, so our data will be sampled, data created by this model.

00:09:01 And this is very interesting. And the whole story of this openSAP course is really that based on this EPM model,

00:09:09 we will demonstrate how you would build now an analytic model with SAP BW powered by SAP HANA.

00:09:16 And all the new concepts and our latest release in SAP BW was 7.4 where we introduced a lot of SAP HANA-specific things.

00:09:26 And we will show you, after each chapter, the demonstrations in the systems,

00:09:34 what you can do with all of these features. So this is mainly the main driver.

00:09:40 Everything is going to be based on the E-Procurement Model. We'll show you in each of the units some overview slides, and maybe give you some overview of the whole concepts.

00:09:50 And then we'll always show you a system demonstration to really get hands on. And as we said, if you have a system of your own, you can work through the exercises yourself as well.

00:09:59 But this is a good point: system of your own. So you will have the chance in parallel to create an own cloud trial system for free.

00:10:11 And in this cloud trial system, you will find our content. So the content we are showing you in this course will be available there.

00:10:18 And the hand out of this course will also include a list of all the objects, all the, you know, persistency objects and joins and stuff like that.

00:10:27 And we're giving that to you, and you can really then work on this trial instance, with the models we have introduced during the course.

00:10:36 For details, you can check the SAP HANA Marketplace. Maybe, since you mentioned for free, of course, nothing is for free in real life.

00:10:43 SAP gives you the software for free for 30 days, but you have to pay for the Amazon Web Services.

00:10:50 Yes, for the Web Services, but it's really a few bucks, and I think...and it's charged hourly, so it's really something which you can try. And if it doesn't fulfill your needs...an exercise are worth to spend this money. Of course.

00:11:02 Okay, to pass the final exam, right? So this was basically unit 1.

00:11:07 I would like to highlight that everything what we will do again will be based on the E-Procurement Model. This is really important because it's really showing

00:11:17 the value of BW. So you will recognize that we really can show nicely all the different data warehousing services based on this model.

00:11:25 And you can reuse it whenever you have a NetWeaver system or our trial version.

00:11:30 The course will, we will start in this week also with the introduction of the key concepts of BW

00:11:37 in general. So what is an InfoObject, for instance, so such important things.

00:11:41 But then we will really fast go to the HANA specifics to the latest innovations.

00:11:47 Okay, so the next week or the next unit will then be focusing on why are we doing data warehousing in general.

WEEK 1, UNIT 2

00:00:12 Hello and welcome to week 1, unit 2: Enterprise Data Warehousing – Why?

00:00:17 This Why is exactly the question we would like to pursue in this unit, so we would like to give you an answer to why data warehousing is necessary in a customer landscape.

00:00:28 We do this with a certain example. Based on that, we will show the usual characteristics and tasks what a data warehouse does in such situations

00:00:39 and we will end with different approaches for implementation. A data warehouse can be built following these approaches.

00:00:47 And obviously, SAP BW powered by SAP HANA is one of those approaches which will then be our target afterward.

00:00:56 So let's jump into a typical customer situation.

00:01:03 And we are now focusing on the graphic here where you've typically got different sources.

00:01:10 This means you have sources with a transactional character like ERP data. You've got supplier data. You've got different kinds of master data,

00:01:20 like I have customer-specific data. Like how much money did I do with one customer?

00:01:27 What are the different characteristics like address, like country of a customer?

00:01:33 To bring such data types together in one environment, to build one common view on my company,

00:01:41 usually a data warehouse is the single point of truth where you have a common model.

00:01:47 This means out of these different sources, we are bringing data up into a common corporate data model where we have a common view based on the data.

00:01:59 One important thing is that such a data model is something which everybody in your organization who is using this system has to agree on.

00:02:06 For example, the definition of a customer as an entity in this model is something which people should agree on.

00:02:14 They should agree on what information about a customer you actually need in this system, what information about a sales order you need, all this kind of stuff.

00:02:21 So you basically define what a sales order means to you. You define what a customer means to you.

00:02:27 And you basically have to find a way to map the data from all your source systems, which might have a completely different data model, to this enterprise-wide data warehouse data model.

00:02:38 So a standardized data model for the whole company. And of course, it should support decision making or analysis with different BI tools and stuff like that.

00:02:48 So really one common approach instead of having multiple local Excel sheets and everybody has their own definition of, for example, revenue.

00:02:59 That's really something which is reflected in the data warehouse. There's one definition of revenue and that's my revenue for the whole company.

00:03:06 That's very important because many discussions in your company rely on the result of your

reports.

00:03:13 So you basically make your decisions based on this. Therefore, everybody has to make sure that they understand the results of your reports in the same way.

00:03:21 That's why standardizing things like key figures and terms like revenue is very, very important.

00:03:26 Right. And the data warehouse is usually like a data hub in the middle of this whole process

00:03:33 where we also feed data into external systems, maybe for legal purposes, maybe even for a smaller data mart purpose.

00:03:44 And a data mart in this context is just a small collection for a specific purpose. For example,

00:03:49 my financial guys, they have an own environment and I will give them a small data mart, a subset of the corporate view, to fulfill their analysis needs.

00:04:00 In theory, and how we would hope it would be, the picture how data warehouse should be placed in the middle of this picture,

00:04:09 feeding all the external systems and a single point of truth for the reporting.

00:04:14 Maybe even providing data back to the original systems for planning purposes.

00:04:20 Or forecasting. This is usually you are doing a calculation in the data warehouse, how many things will I sell? And then this is going back into the transactional systems.

00:04:31 So this was the customer example. There are many definitions, many theories, around data warehousing,

00:04:39 but actually they all pursue the same common goal: to harmonize, to get data, to consolidate data across the landscape.

00:04:49 So here we've listed the two most prominent schools of data warehousing. That's the one by Bill Inmon and another school following the ideas of Ralph Kimball.

00:04:57 Inmon defines the data warehouse as a subject-oriented, integrated, time-variant, non-volatile collection of data.

00:05:03 We've basically seen all of these aspects in the picture before; that's why we started with this introductory picture.

00:05:09 Subject-oriented basically refers to a certain area of interest. Like you're interested in sales data and everything which is related to it, like customer data, product data. That would be one subject.

00:05:20 Integrated basically means the harmonization process and integration of data from various sources and bringing it to one consolidated structure.

00:05:29 Time-variant and non-volatile basically refer to something which I think we have not mentioned yet. That's about historic data.

00:05:36 It's also very important in a data warehouse to be able to keep historic data and maybe even be able to review decisions you've made in the past

00:05:48 to basically see why did I draw this or that conclusion last year?

00:05:52 So basically you would want to be able to review the result of a certain report at an earlier point in time.

00:06:00 And that's important with being independent from the sources. Because a source system like an ERP system maybe archives the data already, so it's not there anymore,

00:06:11 but you would like to make an analysis of how is my business going this year compared to the last five years, or with different market influences.

00:06:21 So basically, in a data warehouse, you would expect when it comes to, say, customer data, you would not update the customer record,

00:06:28 but you would keep multiple versions of the same customer to be able to track the history and the historic changes of this customer, or even for transaction data as well.

00:06:41 The definition by Kimball is maybe a little bit more pragmatic. Kimball just calls a data warehouse a conglomerate of all data marts.

00:06:52 So basically all these subject-oriented views which you build on the database are using the data warehouse too.

00:06:59 Those will form your data warehouse.

00:07:03 But those data marts could also sit in one physical environment, let's say it like this.

00:07:09 And obviously SAP BW is one enterprise data warehouse solution to build all of this, or to follow all of these different theories.

00:07:22 SAP BW is an integrated application. We talked already a bit about the services we have in this environment.

