Chapter 2
New Requirements for Enterprise Computing

When thinking about developing a completely new database management system for enterprise computing, the question whether there is a need for a new database management system arises. And the answer is yes! Modern companies have changed dramatically. Nowadays companies are more data-driven than ever before. For example, during manufacturing a much higher amount of data is produced, e.g. by assembly line sensors or manufacturing robots. Furthermore, companies process data at a much larger scale, e.g. competitor behavior, price trends, etc. to support management decisions. And data volumes will continue to grow in the future. There are two major requirements for a modern database management system:

- Data from various sources have to be combined in a single database management system, and
- This data has to be analyzed in real-time to support interactive decision taking.

The following sections outline use cases for modern enterprises and derive associated requirements for a completely new enterprise data management system.

2.1 Processing of Event Data

Event data influences enterprises today more and more. Event data is characterized by the following aspects:

- Each event dataset itself is small (some bytes or kilobytes) compared to the size of traditional enterprise data, such as all data contained in a single sales order, and
- The number of generated events for a specific entity is high compared to the amount of entities, e.g. hundreds or thousand events are generated for a single product.

In the following, use cases of event data in modern enterprises are outlined.

H. Plattner, *A Course in In-Memory Data Management*,

DOI: 10.1007/978-3-642-36524-9_2,

2.1.1 Sensor Data

Sensors are used to supervise the function of more and more systems today. One example is the tracking and tracing of sensitive goods, such as pharmaceuticals, clothes, or spare parts. Hereby packages are equipped with Radio-Frequency Identification (RFID) tags or two-dimensional bar codes, the so-called data matrix. Each product is virtually represented by an Electronic Product Code (EPC), which describes the manufacturer of a product, the product category, and a unique serial number. As a result, each product can be uniquely identified by its EPC code. In contrast, traditional one-dimensional bar codes can only be used for identification of classes of products due to their limited domain set. Once a product passes through a reader gate, a reading event is captured. The reading event consists of the current reading location, timestamp, the current business step, e.g. receiving, unpacking, repacking or shipping, and further related details. All events are stored in decentralized event repositories.

Real-Time Tracking of Pharmaceuticals

For example, approx. 15 billion prescription-based pharmaceuticals are produced in Europe. Tracking any of them results in approx. 8,000 read event notifications per second. These events build the basis for anti-counterfeiting techniques. For example, the route of a specific pharmaceutical can be reconstructed by analyzing all relevant reading events. The in-memory technology enables tracing of 10 billion events in less than 100 ms.

Formula One Racing Cars

Formula one racing cars are also generating excessive sensor data. These sports cars are equipped with up to 600 individual sensors, each recording tens to hundreds of events per second. Capturing sensor data for a 2 h race produces giga- or even terabytes of sensor data depending on their granularity. The challenge is to capture, process, and analyze the acquired data during the race to optimize the car parameters instantly, e.g. to detect part faults, optimize fuel consumption or top speed.

2.1.2 Analysis of Game Events

Personalized content in online games is a success factor for the gaming industry. The German company Bigpoint is a provider of browser games with more than 200 million active users.[1] Their browser games generate a steady stream of more than

[1] Bigpoint GmbH—http://www.bigpoint.net/

10,000 events per second, such as current level, virtual goods, time spent in the game, etc. Bigpoint tracks more than 800 million events per day. Traditional databases do not support processing of these huge amounts of data in an interactive way, e.g. join and full table scans require complex index structures or data warehouse systems optimized to return some selected aspects in a very fast way. However, individual and flexible queries from developers or marketing experts cannot be answered interactively.

Gamers tend to spend money when virtual goods or promotions are provided in a critical game state, e.g. a lost adventure or a long-running level that needs to be passed. In-game trade promotion management needs to analyze the user data, the current in-game events, and external details, e.g. current discount prices.

In-memory database technology is used to conduct in-game trade promotions and, at the same time, conduct A/B testing. To this end, the gamers are divided into two segments. The promotion is applied to one group. Since the feedback of the users is analyzed in real-time, the decision to roll-out a huge promotion can be taken within seconds after the small test group accepted the promotion.

Furthermore, in-memory technology improves discovery of target groups and testing of beta features, real-time prediction, and evaluation of advert placement.

2.2 Combination of Structured and Unstructured Data

Firstly, we want to understand structured data as any kind of data that is stored in a format, which is automatically processed by computers. Examples for structured data are ERP data stored in relational database tables, tree structures, arrays, etc. Secondly, we want to understand partially or mostly unstructured data, which cannot easily be processed automatically, e.g. all data that is available as raw documents, such as videos or photos. In addition, any kind of unformatted text, such as freely entered text in a text field, document, spreadsheet or database, is considered as unstructured data unless a data model for its interpretation is available, e.g. a possible semantic ontology.

