Designing Data Intensive Applications The Big Ideas Behind Reliable Scalable And Maintainable Systems

This report examines the links between inequality and other major global trends (or megatrends), with a focus on technological change, climate change, urbanization and international migration. The analysis pays particular attention to poverty and labour market trends, as they mediate the distributional impacts of the major trends selected. It also provides policy recommendations to manage these megatrends in an equitable manner and considers the policy implications, so as to reduce inequalities and support their implementation. This invaluable roadmap for startup engineers reveals how to successfully handle web application scalability challenges to meet increasing product and traffic demands. Web Scalability for Startup Engineers shows engineers working at startups and small companies how to plan and implement a comprehensive scalability strategy. It presents broad and holistic view of infrastructure and architecture of a scalable web application. Successful startups often face the challenge of scalability, and the core concepts driving a scalable architecture are language and platform agnostic. The book covers scalability of HTTP-based systems (websites, REST APIs, SaaS, and mobile application backends), starting with a high-level perspective before taking a deep dive into common challenges and issues. This approach builds a holistic view of the problem, helping you see the

big picture, and then introduces different technologies and best practices for solving the problem at hand. The book is enriched with the author's real-world experience and expert advice, saving you precious time and effort by learning from others' mistakes and successes. Language-agnostic approach addresses universally challenging concepts in Web development/scalability—does not require knowledge of a particular language Fills the gap for engineers in startups and smaller companies who have limited means for getting to the next level in terms of accomplishing scalability Strategies presented help to decrease time to market and increase the efficiency of web applications "This book focuses on the challenges of distributed systems imposed by the data intensive applications, and on the different state-of-the-art solutions proposed to overcome these challenges"--Provided by publisher.

Data-intensive systems are a technological building block supporting Big Data and Data Science applications.This book familiarizes readers with core concepts that they should be aware of before continuing with independent work and the more advanced technical reference literature that dominates the current landscape. The material in the book is structured following a problem-based approach. This means that the content in the chapters is focused on developing solutions to simplified, but still realistic problems using data-intensive technologies and approaches. The reader follows one reference scenario through the whole book, that uses an open Apache dataset. The origins of this volume are in lectures from a master's course

in Data-intensive Systems, given at the University of Stavanger. Some chapters were also a base for guest lectures at Purdue University and Lodz University of Technology.

Data-Intensive Computing

Designing Distributed Systems

Elements of Reusable Object-Oriented Software

Design and Deploy Production-Ready Software

A Guide to Modern Databases and the NoSQL Movement

Beginning Database Design Solutions

Principles and Fundamentals using Hadoop and Spark

Software -- Software Engineering.

A new challenge of learning in work organizations--both in business and public administration--is to master entire life cycles of product, production and business concepts. Meeting this challenge calls--at all levels of the organization--for learning that expand the learners' horizon and practical mastery from individual tasks up to the level of the whole system of the collective activity and its transformation. The Change Laboratory is a method for formative intervention in work communities that supports this kind of organizational learning. It is a path breaker in the area of work place learning due to its strong theoretical and research basis and the way that it integrates the change of organizational practices and individuals' learning. It provides a way to develop practitioners' transformative agency and capacity for creating and implementing new conceptual and practical tools for mastering their joint activity. This first comprehensive presentation of the already

widely used method is written for researchers, consultants, agricultural extension and HRD professionals, as well as practitioners involved in developing activities in their professional field. It explains this novel method as well as its theoretical basis on the Cultural Historical Activity Theory providing also practical examples and tools for carrying out a Change Laboratory intervention. A review is also provided of studies concerning various aspects of expansive learning processes in Change Laboratory interventions.

Architect and design data-intensive applications and, in the process, learn how to collect, process, store, govern, and expose data for a variety of use cases Key Features Integrate the data-intensive approach into your application architecture Create a robust application layout with effective messaging and data querying architecture Enable smooth data flow and make the data of your application intensive and fast Book Description Are you an architect or a developer who looks at your own applications gingerly while browsing through Facebook and applauding it silently for its data-intensive, yet fluent and efficient, behaviour? This book is your gateway to build smart data-intensive systems by incorporating the core data-intensive architectural principles, patterns, and techniques directly into your application architecture. This book starts by taking you through the primary design challenges involved with architecting data-intensive applications. You will learn how to implement data curation and data dissemination, depending on the volume of your data. You will then implement your application architecture one step at a time. You will get to grips

with implementing the correct message delivery protocols and creating a data layer that doesn't fail when running high traffic. This book will show you how you can divide your application into layers, each of which adheres to the single responsibility principle. By the end of this book, you will learn to streamline your thoughts and make the right choice in terms of technologies and architectural principles based on the problem at hand. What you will learn Understand how to envision a data-intensive system Identify and compare the non-functional requirements of a data collection component Understand patterns involving data processing, as well as technologies that help to speed up the development of data processing systems Understand how to implement Data Governance policies at design time using various Open Source Tools Recognize the anti-patterns to avoid while designing a data store for applications Understand the different data dissemination technologies available to query the data in an efficient manner Implement a simple data governance policy that can be extended using Apache Falcon Who this book is for This book is for developers and data architects who have to code, test, deploy, and/or maintain large-scale, high data volume applications. It is also useful for system architects who need to understand various non-functional aspects revolving around Data Intensive Systems.

