In-Memory Data Management

Hasso Plattner · Alexander Zeier

In-Memory Data Management

Technology and Applications

Second Edition

Springer

Hasso Plattner
Hasso Plattner Institute
Potsdam, Brandenburg
Germany

Alexander Zeier
Massachusetts Institute of Technology
Cambridge, MA
USA

ISBN 978-3-642-29574-4
ISBN 978-3-642-29575-1 (eBook)
DOI 10.1007/978-3-642-29575-1
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2012935682

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

In Praise of "In-Memory Data Management"

Academia

Prof. David Simchi-Levi (*Massachusetts Institute of Technology, Cambridge, USA*)

The book "In-Memory Data Management—Technology and Applications" by Hasso Plattner and Alexander Zeier describes a revolutionary database technology and many implementation examples for business intelligence and operations. Of particular interest to me are the great opportunities opening up in supply chain management, where the need to balance the speed of planning algorithms with data granularity has been a long time obstacle to performance and usability.

Prof. Karl Max Einhäupl (*CEO, Charité—Universitätsmedizin Berlin, Germany*)

Personalized health care requires the combination of distributed medical data. In-memory technology enables us to analyze millions of patient records within a second and allows us to devise therapies even closer to the patients' individual needs than ever before.

Prof. Christoph Meinel (*Hasso Plattner Institute* (*HPI*)*, Potsdam, Germany*)

I'm proud that HPI and the cooperation between HPI and SAP has provided such an inspirational research environment that enabled the young research team around Hasso Plattner and Alexander Zeier to generate valuable and new scientific insights into the complex world of enterprise computations. Even more than that, they developed groundbreaking innovations that will open the door to a new age, the age in which managers can base their decisions on complex computational real-time analysis of business data, and thus will change the way how businesses are being operated.

Prof. Donald Kossmann (*ETH Zurich, Switzerland*)

This is the first book on in-memory database systems and how this technology can change the whole industry. The book describes how to build in-memory databases:

what is different, what stays the same. Furthermore, the book describes how in-memory databases can become the single source of truth for a business.

Prof. Hector Garcia-Molina (*Stanford University, California, USA*)

Memory resident data can very significantly improve the performance of data intensive applications. This book presents an excellent overview of the issues and challenges related to in-memory data, and is highly recommended for anyone wishing to learn about this important area.

Prof. Hubert Oesterle (*University of St. Gallen, Switzerland*)

Technological innovations have again and again been enablers and drivers of innovative business solutions. As database management systems in the 1970s provided the grounds for ERP systems, which then enabled companies in almost all industries to redesign their business processes, upcoming in-memory databases will improve existing ERP-based business solutions (esp. in analytic processing) and will even lead to business processes and services being redesigned again. Plattner and Zeier describe the technical concepts of column- and row-based databases and encourage the reader to make use of the new technology in order to accomplish business innovation.

Prof. Michael Franklin (*University of California at Berkeley, USA*)

Hardware technology has evolved rapidly over the past decades, but database system architectures have not kept pace. At the same time, competition is forcing organizations to become more and more data-driven. These developments have driven a re-evaluation of fundamental data management techniques and tradeoffs, leading to innovations that can exploit large memories, parallelism, and a deeper understanding of data management requirements. This book explains the powerful and important changes that are brought about by in-memory data processing. Furthermore, the unique combination of business and technological insights that the authors bring to bear provide lessons that extend beyond any particular technology, serving as a guidebook for innovation in this and future Information Technology revolutions.

Prof. Sam Madden (*Massachusetts Institute of Technology, Cambridge, USA*)

Plattner and Zeier's book is a thorough accounting of the need for, and the design of, main memory database systems. By analyzing technology trends, they make a compelling case for the coming dominance of main-memory in database systems. They go on to identify a series of key design elements that main memory database system should have, including a column-oriented design, support for multi-core processor parallelism, and data compression. They also highlight several important requirements imposed by modern business processes, including heavy use of stored procedures and accounting requirements that drive a need for no-overwrite storage.

This is the first book of its kind, and it provides a complete reference for students and database designers alike.

