



Pro Serverless Data Handling with Microsoft Azure

Architecting ETL and Data-Driven
Applications in the Cloud

—

Benjamin Kettner
Frank Geisler

Apress®

Pro Serverless Data Handling with Microsoft Azure

**Architecting ETL and Data-Driven
Applications in the Cloud**

**Benjamin Kettner
Frank Geisler**

Apress®

Pro Serverless Data Handling with Microsoft Azure: Architecting ETL and Data-Driven Applications in the Cloud

Benjamin Kettner
Berlin, Germany

Frank Geisler
Lüdinghausen, Germany

ISBN-13 (pbk): 978-1-4842-8066-9
<https://doi.org/10.1007/978-1-4842-8067-6>

ISBN-13 (electronic): 978-1-4842-8067-6

Copyright © 2022 by Benjamin Kettner and Frank Geisler

This work is subject to copyright. All rights are reserved by the publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

Trademarked names, logos, and images may appear in this book. Rather than use a trademark symbol with every occurrence of a trademarked name, logo, or image we use the names, logos, and images only in an editorial fashion and to the benefit of the trademark owner, with no intention of infringement of the trademark.

The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Managing Director, Apress Media LLC: Welmoed Spahr
Acquisitions Editor: Jonathan Gennick
Development Editor: Laura Berendson
Coordinating Editor: Jill Balzano
Copy Editor: April Rondeau

Cover image designed by Freepik (www.freepik.com)

Distributed to the book trade worldwide by Springer Science + Business Media LLC, 1 New York Plaza, Suite 4600, New York, NY 10004. Phone 1-800-SPRINGER, fax (201) 348-4505, email orders-ny@springer-sbm.com, or visit www.springeronline.com. Apress Media, LLC is a California LLC and the sole member (owner) is Springer Science + Business Media Finance Inc (SSBM Finance Inc). SSBM Finance Inc is a **Delaware** corporation.

For information on translations, please email booktranslations@springernature.com; for reprint, paperback, or audio rights, please email bookpermissions@springernature.com.

Apress titles may be purchased in bulk for academic, corporate, or promotional use. eBook versions and licenses are also available for most titles. For more information, reference our Print and eBook Bulk Sales web page at <http://www.apress.com/bulk-sales>.

Any source code or other supplementary material referenced by the author in this book is available to readers on GitHub at <https://github.com/Apress/pro-serverless-data-handling-w-microsoft-azure>.

Printed on acid-free paper

*This book is dedicated to all who—like us—enjoy learning new things.
Stay curious. To our families, friends, and colleagues.*

Table of Contents

About the Authors	xi
About the Technical Reviewer	xiii
Acknowledgments	xv
Introduction	xvii
Part I: The Basics	1
Chapter 1: Azure Basics	3
The Different Cloud Service Models.....	3
Infrastructure as a Service (IaaS)	4
Platform as a Service (PaaS)	4
Software as a Service (SaaS)	5
Cloud Model Responsibilities	5
The Structure of Microsoft Azure	6
Azure Geographies	6
Azure Regions.....	7
Azure Availability Zones.....	8
Azure Account.....	8
Azure Subscription	8
Azure Resource Groups	8
Azure Resource Manager	9
Creating and Naming the Resources	10
Creating Resources	11
Naming Resources	11

TABLE OF CONTENTS

- Overview of Data Services 12
 - Data Categories 12
 - Azure Data Services 13
- Summary..... 16
- Chapter 2: Serverless Computing 17**
 - Cloud Software Delivery..... 17
 - Serverless Delivery 21
 - The Cost of Perfection..... 26
 - Handling Data 28
- Chapter 3: Data-Driven Applications 31**
 - ETL the Classic Way 31
 - Transformation: What Does That Mean? 32
 - Different Data Models for Different Applications 34
 - OLTP: The Relational Model 34
 - OLAP: Star and Snowflake Schemas 41
 - Modern Data Warehouses and Data Applications 44
- Part II: Hands-On 47**
- Chapter 4: Azure Functions 49**
 - The Flavors of Azure Functions..... 49
 - Triggers and Bindings 50
 - Creating Your First Azure Function..... 54
 - Creating the Azure Resources 54
 - Creating the Function 57
 - A Look at the Code 61
 - Testing the Function 63
 - Deploying Your Function..... 67
 - Handling State 70
 - The Basics 71
 - The Code..... 72
 - Running It in the Cloud 76