Designing Data-Intensive ApplicationsThe Big Ideas Behind Reliable, Scalable, and Maintainable Systems"O'Reilly Media, Inc."
Web Scalability for Startup Engineers
Large Scale Data Analytics under the Hood

API Design Patterns

Designing data-intensive Web applications
System Design Interview - An Insider's Guide
Design Patterns for Cloud Native Applications
Morgan Kaufmann series in data management
systems

'The text provides an interesting history of previous and anticipated accomplishments, ending with a chapter on the relationship of fusion power to nuclear weaponry. They conclude on an optimistic note, well worth being understood by the general public.'CHOICEThe gap between the state of fusion energy research and public understanding is vast. In an entertaining and engaging narrative, this popular science book gives readers the basic tools to understand how fusion works, its potential, and contemporary research problems.Written by two young researchers in the field, The Future of Fusion Energy explains how physical laws and the Earth's energy resources motivate the current fusion program — a program that is approaching a critical point. The world's largest science project and biggest ever fusion reactor, ITER, is nearing completion. Its success could trigger a worldwide race to build a power plant, but failure could delay fusion by decades. To these ends, this book details how ITER's results could be used to design an economically competitive power plant as well as some of the many alternative fusion concepts.

Organizations today often struggle to balance business requirements with ever-increasing volumes of data. Additionally, the demand for leveraging large-scale, real-

time data is growing rapidly among the most competitive digital industries. Conventional system architectures may not be up to the task. With this practical guide, you'll learn how to leverage large-scale data usage across the business units in your organization using the principles of event-driven microservices. Author Adam Bellemare takes you through the process of building an event-driven microservice-powered organization. You'll reconsider how data is produced, accessed, and propagated across your organization. Learn powerful yet simple patterns for unlocking the value of this data. Incorporate event-driven design and architectural principles into your own systems. And completely rethink how your organization delivers value by unlocking near-real-time access to data at scale. You'll learn: How to leverage event-driven architectures to deliver exceptional business value The role of microservices in supporting event-driven designs Architectural patterns to ensure success both within and between teams in your organization Application patterns for developing powerful event-driven microservices Components and tooling required to get your microservice ecosystem off the ground When it comes to choosing, using, and maintaining a database, understanding its internals is essential. But with so many distributed databases and tools available today, it's often difficult to understand what each one offers and how they differ. With this practical guide, Alex Petrov guides developers through the concepts behind modern database and storage engine internals. Throughout the

book, you'll explore relevant material gleaned from numerous books, papers, blog posts, and the source code of several open source databases. These resources are listed at the end of parts one and two. You'll discover that the most significant distinctions among many modern databases reside in subsystems that determine how storage is organized and how data is distributed. This book examines: Storage engines: Explore storage classification and taxonomy, and dive into B-Tree-based and immutable Log Structured storage engines, with differences and use-cases for each Storage building blocks: Learn how database files are organized to build efficient storage, using auxiliary data structures such as Page Cache, Buffer Pool and Write-Ahead Log Distributed systems: Learn step-by-step how nodes and processes connect and build complex communication patterns Database clusters: Which consistency models are commonly used by modern databases and how distributed storage systems achieve consistency Architect scalable, reliable, and maintainable applications for enterprises with Python Key FeaturesExplore various Python design patterns used for enterprise software developmentApply best practices for testing and performance optimization to build stable applicationsLearn about different attacking strategies used on enterprise applications and how to avoid themBook Description Dynamically typed languages like Python are continuously improving. With the addition of exciting new features and a wide selection of modern

libraries and frameworks, Python has emerged as an ideal language for developing enterprise applications. Hands-On Enterprise Application Development with Python will show you how to build effective applications that are stable, secure, and easily scalable. The book is a detailed guide to building an end-to-end enterprise-grade application in Python. You will learn how to effectively implement Python features and design patterns that will positively impact your application lifecycle. The book also covers advanced concurrency techniques that will help you build a RESTful application with an optimized frontend. Given that security and stability are the foundation for an enterprise application, you'll be trained on effective testing, performance analysis, and security practices, and understand how to embed them in your codebase during the initial phase. You'll also be guided in how to move on from a monolithic architecture to one that is service oriented, leveraging microservices and serverless deployment techniques. By the end of the book, you will have become proficient at building efficient enterprise applications in Python. What you will learnUnderstand the purpose of design patterns and their impact on application lifecycleBuild applications that can handle large amounts of data-intensive operationsUncover advanced concurrency techniques and discover how to handle a large number of requests in productionOptimize frontends to improve the client-side experience of your applicationEffective testing and performance profiling techniques to detect issues in

applications early in the development cycleBuild applications with a focus on securityImplement large applications as microservices to improve scalabilityWho this book is for If you're a developer who wants to build enterprise-grade applications, this book is for you. Basic to intermediate-level of programming experience with Python and database systems is required to understand the concepts covered in this book.