00:07:31 So what we are doing there is we have a model-driven modeling experience. So we will see that we are not working directly on database level.

00:07:39 We have a special semantic on top of the data how we treat master data, transactional data.

00:07:45 Just very business oriented.

00:07:47 Right.

00:07:48 So you talk about things like the customer, which is a term that everybody in your company who has no knowledge about databases would understand.

00:07:57 He knows what a customer is and he knows that certain attributes of a customer are relevant. He knows that an address would be relevant, or this kind of stuff.

00:08:05 And that's the level on which you do your modeling in BW. So it's really a level higher than the actual database.

00:08:12 That's an integrated environment. So one application fulfilling the different tasks like loading data, transforming data, doing the reporting, and so on.

00:08:22 We have many tools out of the box with the predelivered content and information models, which are something like templates. We will show this to you.

00:08:32 And of course, the other way around, where you can also build a data warehouse, is that you use loosely coupled tools.

00:08:40 This means you have a strong database and you have a tool for doing ETL, for doing reporting, for doing data quality checks...

00:08:47 and all the modeling stuff.

00:08:50 Yeah, even an external modeling tool.

00:08:52 So there are two different approaches. Both have their pros and their cons. We will show you also this in the next unit with a few examples.

00:09:01 So what is data modeling looking like in the two different ways. They are differentiating in development efforts and maintenance.

00:09:11 And just from the...you know, if you are following the best-of-breed, picking the tools for the different purposes, or you have one integrated enterprise data warehouse application.

00:09:22 In our case, in this course, enterprise data warehousing will be defined with SAP BW as the data warehouse application

00:09:31 with the orchestration tools inside, so integrated stack. And SAP HANA will be the database.

00:09:40 Maybe before we come to the key takeaways, a very, very small example of what a data warehouse does.

00:09:52 You have to imagine we are doing this in a huge corporate use case. This means you are doing this with many, many tables, many different sources.

00:10:02 But the source table, for example, of our model here is looking like that.

00:10:08 So you have a sales order table here, and you can also see that in the system. And this sales order table, if I go to the content of it,

00:10:17 you would see that we are working there with internal keys, which are not very nice to report.

00:10:24 You see there the time. If you go to this column here, CREATED_BY or CHANGED_AT, for example, it's a time stamp.

00:10:32 I think my manager doesn't want to report based on time stamps. that's very hard to understand.

00:10:38 What they would like is a column with year, column with month, something like that.

00:10:43 And the whole transformation, the whole harmonization of this data, this is what you are typically doing in a data warehouse.

00:10:50 And we will do exactly this in the following examples with the following units.

00:10:56 Okay, so let's summarize the key takeaways.

00:11:02 Enterprise data warehousing will harmonize and consolidate different data from many sources.

00:11:09 Within the data warehouse, you will map, you will transform the data to a common view.

00:11:15 The data warehousing should be the data hub for decision making and supplying other external systems or data marts with data.

00:11:23 And we have two principles, or approaches, to build a data warehouse.

00:11:29 You can use an integrated application like SAP BW or you can also follow with SAP HANA as standalone,

00:11:37 a custom-built approach, and build it by yourself. And this is exactly what we would like to compare in the next unit.

WEEK 1, UNIT 3

00:00:12 Hello and welcome back to week 1, unit 3: Value Proposition of SAP BW.

00:00:19 In this unit, we would like to speak about different services of SAP BW as a data warehouse application.

00:00:26 So far we talked about what is a data warehouse, and now we would like to focus on what are the specific services needed to fulfill the service of a data warehouse.

00:00:36 So as we are describing in this slide, data warehousing is more than a database.

00:00:42 Data warehousing requires a database, obviously, to keep the data. But it also requires a lot of services on top of that.

00:00:48 And there are basically two different approaches to data warehousing.

00:00:51 One is that of an integrated set of tools which are closely intertwined and work together very well.

00:00:57 That's what BW basically provides, and BW is maybe the only example of this.

00:01:04 And there's another approach, maybe the more general approach, where you basically pick a database like SAP HANA and a set of tools,

00:01:13 typically best of breed or according to your preferred vendor, and you just use these tools individually to create your data warehouse.

00:01:24 SAP HANA is a very special database, and also fast. But let's consider it in this case really as a database where you would like to custom build a data warehouse application.

00:01:35 And to fulfill these services, you have to bring this different pieces, meaning you have to answer the questions like:

00:01:43 How can I load data into this database? How can I load data within the database? How can I model the data in an abstract way?

00:01:51 You need to answer that with several tools. So this is more a best-of-breed approach. You pick different tools and bring them together.

00:02:00 If you compare this with parcels, it's more like you have to take care that the parcels are fitting together,

00:02:07 whereby BW has so seen the parcels integrated and they are fitting very well together. This is what we would like to highlight now in a few examples.

00:02:17 And this is what we call the integrated data warehousing stack of SAP BW.

00:02:22 So let's talk about these services in a little more detail. The first one which we have here on the list is scheduling and monitoring.

00:02:29 Obviously a data warehouse is like a big machine which has to run in an automated way 24x7.

00:02:37 You don't want to have to do too many manual steps to, for example, load data. That's basically impossible because many of these things happen during the night.

00:02:46 So it's very important to be able to schedule, to define jobs, to schedule jobs, to monitor jobs, to be able to correct errors, all this kind of stuff. That's a very important area.

00:02:56 The next area is the more analytic area. It's the so-called OLAP area where you would like to define a set of analytic functions based on your data.

00:03:10 This means you would like to do, maybe, different kinds of calculations, different kinds of aggregation, behaviors based on your data,

00:03:19 without that you have to manually write the SQL or that the BI tool is doing that. In BW, there is an abstract layer for that which we will cover during this course.

00:03:30 Lifecycle management is also a very interesting point because it has two different flavors.

00:03:34 There's the model lifecycle management, so like transporting things from a development environment to a production environment.

00:03:42 This really happens throughout the stack, so you have this in the OLAP area when it comes to defining and creating reports. They have a certain life cycle.

00:03:50 Also the schedules which you derive or the scheduling artifacts which you should build, they have a certain life cycle. So everything here has a certain life cycle

00:03:59 and therefore needs to be transported. And changes in one area, like if you do model changes, this typically needs changes in transformations.

00:04:07 So the life cycle of the individual objects and individual service areas here is actually intertwined.

00:04:15 So that's one aspect of lifecycle management. But also the life cycle of data itself is important, so you have to be able to phase out aging data,

00:04:25 handle big data volumes, so also this is kind of one aspect of lifecycle management.

00:04:29 Or the question of archiving or consistency of the data. This is something you would like to do there and would like to have help from the application.

00:04:39 Another layer is the Extract, Transform and Load area (ETL), which is classically a database-driven approach. It is done by a standalone tool.

00:04:48 Within BW, we have very easily the possibility to go right from container A to B,

00:04:55 to have this scheduled and aligned with our life cycle, with our scheduling and monitoring.

00:05:00 It's one consistent and transparent process within BW, and you have even the possibility

00:05:07 to choose out of a variety of predefined transformations to change, to manipulate, the data.

00:05:14 Maybe one aspect which we should mention in the ETL area is that BW is very famous and well-known for its optimum integration of SAP ERP and SAP applications in general.

00:05:25 That's one of the very fundamental strengths of BW, actually.

00:05:28 Another area which is very important in all kinds of businesses is planning.

00:05:34 That's also a very specific aspect of BW, that BW comes with a built-in planning engine and planning applications.

00:05:41 Right. So typically you don't want to just do analysis or to just do the as-is reporting.

00:05:49 Typically you would like to do the planning and compare it with your plan data and to your as-is data.