Prof. Terry Winograd (*Stanford University, California, USA*)

There are moments in the development of computer technology when the ongoing evolution of devices changes the tradeoffs to allow a tectonic shift—a radical change in the way we interact with computers. The personal computer, the Web, and the smart phone are all examples where long-term trends reached a tipping point allowing explosive change and growth. Plattner and Zeier present a vision of how this kind of radical shift is coming to enterprise data management. From Plattner's many years of executive experience and development of data management systems, he is able to see the new space of opportunities for users—the potential for a new kind of software to provide managers with a powerful new tool for gaining insight into the workings of an enterprise. Just as the Web and the modern search engine changed our idea of how, why, and when we "retrieve information," large in-memory databases will change our idea of how to organize and use operational data of every kind in every enterprise. In this visionary and valuable book, Plattner and Zeier lay out the path for the future of business.

Prof. Warren B. Powell (*Princeton University, Princeton, New Jersey, USA*)

In this remarkable book, Plattner and Zeier propose a paradigm shift in memory management for modern information systems. While this offers immediate benefits for the storage and retrieval of images, transaction histories and detailed snapshots of people, equipment and products, it is perhaps even more exciting to think of the opportunities that this technology will create for the future. Imagine the fluid graphical display of spatially distributed, dynamic information. Or the ability to move past the flat summaries of inventories of equipment and customer requests to capture the subtle context that communicates urgency and capability. Even more dramatic, we can envision the real-time optimization of business processes working interactively with domain experts, giving us the information-age equivalent of the robots that make our cars and computers in the physical world today.

Prof. Wolfgang Lehner (*Technical University of Dresden, Germany*)

This book shows in an extraordinary way how technology can drive new applications—a fascinating journey from the core characteristics of business applications to topics of leading-edge main-memory database technology.

Industry

Heiko Hubertz (*CEO and Founder, Bigpoint GmbH, Hamburg, Germany*)

It rarely happens that an emerging new technology can be seen as a "game changer" for a whole industry. But I think this is the case when it comes to in-memory database systems. Tasks which took hours or days can now be executed in seconds which is the technical foundation for establishing completely new business ideas. It dramatically speeds up the execution time of data analytics and thereby creates a much greater experience for the end user on our Bigpoint Gaming Platform. The book "In-Memory Data Management—Technology and Applications" from Hasso Plattner and Alexander Zeier describes not only the technical foundations but also the implications for new exciting applications.

Dr. Ralf Schneider (*CIO, Allianz SE, Munich, Germany*)

Being IT savvy and leveraging advances in Information Technology is the most important competitive advantage in today's business world. I see in-memory technology as described in the book "In-Memory Data Management—Technology and Applications" from Plattner and Zeier as one of the most important innovations in the field of IT. The value for the economy as a whole far outweighs the effort of adapting the applications to this new standard. It dramatically speeds up the execution time of business processes, it allows for significantly more detailed and real-time data analytics and thereby creates a much greater experience for the end user. It brings business computing to a level where it is absolutely snappy and fun to work with.

Bill McDermott (*Co-CEO, SAP AG, Newtown Square, Pennsylvania, USA*)

We are witnessing the dawn of a new era in enterprise business computing, defined by the near instantaneous availability of critical information that will drive faster decision making, new levels of business agility, and incredible personal productivity for business users. With the advent of in-memory technology, the promise of real-time computing is now reality, creating a new inflection point in the role IT plays in driving sustainable business value. In their review of in-memory technology, Hasso Plattner and Alexander Zeier articulate how in-memory technology can drive down costs, accelerate business, help companies reap additional value out of their existing IT investments, and open the door to new possibilities in how business applications can be consumed. This work is a "must read" for anyone who leverages IT innovation for competitive advantage.

Falk F. Strascheg (*Founder and General Partner, EXTOREL GmbH, Munich, Germany*)

Since the advent of the Internet we have been witnessing new technologies coming up quickly and frequently. It is, however, rare that these technologies become innovations in the sense that there are big enough market opportunities. Hasso Plattner has proven his ability to match business needs with technical solutions

more than once, and this time he presents the perhaps most significant innovation he has ever been working on: Real-Time Business powered by In-Memory Computing. As the ability for innovation has always been one of the core factors for competitiveness this is a highly advisable piece of reading for all those who aim to be at the cutting edge.