Designing Fine-Grained Systems

The Big Ideas Behind Reliable, Scalable, and Maintainable Systems

Patterns and Paradigms for Scalable, Reliable Services

Database Internals

Data-Intensive Science

Release It!

Foundations of Data Intensive Applications

Data-intensive systems are software applications that process and generate Big Data. Data-intensive systems support the use of large amounts of data strategically and efficiently to provide intelligence. For example, examining industrial sensor data or business process data can enhance production, guide proactive improvements of development processes, or optimize supply chain systems. Designing data-intensive software systems is difficult because distribution of knowledge across stakeholders creates a symmetry of ignorance, because a shared vision of the future requires the development of new knowledge that extends and synthesizes existing knowledge. Knowledge Management

in the Development of Data-Intensive Systems addresses new challenges arising from knowledge management in the development of data-intensive software systems. These challenges concern requirements, architectural design, detailed design, implementation and maintenance. The book covers the current state and future directions of knowledge management in development of data-intensive software systems. The book features both academic and industrial contributions which discuss the role software engineering can play for addressing challenges that confront developing, maintaining and evolving systems;data-intensive software systems of cloud and mobile services; and the scalability requirements they imply. The book features software engineering approaches that can efficiently deal with data-intensive systems as well as applications and use cases benefiting from data-intensive systems. Providing a comprehensive reference on the notion of data-intensive systems from a technical and non-technical perspective, the book focuses uniquely on software engineering and knowledge management in the design and maintenance of data-intensive systems. The book covers constructing, deploying, and maintaining high quality software products and software engineering in and for dynamic and flexible environments. This book provides a holistic guide for those who need to understand the impact of variability on all aspects of the software life cycle. It leverages practical experience and evidence to look

ahead at the challenges faced by organizations in a fast-moving world with increasingly fast-changing customer requirements and expectations.

Data is at the center of many challenges in system design today. Difficult issues need to be figured out, such as scalability, consistency, reliability, efficiency, and maintainability. In addition, we have an overwhelming variety of tools, including relational databases, NoSQL datastores, stream or batch processors, and message brokers. What are the right choices for your application? How do you make sense of all these buzzwords? In this practical and comprehensive guide, author Martin Kleppmann helps you navigate this diverse landscape by examining the pros and cons of various technologies for processing and storing data. Software keeps changing, but the fundamental principles remain the same. With this book, software engineers and architects will learn how to apply those ideas in practice, and how to make full use of data in modern applications. Peer under the hood of the systems you already use, and learn how to use and operate them more effectively Make informed decisions by identifying the strengths and weaknesses of different tools Navigate the trade-offs around consistency, scalability, fault tolerance, and complexity Understand the distributed systems research upon which modern databases are built Peek behind the scenes of major online services, and learn from their architectures
Antennas and propagation are of fundamental

importance to the coverage, capacity and quality of all wireless communication systems. This book provides a solid grounding in antennas and propagation, covering terrestrial and satellite radio systems in both mobile and fixed contexts. Building on the highly successful first edition, this fully updated text features significant new material and brand new exercises and supplementary materials to support course tutors. A vital source of information for practising and aspiring wireless communication engineers as well as for students at postgraduate and senior undergraduate levels, this book provides a fundamental grounding in the principles of antennas and propagation without excessive recourse to mathematics. It also equips the reader with practical prediction techniques for the design and analysis of a very wide range of common wireless communication systems. Including: Overview of the fundamental electromagnetic principles underlying propagation and antennas. Basic concepts of antennas and their application to specific wireless systems. Propagation measurement, modelling and prediction for fixed links, macrocells, microcells, picocells and megacells Narrowband and wideband channel modelling and the effect of the channel on communication system performance. Methods that overcome and transform channel impairments to enhance performance using diversity, adaptive antennas and equalisers. Key second edition updates: New chapters on Antennas for Mobile Systems and Channel Measurements for Mobile Radio

Systems. Coverage of new technologies, including MIMO antenna systems, Ultra Wideband (UWB) and the OFDM technology used in Wi-Fi and WiMax systems. Many new propagation models for macrocells, microcells and picocells. Fully revised and expanded end-of-chapter exercises. The Solutions Manual can be requested from http://www.wiley.com/go/saunders_antennas_2e In the race to compete in today's fast-moving markets, large enterprises are busy adopting new technologies for creating new products, processes, and business models. But one obstacle on the road to digital transformation is placing too much emphasis on technology, and not enough on the types of processes technology enables. What if different lines of business could build their own services and applications—and decision-making was distributed rather than centralized? This report explores the concept of a digital business platform as a way of empowering individual business sectors to act on data in real time. Much innovation in a digital enterprise will increasingly happen at the edge, whether it involves business users (from marketers to data scientists) or IoT devices. To facilitate the process, your core IT team can provide these sectors with the digital tools they need to innovate quickly. This report explores: Key cultural and organizational changes for developing business capabilities through cross-functional product teams A platform for integrating applications, data sources, business partners, clients, mobile apps, social networks,