For years, enterprise data management focused on structured data only. Structured data is stored in a relational database format using tables with specific attributes. However, many documents, papers, reports, web sites, etc. are only available in an unstructured format, e.g. text documents. Information within these documents is typically identified via the document's meta data. However, a detailed search within the content of these documents or the extraction of specific facts is not possible by using the meta data. As a result, there is a need to harvest information buried within unstructured enterprise data. Searching any kind of data—structured or unstructured—needs to be equally flexible and fast.

2.2.1 Patient Data

In the course of the patient treatment process, e.g. in hospitals, structured and unstructured data is generated. Examples of unstructured data are diagnosis reports, histologies, and tumor documentations. Examples of structured data are results of the erythrogram, blood pressure, temperature measurements, or the patient's gender. The in-memory technology enables the combination of both classes of patient data with additional external sources, such as clinical trials, pharmacological combinations or side-effects. As a result, physicians can prove their hypotheses by interactively combing data and reduce necessary manual and time-consuming searches. Physicians are able to access all relevant patient data and to take their decision on latest available patient details.

Due to their high fluctuation of unexpected events, such as emergencies or delayed surgeries, the daily time schedule of physicians is very time-optimized. In addition to certain technical requirements of their tools, they have also very strict response time requirements. For example, the HANA Oncolyzer, an application for physicians and researchers was designed for mobile devices. The mobile application supports the use-as-you-go factor, i.e., the required patient data is available at any location on the hospital campus and the physician is no longer forced to go to a certain desktop computer for checking a certain aspect. In addition, if the required detail is not available in real-time for the physician, she/he will no longer use the application. Thus, all analyses performed by the in-memory database are running on a server landscape in the IT department while the mobile application is the remote user interface for it.

Having the flexibility to request arbitrary analyses and getting the results within milliseconds back to the mobile application makes in-memory technology a perfect technology for the requirements of physicians. Furthermore, the mobility aspect bridges the gap between the IT department where the data are stored and the physician that visits multiple work places throughout the hospital every day.

2.2.2 Airplane Maintenance Reports

Airport maintenance logs are documented during exchange of any spare parts at Boeing. These reports contain structured data, such as date and time of the replacement or order number of the spare part, and unstructured data, e.g. kind of damage, location, and observations in the spacial context of the part. By combining structured and unstructured data, in-memory technology supports the detection of correlations, e.g. how often a specific part was replaced in a specific aircraft or location. As a result, maintenance managers are able to discover risks for damages before a certain risk for human-beings occurs.

2.3 Social Networks and the Web

Social networks are very popular today. Meanwhile, the time when they were only used to update friends about current activities are long gone. Nowadays, they are also used by enterprises for global branding, marketing and recruiting.

Additionally, they generate a huge amount of data, e.g. Twitter deals with one billion new tweets in five days. This data is analyzed, e.g. to detect messages about a new product, competitor activities, or to prevent service abuses. Combining social media data with external details, e.g. sales campaigns or seasonal weather details, market trends for certain products or product classes can be derived. These insights are valuable, e.g. for marketing campaigns or even to control the manufacturing rate.

Another example for extracting business relevant information from the Web is monitoring search terms. The search engine Google analyzes regional and global search trends. For example, searches for "influenza" and flu related terms can be interpreted as a indicator for a spread out of the influenza disease. By combining location data and search terms, Google is able to draw a map of regions that might be affected from an influenza epidemic.

2.4 Operating Cloud Environments

Operating software systems in the cloud requires a perfect data integration strategy. Assume, you process all your company's human resources (HR) tasks in an on-demand HR system provided by provider A. Consider a change of the provider to cloud provider B. Of course, a standardized data format for HR records can be used to export data from A and import it at B. However, what happens if there is no compatible standard for your application? Then, the data exported from A needs to be migrated, respectively remodeled, before it can imported by B. Data transformation is a complex and time-consuming task which often has to be done manually due to the required knowledge about source and target formats and many exceptions which have to be solved separately.

In-memory technology provides a transparent view concept. Views deta scribe how input values are transformed to the desired output format. The required transformations are performed automatically when the view is called. For example, consider the attributes `first name` and `last name` that need to be transformed into a single attribute `contact name`. A possible view `contact name` performs the concatenation of both attributes by performing `concat(first name, last name)`.

Thus, in-memory technology does not change the input data, while offering the required data formats by transparent processing of the view functions. This enables a transparent data integration compared to the traditional Extract Transform and Load (ETL) process used for (BI) systems.