Gerhard Oswald (*COO, SAP AG, Walldorf, Germany*)

In my role as COO of SAP it is extremely important to react quickly to events and to have instant access to the current state of the business. At SAP, we have already moved a couple of processes to the new in-memory technology described in the book by Hasso Plattner and Alexander Zeier. I'm very excited about the recently achieved improvements utilizing the concepts described in this book. For example, I monitor our customer support messaging system everyday using in-memory technology to make sure that we provide our customers with the timely responses they deserve. I like that this book provides an outlook of how companies can smoothly adopt the new database technology. This transition concept, called the bypass solution, gives our existing customer base the opportunity to benefit from this fascinating technology, even for older releases of SAP software.

Hermann-Josef Lamberti (*COO, Deutsche Bank AG, Frankfurt, Germany*)

Deutsche Bank has run a prototype with an early version of the in-memory technology described in the book by Hasso Plattner and Alexander Zeier. In particular, we were able to speed up the data analysis process to detect cross-selling opportunities in our customer database, from previously 45 min to 5 s. In-memory is a powerful new dimension of applied compute power.

Jim Watson (*Managing General Partner, CMEA Capital, San Francisco, California, USA*)

During the last 50 years, every IT era has brought us a major substantial advancement, ranging from mainframe computers to cloud infrastructures and smart phones. In certain decades the strategic importance of one technology versus the other is dramatically different and it may fundamentally change the way in which people do business. This is what a Venture Capitalist has to bear in mind when identifying new trends that are along for the long haul. In their book, Hasso and Alex not only describe a market-driven innovation from Germany, that has the potential to change the enterprise software market as a whole, but they also present a working prototype.

Martin Petry (*CIO, Hilti AG, Schaan, Liechtenstein*)

Hilti is a very early adopter of the in-memory technology described in the book by Hasso Plattner and Alexander Zeier. Together with SAP, we have worked on developing prototypical new applications using in-memory technology. By merging the transactional world with the analytical world these applications will allow us to gain real-time insight into our operations and allow us to use this insight in our

interaction with customers. The benefit for Hilti applying SAP's in-memory technology is not only seen in a dramatic improvement of reporting execution speed—for example, we were able to speed up a reporting batch job from 3 hours to seconds—but even more in the opportunity to bring the way we work with information and ultimately how we service our customers on a new level.

Prof. Norbert Walter (*former Chief Economist of Deutsche Bank AG, Frankfurt, Germany*)

Imagine you feel hungry. But instead of just opening the fridge (imagine you don't have one) to get hold of, say, some butter and cheese, you would have to leave the house for the nearest dairy farm. Each time you feel hungry. This is what we do today with most company data: We keep them far away from where we process them. In their highly accessible book, Hasso Plattner and Alexander Zeier show how in-memory technology moves data where they belong, promising massive productivity gains for the modern firm. Decision makers, get up to speed!

Paul Polman (*CEO, Unilever PLC, London, UK*)

There are big opportunities right across our value chain to use real-time information more imaginatively. Deeper, real time insight into consumer and shopper behavior will allow us to work even more closely and effectively with our customers, meeting the needs of today's consumers. It will also transform the way in which we serve our customers and consumers and the speed with which we do it. I am therefore very excited about the potential that the in-memory database technology offers to my business.

Tom Greene (*CIO, Colgate-Palmolive Company, New York City, USA*)

In their book, Hasso Plattner and Alexander Zeier not only describe the technical foundations of the new data processing capabilities coming from in-memory, but they also provide examples for new applications that can now be built on top. For a company like Colgate-Palmolive, these new applications are of strategic importance, as they allow for new ways of analyzing our transactional data in real time, which can give us a competitive advantage.