and IoT devices Creating internal API programs for building innovative edge services in low-code or no-code environments Tools including Integration Platform as a Service, Application Platform as a Service, and Integration Software as a Service The challenge of integrating microservices and serverless architectures Event-driven architectures for processing and reacting to events in real time You'll also learn about a complete pervasive integration solution as a core component of a digital business platform to serve every audience in your organization.

Architectures, Algorithms, and Applications

Building Microservices

Normal Forms and All That Jazz

The What, Where, When, and How of Large-Scale Data Processing

Seven Databases in Seven Weeks

Data-intensive Text Processing with MapReduce

Antennas and Propagation for Wireless Communication Systems

A single dramatic software failure can cost a company millions of dollars - but can be avoided with simple changes to design and architecture. This new edition of the best-selling industry standard shows you how to create systems that run longer, with fewer failures, and recover better when bad things happen. New coverage includes DevOps, microservices, and cloud-native architecture. Stability antipatterns have grown to include systemic problems in large-scale systems. This is a must-have pragmatic guide to engineering for production

systems. If you're a software developer, and you don't want to get alerts every night for the rest of your life, help is here. With a combination of case studies about huge losses - lost revenue, lost reputation, lost time, lost opportunity - and practical, down-to-earth advice that was all gained through painful experience, this book helps you avoid the pitfalls that cost companies millions of dollars in downtime and reputation. Eighty percent of project life-cycle cost is in production, yet few books address this topic. This updated edition deals with the production of today's systems - larger, more complex, and heavily virtualized - and includes information on chaos engineering, the discipline of applying randomness and deliberate stress to reveal systematic problems. Build systems that survive the real world, avoid downtime, implement zero-downtime upgrades and continuous delivery, and make cloud-native applications resilient. Examine ways to architect, design, and build software - particularly distributed systems - that stands up to the typhoon winds of a flash mob, a Slashdotting, or a link on Reddit. Take a hard look at software that failed the test and find ways to make sure your software survives. To skip the pain and get the experience...get this book.

The system design interview is considered to be the most complex and most difficult technical job interview by many. Those questions are intimidating, but don't worry. It's just that nobody has taken the time to prepare you systematically. We take the time. We go slow. We draw lots of diagrams and use lots of examples. You'll learn step-by-step, one question at a time.Don't miss out.What's inside?- An insider's take on what interviewers really look for and why.- A 4-step framework for solving any system design interview question.- 16 real system

design interview questions with detailed solutions.- 188 diagrams to visually explain how different systems work. Every enterprise application creates data, whether it⊠s log messages, metrics, user activity, outgoing messages, or something else. And how to move all of this data becomes nearly as important as the data itself. If you⊠re an application architect, developer, or production engineer new to Apache Kafka, this practical guide shows you how to use this open source streaming platform to handle real-time data feeds. Engineers from Confluent and LinkedIn who are responsible for developing Kafka explain how to deploy production Kafka clusters, write reliable event-driven microservices, and build scalable stream-processing applications with this platform. Through detailed examples, you⊠ll learn Kafka⊠s design principles, reliability guarantees, key APIs, and architecture details, including the replication protocol, the controller, and the storage layer. Understand publish-subscribe messaging and how it fits in the big data ecosystem. Explore Kafka producers and consumers for writing and reading messages Understand Kafka patterns and use-case requirements to ensure reliable data delivery Get best practices for building data pipelines and applications with Kafka Manage Kafka in production, and learn to perform monitoring, tuning, and maintenance tasks Learn the most critical metrics among Kafka⊠s operational measurements Explore how Kafka⊠s stream delivery capabilities make it a perfect source for stream processing systems

Today, software engineers need to know not only how to program effectively but also how to develop proper engineering practices to make their codebase sustainable and healthy. This book emphasizes this difference between

programming and software engineering. How can software engineers manage a living codebase that evolves and responds to changing requirements and demands over the length of its life? Based on their experience at Google, software engineers Titus Winters and Hyrum Wright, along with technical writer Tom Manshreck, present a candid and insightful look at how some of the world's leading practitioners construct and maintain software. This book covers Google's unique engineering culture, processes, and tools and how these aspects contribute to the effectiveness of an engineering organization. You'll explore three fundamental principles that software organizations should keep in mind when designing, architecting, writing, and maintaining code: How time affects the sustainability of software and how to make your code resilient over time How scale affects the viability of software practices within an engineering organization What trade-offs a typical engineer needs to make when evaluating design and development decisions