00:05:55 And this is really a possibility there without additional tools necessary. It's in-built. And we will

also give you some examples in the last week of the course.

00:06:05 The last aspect here would be modeling.

00:06:08 Modeling basically, at least as we define it here, refers to the layout of data containers, of InfoProviders, in BW for transaction data, for master data.

00:06:17 BW actually has also a very specific modeling paradigm which is more abstract than creating things and tables on the database.

00:06:27 The idea behind modeling in BW is that you actually model in terms of business semantics.

00:06:34 So what you actually model in BW are business entities, like you model an object for customer data, you model an object for product data.

00:06:44 Material could be another example. So all the entities which play an important role in your organization.

00:06:51 And out of these objects, which are—if you compare to what happens on database level—much more complicated because

00:07:00 for example, the product has a number of attributes which would be kept in a table. It also has text because you probably, if you have

00:07:08 an international company, a global player, then you will have descriptions for the products in many languages

00:07:16 because the users in each location will want to see the description in their local language.

00:07:22 So this basically means that the text descriptions need an own table. But you don't want to care about this dependency between the attributes and the text every time.

00:07:32 The strength of BW is that it has one object. In this case you would create one object for products which encapsulates both of these.

00:07:41 And on top of that, even hierarchies. And product would be a very good example for hierarchies.

00:07:46 And you see that here in the example. It's not about, with these few examples we are providing here, that the one or the other approach is better or worse than the other.

00:07:59 It's just really about comparing how would you do it on a database level with different tools and how would you do it with BW.

00:08:05 And then we from SAP, we would also like to encourage you to pick the best-fitting approach.

00:08:12 And this is really a very strong one, so you could define everything on database level by your own, by defining associations and joins.

00:08:19 Or you do it in a more abstract way with BW objects. We will not mention them now, but we will do it in a later unit.

00:08:25 And then it's application, it's model-driven.

00:08:30 Another aspect which is very important when it comes to loading data in a data warehouse is the handling of data changes, of delta.

00:08:37 Because obviously you can't process the full amount of all historic data every time, but you only want to process, for example during a nightly load, you only want to process the changed data.

00:08:48 So therefore, it's very important to have some means to detect which data has changed in a certain layer before you load this changed data to another layer.

00:08:58 Obviously, it's possible and there are methods described in the literature how you can do this in the classical database way.

00:09:06 BW, on the other hand, provides an out-of-the-box solution for this

00:09:11 which basically is able to, while you load, detect the changes of the newly loaded data compared to the existing data

00:09:18 and then allows you, when you load out of a given container, to only propagate the changed values.

00:09:25 And that without that you have the need to contain a time stamp or a criteria for it.

00:09:30 So we have services in-built in the application. We have a specific object for that, for determining the data also in relation to

00:09:40 how was the data delivered into my application, into BW. And then we have an in-built calculation to do summations, to do overrides, and this is exactly the described service here.

00:09:52 In the OLAP world—I briefly summarized that—it's really basically the same.

00:09:58 So you can do, in terms of if you would like to achieve complex aggregation behavior which is not just a summation.

00:10:05 Maybe you would like to do a counting. Count my different customers based on a different criteria.

00:10:11 And maybe even you have a complex structure to do so. Not just one key figure. Maybe you have a bunch of key figures

00:10:18 and you would like to have it exactly in this layout all the time with this specific currency for reporting.

00:10:24 This is exactly where in the OLAP area we have an engine with predefined functions and an object model

00:10:34 way that you can model a specific business reporting case. And basically, in the database we are translating this that the data is being calculated

00:10:44 without having to write SQL statements or the need for an additional BI tool.

00:10:50 Of course, you can consume the data with a BI tool out of BW, but it's something optional

00:10:56 where we encapsulated again all this functionality in this OLAP area already.

00:11:02 In the handout you will also find a few other examples. This was just a few.

00:11:08 But we wanted really to highlight that SAP BW offers an integrated, out-of-the-box way to do data warehousing.

00:11:17 So if somebody asks you is BW still necessary?

00:11:22 The answer is yes. If you would like to follow an application-driven data warehousing approach, our answer would be SAP BW.

00:11:31 The combination of SAP BW and SAP HANA as database will then be the key and the main interest for the next units.

00:11:40 In the next unit, we will first focus on the terminology within BW, on the basics and building blocks.

00:11:46 So we're done for today and you should get ready for your self-test now.

WEEK 1, UNIT 4

00:00:12 Hello and welcome back to week 1, unit 4: Basics & Building Blocks.

00:00:18 This will be a very interesting unit because we would like to introduce the common methodology

00:00:26 about the terms and the objects will speak about during this course.

00:00:32 We will start with an explanation about the multidimensional objects for analytical purposes

00:00:38 and like we discussed previously, BW offers now services to represent this multidimensional model in the application.

00:00:48 This is exactly what our objects, our different building blocks, are doing.

00:00:55 Multidimensional models are sometimes also referred to as star schemas. Basically turned out to be this de-facto standard for organizing data for analytic purposes.

00:01:05 What normally happens is you organize data in a central table which contains the facts, which is basically the key values for products, business partners,

00:01:16 and measures or key figures, which are basically the numbers which you want to sum up, which would be revenue in this case, for example.

00:01:25 Or tax amount or net amount or gross amount.

00:01:27 Things related to a currency, for example. All amounts with the currency units and stuff like that.

00:01:33 That's the central part of the so-called star—that's where the name comes from—and then you have surrounding additional tables which enrich the information of the facts.

00:01:41 And the facts, you might only have the key of the business partner. And then you have a second table, a dimension table, for the business partner, which contains detailed information about business partners

00:01:51 like e-mail addresses, phone numbers, fax numbers, legal forms, all this kind of stuff.

00:01:58 The combination is basically what gives the flexibility and gives you the insight because you might be interested, for example, in

00:02:06 how does my revenue change or how is my revenue organized according to regions.

00:02:12 For such an analysis, you need more than the facts. You also need the information from the business partner dimension.

00:02:18 And the interesting part is now in BW we have an approach to say based on the character of data, we are treating it in different objects.

00:02:29 We have objects that are focusing on master data treatment.

00:02:33 Which would be the dimensions.

00:02:34 Right. We have objects to store the transactional, means the fact, data. And this is exactly how you start then to build a model within BW.

00:02:44 Our smallest object representing master data or key figure values is the so-called InfoObject.

00:02:52 It's the smallest building block we have in BW.

00:02:55 This is representing...it's actually starting with the idea that we represent a field of a table with an InfoObject. But it's much more than that

00:03:04 because one field, like an ID of a product, can have many, many things attached to it like attributes,

00:03:13 like hierarchies, and this is exactly the object where we encapsulate all those different services based on one business object.

00:03:21 So the product would basically be one InfoObject which contains all the information which is relevant for the business object, for the product.

00:03:30 So let's have a look at this in the system. If you look at the Product InfoObject of the EPM Demo scenario,

00:03:38 you see that first of all we have a key definition, which is basically the key we see in the facts. This has technical details. It has a data type, a length.

00:03:49 And then there's all the additional stuff which is logically contained in the dimensions.

00:03:55 There's text information, which normally has a separate persistence because they have a different key, they are language-dependent, therefore they have an extended key.

00:04:05 There's also the master data, the attributes, which we just saw in the star schema, and there's even more. There's the hierarchy information. Products typically have a hierarchy.

00:04:14 You have different sorts of products and subtypes of products. Products more or less naturally fall into a hierarchy.

00:04:20 And all this is basically combined in one InfoObject and as soon as you use this InfoObject in a certain context, all this information is automatically available.

00:04:39 That's the strength of the BW modeling approach. That you have one central object and wherever you use it, all this stuff comes automatically.