Dr. Vishal Sikka (*CTO, Executive Board Member, SAP AG, Palo Alto, California, USA*)

Hasso Plattner is not only an amazing entrepreneur, he is an incredible teacher. His work and his teaching have inspired two generations of students, leaders, professionals, and entrepreneurs. Over the last 5 years, we have been on a fantastic journey with him, from his early ideas on rethinking our core financials applications, to conceiving and implementing a completely new data management foundation for all our SAP products. This book by Hasso and Alexander, captures these experiences and I encourage everyone in enterprise IT to read this book and take advantage of these learnings, just as I have endeavored to embody these in our products at SAP.

To Annabelle and my family
AZ

Foreword

Is anyone else in the world both as well-qualified as Hasso Plattner to make a strong business case for real-time data analytics and describe the technical details for a solution based on insights in database design for Enterprise Resource Planning that leverage recent hardware technology trends?

The P of SAP has been both the CEO of a major corporation and a Professor of Computer Science at a leading research institute, where he and his colleagues built a working prototype of a main memory database for ERP. Taking advantage of rapid increases in DRAM capacity and in the number of the processors per chip, SanssouciDB demonstrates that the traditional split of separate systems for Online Transaction Processing (OLTP) and for Online Analytical Processing (OLAP) is no longer necessary for enterprise systems.

Business leaders now can ask ad hoc questions of the production transaction database and get the answer back in seconds. With the traditional divided OLTP/OLAP systems, it can take a week to write the query and receive the answer. In addition to showing how software can use concepts from shared nothing databases to scale across blade servers and use concepts from shared everything databases to take advantage of the large memory and many processors inside a single blade, this book touches on the role of Cloud Computing to achieve a single system for transactions and analytics.

Equally as important as the technical achievement, the "Bill Gates of Germany" shows how businesses can integrate this newfound ability to improve the efficiency and profitability of business. Moreover, if this ability is embraced and widely used, perhaps business leaders can quickly and finely adjust enterprise resources to meet rapidly varying demands so that the next economic downturn will not be as devastating to the world's economy as the last one.

Stanford University, CA, USA — Prof. John L. Hennessy
University of California at Berkeley, CA, USA — Prof. David A. Patterson

Preface

Preface to the Second Edition

About one year ago we published the first edition of this book. Within this last year, in-memory technology had such a big impact on the enterprise computing and application market that it truly marked an inflection point. This progress on the one hand, but also the resulting new questions on the other hand, convinced us that it is time for an extended second edition of our book.

The new content in the second edition targets the development and deployment of data-intensive applications that are designed for leveraging the capabilities of in-memory database systems. Among other new content, Sect. 6.1.1 introduces an in-memory application programming model that includes the most important aspects and guidelines for developing in-memory applications. To ease the tasks of application developers and database administrators, we discuss the graphical creation of database views in Sect. 6.1.5. Finally, we also elaborate on new features on application level, e.g., in Sect. 6.2.4 through the combination of data analytics and text search and by presenting two industry case studies in Sect. 9.2.

Of course, we could not have written a second edition of this book in such a short time without the help of our students at our Enterprise Platform and Integration Concepts chair. Therefore, we want to thank them in addition to the acknowledgement in the following preface of the first edition for their hard work and efforts.

Potsdam, 1 March 2012

Hasso Plattner
Alexander Zeier

Preface to the First Edition

We wrote this book because we think that the use of in-memory technology marks an inflection point for enterprise applications. The capacity per dollar and the availability of main memory has increased markedly in the last few years. This has led to a rethinking of how mass data should be stored. Instead of using mechanical disk drives it is now possible to store the primary data copy of a database in silicon-based main memory resulting in an orders of magnitude improvement in performance and allowing completely new applications to be developed. This change in the way data are stored is having, and will continue to have a significant impact on enterprise applications and ultimately on the way businesses are run. Having real-time information available at the speed of thought provides decision makers in an organization with insights that have, until now, not existed.

This book serves the interests of specific reader groups. Generally, the book is intended for anyone who wishes to find out how this fundamental shift in the way data is managed is affecting, and will continue to affect enterprise applications. In particular, we hope that university students, IT professionals, and IT managers, as well as senior management, who wish to create new business processes by leveraging in-memory computing, will find this book inspiring.