The Future Of Fusion Energy

Knowledge Management in the Development of Data-Intensive Systems

Designing Data-Intensive Applications

Software Engineering at Google

Designing Cloud Data Platforms

Data Pipelines Pocket Reference

Monolith to Microservices

Our world is being revolutionized by data-driven methods: access to large amounts of data has generated new insights and opened exciting new opportunities in commerce, science, and computing applications. Processing the enormous quantities of

data necessary for these advances requires large clusters, making distributed computing paradigms more crucial than ever. MapReduce is a programming model for expressing distributed computations on massive datasets and an execution framework for large-scale data processing on clusters of commodity servers. The programming model provides an easy-to-understand abstraction for designing scalable algorithms, while the execution framework transparently handles many system-level details, ranging from scheduling to synchronization to fault tolerance. This book focuses on MapReduce algorithm design, with an emphasis on text processing algorithms common in natural language processing, information retrieval, and machine learning. We introduce the notion of MapReduce design patterns, which represent general reusable solutions to commonly occurring problems across a variety of problem domains. This book not only intends to help the reader "think in MapReduce", but also discusses limitations of the programming model as well. This volume is a printed version of a work that appears in the Synthesis Digital Library of Engineering and Computer Science. Synthesis Lectures provide concise, original presentations of important research and development topics, published quickly, in digital and print formats. For more information visit www.morganclaypool.com Salary surveys worldwide regularly place software

architect in the top 10 best jobs, yet no real guide exists to help developers become architects. Until now. This book provides the first comprehensive overview of software architecture's many aspects. Aspiring and existing architects alike will examine architectural characteristics, architectural patterns, component determination, diagramming and presenting architecture, evolutionary architecture, and many other topics. Mark Richards and Neal Ford—hands-on practitioners who have taught software architecture classes professionally for years—focus on architecture principles that apply across all technology stacks. You'll explore software architecture in a modern light, taking into account all the innovations of the past decade. This book examines: Architecture patterns: The technical basis for many architectural decisions Components: Identification, coupling, cohesion, partitioning, and granularity Soft skills: Effective team management, meetings, negotiation, presentations, and more Modernity: Engineering practices and operational approaches that have changed radically in the past few years Architecture as an engineering discipline: Repeatable results, metrics, and concrete valuations that add rigor to software architecture With the immense cost savings and scalability the cloud provides, the rationale for building cloud native applications is no longer in question. The real issue is how. With this practical guide, developers will learn

about the most commonly used design patterns for building cloud native applications using APIs, data, events, and streams in both greenfield and brownfield development. You'll learn how to incrementally design, develop, and deploy large and effective cloud native applications that you can manage and maintain at scale with minimal cost, time, and effort. Authors Kasun Indrasiri and Sriskandarajah Suhothayan highlight use cases that effectively demonstrate the challenges you might encounter at each step. Learn the fundamentals of cloud native applications Explore key cloud native communication, connectivity, and composition patterns Learn decentralized data management techniques Use event-driven architecture to build distributed and scalable cloud native applications Explore the most commonly used patterns for API management and consumption Examine some of the tools and technologies you'll need for building cloud native systems

Describes principles of the emerging field of data-intensive computing, along with methods for designing, managing and analyzing the big data sets of today.

Fundamentals of Software Architecture
World Social Report 2020
An Engineering Approach
Evolutionary Patterns to Transform Your Monolith
Second Edition

Streaming Systems

A Tool for Collaborative Development of Work and Education

PEEK "UNDER THE HOOD" OF BIG DATA ANALYTICS The world of big data analytics grows ever more complex. And while many people can work superficially with specific frameworks, far fewer understand the fundamental principles of large-scale, distributed data processing systems and how they operate. In Foundations of Data Intensive Applications: Large Scale Data Analytics under the Hood, renowned big-data experts and computer scientists Drs. Supun Kamburugamuve and Saliya Ekanayake deliver a practical guide to applying the principles of big data to software development for optimal performance. The authors discuss foundational components of large-scale data systems and walk readers through the major software design decisions that define performance, application type, and usability. You???ll learn how to recognize problems in your applications resulting in performance and distributed operation issues, diagnose them, and effectively eliminate them by relying on the bedrock big data principles explained within. Moving beyond individual frameworks and APIs for data processing, this book unlocks the theoretical ideas that operate under the hood of every big data processing system. Ideal for data scientists, data architects, dev-ops engineers, and

developers, Foundations of Data Intensive Applications: Large Scale Data Analytics under the Hood shows readers how to: Identify the foundations of large-scale, distributed data processing systems Make major software design decisions that optimize performance Diagnose performance problems and distributed operation issues Understand state-of-the-art research in big data Explain and use the major big data frameworks and understand what underpins them Use big data analytics in the real world to solve practical problems

