Contents

Part I An Inflection Point for Enterprise Applications

1 Desirability, Feasibility, Viability: The Impact of In-Memory . . . . 3
  1.1 Information in Real Time: Anything, Anytime, Anywhere . . . . 3
  1.1.1 Response Time at the Speed of Thought . . . . . . . . . . . 4
  1.1.2 Real-Time Analytics and Computation on the Fly . . . . 6
  1.2 The Impact of Recent Hardware Trends . . . . . . . . . . . . . 6
     1.2.1 Database Management Systems for
           Enterprise Applications . . . . . . . . . . . . . . . . . . . . . 7
     1.2.2 Main Memory is the New Disk . . . . . . . . . . . . . . . . 10
     1.2.3 From Maximizing CPU Speed
           to Multi-Core Processors . . . . . . . . . . . . . . . . . . . . 11
     1.2.4 Increased Bandwidth Between CPU
           and Main Memory . . . . . . . . . . . . . . . . . . . . . . . . 13
  1.3 Reducing Cost Through In-Memory Data Management . . . . . 16
     1.3.1 Total Cost of Ownership . . . . . . . . . . . . . . . . . . . . 16
     1.3.2 Cost Factors in Enterprise Systems . . . . . . . . . . . . . 17
     1.3.3 In-Memory Performance Boosts Cost Reduction . . . . 18
  1.4 Conclusion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 19

2 Why Are Enterprise Applications So Diverse? . . . . . . . . . . 21
  2.1 Current Enterprise Applications . . . . . . . . . . . . . . . . . . . . 22
  2.2 Examples of Enterprise Applications . . . . . . . . . . . . . . . . . 23
  2.3 Enterprise Application Architecture . . . . . . . . . . . . . . . . . 25
  2.4 Data Processing in Enterprise Applications . . . . . . . . . . . . . 26
  2.5 Data Access Patterns in Enterprise Applications . . . . . . . . 27
  2.6 Conclusion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 27
3 SanssouciDB: Blueprint for an In-Memory Enterprise

Database System .............................................. 29
3.1 Targeting Multi-Core and Main Memory .................. 30
3.2 Designing an In-Memory Database System ................. 31
3.3 Organizing and Accessing Data in SanssouciDB .......... 33
3.4 Conclusion ................................................. 35

Part II SanssouciDB: A Single Source of Truth Through In-Memory

4 The Technical Foundations of SanssouciDB .................. 39
4.1 Understanding Memory Hierarchies ....................... 40
  4.1.1 Introduction to Main Memory ......................... 40
  4.1.2 Organization of the Memory Hierarchy ............... 44
  4.1.3 Trends in Memory Hierarchies ...................... 45
  4.1.4 Memory from a Programmer’s Point of View .......... 46
4.2 Parallel Data Processing Using Multi-Core and Across Servers. ........................................... 53
  4.2.1 Increasing Capacity by Adding Resources ............ 54
  4.2.2 Parallel System Architectures ....................... 56
  4.2.3 Parallelization in Databases for Enterprise Applications ........................................... 58
  4.2.4 Parallel Data Processing in SanssouciDB .......... 60
4.3 Compression for Speed and Memory Consumption ........... 64
  4.3.1 Light-Weight Compression ............................ 65
  4.3.2 Heavy-Weight Compression ............................ 69
  4.3.3 Data-Dependent Optimization ....................... 70
  4.3.4 Compression-Aware Query Execution .................. 70
  4.3.5 Compression Analysis on Real Data .................. 71
4.4 Column, Row, Hybrid: Optimizing the Data Layout .......... 72
  4.4.1 Vertical Partitioning .................................. 72
  4.4.2 Finding the Best Layout ............................. 76
  4.4.3 Challenges for Hybrid Databases .................... 78
  4.4.4 Application Scenarios ................................ 79
4.5 The Impact of Virtualization ............................... 79
  4.5.1 Virtualizing Analytical Workloads .................... 80
  4.5.2 Data Model and Benchmarking Environment .......... 80
  4.5.3 Virtual Versus Native Execution ..................... 81
  4.5.4 Response Time Degradation with Concurrent VMs ... 82
4.6 Summarizing the Technical Concepts ....................... 84
4.7 Conclusion ................................................. 95
5 Organizing and Accessing Data in SanssouciDB

5.1 SQL for Accessing In-Memory Data

5.1.1 The Role of SQL

5.1.2 The Lifecycle of a Query

5.1.3 Stored Procedures

5.1.4 Data Organization and Indices

5.1.5 Any Attributes as Index

5.2 Increasing Performance with Data Aging

5.2.1 Active and Passive Data

5.2.2 Implementation Considerations for an Aging Process

5.2.3 The Use Case for Horizontal Partitioning of Leads

5.3 Efficient Retrieval of Business Objects

5.3.1 Retrieving Business Data from a Database

5.3.2 Object Data Guide

5.4 Efficient Execution of Business Functions

5.4.1 Separating Business Functions from Application Functions

5.4.2 Comparing Business Functions

5.5 Handling Data Changes in Read-Optimized Databases

5.5.1 The Impact on SanssouciDB

5.5.2 The Merge Process

5.5.3 Improving Performance with Single Column Merge

5.6 Append, Never Delete, to Keep the History Complete

5.6.1 Insert-Only Implementation Strategies

5.6.2 Minimizing Locking Through Insert-Only

5.6.3 The Impact on Enterprise Applications

5.6.4 Feasibility of the Insert-Only Approach

5.7 Enabling Analytics on Transactional Data

5.7.1 Aggregation on the Fly

5.7.2 Analytical Queries without a Star Schema

5.8 Extending Data Layout Without Downtime

5.8.1 Reorganization in a Row Store

5.8.2 On-The-Fly Addition in a Column Store

5.9 Business Resilience Through Advanced Logging Techniques

5.9.1 Recovery in Column Stores

5.9.2 Differential Logging for Row-Oriented Databases

5.9.3 Providing High Availability

5.10 The Importance of Optimal Scheduling for Mixed Workloads

5.10.1 Introduction to Scheduling

5.10.2 Characteristics of a Mixed Workload