pingcap / awesome-database-learning Public

A list of learning materials to understand databases internals

⭐ 5.5k stars   🍴 723 forks

☆ Star   ▾

🔔 Notifications

<> Code   ⊙ Issues 2   ⑈ Pull requests 3   ⊘ Security   ⬈ Insights

⑂ master ▾

SahilTalrejacoder Added a new resource (#51) …   Sep 2, 2022   🕑 48

View code

≡ README.md

🔗 Awesome Database Learning

A list of learning materials to understand databases internals, including but not limited to:

- papers
- blogs
- courses
- talks

Please submit a pull request if there is any material that you think should be included in this collection.

🔗 Table of Contents

- Recommended Courses, Books and Talks
  - Courses
  - Books
  - Talks
  - Blogs
- SQL & Relation Algebra

🔗 Recommended Courses, Books and Talks

🔗 Courses

- CMU Database Systems (15-445/645), thanks to Andy Pavlo
- CMU Advanced Database Systems (15-721), thanks to Andy Pavlo
- UC Berkeley Introduction to Database Systems
- Stanford Database System Implementation
- Cornell Introduction to Database Systems by Prof. Trummer
- Let's Build a Simple Database, thanks to cstack

🔗 Books

- Stanford Database Systems: The Complete Book
- Designing Data-Intensive Applications, 中文翻译
- Database Internals
- Foundations of Databases
- Readings in Database Systems, 5th Edition
- Database Design and Implementation: Second Edition (Data-Centric Systems and Applications)
- Principles of Distributed Database Systems, 4th ed
- Inside SQLite
- Architecture of a Database System
- Relational Database Index Design and the Optimizers
- Transactional Information Systems: Theory, Algorithms, and the Practice of Concurrency Control

🔗 Talks

- Data Structures and Algorithms for Big Databases
- A Journey From A Quick HackTo A High-Reliability Database Engine

🔗 Blogs

- How does a relational database work
- The Internals of PostgreSQL
- Books propose
- what is database and its types

🔗 SQL & Relation Algebra

Courses:

- CMU Database Systems (15-445/645), thanks to Andy Pavlo

    - Course Introduction and the Relational Model

    - Advanced SQL

- UC Berkeley Introduction to Database Systems

    - Introduction + SQL I
    - SQL II
    - Relational Algebra

🔗 Query Optimizer

Courses:

- CMU Database Systems (15-445/645), thanks to Andy Pavlo
    - Query Planning & Optimization I
    - Query Planning & Optimization II

Blogs:

- 数据库内核杂谈, thanks to 顾仲贤
    - 数据库内核杂谈（七）：数据库优化器（上）
    - 数据库内核杂谈（八）：数据库优化器（下）
- SQL优化器原理 - 查询优化器综述, thanks to 勿烦

🔗 Planner Models

Blogs:

- 数据库内核杂谈, thanks to 顾仲贤
    - 数据库内核杂谈（九）：开源优化器 ORCA
- SQL 查询优化原理与 Volcano Optimizer 介绍, thanks to 张茄子
- Cascades Optimizer, thanks to hellocode

Papers:

- 1979, Access Path Selection in a Relational Database Management System, SIGMOD
- 1979, Query Processing in Main Memory Database Management Systems, VLDB
- 1987, Query Optimization by Simulated Annealing, SIGMOD
- 1988, Grammar-like Functional Rules for Representing Query Optimization Alternatives, SIGMOD
- 1993, The Volcano Optimizer Generator- Extensibility and Efficient Search, ICDE
- 1995, The Cascades Framework for Query Optimization, IEEE Data engineering

Bulltin

- 1998, [An Overview of Query Optimization in Relational Systems](), PODS
- 2001, [LEO – DB2's LEarning Optimizer](), VLDB
- 2004, [Robust Query Processing through Progressive Optimization](), SIGMOD
- 2014, [Orca: A Modular Query Optimizer Architecture for Big Data](), SIGMOD
- 2016, [Parallelizing Query Optimization on Shared-Nothing Architectures](), VLDB
- 2016, [The MemSQL Query Optimizer: A modern optimizer for real-time analytics in a distributed database](), VLDB

🔗 Subquery Optimization

Blogs:

- [SQL 子查询的优化](), thanks to [Eric Fu]()
- [Calcite 子查询处理 - I (RemoveSubQuery)](), thanks to [一只无情的小猫咪]()
- [Calcite 子查询处理 - II (Decorrelate)](), thanks to [一只无情的小猫咪]()

Papers:

- 2001, [Orthogonal Optimization of Subqueries and Aggregation](), SIGMOD
- 2009, [Enhanced subquery optimizations in Oracle](), VLDB
- 2015, [Unnesting Arbitrary Queries](), BTW

🔗 Join Order Optimization

Papers:

- 2006, [Analysis of Two Existing and One New Dynamic Programming Algorithm for the Generation of Optimal Bushy Join Trees without Cross Products](), VLDB
- 2015, [How Good Are Query Optimizers, Really?](), VLDB
- 2018, [Adaptive Optimization of Very Large Join Queries](), SIGMOD

🔗 Functional Dependency & Physical Properties

Thesis:

- 2000, [Exploiting Functional Dependence in Query Optimization]()

Papers:

- 1996, [Fundamental Techniques for Order Optimization](), SIGMOD
- 2004, [An Efficient Framework for Order Optimization](), ICDE
- 2010, [Incorporating Partitioning and Parallel Plans into the SCOPE Optimizer](), ICDE

🔗 Cost Model

Papers:

- 1996, [Modelling Costs for a MM-DBMS](#), in Real-Time Databases
- 2014, [Approximation Schemes for Many-Objective Query Optimization](#), SIGMOD
- 2015, [Multi-Objective Parametric Query Optimization](#), VLDB

🔗 Statistics

Papers:

- 1984, [Accurate Estimation of the Number of Tuples Satisfying a Condition](#), SIGMOD
- 1993, [Optimal Histograms for Limiting Worst-Case Error Propagation in the Size of Join Results](#), ACM Trans. on Database Systems
- 1993, [Universality of Serial Histograms](#), VLDB
- 1995, [Balancing Histogram Optimality and Practicality for Query Result Size Estimation](#), SIGMOD
- 1996, [Improved Histograms for Selectivity Estimation of Range Predicates](#), SIGMOD
- 1997, [SEEKing the truth about ad hoc join costs](#), VLDB
- 2000, [Towards Estimation Error Guarantees for Distinct Values](#), SIGMOD/PODS
- 2001, [Distinct Sampling for Highly-Accurate Answers to Distinct Values Queries and Event Reports](#), VLDB
- 2003, [The History of Histograms](#), VLDB
- 2005, [An Improved Data Stream Summary: The Count-Min Sketch and its Applications](#), Journal of Algorithms
- 2007, [New Estimation Algorithms for Streaming Data: Count-min Can Do More](#)
- 2009, [Preventing Bad Plans by Bounding the Impact of Cardinality Estimation Errors](#), VLDB
- 2010, [Histograms Reloaded: The Merits of Bucket Diversity](#), SIGMOD
- 2014, [Exploiting Ordered Dictionaries to Efficiently Construct Histograms with Q-Error Guarantees in SAP HANA](#), SIGMOD
- 2017, [Adaptive Statistics in Oracle 12c](#), VLDB
- 2019, [Pessimistic Cardinality Estimation: Tighter Upper Bounds for Intermediate Join Cardinalities](#), SIGMOD
- 2019, [Deep Unsupervised Cardinality Estimation](#), VLDB
- 2020, [NeuroCard: One Cardinality Estimator for All Tables](#), VLDB

Books:

- [Synopses for Massive Data: Samples, Histograms, Wavelets, Sketches](#)

🔗 Query Execution

Courses:

- CMU Database Systems (15-445/645), thanks to Andy Pavlo
  - Query Execution I
  - Query Execution II

🔗 Execution Framework

Papers:

- 1994, Volcano-An Extensible and Parallel Query Evaluation System, IEEE Transactions on Knowledge and Data EngineeringFebruary
- 2014, Morsel-Driven Parallelism: A NUMA-Aware Query Evaluation Framework for the Many-Core Age, SIGMOD

🔗 Vectorization vs Compilization

Blogs:

- Overhead of a Generalized Query Execution Engine, from The Pivotal Engineering Journal, thanks to the Pivotal Engineering team

Papers:

- 2005, MonetDB/X100: Hyper-Pipelining Query Execution, CIDR
- 2011, Efficiently Compiling Efficient Query Plans for Modern Hardware, VLDB
- 2017, Relaxed Operator Fusion for In-Memory Databases: Making Compilation, Vectorization, and Prefetching Work Together At Last, VLDB
- 2018, Everything You Always Wanted to Know About Compiled and Vectorized Queries But Were Afraid to Ask, VLDB
- 2018, Adaptive Execution of Compiled Queries, ICDE

🔗 Join

Papers:

- 2013, Multi-Core, Main-Memory Joins: Sort vs. Hash Revisited, VLDB
- 2017, Looking Ahead Makes Query Plans Robust, VLDB

🔗 Hash Table

Courses:

- CMU Database Systems (15-445/645), thanks to Andy Pavlo
  - Hash Tables

Blogs:

- Fibonacci Hashing: The Optimization that the World Forgot (or: a Better Alternative to Integer Modulo), thanks to Malte Skarupke
- All hash table sizes you will ever need, thanks to Database Architects - Thomas Neumann

🔗 Bloom Filter

Papers:

- 2018, SuRF: Practical Range Query Filtering with Fast Succinct Tries, SIGMOD

🔗 DDL

- 2013, Online, Asynchronous Schema Change in F1, VLDB

🔗 Relational Model

Blogs:

- What is a Relational Database?, thanks to Caleb Curry
- What is a Relational Database?,thank to JAMES GALLAGHER

🔗 Codd's Rules

Blogs:

- Codd's Rules for Relational Database Systems, thanks to Kevin Kline

🔗 Relational Data Model

Blogs:

- Relational model, thanks to Wikipedia

🔗 Relational Algebra

Blogs:

- Introduction of Relational Algebra in DBMS, thanks to GeeksforGeeks

🔗 ER to Relational Model

Blogs:

- ER Model to Relational Model, thanks to tutorialspoint

🔗 SQL - Overview

Blogs:

- [An Overview of SQL Text Functions](#), thanks to [Zahin Rahman](#)

🔗 Transaction

🔗 Isolation Levels

Blogs:

- [一致性模型](#), thanks to [siddontang](#)

Papers:

- 1995, [A Critique of ANSI SQL Isolation Levels](#), SIGMOD
- 2000, [Generalized Isolation Level Definitions](#), Proceedings of 16th International Conference on Data Engineering

🔗 Concurrency Control

Courses:

- CMU [Database Systems (15-445/645)](#), thanks to [Andy Pavlo](#)

  - [Concurrency Control Theory](#)
  - [Two-Phase Locking Concurrency Control](#)
  - [Timestamp Ordering Concurrency Control](#)
  - [Multi-Version Concurrency Control](#)

- CMU [Advanced Database Systems (15-721)](#), thanks to [Andy Pavlo](#)

  - [Multi-Version Concurrency Control (Design Decisions)](#)
  - [Multi-Version Concurrency Control (Protocols)](#)
  - [Multi-Version Concurrency Control (Garbage Collection)](#)

Papers:

- 1976, [The Notions of Consistency and Predicate Locks in a Database System](#), Communications of the ACM
- 1981, [Concurrency Control in Distributed Database Systems](#), ACM Computing Surveys
- 1981, [On Optimistic Methods for Concurrency Control](#), ACM Transactions on Database Systems
- 1983, [Multiversion Concurrency Control - Theory and Algorithms](#), ACM Transactions on Database Systems

- 2012, Serializable Snapshot Isolation in PostgreSQL, VLDB
- 2012, Calvin: Fast Distributed Transactions for Partitioned Database Systems, SIGMOD
- 2014, MaaT: effective and scalable coordination of distributed transactions in the cloud, VLDB
- 2014, Staring into the Abyss: An Evaluation of Concurrency Control with One Thousand Cores, VLDB
- 2014, An Evaluation of the Advantages and Disadvantages of Deterministic Database Systems, VLDB
- 2015, Fast Serializable Multi-Version Concurrency Control for Main-Memory Database Systems, SIGMOD
- 2017, An Empirical Evaluation of In-Memory Multi-Version Concurrency Control, VLDB
- 2017, An Evaluation of Distributed Concurrency Control, VLDB
- 2019, Scalable Garbage Collection for In-Memory MVCC Systems, VLDB

🔗 Network

Courses:

- CMU Advanced Database Systems (15-721), thanks to Andy Pavlo
  - Networking Protocols

Papers:

- 2016, The End of Slow Networks: It's Time for a Redesign, VLDB
- 2016, Accelerating Relational Databases by Leveraging Remote Memory and RDMA, SIGMOD
- 2017, Don't Hold My Data Hostage: A Case for Client Protocol Redesign, VLDB

🔗 Storage

🔗 NoSQL Systems

Papers:

- 2006, Bigtable: A Distributed Storage System for Structured Data, OSDI
- 2007, Dynamo: Amazon's Highly Available Key-value Store, SOSP
- 2008, PNUTS: Yahoo!'s Hosted Data Serving Platform, VLDB
- 2010, Cassandra - A Decentralized Structured Storage System, SOSP
- 2019, PNUTS to Sherpa: Lessons from Yahoo!'s Cloud Database , VLDB

🔗 Buffer Management

Courses:

- CMU Database Systems (15-445/645), thanks to Andy Pavlo
  - Buffer Pools

Papers:

- 1987, The 5 Minute Rule for Trading Memory for Disc Accesses and the 5 Byte Rule for Trading Memory for CPU Time, SIGMOD
- 2008, The Five Minute Rule 20 Years Later and How Flash Memory Changes the Rules, ACM Queue
- 2018, Managing Non-Volatile Memory in Database Systems, SIGMOD
- 2018, LeanStore: In-Memory Data Management Beyond Main Memory, ICDE
- 2020, Umbra: A Disk-Based System with In-Memory Performance, CIDR

