

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Motivation	3
1.2	Research Questions and Scope	4
1.3	Outline	5
<b>2</b>	<b>Related Work and Background</b>	<b>7</b>
2.1	Current Computing Hardware Trends	7
2.1.1	Larger and Cheaper Main Memory Capacities	9
2.1.2	Multi-Core Processors and the Memory Wall	10
2.1.3	Switch Fabric Network and Remote Direct Memory Access	12
2.2	In-Memory Database Management Systems	16
2.2.1	Column- and Row-Oriented Data Layout	17
2.2.2	Transactional Versus Analytical Versus Mixed Workload Processing	19
2.2.3	State-of-the-Art In-Memory Database Management Systems	21
2.3	Parallel Database Management Systems	22
2.3.1	Shared-Memory Versus Shared-Disk Versus Shared-Nothing	23
2.3.2	State-of-the-Art Parallel Database Management Systems	26
2.3.3	Database-Aware Storage Systems	29
2.3.4	Operator Placement for Distributed Query Processing	30
2.4	Cloud Storage Systems	34
2.4.1	State-of-the-Art Cloud Storage Systems	38
2.4.2	Combining Database Management and Cloud Storage Systems	39
2.5	Classification	41

## **Part I A Database System Architecture for a Shared Main Memory-Based Storage**

<b>3</b>	<b>System Architecture</b> . . . . .	45
3.1	System Architecture—Requirements, Assumptions, and Overview . . . . .	45
3.2	AnalyticsDB . . . . .	48
3.3	Stanford’s RAMCloud. . . . .	50
<b>4</b>	<b>Data Storage</b> . . . . .	53
4.1	Mapping from Columnar Data to RAMCloud Objects. . . . .	53
4.2	Main Memory Access Costs and Object Sizing . . . . .	54
<b>5</b>	<b>Data Processing</b> . . . . .	63
5.1	Database Operators in AnalyticsDB . . . . .	63
5.2	Operator Push-Down into RAMCloud. . . . .	64
5.3	From SQL Statement to Main Memory Access. . . . .	65

## **Part II Database Operator Execution on a Shared Main Memory-Based Storage**

<b>6</b>	<b>Operator Execution on One Relation</b> . . . . .	69
6.1	Evaluating Operator Execution Strategies. . . . .	72
6.2	Optimizing Operator Execution . . . . .	75
6.3	Implications of Data Partitioning . . . . .	75
<b>7</b>	<b>Operator Execution on Two Relations</b> . . . . .	79
7.1	Grace Join . . . . .	81
7.2	Distributed Block Nested Loop Join . . . . .	83
7.3	Cyclo Join . . . . .	85
7.4	Join Algorithm Comparison . . . . .	87
7.5	Parallel Join Executions. . . . .	90

## **Part III Evaluation**

<b>8</b>	<b>Performance Evaluation</b> . . . . .	95
8.1	Analytical Workload: Star Schema Benchmark. . . . .	95
8.2	Mixed Workload: Point-of-Sales Customer Data. . . . .	98
<b>9</b>	<b>High-Availability Evaluation</b> . . . . .	101
<b>10</b>	<b>Elasticity Evaluation</b> . . . . .	103

**Part IV   Conclusions**

<b>11   Conclusions . . . . .</b>	<b>111</b>
<b>Appendix . . . . .</b>	<b>113</b>
<b>Glossary . . . . .</b>	<b>115</b>
<b>References. . . . .</b>	<b>119</b>

Building a Columnar Database on RAMCloud  
Database Design for the Low-Latency Enabled Data  
Center

Tinnefeld, C.

2016, XIX, 130 p. 37 illus., Hardcover

ISBN: 978-3-319-20710-0