00:04:38 Let's say you are now reporting on a transactional model. We come to that. And you have the product there.

00:04:45 This means you can access all the information, all the attributes attached to it. For example, if the product has a size,

00:04:53 we are able in the report to drill down by size or by color or by region or by whatever. This is how we model it.

00:05:01 In the past, our InfoObject was really mandatory for everything we are doing in BW.

00:05:07 And we can now already tell you this is something which is not the case anymore.

00:05:13 With SAP BW 7.4 powered by SAP HANA, we also have the possibility to work without these InfoObjects. But that will be covered in a later unit.

00:05:23 Yeah, we'll have a detailed session about this later.

00:05:28 I think the second part of this InfoObjects were the key figures. So far it was master-data-related for the characteristics, but we also have a few only for key figures.

00:05:40 Key figures again, as in the case of characteristics, they are not only the numbers which you might want to add, but they also contain other relevant information which is important to add, for example, sum up numbers.

00:05:51 First of all, these numbers might come with units or with currencies. And it's very important to know this relation

00:05:59 because it doesn't make sense to simply add up different currency values. It doesn't makes sense to sum up euros and dollars.

00:06:07 This relation, for example, is encapsulated in a key figure in BW. Also the default aggregation is something which you can describe in a key figure.

00:06:17 For example, some key figures, for some cases, it doesn't make sense to sum them up. You typically would want to see the maximum.

00:06:23 That's something you can store in the definition of a key figure. And again, as in the case of characteristics, whenever you use it,

00:06:30 the system knows all these properties and how they belong together and behaves accordingly.

00:06:37 And this adds a sort of intelligence based on a data model.

00:06:44 These InfoObjects are combined together representing a so-called InfoProvider. An InfoProvider is really our data container in SAP BW and we have two types of it.

00:06:55 So it could be a data container which stores physically the data in the database, in our case SAP HANA,

00:07:02 or it could be a virtual container accessing the data remotely during runtime.

00:07:09 Maybe just one more thing when it comes to modeling such physical InfoProviders. When you model these InfoProviders, you basically use the InfoObject which we have described earlier.

00:07:20 So you would, for example, take a product ID and say okay, I want to have the product ID as part of my DataStore object.

00:07:27 Then the system, first of all knows, all tables which it's going to create for you in the database need a field with the corresponding data type, which the InfoObject describes.

00:07:39 And it also knows that there's a natural relation to, for example, the dimension table, to the attribute table, to the text table,

00:07:47 and that hierarchies are attached to this. So all this information is then contained in the combination of the InfoProvider and the InfoObjects which it's made of.

00:07:55 What you see here is then really an overview of what is new now with SAP BW 7.4 powered by SAP HANA.

00:08:02 It's not everything. We tried to bring every object we have, but we will really have a focus during this course

00:08:10 on objects which have been re-innovated or especially are available with the usage of SAP HANA as a database.

00:08:17 So let's start briefly with objects with persistency. Basically, we have two of our most famous objects. This is the InfoCube,

00:08:26 which is representing in the end a star schema on the database. But this InfoCube contains then aggregated data.

00:08:32 And we have an object, the DataStore object, where we store granular data and are able to do a delta calculation within the application.

00:08:40 So based on what delta is delivered from the source, this object can calculate a delta and can override values by creating a proper delta value.

00:08:50 One more elaborate object which you also see here in the slide is the semantically partitioned object.

00:08:56 That's actually something which offers services on top of the classical DataStore object or InfoCube.

00:09:03 Namely, it takes care of partitioning, which is necessary in many situations, for example, if you have different regions reporting on similar data,

00:09:11 you have different load schedules, and the semantically partitioned object allows you to partition data according to such characteristics.

00:09:18 Again, automatically, without you having to do much. The application is handling it.

00:09:22 There is an interesting change coming up, and I'm really happy to make this announcement for later in the course.

00:09:28 The DataStore object has been re-innovated. So we have a new DataStore object available. It's the advanced DataStore object,

00:09:36 which is a pure SAP HANA-related and optimized object. And we will cover that in its own unit. You will see the demonstrations

00:09:44 sometimes based on a standard DataStore object which you are aware of, but we will also show the new advanced DataStore object, which will be introduced with BW 7.4 SP8.

00:09:54 Okay. Let's come to the objects without persistency. Maybe the most classical one which people with affinity to BW know already is the MultiProvider.

00:10:04 A MultiProvider is classically used first of all to decouple the persistent InfoProviders, which we just discussed, from the query consumption.

00:10:13 And it's also used to do something like union of physical providers and that.

00:10:19 With 7.4 SP5, we have introduced basically the successor of the MultiProvider and several others. That's the CompositeProvider.

00:10:27 The CompositeProvider is also able to do this union functionality, but it can do more than just union. It can also do joins, inner joins, outer joins. It will be able to do temporal joins.

00:10:36 So it can also be seen as a replacement, for example, of the InfoSet.

00:10:42 That's one step in the direction of simplification, which we are going to cover in a later unit. Namely that we consolidate many different types of InfoProviders into one object,

00:10:52 which makes the choice of choosing the right object much simpler.

00:10:56 And because these objects are so important now with SAP HANA, we will cover them in an own specific unit. For now,

00:11:05 remember that the CompositeProvider and the MultiProvider are all objects to define joins in the BW application.

00:11:11 And we have a very new object, the Open ODS view, which is really a view-based object allowing us to work virtually in view-based modeling.

00:11:22 And it's really one of those objects where we allow field-based modeling.

00:11:28 Field-based modeling will also be a very important part during the course. That means InfoObjects are optional. You can define the field definition.

00:11:36 But with that, the Open ODS view is our object, a really strong object, to access external sources, which could even be a Hadoop cluster. So we are very, very flexible there.

00:11:47 The whole point of this is that it's very agile. You can very easily adapt external data and consume external data without a lot of modeling effort beforehand.

00:11:56 There is one interesting part. You see the SAP HANA models also there. So people with previous BW knowledge are now wondering.

00:12:06 And yes, you are right. It's not a classic InfoProvider but we would already here highlight that due to the combination of SAP BW with SAP HANA as a database,

00:12:16 we are able to model in a hybrid mode. We also call it sometimes mixed scenario.

00:12:23 This means we are able to combine the modeling approaches coming from the SAP HANA view world together with our BW InfoProviders,

00:12:31 which will again be an own unit because it's very important for us to see that as a common model environment.

00:12:40 It really shows how BW and HANA together basically allow you to use and leverage the best of both worlds,

00:12:48 depending on the situation. Whatever is useful in the current situation.

00:12:53 So far we discussed the smallest objects in BW: our InfoObjects, the data containers, our InfoProvider. And now it's time to fill it with data.

00:13:02 We do this by defining so-called data flows. The data flow is, again, a sequence of different steps. And for that, we also have different options available.

00:13:13 Maybe let's start bottom-up in the picture on the right-hand side.

00:13:18 The DataSource is basically the interface to the source data. So if you have, for example, an ERP system which you connect to your BW system,

00:13:27 and the ERP system contains financial data—for example, accounts receivable—then there would be a corresponding DataSource which basically knows how to connect to the ERP system,

00:13:37 which gives you the structure of the accounts receivable data which you can access in the ERP system,

00:13:43 and also is able to, for example, even execute the load of this.

00:13:48 This is also our object where the different data provisioning options will be attached to.

00:13:53 The DataSource is really just our interface object, and then we can connect via this interface many, many different sources. You will see that also in its own chapter.

00:14:03 So this is the first layer. It could have an own persistency, this DataSource.

00:14:09 It could also work—again, this is new with 7.4—work without a PSA, a persistent staging area, which is the persistency. We will also cover that.