The book is divided into three parts:

- Part I gives an overview of our vision of how in-memory technology will change enterprise applications. This part will be of interest to all readers.
- Part II provides a more in-depth description of how we intend to realize our vision, and addresses students and developers, who want a deeper technical understanding of in-memory data management.
- Part III describes the resulting implications on the development and capabilities of enterprise applications, and is suited for technical as well as business-oriented readers.

Writing a book like this always involves more people than just the authors. We would like to thank the members of our Enterprise Platform and Integration Concepts group at the Hasso Plattner Institute at the University of Potsdam in Germany. Anja Bog, Martin Grund, Jens Krüger, Stephan Müller, Jan Schaffner, and Christian Tinnefeld are part of the HANA research group and their work over the last 5 years in the field of in-memory applications is the foundation for our book. Vadym Borovskiy, Thomas Kowark, Ralph Kühne, Martin Lorenz, Jürgen Müller, Oleksandr Panchenko, Matthieu Schapranow, Christian Schwarz, Matthias Uflacker, and Johannes Wust also made significant contributions to the book and our assistant Andrea Lange helped with the necessary coordination. Additionally, writing this book would not have been possible without the help of many colleagues at SAP. Cafer Tosun in his role as the link between HPI and SAP not only coordinates our partnership with SAP, but also actively provided sections for our book. His team members Andrew McCormick-Smith and Christian Mathis added important text passages to the book. We are grateful for the work of Joos-Hendrik Boese, Bernhard Fischer, Enno Folkerts, Andreas Herschel, Sarah Kappes,

Christian Münkel, Frank Renkes, Frederik Transier, and other members of his extended team. We would like to thank Paul Hofmann for his input and for his valuable help in managing our research projects with American Universities. The results we achieved in our research efforts would also not have been possible without the outstanding help of many other colleagues at SAP. We would particularly like to thank Franz Färber and his team for their feedback and their outstanding contributions to our research results over the past years. Many ideas that we describe throughout the book were originally Franz's, and he is also responsible for their implementation within SAP. We especially want to emphasize his efforts.

Finally, we want to express our gratitude to SAP CTO Vishal Sikka for his sponsorship of our research and his personal involvement in our work. In addition, we are grateful to SAP COO Gerhard Oswald and SAP Co-CEOs Jim Hagemann Snabe and Bill McDermott for their ongoing support of our projects.

We encourage you to visit the official website of this book. The website contains updates about the book, reviews, blog entries about in-memory data management, and examination questions for students.

no-disk.com

Potsdam, 1 February 2011

Hasso Plattner
Alexander Zeier

The Essence of In-Memory Data Management

Imagine you live in a major US city. Now, imagine that every time you want a glass of water, instead of getting it from the kitchen, you need to drive to the airport, get on a plane and fly to Germany, and pick up your water there. From the perspective of a modern CPU, accessing data which is in-memory is like getting water from the kitchen. Accessing a piece of data from the computer's hard disk is like flying to Germany for your glass of water. In the past the prohibitive cost of main memory has made the flight to Germany necessary. The last few years, however, have seen a dramatic reduction in the cost per megabyte of main memory, finally making the glass of water in the kitchen a cost effective and much more convenient option.

This orders-of-magnitude difference in access times has profound implications for all enterprise applications. Things that in the past were not even considered because they took so long, now become possible, allowing businesses concrete insight into the workings of their company that previously were the subject of speculation and guess work.

The in-memory revolution that we describe in this book is not simply about putting data into memory and thus being able to work with it "faster". We also show how the convergence of two other major trends in the IT industry: (a) the advent of multi-core CPUs and the necessity of exploiting this parallelism in software and (b) the stalling access latency for DRAM, requiring software to cleverly balance between CPU and memory activity; have to be harnessed to truly exploit the potential performance benefits. Another key aspect of the vision of in-memory data management that we present, is a change in the *way* data are stored in the underlying database. As we will see in the next section, this is of particular relevance for the enterprise applications that are our focus. The power of in-memory data management is in connecting all these dots.