🔗 Disk IO

Blogs:

- On Disk IO, Part 1: Flavors of IO, thanks to Alex
- On Disk IO, Part 2: More Flavours of IO, thanks to Alex
- On Disk IO, Part 3: LSM Trees, thanks to Alex
- On Disk IO, Part 4: B-Trees and RUM Conjecture, thanks to Alex
- On Disk IO, Part 5: Access Patterns in LSM Trees, thanks to Alex
- Ensuring data reaches disk(LWN)
- Read, write & space amplification - pick 2, thanks to Mark Callaghan

Papers:

- 2016, Design Tradeoffs of Data Access Methods, SIGMOD
- 2016, Designing Access Methods: The RUM Conjecture, EDBT

🔗 B-Tree

Blogs:

- B树、B+树索引算法原理（上） thanks to codedump
- B树、B+树索引算法原理（下）

Courses:

- CMU Database Systems (15-445/645), thanks to Andy Pavlo

  - Trees Indexes I
  - Trees Indexes II

- CMU [Advanced Database Systems (15-721)](), thanks to [Andy Pavlo]()

  - [OLTP Indexes (B+Tree Data Structures)]()

Papers:

- 1979, [The Ubiquitous B-Tree]()

🔗 LSM-Tree

Papers:

- 1996, [The Log-Structured Merge-Tree (LSM-Tree),]()
- 2014, [A Comparison of Fractal Trees to Log-Structured Merge (LSM) Trees]()
- 2017, [WiscKey: Separating Keys from Values in SSD-conscious Storage](), TOS
- 2019, [LSM-based Storage Techniques: A Survey]()

🔗 Learned Indexes Structures

Papers:

- 2018, [The Case for Learned Index Structures]()
- 2019, [Learning Multi-dimensional Indexes]()
- 2020, [XIndex: A Scalable Learned Index for Multicore Data Storage]()
- 2020, [RadixSpline: A Single-Pass Learned Index](), [Source Code](), aiDM@SIGMOD
- 2020, [The PGM-index: a fully-dynamic compressed learned index with provable worst-case bounds](), [Source Code](), VLDB
- 2020, [From WiscKey to Bourbon: A Learned Index for Log-Structured Merge Trees]()

🔗 Serializing & RPC

- [Protocol Buffers Developer Guide]()
- [gRPC Documntation]()

🔗 Data Partitioning

Blogs:

- [TiDB Internal (I) - Data Storage]()
- [Partitioning Behavior of DynamoDB](), thanks to [Parth Modi]()

Papers:

- 2007, [Dynamo: Amazon's Highly Available Key-value Store](), SOSP

🔗 Replication & Consistency

Blogs:

- [Tick or Tock? Keeping Time and Order in Distributed Databases](), thanks to [Liu Tang]()

Papers:

- 2012, [Consistency Tradeoffs in Modern Distributed Database System Design]()
- 2020, [Strong and Efficient Consistency with Consistency-Aware Durability]() , FAST 2020

🔗 Consensus

Technical report:

- University of Cambridge [Distributed consensus revised](), a great paper about Consenssus especially Paxos and Paxos-Related algorithms, by Heidi Howard

Papers:

- 2014, [Ark: A Real-World Consensus Implementation](), CoRR

🔗 Scheduling

Blogs:

- [Building a Large-scale Distributed Storage System Based on Raft](), by Ed Huang

Papers:

- 2016, [Automated Demand-driven Resource Scaling in Relational Database-as-a-Service](), SIGMOD
- 2019, [Autoscaling Tiered Cloud Storage in Anna](), VLDB
- 2020, [Adaptive HTAP through Elastic Resource Scheduling](), SIGMOD
- 2020, [MorphoSys: Automatic Physical Design Metamorphosis for Distributed Database Systems](), VLDB

🔗 Benchmark & Testing

Blogs:

- [Use go-ycsb to benchmark different databases (1)](), thanks to [siddontang]()
- [Chaos Tools and Techniques for Testing the TiDB Distributed NewSQL Database](), thanks to [Liu Tang]()
- [Creating Custom Sysbench Scripts](), thanks to [Matthew Boehm]()

Papers:

- 2010, Benchmarking Cloud Serving Systems with YCSB, SOCC

🔗 HTAP

Papers:

- 2020, TiDB: A Raft-based HTAP Database, VLDB
- 2020, F1 Lightning: HTAP as a Service, VLDB
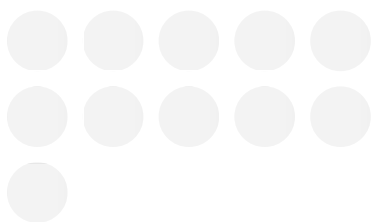
🔗 TLA+

Talks:

- The TLA+ Video Course

Releases

No releases published

Packages

No packages published

Contributors 24

+ 13 contributors