00:14:18 But the next object then would be the transformation.

00:14:22 The transformation is basically the logic where you define how to transform the data from the sources to bring it into the shape of your consolidated data warehouse.

00:14:34 Different sources might provide similar data in different ways. And of course, you want one consolidated way. And therefore you need transformations,

00:14:44 which basically allow you to adapt keys, maybe enrich data from one source with data from somewhere else because data from a third source has additional key fields and you want one unified source.

00:14:55 So that where you model transformation to bring data into a unified form.

00:14:59 We spoke in the data warehousing context about consolidation, harmonization, data cleansing. That's your object.

00:15:06 This transformation is being executed or adopted by the so-called data transfer process (DTP). So this is the scheduled process bringing

00:15:57 or loading data from A to B. This could be out of a source system into a data container or it could be between two InfoProviders.

00:15:26 The data transfer process also takes care of monitoring, error handling, all this kind of stuff. So all this kind of infrastructure is contained in the DTP.

00:15:34 Everything together can be then scheduled and re-executed daily, once a week, over the weekend,

00:15:42 by a process chain, which is then our batch job to execute this whole data flow.

00:15:50 Data flow can, of course, be much more complicated than the example here. You might have to load multiple InfoProviders before you can load your last InfoProvider because you are doing a lookup.

00:15:59 Before you do a lookup, obviously, the data which you want to look up has to be loaded first. So this type of synchronization, which job has to run first, is something you can model in a process chain.

00:16:09 And again, the process chain can be scheduled, say, in a daily manner or weekly or even hourly

00:16:16 in order to make sure that all the different containers which are relevant in this scenario are loaded in the right order in order to make sure you get consistent information.

00:16:24 What you see here are basically the different features you would expect from an ETL tool. And that's in-built in SAP BW.

00:16:30 So this is exactly what you typically do in an ETL tool with data quality checks and so on, and we have encapsulated it in different objects within BW.

00:16:41 The last area is then the consumption of the data via different BI tools.

00:16:48 Also we from SAP, we do have a range of tools within this openSAP course. We will use various tools.

00:16:58 So for different for use cases we will show dashboarding, we will show Microsoft Excel-based analysis. This will be done via the BI tools.

00:17:07 And the layer how you can consume, or you should consume because the OLAP logic means our analytic manager logic where we really do the aggregation types,

00:17:17 the temporary calculations and stuff like that, is done via the BEx Query Designer, which is our modeling tool so seen for the OLAP engine.

00:17:25 And you can predefine a report which can then be displayed or consumed by a BI tool.

00:17:33 Also this will be in its own unit, where we show the different functionalities of such a BW query, which can be the last layer, but it doesn't have to be.

00:17:45 And this is again something interesting which is new. We are also open for SQL consumption.

00:17:52 In general, BW has various interfaces for data consumption. We have some interfaces which are only available for SAP tools.

00:18:03 We have more open interfaces like MDX. And we are now able also to have SQL consumption which will then be in its own unit.

00:18:14 So maybe let's show the small traditional data view, data flow.

00:18:23 This is the standard content how right now the EPM model looks. So you have all the older objects like MultiProvider combining two InfoCubes for aggregation.

00:18:34 The data has been aggregated from granular DataStore objects. You will find that also in the trial system.

00:18:40 During that course, we will show you how you would model exactly this analytical purpose, new, with technologies coming with SAP BW 7.4 powered by SAP HANA.

00:18:53 So this was just a short excursus. So let's summarize.

00:19:00 In general, BW is again the application. We have objects for an object-driven modeling approach.

00:19:10 We have different InfoObjects, characteristics, and key figures. We do have InfoProviders for storing the data physically or virtually.

00:19:20 And we have the query which encapsulates the logic for the OLAP engine, which is an important part for doing interesting calculations.

00:19:28 In the next unit—this was the basics—we will show you the simplifications coming with SAP HANA.

00:19:34 Don't forget to do your self-test.

WEEK 1, UNIT 5

00:00:11 Hello and welcome to week 1, unit 5: Simplification with SAP BW 7.4.

00:00:21 And as you see, it's part 1. So that means we also have part 2.

00:00:26 In part 1, we'll start with an explanation of the journey we had in SAP BW with SAP HANA as database. So there will be a few principles,

00:00:36 a few ideas which were guiding us in everything we changed and implemented in the combination of both.

00:00:43 We will talk about the optimizations of SAP HANA in the application.

00:00:49 This means we aligned the application with the database. It's very interesting to see what we did there.

00:00:56 We will give you a few customer experiences, customer cases we had out of the value of this combination.

00:01:03 And then it's not ending there that we optimized existing stuff. So you will see that

00:01:09 also with HANA we introduced many new functionalities with SAP BW as kind of a platform, as a modeling platform.

00:00:17 Maybe we should mention that of course the optimizations are still ongoing.

00:01:21 But besides the optimizations, that's what you were just saying, we are also going for a second approach, namely leveraging all the functionality which HANA provides

00:01:29 which is relevant for data warehousing or analytics directly within BW to tighter integrate both platforms.

00:01:36 So we took, as I mentioned, several steps. The first step—and this is kind of reflected here—

00:01:43 when you look at the difference between an SAP BW installation based on any database and, in comparison, powered by SAP HANA,

00:01:53 that's basically that we took the different key areas we have in BW, like the data modeling area, the planning area, analytical data management.

00:02:04 This is something in a traditional installation usually being executed in the application server.

00:02:11 So most of the time we are spending in the application server, doing the added calculations, and then we are writing data back to the database.

00:02:18 With SAP HANA, the approach is totally different.

00:02:22 You see that as an indicator in the bars of this area on the right-hand side, that we are pushing many functionalities down to HANA directly.

00:02:32 One of the bars where you see that there's no change is the data modeling part, and that's actually a huge benefit from a customer perspective

00:02:39 because what that actually means is that your existing models, which you have created maybe years ago in BW on any DB, will still be running once you migrate your BW to HANA.

00:02:49 So you don't have to remodel everything. But these optimizations, this pushdown of expensive logic, will happen immediately and without modification of the models.

00:02:59 I think the term usually used there is nondisruptive. This means we take the existing

application, also your logic there,

00:03:07 bring it toward HANA and then we are—and I think this is what we are going to show you next—we are going to change a lot of things in the background

00:03:17 and optimizing the processes toward SAP HANA.

00:03:21 And with that, we see an excellent query performance. This means we see a huge benefit in reading the data.

00:03:28 We see benefit with latest release of BW also on the data loading side. This means how fast you can load data through the Business Warehouse.

00:03:37 This goes along in planning we see a huge change in how you build models. So it will get definitely simpler.

00:03:45 It's getting more flexible. You can drive agile data modeling approaches, stuff like that.

00:03:51 With the optimizations, we will see that data persistency layers got reduced. So it will get, simply said, a smaller data model.

00:04:01 And I think all over this is really a good combination and where we did a lot of investment in BW as a data warehouse.

00:04:10 For that, step 1 of the journey was really to take the traditional objects we have in BW—

00:04:17 the InfoCube for reporting, the DataStore object for data persistency—because all data models are really built out of them.

00:04:26 So when you look at the details here on the right-hand side, you see that the table layer of the InfoCube on the database changed

00:04:34 from BW running on any DB to BW running on HANA.

00:04:38 As I said before, the model which you develop on ABAP side in your BW system does not change.

00:04:44 You still have the same InfoCube, but activating this InfoCube on BW side, we created different table layouts than we used to create before.

00:04:55 Funny story. I have always to think about that. A customer is really asking me, when we introduced this the first time, what is the InfoCube looking like in the application?

00:05:04 How is the icon looking like for such a HANA-optimized Infocube. But it's still the same.

