

# Fusion and Integration of Clouds, Edges, and Devices

This book provides an in-depth examination of recent research advances in cloud-edge-end computing, covering theory, technologies, architectures, methods, applications, and future research directions. It aims to present state-of-the-art models and optimization methods for fusing and integrating clouds, edges, and devices.

Cloud-edge-end computing provides users with low-latency, high-reliability, and cost-effective services through the fusion and integration of clouds, edges, and devices. As a result, it is now widely used in various application scenarios. The book introduces the background and fundamental concepts of clouds, edges, and devices, and details the evolution, concepts, enabling technologies, architectures, and implementations of cloud-edge-end computing. It also examines different types of cloud-edge-end orchestrated systems and applications and discusses advanced performance modeling approaches, as well as the latest research on offloading and scheduling policies. It also covers resource management methods for optimizing application performance on cloud-edge-end orchestrated systems.

The intended readers of this book are researchers, undergraduate and graduate students, and engineers interested in cloud computing, edge computing, and the Internet of Things. The knowledge of this book will enrich our readers to be at the forefront of cloud-edge-end computing.

**Junlong Zhou** is an Associate Professor at the School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, China. His research interests include edge computing, cloud computing, and embedded systems.

**Kun Cao** is an Associate Professor at the Jinan University, Guangzhou, China. His research interests include the Internet of Things, edge/fog/cloud computing, and cyber-physical systems.

**Jin Sun** is a Professor at the School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, China. His research interests include computer architecture and high-performance computing.

**Keqin Li** is a SUNY Distinguished Professor at the State University of New York, a National Distinguished Professor at Hunan University, China, and a Member of Academia Europaea (Academician of the Academy of Europe). He is among the world's top five most influential scientists in parallel and distributed computing in terms of single-year and career-long impacts based on a composite indicator of the Scopus citation database. He is an AAAS Fellow, an IEEE Fellow, an AAIA Fellow, an ACIS Fellow, and an AIIA Fellow.



# Fusion and Integration of Clouds, Edges, and Devices

Junlong Zhou, Kun Cao, Jin Sun, and Keqin Li



CRC Press

Taylor & Francis Group

Boca Raton London New York

---

CRC Press is an imprint of the  
Taylor & Francis Group, an **informa** business

This book is funded by the National Natural Science Foundation of China under Grants 62172224 and 62102164, the Natural Science Foundation of Jiangsu Province under Grant BK20220138, the China Postdoctoral Science Foundation under Grants 2021T140272 and 2021M691240, Guangdong Basic and Applied Basic Research Foundation under Grant 2024A1515010232, the Fundamental Research Funds for the Central Universities under Grant 30922010318.

Cover image designed by the authors.

First edition published 2025

by CRC Press

2385 NW Executive Center Drive, Suite 320, Boca Raton FL 33431

and by CRC Press

4 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

*CRC Press is an imprint of Taylor & Francis Group, LLC*

© 2025 Junlong Zhou, Kun Cao, Jin Sun, and Keqin Li

Reasonable efforts have been made to publish reliable data and information, but the author and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged, please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, access [www.copyright.com](http://www.copyright.com) or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. For works that are not available on CCC, please contact [mpkbookspermissions@tandf.co.uk](mailto:mpkbookspermissions@tandf.co.uk)

*Trademark notice:* Product or corporate names may be trademarks or registered trademarks and are used only for identification and explanation without intent to infringe.

*Library of Congress Cataloging-in-Publication Data*

ISBN: 978-1-032-88457-8 (hbk)

ISBN: 978-1-032-88911-5 (pbk)

ISBN: 978-1-003-54028-1 (ebk)

DOI: 10.1201/9781003540281

Typeset in Latin Modern Roman

by KnowledgeWorks Global Ltd.