In-Memory Data Management in Combination with Columnar Storage

Our experience has shown us that many enterprise applications work with databases in a similar way. They process large numbers of rows during their execution, but

crucially, only a small number of columns in a table might be of interest in a particular query. The columnar storage model that we describe in this book allows only the required columns to be read while the rest of the table can be ignored. This is in contrast to the more traditional row-oriented model, where all columns of a table—even those that are not necessary for the result—must be accessed.

The columnar storage model also means that the elements of a given column are stored together. This makes the common enterprise operation of aggregation much faster than in a row-oriented model where the data from a given column are stored amongst the other data in the row.

Parallelization Across Multiple Cores and Machines

Single CPU cores are no longer getting any faster but the number of CPU cores is still expected to double every 18 months. This makes exploiting the parallel processing capabilities of multi-core CPUs of central importance to all future software development. As we saw above, in-memory columnar storage places all the data from a given column together in memory making it easy to assign one or more cores to process a single column. This is called vertical fragmentation.

Tables can also be split into sets of rows and distributed to different processors, in a process called horizontal fragmentation. This is particularly important as data volumes continue to grow and have been used with some success to achieve parallelism in data warehousing applications. Both these methods can be applied, not only across multiple cores in a single machine, but across multiple machines in a cluster or in a data center.

Using Compression for Performance and to Save Space

Data compression techniques exploit redundancy within data and knowledge about the data domain. Compression applies particularly well to columnar storage in an enterprise data management scenario, since all data within a column (a) have the same data type and (b) in many cases there are few distinct values, for example in a country column or a status column. In column stores, compression is used for two reasons: to save space and to increase performance.

Efficient use of space is of particular importance to in-memory data management because, even though the cost of main memory has dropped considerably, it is still relatively expensive compared to disk. Due to the compression within the columns, the density of information in relation to the space consumed is increased. As a result more relevant information can be loaded for processing at a time thereby increasing performance. Fewer load actions are necessary in comparison to row storage, where even columns of no relevance to the query are loaded without being used.

Conclusion

In-memory data management is not only a technology but a different way of thinking about software development: we must take fundamental hardware factors

into account, such as access times to main memory versus disk and the potential parallelism that can be achieved with multi-core CPUs. Taking this new world of hardware into account, we must write software that explicitly makes the best possible use of it. On the positive side for developers of enterprise applications, this book lays the technological foundations for a database layer tailored specifically to all these issues. On the negative side, however, the database will not take care of all the issues on its own. Developers must understand the underlying layers of software and hardware to best take advantage of the potential for performance gains. The goal of this book is to help build such understanding.

Contents

Acronyms

ACID	Atomicity, Consistency, Isolation, Durability
ALU	Arithmetic Logic Unit
AMD	Advanced Micro Devices
API	Application Programming Interface
ASP	Application Service Provider
ATP	Available-to-Promise
BI	Business Intelligence
BWA	Business Warehouse Accelerator
CAPEX	Capital Expenditures
CBTR	Combined Benchmark for Transactions and Reporting
CC	Concurrency Control
CID	CommitID
ccNUMA	Cache-Coherent NUMA
CPU	Central Processing Unit
CRM	Customer Relationship Management
DBMS	Database Management System
DRAM	Dynamic Random Access Memory
ELT	Extract, Load and Transform
ERP	Enterprise Resource Planning
ETL	Extract, Transform and Load
FIFO	First In, First Out
FSB	Front Side Bus
I/O	Input/Output
IaaS	Infrastructure-as-a-Service
IMDB	In-Memory Database
IMS	Information Management System
IT	Information Technology
LIFO	Last In, First Out
LRU	Least Recently Used
MDX	Multidimensional Expressions
MIS	Management Information System