00:05:07 It's still the same object, but physically we are converting this very complex star schema,

00:05:13 which was tailor-built for traditional databases, we are splitting it up to a flat structure and pushing it toward HANA.

00:05:21 So you see that the dimension tables are gone. Now you might wonder what happens if I have an existing InfoCube with data in it?

00:05:27 How can I bring this to the new table layout? This is done by BW automatically.

00:05:33 There is a conversion mechanism which does that for you and you benefit from this simpler layout.

00:05:42 If you create new InfoCubes, then of course the new model will immediately be applied.

00:05:48 Getting rid of the dimension tables is not only a simplification of the whole structure of the star

schema,

00:05:55 but also makes the modeling easier because it really means that the modeling of dimensions—and you might remember that

00:06:04 this is a very critical task from a performance perspective if you run BW on any DB—is completely uncritical from now.

00:06:11 And you can basically group your characteristics in a semantic way and you don't have to think about the performance impacts anymore because there is no performance impact.

00:06:20 This used to be a very tricky thing. You had to figure out what combinations of characteristic values might occur

00:06:28 to estimate the size of the dimension tables.

00:06:32 All these problems are basically gone. This whole optimization of dimension tables and grouping of characteristics in dimensions is not relevant anymore in HANA.

00:06:41 This is really a side story. It took me ages as a consultant to optimize dimension layouts on customer systems.

00:06:49 So it's really something where on the traditional database, definitely a bottleneck occurs if it's not modeled properly. It cannot happen on HANA, so really a big simplification

00:06:59 without having a big impact on how you're using the object. It's just getting faster by reading out of the data.

00:07:06 So this was the InfoCube and the optimization we did in sourcing and background.

00:07:12 With a DataStore object, it's a bit different there. We had there the activation process, which I think was the most critical and most time-consuming process in BW.

00:07:21 What we realized from our analysis on running customer systems was that 80% of the time consumed by data flow is consumed by DataStore activation.

00:07:30 In the application server. So in the application server, just doing the lookups and the calculations there.

00:07:36 So we figured that optimizing this step would provide a huge benefit for the whole data flow.

00:07:41 And so we took the approach to not modify the table layout, as we did in the case of an InfoCube,

00:07:49 but actually we took the activation logic, re-implemented it on HANA side, and now after migrating your BW from any DB to HANA,

00:07:59 BW still triggers the DataStore activation but it's not executed on the ABAP layer anymore,

00:08:06 which meant that a lot of data had to be transferred from the database up to the application server and backward after that.

00:08:13 But now the whole activation procedure—all the lookups, all the inserts, updates, and deletes—

00:08:21 are immediately executed on the database with far less traffic between application server and database. And this is much, much more performant.

00:08:29 This is really a tremendous speedup, the whole data loading and the whole data activation process. It's easily—and really, I have seen many customer situations and systems there

00:08:39 —it's usually a factor of 5 to 10. When we adjust you to this speedup and activation, we accelerate the whole process.

00:08:48 If you look now—and remember, this was step 1—

00:08:53 we took existing objects, we took the InfoCube and the DataStore object, and made them HANA-optimized in background.

00:09:01 That means after conversion, they are working as they were working before, definitely in a faster manner.

00:09:10 But they are still fulfilling the same purpose. So this was really step 1.

00:09:14 Maybe there's one thing we should mention. That info that aggregates, which we still see on the left-hand side, if you look at the little InfoCubes we have there,

00:09:23 they are completely gone in HANA. We don't need any aggregates on HANA. Actually, I think you can't build them anymore.

00:09:30 We don't have aggregates anymore because they're not necessary because the columnar store of HANA

00:09:37 really allows us to perform queries on line item level and we don't have to build and maintain preaggregates in order to guarantee query performance.

00:09:48 Using this layout, the columnar idea for SAP HANA, we can work on very large and very broad table structures with a lot of records much more efficiently

00:09:57 than with any other database. This is exactly the reason why we took this optimization. There is also, besides these aggregates, there are no index information there.

00:10:07 There are no statistic updates. This is all gone.

00:10:10 But now if you think hey, this is looking boring because the data flow is still looking basically the same but it's HANA-optimized, this is exactly what we show you in step 2

00:10:19 because with the latest releases, it's looking completely different.

00:10:23 Maybe one thing we should also mention is that at this stage we already realized that running a query on a DataStore object gives you about the same performance as running a query on an InfoCube.

00:10:33 So there was the potential already to remove InfoCubes in certain cases, if you had the same data and the same granularity–or even a finer granularity—

00:10:42 in a DataStore object in a lower layer. But of course, this is a little bit more tedious because you would really have to modify queries and all that stuff.

00:10:52 So that's maybe a second step of optimization.

00:10:55 The first step of optimization is done by the system automatically, by pushing down the DataStore activation and adjusting the layout of the InfoCube.

00:11:05 Afterward, it's still possible to optimize data flows and speed them up, especially in the loading area, by removing layers.

00:11:13 This is something that always seemed from a too technology-driven perspective from my opinion.

00:11:22 So what we definitely achieve with SAP HANA as a database is a compression because with a

columnar database as HANA,

00:11:30 we reduce the overall data size of the database just due to it's compression, to the way we store the data.

00:11:37 Besides, we are getting rid of index information and stuff like that and this is giving us this huge and impressive reduction factor.

00:11:45 Besides that, we see that the data loading processes are getting faster, that the activation is getting accelerated,

00:11:53 and reading data out of InfoCubes, out of DataStore objects, for analytic queries in general, also like—we will see that later—

00:12:03 like operations, like exceptional aggregation, count something. It will definitely get faster and we will show you the examples during this course.

00:12:12 But the technology value is of course the one side. It comes along with the business value.

00:12:18 So it's not just that something is getting faster. It's also relating to business value.

00:12:25 If the data is earlier available, you are earlier available to do reporting.

00:12:31 Or you can do updates more often. You can load data more frequently.

00:12:35 Right. You can drive analysis much more deeper and this will definitely lead to business value.

00:12:42 Therefore we always recommend to have an eye on both sides. It's not just the technology; definitely there is business value.

00:12:49 I've seen many customer cases. Usually there is definitely a business value coming along with it.

00:12:56 This was, like I said, step 1. Now we are, from the idea, in the second step.

00:13:03 SAP HANA is a database. Of course, you can also use SAP HANA in other contexts, like in a native environment to build analytical models there.

00:13:13 These capabilities which are coming along with SAP HANA are something we also can—and we would like—to leverage in BW.

00:13:21 So our way and our strategy is definitely to align BW toward SAP HANA. This is what we are doing with the pushdown.

00:13:30 But on the other hand, it's also that we are bringing up new functionalities coming from SAP HANA into the BW application world.

00:13:39 So you see the green bar here.

00:13:41 We definitely have features—and they will get introduced to you during this course—where we see platform-related features.

00:13:49 This means a flexible way how you can model data in HANA and then combine it with BW.

00:13:54 The key innovation here is the combination of the strengths of both of them.

00:13:59 Another example would be certain business functionality or libraries which are implemented in HANA

00:14:05 which you want to leverage from BW, for example, within a transformation. This kind of stuff.

00:14:10 This is kind of the middle of the journey. We are not there, we're not at the end of the journey. Step 1 was really—to summarize this one here—

00:14:19 that we pushed SAP HANA as primary database underneath the BW to accelerate

00:14:26 the existing loading, the existing reporting, and every traditional process there.

00:14:33 The next step would be that also new functionalities—besides that we optimized the existing objects—that new functionalities are getting used

00:14:42 and new innovations developed by the BW application will natively and directly be integrated, aligned with SAP HANA.