"A concept-rich book on API design patterns. Deeply engrossing and fun to read." - Satej Sahu, Honeywell API Design Patterns lays out a set of design principles for building internal and public-facing APIs. In API Design Patterns you will learn: Guiding principles for API patterns Fundamentals of resource layout and naming Handling data types for any programming language Standard methods that ensure predictability Field masks for targeted partial updates Authentication and validation methods for secure APIs Collective operations for moving, managing, and deleting data Advanced patterns for special interactions and data transformations API Design Patterns reveals best practices for building stable, user-friendly APIs. These design patterns can be applied to solve common API problems and flexibly altered to fit specific needs. Hands-on examples and relevant cases illustrate patterns for

API fundamentals, advanced functionalities, and uncommon scenarios. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the technology APIs are contracts that define how applications, services, and components communicate. API design patterns provide a shared set of best practices, specifications and standards that ensure APIs are reliable and simple for other developers. This book collects and explains the most important patterns from both the API design community and the experts at Google. About the book API Design Patterns lays out a set of principles for building internal and public-facing APIs. Google API expert JJ Geewax presents patterns that ensure your APIs are consistent, scalable, and flexible. You'll improve the design of the most common APIs, plus discover techniques for tricky edge cases. Precise illustrations, relevant examples, and detailed scenarios make every pattern clear and easy to understand. What's inside Guiding principles for API patterns Fundamentals of resource layout and naming Advanced patterns for special interactions and data transformations A detailed case-study on building an API and adding features About the reader For developers building web and internal APIs in any language. About the author JJ Geewax is a software engineer at Google, focusing on Google Cloud Platform, API design, and real-time

payment systems. He is also the author of Manning's Google Cloud Platform in Action. Table of Contents PART 1 INTRODUCTION 1 Introduction to APIs 2 Introduction to API design patterns PART 2 DESIGN PRINCIPLES 3 Naming 4 Resource scope and hierarchy 5 Data types and defaults PART 3 FUNDAMENTALS 6 Resource identification 7 Standard methods 8 Partial updates and retrievals 9 Custom methods 10 Long-running operations 11 Rerunnable jobs PART 4 RESOURCE RELATIONSHIPS 12 Singleton sub-resources 13 Cross references 14 Association resources 15 Add and remove custom methods 16 Polymorphism PART 5 COLLECTIVE OPERATIONS 17 Copy and move 18 Batch operations 19 Criteria-based deletion 20 Anonymous writes 21 Pagination 22 Filtering 23 Importing and exporting PART 6 SAFETY AND SECURITY 24 Versioning and compatibility 25 Soft deletion 26 Request deduplication 27 Request validation 28 Resource revisions 29 Request retrial 30 Request authentication

How do you detangle a monolithic system and migrate it to a microservice architecture? How do you do it while maintaining business-as-usual? As a companion to Sam Newman's extremely popular Building Microservices, this new book details a proven method for transitioning an existing monolithic system to a microservice architecture. With many illustrative examples, insightful migration

patterns, and a bevy of practical advice to transition your monolith enterprise into a microservice operation, this practical guide covers multiple scenarios and strategies for a successful migration, from initial planning all the way through application and database decomposition. You'll learn several tried and tested patterns and techniques that you can use as you migrate your existing architecture. Ideal for organizations looking to transition to microservices, rather than rebuild Helps companies determine whether to migrate, when to migrate, and where to begin Addresses communication, integration, and the migration of legacy systems Discusses multiple migration patterns and where they apply Provides database migration examples, along with synchronization strategies Explores application decomposition, including several architectural refactoring patterns Delves into details of database decomposition, including the impact of breaking referential and transactional integrity, new failure modes, and more

Streaming data is a big deal in big data these days. As more and more businesses seek to tame the massive unbounded data sets that pervade our world, streaming systems have finally reached a level of maturity sufficient for mainstream adoption. With this practical guide, data engineers, data scientists, and developers will learn how to work with streaming data in a conceptual and platform-

agnostic way. Expanded from Tyler Akidau's popular blog posts "Streaming 101" and "Streaming 102", this book takes you from an introductory level to a nuanced understanding of the what, where, when, and how of processing real-time data streams. You'll also dive deep into watermarks and exactly-once processing with co-authors Slava Chernyak and Reuven Lax. You'll explore: How streaming and batch data processing patterns compare The core principles and concepts behind robust out-of-order data processing How watermarks track progress and completeness in infinite datasets How exactly-once data processing techniques ensure correctness How the concepts of streams and tables form the foundations of both batch and streaming data processing The practical motivations behind a powerful persistent state mechanism, driven by a real-world example How time-varying relations provide a link between stream processing and the world of SQL and relational algebra
Inequality in a Rapidly Changing World
Database Design and Implementation
The Change Laboratory
Data Intensive Distributed Computing: Challenges and Solutions for Large-scale Information Management
Design Patterns
2nd Edition
A Deep Dive into How Distributed Data Systems