MOLAP	Multidimensional OLAP
MVCC	Multiversion Concurrency Control
NUMA	Non-uniform Memory Access
ODS	Operational Data Stores
OLAP	Online Analytical Processing
OLTP	Online Transaction Processing
OPEX	Operational Expenditures
OS	Operating System
PaaS	Platform-as-a-Service
PADD	Parallel Add
PCM	Phase Change Memory
PFOR	Patched Frame-of-Reference
QPI	Quick Path Interconnect
RAM	Random Access Memory
RDBMS	Relational Database Management System
RLE	Run-Length Encoding
ROLAP	Relational OLAP
SaaS	Software-as-a-Service
SCM	Supply Chain Management
SQL	Structured English Query Language
SID	Surrogate Identifier
SIMD	Single Instruction Multiple Data
SPEC	Standard Performance Evaluation Corporation
SQL	Structured Query Language
SRM	Supplier Relationship Management
SSB	Star Schema Benchmark
SSD	Solid State Drive
SSE	Streaming SIMD Extensions
TCO	Total Cost of Ownership
TLB	Translation Lookaside Buffer
TPC	Transaction Processing Performance Council
UMA	Uniform Memory Access
VM	Virtual Machine
XML	Extensible Markup Language

Introduction

Over the last 50 years, advances in Information Technology (IT) have had a significant impact on the success of companies across all industries. The foundation for this success is the successful leveraging of the strengths of IT systems in the rapid and accurate processing of repetitive tasks and the integration of these systems into the business processes of a company. This combination allows a more accurate and complete picture of an organization to be created. Another key aspect to this is that the speed at which IT systems have been able to create this picture means that it can be based on the most recent data available. This aspect has often been described and associated with the term "real-time" as it suggests that every change that happens within a company is instantly visible.

Significant milestones have been reached in the pursuit of this goal throughout the history of enterprise computing. Examples are the development of relational databases and the introduction of SAP's R/3 ERP (Enterprise Resource Planning) system. However, as the sheer volume of data that needs to be processed has increased the milestones have shifted. Currently, most of the data within a company are still distributed throughout a wide range of applications and stored in several disjoint silos. Creating a unified view of these data is a cumbersome and time-consuming process. Additionally, business analysis reports typically do not run directly on operational data, but on aggregated data from a data warehouse. Operational data is transferred into this data warehouse in batch jobs, which makes flexible, ad hoc reporting on the most up-to-date data impossible. As a consequence, company leaders have to make decisions based on data which are either out of data or incomplete. This is obviously not a true "real-time" solution.

We predict that this is about to change as hardware architectures continue the rapid evolution we have seen since the introduction of the microprocessor. This has become especially significant in the last decade where multi-core processors and the availability of large amounts of main memory at low cost are now enabling new breakthroughs in the software industry. It has become possible to store data sets of whole companies entirely in main memory, offering performance that is orders of magnitudes faster than traditional disk-based systems. Hard disks, the only remaining mechanical device in a world of silicon, will soon only be

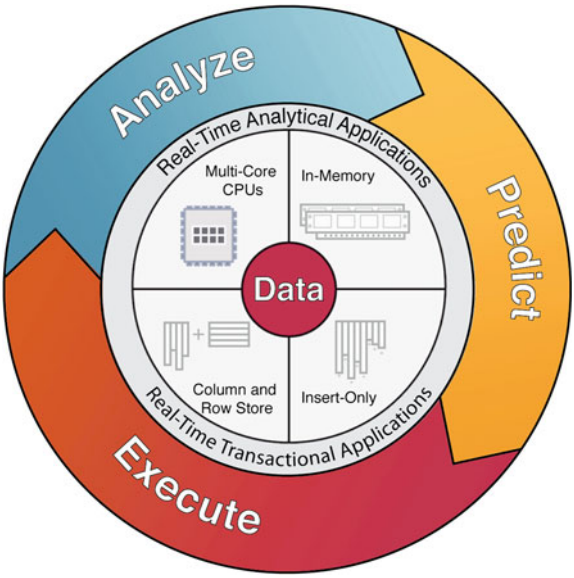

Fig. 1 Enterprise Performance In-memory Circle (EPIC)

necessary for backing up data. With in-memory computing and insert-only databases using row- and column-oriented storage, transactional and analytical processing can be unified. This development has the potential to change how enterprises work and finally offer the promise of real time computing.

As summarized in Fig. 1, the combination of the technologies mentioned above finally enables an iterative link between the instant analysis of data, the prediction of business trends, and the execution of business decisions without delays.