00:14:52 And that's definitely the core direction and the strategy here. This is what we will see in unit 2.

WEEK 1, UNIT 6

00:00:12 Hello and welcome to week 1, unit 6: Simplification with SAP BW 7.4.

00:00:21 It's the second part, so we had already one session where we talked about the optimization of existing objects in BW powered by SAP HANA.

00:00:33 In this unit, we would like to focus on new features, new objects coming with the latest release of BW

00:00:41 and of simplifications, especially in the data modeling and the architecture perspective.

00:00:49 And we will introduce new scenarios with new objects and a totally new

00:00:54 modeling environment which will be an appetizer for the rest of this course.

00:01:00 We talked already about BW and HANA and the object optimizations.

00:01:07 But there is, beside the fact that with these optimizations in the background and the performance we gain,

00:01:14 there's also that the combination of BW and HANA enables us really to simplify the whole data modeling process.

00:01:22 This means we increase the agility of the data warehouse by leveraging HANA and BW as a platform for data modeling,

00:01:35 reducing on one hand the complexity of the landscape because we have now a strong core environment where you can do a lot of things

00:01:44 which is combining the strength of BW as application, of ABAP-based application

00:01:51 together with SAP HANA as SQL-based database and, of course, a SQL-oriented approach

00:02:00 for modeling different data models.

00:02:05 This all will happen in one common modeling. You will see that we have a seamless integration in the meanwhile with the latest release

00:02:13 that we can process much more data with new processes which are totally HANA-optimized

00:02:20 and that the existing BW services, the core stuff we have in BW, is still a very important factor in this whole game.

00:02:31 So looking a bit at the details. So don't bother too much on the right-hand side.

00:02:37 It should just indicate that we have a data warehouse architecture with different kinds of data marts within SAP BW

00:02:46 we have a layer there for storing the data. We are transforming the data, and then we create different architected data marts

00:02:55 and besides that—and this is now really the new thing and the very interesting thing left inside.

00:03:01 It's the operational area we are receiving from SAP HANA as modeling environment.

00:03:07 The nice thing is—and that's very important—we want to integrate both worlds. We don't want to have them sitting side by side with no integration

00:03:15 but we want to be able—and that's what BW 7.4 brings you—we want to be able to leverage all the data which you have in the native HANA side

00:03:25 within BW for all purposes you can imagine. For query purposes, but also for ETL purposes so you can integrate data from the operational side

00:03:35 directly in queries. You can also integrate it into a data flow if you have the need.

00:03:42 That's one way of integration, one direction, from HANA native into BW.

00:03:48 But there's also the other way around so that we can basically expose data which you have prepared in BW,

00:03:53 cleansed master data, for example, or consolidated transaction data, consolidated sales data from many systems.

00:03:59 You can publish on the native HANA side and give it to end users for agile modeling on top of consolidated data.

00:04:10 I think that's a very nice example. So imagine you have your core data warehouse with a reference architecture and all the stuff around it,

00:04:17 harmonized master data there, but you would also like to have a very agile world where you can very easily build a SQL query

00:04:24 and create HANA-based standalone models. So I even have, on the left-hand side in HANA, SAP HANA Live scenario, means content there for operational reporting.

00:04:35 And then you would like to integrate this operational data with your data warehouse.

00:04:39 This is exactly the area of integration we are speaking about.

00:04:44 There's another aspect which is the integration of algorithms which exist within HANA. HANA comes with a lot of libraries for...

00:04:54 Like predictive. Take predictive as an example. In BW, it was very, very hard before to implement a predictive case,

00:05:03 to really make predictions based on data.

00:05:07 HANA is, in the data mart area, very strong in that we have different libraries, algorithms there,

00:05:13 and this is of course something we would like to leverage, to implement, and to integrate in the BW area.

00:05:19 Still, BW brings you the benefit of a scheduling mechanism, of monitoring functionality.

00:05:27 So basically if you have such a scenario which leverages a predictive functionality

00:05:32 on either BW data or data on your native HANA side,

00:05:38 you can still use all of these services which BW brings to make this basically an industrialized solution.

00:05:46 So what we will do in this course, and this is really a trend we see on customer side, is we will consider both worlds as one,

00:05:55 let's say, optimized aligned on each other modeling world where you can pick the right approach where you need it.

00:06:01 This is exactly what we are going to show in the following units. Also a lot of examples.

00:06:09 But for now, let's also take a look in our journey, in our optimization journey.

00:16:16 And we started, you remember, with the picture of what a traditional data flow looks like on a traditional database, the left-hand side.

00:06:24 And now you see there the full-blown picture with BW 7.4 especially tailored and powered by SAP HANA,

00:06:32 so there are much more and new components in there. So you see that

00:06:37 instead of having a MultiProvider, for instance, doing the join, we have a new object which is completely and exclusively available with SAP HANA. It's the CompositeProvider:

00:06:46 our strong new object to do the joining operations you see in the data persistence area.

00:06:53 You see we still have a DataStore object, but also there we have a new kind of DataStore object, the advanced DataStore object,

00:07:00 which is a new approach how we manage the persistency in BW with much more flexibility than before.

00:07:08 And even the data source means the source system has changed. This means that with 7.4, we even have new opportunities in data loading.

00:07:19 And this is exactly the scenario we will show you in the rest of the course. But for now,

00:07:23 we will also focus on the single objects to give you kind of an appetizer what 7.4 powered by SAP HANA means.

00:07:30 One thing you also see on the right-hand side. There are no InfoCubes anymore. Basically, one of the things that comes with BW 7.4 SP8

00:07:39 is, as Marc said, the advanced DataStore object, which basically consolidates the

00:07:47 classic DataStore object and all the subtypes of DataStore objects and the InfoCube in one object,

00:07:52 which makes it much easier to decide which object to take, because there is only one.

00:07:57 The different flavors which you have in behavior of this object—like standard DataStore object behave like a right-optimized DataStore object

00:08:07 or behave like an InfoCube—can be switched easily at any time.

00:08:11 So this is the main object for managing the data persistency in BW.

00:08:17 So the advanced DataStore object is really the future and in that area, where we also see a reduction of the amount of InfoProviders we had, and the other side is the new CompositeProvider.

00:08:28 the CompositeProvider is basically the virtualization layer of BW.

00:08:32 It's basically the successor of the MultiProvider, but it's also the successor of objects like the InfoSet for joining

00:08:41 or for temporal joins, this kind of stuff. So all of this can now be done with one modeling artifact, which is the new CompositeProvider.

00:08:50 And for both objects, we will show you in a corresponding unit. So each object, because it's one of the key aspects which are new,

00:08:59 we will have an own unit for this and we'll show you really the optimization and detail in the

example.

00:09:05 But both are exclusively available for HANA and they are optimized in background, again, for HANA. But keep in mind here we have one object for joining,

00:09:13 one for persistence, management, this is the advanced DataStore object and Composite Provider.

00:09:18 And I think there is another interesting one for accessing external sources.

00:09:24 For accessing external sources, we have developed a view type object, which is in that respect similar to the CompositeProvider,

00:09:33 which allows us to adapt external data sources and enrich them by the specific semantics which you need to leverage it within BW.

00:09:41 For example, if you have a table inside HANA and you want to use a BW query on top of this, you can do this very easily using Open ODS views

00:09:49 where you basically enrich the metadata which comes from the table with information like which of the fields are key figures, which of the fields are characteristics.

00:09:58 You can also create links between different tables in order to associate master data to transaction data. All this kind of stuff.

00:10:07 So it's very easy to leverage star schemas on the HANA native side within BW.

00:10:14 This is really the first object, and this is a really important key feature here, working with field definitions.