---

# Contents

---

<b>Preface</b>	<b>vii</b>
<b>Foreword</b>	<b>ix</b>
<b>1 Introduction to Clouds, Edges, and Devices</b>	<b>1</b>
1.1 Cloud Computing . . . . .	1
1.1.1 Understanding Cloud Computing . . . . .	1
1.1.2 Cloud Computing Architectures . . . . .	4
1.1.3 Cloud Service Quality Metrics . . . . .	9
1.1.4 Types of Clouds . . . . .	13
1.1.5 Conclusions and Discussions . . . . .	17
1.2 Edge Computing . . . . .	18
1.2.1 Understanding Edge Computing . . . . .	18
1.2.2 Edge Offloading . . . . .	23
1.2.3 Edge Caching . . . . .	26
1.2.4 Edge Inference . . . . .	28
1.2.5 Edge Training . . . . .	30
1.2.6 Conclusions and Discussions . . . . .	32
1.3 End Device in IoT . . . . .	33
1.3.1 Internet of Things . . . . .	33
1.3.2 Multicore Processors of IoT End Devices . . . . .	36
1.3.3 Performance Metrics of IoT End Devices . . . . .	39
1.3.4 Conclusions and Discussions . . . . .	47
<b>2 Introduction to Cloud-Edge-End Computing</b>	<b>48</b>
2.1 Evolution of Cloud-Edge-End Computing . . . . .	48
2.1.1 Mobile Cloud Computing . . . . .	48
2.1.2 Mobile Edge Computing . . . . .	50
2.1.3 Cloud-Edge Computing . . . . .	52
2.1.4 Conclusions and Discussions . . . . .	58
2.2 Cloud-Edge-End Computing Concepts and Architectures . . . . .	59
2.2.1 Understanding Cloud-Edge-End Computing . . . . .	59
2.2.2 Enabling Technologies for Cloud-Edge-End Computing . . . . .	62
2.2.3 Cloud-Edge-End Computing Architectures . . . . .	65
2.2.4 Conclusions and Discussions . . . . .	68
2.3 Cloud-Edge-End Computing Implementations . . . . .	69
2.3.1 Server Placement in Cloud-Edge-End Architectures . . . . .	69
2.3.2 Data Training in Cloud-Edge-End Architectures . . . . .	73
2.3.3 Resource Management in Cloud-Edge-End Architectures . . . . .	75
2.3.4 Conclusions and Discussions . . . . .	79

- 3 Cloud-Edge-End Orchestrated Systems 80**

  - 3.1 Collaborative Cloud-Edge-End Systems 80
    - 3.1.1 The Motivation of Latency Optimization 80
    - 3.1.2 The Motivation of Energy Optimization 82
    - 3.1.3 The Motivation of Security and Privacy Optimization 84
    - 3.1.4 The Motivation of Reliability Optimization 86
    - 3.1.5 Conclusions and Discussions 87
  - 3.2 Performance Modeling for Cloud-Edge-End Systems 87
    - 3.2.1 Latency of Cloud-Edge-End Systems 87
    - 3.2.2 Energy of Cloud-Edge-End Systems 92
    - 3.2.3 Security and Privacy of Cloud-Edge-End Systems 95
    - 3.2.4 Reliability of Cloud-Edge-End Systems 99
    - 3.2.5 Conclusions and Discussions 102
  - 3.3 Performance Optimization for Cloud-Edge-End Systems 102
    - 3.3.1 Optimization for Latency of Cloud-Edge-End Systems 103
    - 3.3.2 Optimization for Energy of Cloud-Edge-End Systems 109
    - 3.3.3 Optimization for Security and Privacy of Cloud-Edge-End Systems 118
    - 3.3.4 Optimization for Reliability of Cloud-Edge-End Systems 127
    - 3.3.5 Conclusions and Discussions 136