How can companies take advantage of in-memory applications to improve the efficiency and profitability of their business? We predict that this break through innovation will lead to fundamentally improved business processes, better decision-making, and new performance standards for enterprise applications across industries and organizational hierarchies. We are convinced that in-memory technology is a catalyst for innovation, and the enabler for a level of information quality that has not been possible until now. In-memory enterprise data management provides the necessary equipment to excel in a future where businesses face ever-growing demands from customers, partners, and shareholders. With billions of users and a hundred times as many sensors and devices on the Internet, the amount of data we are confronted with is growing exponentially. Being able to quickly extract business-relevant information not only provides unique opportunities for businesses; it will be a critical dfferentiator in future competitive markets.

With in-memory technology, companies can turn the massive amounts of data available into information to create strategic advantage. Operational business data

can be interactively analyzed and queried directly by decision makers, opening up completely new scenarios and opportunities.

Consider the area of financial accounting, where data need to be frequently aggregated for reporting on a daily, weekly, monthly, or annual basis. With in-memory data management, the necessary filtering and aggregation can happen in real time. Accounting can be done anytime and in an ad hoc manner. Financial applications will not only be significantly faster, they will also be less complex and easier to use. Every user of the system will be able to directly analyze massive amounts of data. New data are available for analysis as soon as they are entered into the operational system. Simulations, forecasts, and what-if scenarios can be done on demand, anytime and anywhere. What took days or weeks in traditional disk-based systems can now happen in the blink of an eye. Users of in-memory enterprise systems will be more productive and responsive.

The concepts presented in this book describe new opportunities and areas for improvement across all industries. Below, we present a few examples:

- *Daily Operations*:
  Gain real-time insight into daily revenue, margin, and labor expenses.
- *Competitive Pricing*:
  Intuitively explore the impact of competition on product pricing to instantly understand the impact on profit contribution.
- *Risk Management*:
  Immediately identify high-risk areas across multiple products and services and run what-if scenarios on the fly.
- *Brand and Category Performance*:
  Evaluate the distribution and revenue performance of brands and product categories by customer, region, and channel at any time.
- *Product Lifecycle and Cost Management*:
  Get immediate insight into yield performance versus customer demand.
- *Inventory Management*:
  Optimize inventory and reduce out-of-stock scenarios based on live business events.
- *Financial Asset Management*:
  Gain a more up-to-date picture of financial markets to manage exposure to currencies, equities, derivatives, and other instruments.
- *Real-Time Warranty and Defect Analysis*:
  Get live insight into defective products to identify deviation in production processes or handling.

In summary, we foresee in-memory technology triggering improvements in the following three interrelated strategic areas:

- *Reduced Total Cost of Ownership*:
  The in-memory data management concepts described in this book enable the required analytical capabilities to be directly incorporated into the operational enterprise systems. Dedicated analytical systems can become a thing of the past.

This will allow enterprise systems to become less complex and easier to maintain, resulting in less hardware maintenance and fewer IT resource requirements.

- *Innovative Applications*:
  In-memory data management combines highvolume transaction processing capabilities with analytics, directly in the operational system. Planning, forecasting, pricing optimization, and other processes that previously had to be done on separate analytical systems can be dramatically improved and supported with new applications that were not possible before.
- *Better and Faster Decisions*:
  In-memory enterprise systems allow quick and easy access to information that decision makers need, providing them with new ways to look at the business. Simulation, what-if analyses, and planning can be performed interactively on operational data. Relevant information is instantly accessible. Collaboration within and across organizational units in a company is simplified. This can lead to a much more dynamic management style where problems can be dealt with as they happen.

At the research group, "Enterprise Platform and Integration Concepts", under the supervision of Prof. Dr. Hasso Plattner and Dr. Alexander Zeier at the Hasso Plattner Institute (HPI), we have been working since 2006 on research projects aimed at revolutionizing enterprise systems and applications. Our vision is that in-memory computing will enable completely new ways of operating a business and fulfill the promise of real-time data processing. This book serves to explain in-memory data management and how it is an enabler for this vision.