00:10:23 So when we say it's a view, it's really a view, means the definition can be done via fields.

00:10:31 This means you are typing in this is now a field character 20 and it's related to this field in my source.

00:10:38 So it's really a new way how we can consume and integrate external data models in an agile way, for example, for prototyping, to see

00:10:46 is this external data model fitting to my BW world, to the master data I have there.

00:10:52 And you can really very nicely play around and give it semantics, use the field-based modeling approach, one key aspect.

00:11:00 And like I said, we will have a detailed example for that.

00:11:06 To sum up the area of the objects which are new with 7.4 and mainly starting with Support Package 8 in BW

00:11:17 and now the view again what is powered by HANA specific and what is still there for any DB.

00:11:25 You see the left-hand side with all the available objects we have in BW.

00:11:29 We split it into a virtual layer and a persistence layer, so really where is data behind and which is just a virtual consumption.

00:11:37 And there it's clearly our...this should just show our strategy there. That for all the virtual objects and

00:11:47 for combining data sets, whether they are in BW, whether they are in HANA,

00:11:52 and should be integrated with BW models, we see the CompositeProvider as the future object

where we see the most innovations happening.

00:12:01 We see for direct accessing but with a much more enriched semantic on top of it as it was possible with the VirtualProvider.

00:12:11 We see the Open ODS view for that. Also supporting field-based modeling. This means really a modeling experience without creating InfoObjects.

00:12:20 Thereby very easy and very fast.

00:12:23 Right. We still see the InfoObject.

00:12:29 If you need mature master data services, the InfoObjects with all the master data stuff we see there with all the attributes,

00:12:40 it's still the main object there.

00:12:43 We see for data persistency we see the advanced DataStore object, which is a totally new object built exclusively for HANA.

00:12:55 where you also are able to combine InfoObjects and fields. So this is a new paradigm in BW.

00:13:01 We would like to leverage this also in this course and to show you extensively what this means because it's really a new way to create data models in BW.

00:13:09 But these are clearly our four most innovative objects we have exclusively for HANA.

00:13:18 Again, the old story, we are still supporting, of course, the left-hand side also in a powered by SAP HANA system.

00:13:25 But this is where we see the most innovation is going and if you ask me how should I implement now a data flow in BW,

00:13:34 this would be my answer in terms of objects.

00:13:39 From a modeling environment, there's also something new and we should introduce this now because everything you will see

00:13:49 will mainly be going in the Eclipse-based modeling world. This means we will leverage Eclipse as a platform for data modeling.

00:13:58 Eclipse is already there for SAP HANA studio modeling. This means the model that you know in the HANA native world.

00:14:06 But we will also introduce a BW-specific modeling tool in Eclipse

00:14:12 to bring new innovations like the objects we mentioned, like in CompositeProvider and Open ODS view and advanced DataStore object.

00:14:18 Also in a new-fashioned way of how we should model it. It's a new experience. It's much easier.

00:14:27 It's looking differently and we are going to demonstrate this.

00:14:31 I think we're going to have a quick look into the system right now.

00:14:33 I think that's the next demo. So our first demonstration.

00:14:38 So we will jump in the modeling tool perspective. This is something you have to download and then install in your HANA studio,

00:14:48 so you will find a corresponding SAP Note in the handout of this example here.

00:14:55 You will see that we are doing also the other way around and integrate the SAP GUI also in Eclipse. And a new modeling environment.

00:15:04 Let's have a quick look into this.

00:15:07 Here you see the SAP HANA studio, which you are probably familiar with.

00:15:11 And you see the different perspectives which Eclipse comes with in the HANA area.

00:15:18 So just for clarification, all the HANA models would be here in the content area? My HANA views and stuff like that,

00:15:26 which are known from the HANA native modeling environment.

00:15:30 So we have a Catalog which contains all the database objects. The Content contains all the analytical calculation and attribute views, this kind of stuff.

00:15:39 The Security contains the user management and management of privileges.

00:15:45 And now you see that besides the SAP HANA Administration Console, there's another perspective which is the BW Modeling perspective.

00:15:56 When you click on this BW Modeling perspective, you can actually see that we have connected a BW system to this Eclipse installation already.

00:16:04 And when we open this, I have to type my password. This is really the back-end password for an ABAP application server.

00:16:12 This was really interesting also for me when I was doing this the first time.

00:16:17 So when you add a system here, it will take the information, the connection parameters and stuff like that, from your SAP Logon.

00:16:25 So this means the instance number you have there will be reused here and we just used our BW user

00:16:36 to connect to a traditional BW system, but not via the ABAP perspective. We are now working in SAP HANA studio in the BW modeling perspective.

00:16:47 When you open, for example, your Back End Favorites, you see that we have created an InfoArea. So this is really an InfoArea as you know it from RSA1 in the back end

00:16:57 for the NetWeaver EPM model. And if I open this,

00:17:02 you further see that we have some sub-InfoAreas. And then we have, grouped by object type,

00:17:08 certain folders for the different BW objects. In here, you would find advanced DataStore objects and let's maybe do this and open one.

00:17:19 So this is an advanced DataStore object. You will see that an Eclipse-based editor opens where you can model a DataStore Object. So this is obviously a very simple one.

00:17:28 Not sure it really created it, but you can probably check this.

00:17:32 But down here it says...this was you.

00:17:35 It was me. It was a test.

00:17:37 But it's really the interesting point, and this is a question we are receiving many times.

00:17:44 What you are now seeing here is really actually your RSA1 perspective. So you are seeing the InfoAreas maintained there.

00:17:52 You are seeing the objects there. Maybe we can open a traditional object.

00:17:56 Yes, let's maybe go to a classic DataStore object.

00:18:00 Oops. See what happens? Because there is no Eclipse-based editor for the classic DataStore object, the SAP GUI editor opens

00:18:10 which is a very nice integration. So whichever object you take, either the model Eclipse-based editor opens up

00:18:16 or if such an editor doesn't exist because it's a classic object, the old editor in SAP GUI opens.

00:18:23 And you can still maintain and do everything with your object which you want to do.

00:18:28 This is coming from the embedded SAP GUI. There is also another perspective on the right-hand side above there, where you can write ABAP code and stuff like that.

00:18:37 But via the BW Modeling tools, clearly new features, new objects, will be available there.

00:18:44 And if you now ask yourself about the strategy, maybe we can switch back to the slides.

00:18:51 The strategy is clearly that we introduce new features via the Eclipse-based modeling platform.

00:19:02 New innovations will clearly come in this direction and we are starting to develop more and more of the existing stuff toward this Eclipse-based modeling.

00:19:13 One nice aspect of this is if you do what we explained earlier, that you use both the native HANA modeling and BW together on one instance,

00:19:23 you don't have to leave the modeling environment. Everything happens within one Eclipse environment. All you have to do is switch perspectives.

00:19:29 Right. So you will see...for us we are really used to this already. For us, we are working in the SAP HANA studio.

00:19:38 We have a BW perspective for the ABAP world. And the rest for the new objects we'll do this clearly Eclipse-based because

00:19:48 I really enjoy working like that and I think it's a very nice way forward.

00:19:54 This was our simplification part 2.

00:20:00 We introduced that the new architectures are now really possible with SAP HANA as database.

00:20:07 We will leverage SAP HANA and BW to model in a common way.

00:20:13 This is done because we have the same tool, but also we are leveraging the modeling environment from HANA in BW and vice versa.

00:20:24 You will see during the course—and I hope that you stay with us and you are interested to see the new InfoProviders we introduced with SAP BW 7.4,

00:20:34 especially with SP8. Totally HANA exclusive.

00:20:40 For now, we would like to wish you all the best for your weekly assignment and see you in week 2.

www.sap.com