In Designing Cloud Data Platforms, Danil Zburivsky and Lynda Partner reveal a six-layer approach that increases flexibility and reduces costs. Discover patterns for ingesting data from a variety of sources, then learn to harness pre-built services provided by cloud vendors. Summary Centralized data warehouses, the long-time defacto standard for housing data for analytics, are rapidly giving way to multi-faceted cloud data platforms. Companies that embrace modern cloud data platforms benefit from an integrated view of their business using all of their data and can take advantage of advanced analytic practices to drive predictions and as yet unimagined data services. Designing Cloud Data Platforms is a hands-on guide to envisioning and designing a modern scalable data platform that takes full advantage of the flexibility of the cloud. As you read, you'll learn the core components of a cloud data platform design, along with the role of key technologies like Spark and Kafka Streams. You'll also explore setting up processes to manage cloud-based data, keep it secure, and using advanced analytic and BI tools to analyze it. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the technology Well-designed pipelines, storage systems, and APIs eliminate the complicated scaling and maintenance required with on-prem data centers. Once you learn the patterns for

designing cloud data platforms, you'll maximize performance no matter which cloud vendor you use. About the book In Designing Cloud Data Platforms, Danil Zburivsky and Lynda Partner reveal a six-layer approach that increases flexibility and reduces costs. Discover patterns for ingesting data from a variety of sources, then learn to harness pre-built services provided by cloud vendors. What's inside Best practices for structured and unstructured data sets Cloud-ready machine learning tools Metadata and real-time analytics Defensive architecture, access, and security About the reader For data professionals familiar with the basics of cloud computing, and Hadoop or Spark. About the author Danil Zburivsky has over 10 years of experience designing and supporting large-scale data infrastructure for enterprises across the globe. Lynda Partner is the VP of Analytics-as-a-Service at Pythian, and has been on the business side of data for over 20 years. Table of Contents 1 Introducing the data platform 2 Why a data platform and not just a data warehouse 3 Getting bigger and leveraging the Big 3: Amazon, Microsoft Azure, and Google 4 Getting data into the platform 5 Organizing and processing data 6 Real-time data processing and analytics 7 Metadata layer architecture 8 Schema management 9 Data access and security 10 Fueling business value with data platforms

There are no easy decisions in software architecture. Instead, there are many hard parts--difficult problems or issues with no best

practices--that force you to choose among various compromises. With this book, you'll learn how to think critically about the trade-offs involved with distributed architectures. Architecture veterans and practicing consultants Neal Ford, Mark Richards, Pramod Sadalage, and Zhamak Dehghani discuss strategies for choosing an appropriate architecture. By interweaving a story about a fictional group of technology professionals--the Sysops Squad--they examine everything from how to determine service granularity, manage workflows and orchestration, manage and decouple contracts, and manage distributed transactions to how to optimize operational characteristics, such as scalability, elasticity, and performance. By focusing on commonly asked questions, this book provides techniques to help you discover and weigh the trade-offs as you confront the issues you face as an architect. Analyze trade-offs and effectively document your decisions Make better decisions regarding service granularity Understand the complexities of breaking apart monolithic applications Manage and decouple contracts between services Handle data in a highly distributed architecture Learn patterns to manage workflow and transactions when breaking apart applications Annotation Over the past 10 years, distributed systems have become more fine-grained. From the large multi-million line long monolithic applications, we are now seeing the benefits of smaller self-contained services. Rather than

heavy-weight, hard to change Service Oriented Architectures, we are now seeing systems consisting of collaborating microservices. Easier to change, deploy, and if required retire, organizations which are in the right position to take advantage of them are yielding significant benefits. This book takes an holistic view of the things you need to be cognizant of in order to pull this off. It covers just enough understanding of technology, architecture, operations and organization to show you how to move towards finer-grained systems. The vast majority of software applications use relational databases that virtually every application developer must work with. This book introduces you to database design, whether you're a DBA or database developer. You'll discover what databases are, their goals, and why proper design is necessary to achieve those goals. Additionally, you'll master how to structure the database so it gives good performance while minimizing the chance for error. You will learn how to decide what should be in a database to meet the application's requirements.

Real-Time Data and Stream Processing at Scale

Hands-On Enterprise Application Development with Python

Building Event-Driven Microservices

Data-intensive Systems

Database Design and Relational Theory

Software Architecture: The Hard Parts

Data is getting bigger and more complex by the day, and so

are your choices in handling it. Explore some of the most cutting-edge databases available - from a traditional relational database to newer NoSQL approaches - and make informed decisions about challenging data storage problems. This is the only comprehensive guide to the world of NoSQL databases, with in-depth practical and conceptual introductions to seven different technologies: Redis, Neo4J, CouchDB, MongoDB, HBase, Postgres, and DynamoDB. This second edition includes a new chapter on DynamoDB and updated content for each chapter. While relational databases such as MySQL remain as relevant as ever, the alternative, NoSQL paradigm has opened up new horizons in performance and scalability and changed the way we approach data-centric problems. This book presents the essential concepts behind each database alongside hands-on examples that make each technology come alive. With each database, tackle a real-world problem that highlights the concepts and features that make it shine. Along the way, explore five database models - relational, key/value, columnar, document, and graph - from the perspective of challenges faced by real applications. Learn how MongoDB and CouchDB are strikingly different, make your applications faster with Redis and more connected with Neo4J, build a cluster of HBase servers using cloud services such as Amazon's Elastic MapReduce, and more. This new edition brings a brand new chapter on DynamoDB, updated code samples and exercises, and a more up-to-date account of each database's feature set. Whether you're a programmer building the next big thing, a data scientist seeking solutions to thorny problems, or a technology enthusiast venturing into new territory, you will find something to inspire you in this book. What You Need: You'll need a *nix shell (Mac OS or Linux preferred, Windows users will need Cygwin), Java 6 (or greater), and Ruby 1.8.7 (or greater). Each chapter will list the

downloads required for that database.