Chapter 5: Logic Apps.....	79
Principles of Code-Free Implementation.....	80
Creating a Logic App	81
The Logic Apps UI.....	84
Chapter 6: Azure Data Factory.....	93
The Building Blocks of ADF	94
Working with Azure Data Factory.....	95
Creating an ADF Using Azure CLI.....	95
Preparing Resources.....	100
Creating a Pipeline.....	104
Parametrizing Your Pipeline.....	110
Creating a Data Flow.....	115
Best Practices	117
Using Git	117
Using Azure Key Vault.....	119
Chapter 7: Database and Storage Options.....	121
Relational and Non-Relational Data Explained.....	121
Storage Accounts.....	123
Storage Account Basics.....	123
Creating a Storage Account.....	126
Using Azure Table Storage.....	128
Azure Queue Storage.....	131
Cosmos DB.....	135
Use Cases for Cosmos DB Accounts.....	139
Azure SQL DB Serverless	143
Creating a Serverless SQL Database.....	146
When to Choose What?	150

TABLE OF CONTENTS

- Chapter 8: IoT Hub, Event Hub, and Streaming Data..... 153**
 - IoT Hub..... 154
 - Event Hub..... 159
 - Service Bus..... 162
 - Stream Analytics 164
- Chapter 9: Power BI..... 169**
 - Power BI Service and Power BI Desktop..... 170
 - Building Data Visualizations with Power BI Reports 177
 - Visualizing Data Streams 184
 - Sharing Content 186
 - Licensing of Power BI 190
- Part III: Design Practices 193**
- Chapter 10: Achieving Resiliency 195**
 - What Is Resiliency? 195
 - How Is Resiliency Ensured? 199
 - Different Areas to Be Resilient 200
 - Patterns That Support Resiliency 203
 - Choosing the Right Services for Resiliency 207
 - Achieving Resiliency 210
- Chapter 11: Queues, Messages, and Commands 213**
 - Messages..... 213
 - Events 214
 - Commands..... 216
 - Scenarios for Events and Commands 217
 - Implementing the Scenario..... 220
- Chapter 12: Processing Streams of Data..... 231**
 - Streaming Data—What Is It About? 231
 - Stream Processing: Lambda Architecture..... 234

Implementing a Lambda Architecture in Azure	235
There's More	239
Chapter 13: Monitoring Serverless Applications	241
Monitoring and Alerting.....	241
Serverless and Monitoring.....	244
Implementing Monitoring.....	245
Implementing Alerting.....	248
Part IV: Putting It All Together	251
Chapter 14: Tools and Helpers	253
Visual Studio Code	253
Azure Data Studio	254
Docker / Docker Desktop	255
Azure CLI	255
PowerShell	256
Bicep / ARM Templates	256
Azure Storage Explorer	257
Azure DevOps.....	257
dbatools	258
Azure Quickstart Templates	258
Git.....	258
Git Kraken	259
Chocolatey	259
Azure Data Community	260
Useful Visual Studio Code Plugins.....	260
Chapter 15: Data-Loading Patterns	263
Data-Loading Patterns for Flat Files	265
Data-Loading Patterns for REST APIs.....	269
Data-Loading Patterns for Databases.....	272
Data-Loading Patterns for Data Streams	275

TABLE OF CONTENTS

- Chapter 16: Data Storage Patterns 279**
 - Relational Databases 280
 - Storage Accounts 283
 - Non-Relational Databases..... 288
- Chapter 17: Architecture for a Modern Data-Driven Application..... 297**
 - REST API, Tracking & Transaction Data 299
 - Communicating with the Shops 303
 - Data Warehousing and Analytics..... 305
- Index..... 309**