- 4 Cloud-Edge-End Orchestrated Applications 137**

  - 4.1 Cloud-Edge-End Computing for IoT Applications 137
    - 4.1.1 Cloud-Edge-End Computing for Intelligent IoT 138
    - 4.1.2 Cloud-Edge-End Computing for Time-Sensitive IoT 141
    - 4.1.3 Cloud-Edge-End Computing for Internet of Vehicles 143
    - 4.1.4 Conclusions and Discussions 146
  - 4.2 Cloud-Edge-End Computing for CPS Applications 147
    - 4.2.1 Cloud-Edge-End Computing for Mobile CPS 150
    - 4.2.2 Cloud-Edge-End Computing for Medical CPS 151
    - 4.2.3 Cloud-Edge-End Computing for Automotive CPS 154
    - 4.2.4 Cloud-Edge-End Computing for Social CPS 156
    - 4.2.5 Cloud-Edge-End Computing for Agricultural CPS 158
    - 4.2.6 Conclusions and Discussions 159
  - 4.3 Cloud-Edge-End Computing for Smart Cities 160
    - 4.3.1 Cloud-Edge-End Computing for Smart Buildings 160
    - 4.3.2 Cloud-Edge-End Computing for AI Applications in Smart Cities 162
    - 4.3.3 Cloud-Edge-End Computing for Intelligent Transportation Systems 164
    - 4.3.4 Cloud-Edge-End Computing for Smart Grids 167
    - 4.3.5 Cloud-Edge-End Computing Empowering Data Analytics for Smart Cities 169
    - 4.3.6 Conclusions and Discussions 172

- 5 Summary and Future Research 174**

  - 5.1 Summary 174
  - 5.2 Future Research 174

- Bibliography 177**

---

# *Preface*

---

## **Motivation of the Book**

With the rapid advancement of Internet of Things (IoT) technology, the amount of data generated by end devices has grown explosively. The transmission of such massive data would inevitably induce heavy network traffic and consume excessive network resources, resulting in high network latency and even service interruption. Emerging applications such as autonomous driving and augmented reality usually demand low network latency, but the quick growth in the number of end devices has made it challenging for cloud servers to execute a large volume of tasks in time. Although cloud servers have strong computing power, their quantity is limited, and they are commonly located far from end devices, resulting in significant transmission delays. To address this issue, edge computing has emerged, where computation can happen at the network edge. Edge servers are not as powerful as cloud servers, but their number is significantly more, and they are closer to end devices, reducing task response latency. However, complex tasks such as deep neural network training are difficult to execute on edge servers. Therefore, cloud-edge-end orchestration, integrating clouds, edges, and devices, has gained popularity in both academia and industry. This orchestration can effectively improve and balance system performance from many aspects, such as energy consumption, latency, security, and reliability.

## **Summary of Contents**

Chapter 1 introduces the supporting technologies of cloud-edge-end orchestration, including cloud computing, edge computing, and IoT. For each supporting technology, we first present its basic concepts for better understanding. Regarding cloud computing, we also discuss its architectures, service quality metrics, and types. Edge intelligence has emerged as a prominent research topic that integrates edge computing and artificial intelligence (AI). We also study four key research dimensions of edge intelligence: edge offloading, edge caching, edge inference, and edge training. Concerning IoT, we mainly focus on the computing hardware and performance metrics of end devices.

Chapter 2 introduces the evolution of cloud-edge-end computing. It then presents the fundamental concepts, enabling technologies, and hierarchical/horizontal models of cloud-edge-end computing. The chapter finally concludes by discussing the research efforts of cloud-edge-end computing, including service and server placement methods, data analysis and training approaches, and resource management strategies.

Chapter 3 studies the performance models and optimization methods for cloud-edge-end orchestrated systems. The chapter begins with a discussion of the challenges faced in optimizing various aspects of cloud-edge-end orchestrated systems such as latency, energy consumption, security and privacy, and reliability. It then examines the various performance models associated with these systems and how they impact the overall performance. This includes models related to latency, energy consumption, security and privacy, and reliability. Finally, the chapter describes performance optimization methodologies that can improve

the latency, energy consumption, security, and reliability of cloud-edge-end orchestrated systems.

Chapter 4 discusses the design and implementation of cloud-edge-end orchestrated applications in various domains such as IoT, cyber-physical systems (CPS), and smart cities. The IoT scenarios that are covered in this chapter include Intelligent IoT, time-sensitive IoT, and Internet of Vehicles. In addition to this, the CPS scenarios that are discussed include mobile CPS, medical CPS, automotive CPS, social CPS, and agricultural CPS. Furthermore, this chapter also presents various aspects of smart cities such as smart buildings, AI applications in smart cities, intelligent transportation systems, smart grids, and data analytics for smart cities.