This text represents a breakthrough in the process underlying the design of the increasingly common and important data-driven Web applications.

This textbook examines database systems from the viewpoint of a software developer. This perspective makes it possible to investigate why database systems are the way they are. It is of course important to be able to write queries, but it is equally important to know how they are processed. We e.g. don't want to just use JDBC; we also want to know why the API contains the classes and methods that it does. We need a sense of how hard is it to write a disk cache or logging facility. And what exactly is a database driver, anyway? The first two chapters provide a brief overview of database systems and their use. Chapter 1 discusses the purpose and features of a database system and introduces the Derby and SimpleDB systems. Chapter 2 explains how to write a database application using Java. It presents the basics of JDBC, which is the fundamental API for Java programs that interact with a database. In turn, Chapters 3-11 examine the internals of a typical database engine. Each chapter covers a different database component, starting with the lowest level of abstraction (the disk and file manager) and ending with the highest (the JDBC client interface); further, the respective chapter explains the main issues concerning the component, and considers possible design decisions. As a result, the reader can see exactly what services each component provides and how it interacts with the other components in the system. By the end of this part, s/he will have witnessed the gradual development of a simple but completely functional system. The remaining four chapters then focus on efficient query processing, and focus on the sophisticated techniques and algorithms that can replace the simple design choices described earlier. Topics include indexing, sorting,

intelligent buffer usage, and query optimization. This text is intended for upper-level undergraduate or beginning graduate courses in Computer Science. It assumes that the reader is comfortable with basic Java programming; advanced Java concepts (such as RMI and JDBC) are fully explained in the text. The respective chapters are complemented by "end-of-chapter readings" that discuss interesting ideas and research directions that went unmentioned in the text, and provide references to relevant web pages, research articles, reference manuals, and books. Conceptual and programming exercises are also included at the end of each chapter. Students can apply their conceptual knowledge by examining the SimpleDB (a simple but fully functional database system created by the author and provided online) code and modifying it.

Because databases often stay in production for decades, careful design is critical to making the database serve the needs of your users over years, and to avoid subtle errors or performance problems. In this book, C.J. Date, a leading exponent of relational databases, lays out the principles of good database design.

Design data-intensive Application with Python 3
Kafka: The Definitive Guide
Understanding Distributed Systems
Lessons Learned from Programming Over Time
Architecting Data Intensive Applications
Challenges and Solutions for Large-scale Information Management

Learning to build distributed systems is hard, especially if they are large scale. It's not that there is a lack of information out there. You can find academic papers, engineering blogs, and even books on the subject. The problem is that the

available information is spread out all over the place, and if you were to put it on a spectrum from theory to practice, you would find a lot of material at the two ends, but not much in the middle. That is why I decided to write a book to teach the fundamentals of distributed systems so that you don't have to spend countless hours scratching your head to understand how everything fits together. This is the guide I wished existed when I first started out, and it's based on my experience building large distributed systems that scale to millions of requests per second and billions of devices. If you develop the back-end of web or mobile applications (or would like to!), this book is for you. When building distributed systems, you need to be familiar with the network stack, data consistency models, scalability and reliability patterns, and much more. Although you can build applications without knowing any of that, you will end up spending hours debugging and re-designing their architecture, learning lessons that you could have acquired in a much faster and less painful way.

Data pipelines are the foundation for success in data analytics. Moving data from numerous diverse sources and transforming it to provide context is the difference between having data and actually gaining value from it. This pocket reference defines data pipelines and explains how they work in

today's modern data stack. You'll learn common considerations and key decision points when implementing pipelines, such as batch versus streaming data ingestion and build versus buy. This book addresses the most common decisions made by data professionals and discusses foundational concepts that apply to open source frameworks, commercial products, and homegrown solutions. You'll learn: What a data pipeline is and how it works How data is moved and processed on modern data infrastructure, including cloud platforms Common tools and products used by data engineers to build pipelines How pipelines support analytics and reporting needs Considerations for pipeline maintenance, testing, and alerting

Data-intensive science has the potential to transform scientific research and quickly translate scientific progress into complete solutions, policies, and economic success. But this collaborative science is still lacking the effective access and exchange of knowledge among scientists, researchers, and policy makers across a range of disciplines. Br