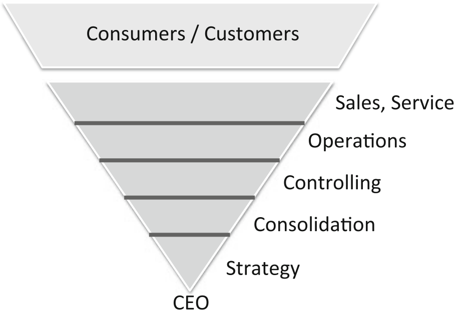

Fig. 2.1 Inversion of corporate structures

2.5 Mobile Applications

The wide-spread of mobile applications fundamentally changed the way enterprises process information. First BI systems were designed to provide detailed business insights for CEOs and controllers only. Nowadays, every employee is getting insights by the use of BI systems. However, for decades information retrieval was bound to stationary desktop computers. With the wide-spread of mobile devices, e.g. PDAs, smartphones, etc., even field workers are able to analyze sales reports or retrieve the latest sales funnel for a certain product or region.

Figure 2.1 depicts the new design of BI systems, which is no longer top-down but bottom-up. Modern BI systems provide all required information to sales representatives directly talking to customers. Thus, customers and sales representatives build the top of the pyramid.

In-memory databases build the foundation for this new corporate structure. On mobile devices, people are eager to get a response within a few seconds [Oul05, OTRK05, RO05]. With the ability to perform complex and freely formulated queries with sub-second responds, in-memory databases can revolutionize the way employees communicate with customers. An example of the radical improvements through in-memory databases is the dunning run. A traditional dunning process took 20 min on an average SAP system, but by rewriting the dunning run on in-memory technology it now takes less than 1 s.

2.6 Production and Distribution Planning

Two further prominent use cases for in-memory databases are complex and long-running processes such as production planning and availability checking.

2.6.1 Production Planning

Production planning identifies the current demand for certain products and consequently adjusts the production rate. It analyzes several indicators, such as the users' historic buying behavior, upcoming promotions, stock levels at manufacturers and whole-sellers. Production planning algorithms are complex due to required calculations, which are comparable to those found in BI systems. With an in-memory database, these calculations are now performed directly on latest transactional data. Thus, algorithms are more accurate with respect to current stock levels or production issues, allowing faster reactions to unexpected incidents.

2.6.2 Available to Promise Check

The Available-to-Promise (ATP) check validates the availability of certain goods. It analyzes whether the amount of sold and manufactured goods are in balance. With raising numbers of products and sold goods, the complexity of the check increases. In certain situations it can be advantageous to withdraw already agreed goods from certain customers and reschedule them to customers with a higher priority. ATP checks can also take additional data into account, e.g. fees for delayed or canceled deliveries or costs for express delivery if the manufacturer is not able to sent out all goods in time.

Due to the long processing time, ATP checks are executed on top of pre-aggregated totals, e.g. stock level aggregates per day. Using in-memory databases enables ATP checks to be performed on the latest data without using pre-aggregated totals. Thus, manufacturing and Scheduling rescheduling decisions can be taken on real-time data. Furthermore, removing aggregates simplifies the overall system architecture significantly, while adding flexibility.

2.7 Self Test Questions

1. **Compression Factor**
   What is the average compression factor for accounting data in an in-memory column-oriented database?

   (a) 100x
   (b) 10x
   (c) 50x
   (d) 5x

2. **Data explosion**
   Consider the formula 1 race car tracking example, with each race car having 512 sensors, each sensor records 32 events per second whereby each event is 64 byte in size.
   How much data is produced by a F1 team, if a team has two cars in the race and the race takes 2 h?
   For easier calculation, assume 1,000 byte = 1 kB, 1,000 kB = 1 MB, 1,000 MB = 1 GB.

   (a) 14 GB
   (b) 15.1 GB
   (c) 32 GB
   (d) 7.7 GB

References

[OTRK05] A. Oulasvirta, S. Tamminen, V. Roto, J. Kuorelahti, Interaction in 4-second bursts: the fragmented nature of attentional resources in mobile hci, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '05* (ACM, New York, 2005), pp. 919–928

[Oul05] A. Oulasvirta, The fragmentation of attention in mobile interaction, and what to do with it. Interactions **12**(6), 16–18 (2005)

[RO05] V. Roto, A. Oulasvirta, Need for non-visual feedback with long response times in mobile hci, in *Special Interest Tracks and Posters of the 14th International Conference on World Wide Web, WWW '05* (ACM, New York, 2005), pp. 775–781

http://www.springer.com/978-3-642-36523-2

A Course in In-Memory Data Management
The Inner Mechanics of In-Memory Databases
Plattner, H.
2013, XXIII, 297 p., Hardcover
ISBN: 978-3-642-36523-2