Chapter 5 provides a summary of the book and suggests potential avenues for future research.

This book aims to provide a comprehensive introduction to cloud-edge-end computing, covering its background, fundamental concepts, supporting technologies, architectures and implementations, and the latest research. It should be a useful reference for researchers and engineers interested in cloud computing, edge computing, and IoT. We hope readers can find inspiration and value for their own studies.

## Acknowledgments

We wish to express our thanks to students Yufan Shen, Xiangpeng Hou, Lu Yin, Linhua Ma, Tianjian Gong, Yizhou Shi, Xiaoyong Kou, and Tian Wang for their help in collecting related papers, to students Mingzhou Zhao, Sijie Lin, Youling Zeng, Changle Tao, Lan Lan, Jie Niu, Xiaozhu Song, Linhui Wang, and Shilong Zhu for their assistance with drawing pictures, and to students Yingying Zheng, Qixi Yin, Peixin Xu, and Shengchen Cai for their help with editing and proofreading.

This book is funded by the National Natural Science Foundation of China under Grants 62172224 and 62102164, the Natural Science Foundation of Jiangsu Province under Grant BK20220138, the China Postdoctoral Science Foundation under Grants 2021T140272 and 2021M691240, the Guangdong Basic and Applied Basic Research Foundation under Grant 2024A1515010232, and the Fundamental Research Funds for the Central Universities under Grant 30922010318.

Nanjing, Jiangsu, China  
Guangzhou, Guangdong, China  
Nanjing, Jiangsu, China  
New Paltz, NY, USA

Junlong Zhou  
Kun Cao  
Jin Sun  
Keqin Li



---

# *Foreword*

---

Cloud computing has become a dominant computing platform in recent years due to its cost-efficient pay-as-you-go pricing, flexible on-demand resource delivery, and convenient service deployment and maintenance. It offers tremendous advantages to various application domains, such as artificial intelligence, big data analysis, industrial automation, and medical fields. The rapid development of the Internet of Things, cyber-physical systems, embedded computing, and edge computing has led to the creation of a new hybrid computing technology called cloud-edge-end computing. This technology has gained popularity in both academia and industry. Cloud-edge-end computing combines clouds, edges, and devices to fully utilize their different capabilities and improve the quality of services for various application scenarios. It is considered to be the next-generation dominant computing paradigm. However, many challenges must be addressed, such as scalable architecture design, lightweight network virtualization, energy-aware sustainable resource allocation, fault-tolerant computation offloading, and security and privacy protection. These issues have yet to be fully considered in current research work. This book aims to provide solutions to system performance optimization problems and in-depth case studies that demonstrate how to model and optimize the system performance of clouds, edges, and devices for cloud-edge-end computing.

This book first introduces the background and fundamental concepts of clouds, edges, and devices. It then presents the evolution, basic concepts, enabling technologies, architectures, and implementations of cloud-edge-end computing. Cloud-edge-end computing can offer users low-latency, high-reliability, and cost-efficient services by fusion and integration of clouds, edges, and devices. As a result, it has become widely used in various application scenarios. This book further explores different types of cloud-edge-end orchestrated systems and applications, and discusses various advanced performance modeling approaches. The book also contains the latest research on offloading and scheduling policies and resource management methodologies for optimizing the performance of applications on cloud-edge-end orchestrated systems. Finally, the book is rich in content and detailed in graphics. It is indeed a significant contribution to the field of cloud-edge-end computing.

This book is a joint effort and creation of four energetic and enthusiastic scholars in the international distributed computing community. The authors have published extensively on cloud computing, edge computing, and embedded systems in the last few years. They are undoubtedly well-regarded leading scientists in cloud-edge-end computing. I congratulate the authors on an excellent job and look forward to seeing the book in print.

Albert Y. Zomaya  
Peter Nicol Russell Chair Professor of Computer Science  
School of Computer Science  
University of Sydney  
Sydney, NSW, Australia

