

数据库开发学习资料

必读课程

- CMU [Database Systems \(15-445/645\)](#), Andy Pavlo 的数据库入门课程
- CMU [Advanced Database Systems \(15-721\)](#), Andy Pavlo 的数据库进阶课程
 - 课程视频均在 YouTube 上: <https://www.youtube.com/@CMUDatabaseGroup>
 - 课程的配套实验代码: <https://github.com/cmu-db/bustub>
- 伯克利大学 [Introduction to Database Systems](#)
- 斯坦福大学 [Database System Implementation](#)
- 康纳尔大学 [Introduction to Database Systems](#)
- 斯坦福大学 [MIT.6824 分布式系统](#)

实战教程

- [Let's Build A Simple Database](#), 构建一个极简的数据库, C 语言实现
- [PingCAP TinyKV](#), 构建一个分布式 KV 存储系统, Go 语言实现
- [PingCAP TinySQL](#), 构建一个分布式数据库, Go 语言实现
- [从零开始写时序数据库](#), Go 语言实现
- OceanBase 的 [miniob 教程](#), C++ 实现
- [200 行代码实现 Paxos KV 存储](#), Go 语言实现
- [关系型数据库从 0 到 1](#), 基于 Java 的简易数据库
- [从零实现极简的 bitcask KV 存储引擎](#), Go 语言实现
- [mini-lsm](#), 迷你 LSM Tree 存储引擎, Rust 语言实现
- [go-sqlldb](#), Go 语言实现的简单的关系型数据库
- [NYADB2](#), Go 语言实现的简单数据库, 用于学习
- [nessDB](#), 事务型 KV 存储, 基于 Fractal-Tree, C 语言实现

书籍

- 斯坦福大学数据库教程: [Database Systems: The Complete Book](#)
- [数据密集型应用系统设计 \(DDIA\)](#)
- [数据库系统内幕](#)
- [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](#), SQLite 内幕
- [Architecture of a Database System](#), 数据库系统架构
- [Relational Database Index Design and the Optimizers](#)
- [Transactional Information Systems](#)

博客/专栏

- [分布式和存储的那些事](#)
 - [CatKang 的博客](#)
 - [CodingHusky 的博客](#)
 - [Codedump 的网络日志](#)
 - [数据库内核月报](#)
 - [db-readings](#), 关于数据库的一些论文
 - [PostgreSQL 内核系列文章](#)
 - [PingCAP 官方博客](#)
 - [数据库内核杂谈](#)
 - [木鸟杂记](#)
 - [虎哥的博客](#)
 - [数据系统论文阅读小组](#)
 - [Presto 专栏](#)
 - [ClickHouse 分享 PPT](#)
 - [PostgreSQL 数据库学习](#)
-

SQL 简介

- CMU 数据库课程 [Database Systems \(15-445/645\)](#)
 - [Course Introduction and the Relational Model](#)
 - [Advanced SQL](#)
- UC Berkeley 数据库课程 [Introduction to Database Systems](#)
 - [Introduction + SQL I](#)
 - [SQL II](#)
 - [Relational Algebra](#)
- [SQL Overview](#), learn SQL 网站
- [SQL 语法教程](#), w3schools 教程

关系模型

博客

- [What is a relational database](#), by Oracle
- <https://www.ibm.com/topics/relational-databases>, by IBM
- <https://careerkarma.com/blog/relational-database>
- [Relation Model in DBMS](#), by Geeks for Geeks
- [ER Model to Relation Model](#)

维基百科

- [什么是关系型数据库](#)

优化器

课程

- CMU Database Systems (15-445/645)
 - [Query Planning & Optimization I](#)
 - [Query Planning & Optimization II](#)

博客

- [数据库内核杂谈](#)
 - [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）](#)
 - [数据库内核杂谈（九）：开源优化器 ORCA](#)
- [SQL优化器原理 – 查询优化器综述](#)
- [深入浅出查询优化器](#)
- [学习数据库优化器如何入手](#), 知乎 henry liang
- [优化器技术论文学习](#)

Planner Models

博客

- [数据库内核杂谈](#)
- [SQL 查询优化原理与 Volcano Optimizer 介绍](#), by 张茄子
- [Cascades Optimizer](#)

论文

- 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

博客

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

论文

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

Join Order Optimization

论文

- 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

论文

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

论文

- 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

论文

- 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

书籍

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

执行引擎

课程

- CMU 数据库 [Introduction to Database Systems \(15–445/645\)](#), by [Andy Pavlo](#)
 - [Query Execution I](#)
 - [Query Execution II](#)

Execution Framework

博客

- [数据库查询引擎的进化之路](#), OceanBase
- [物理执行引擎之火山引擎](#)

论文

- 1994, [Volcano—An Extensible and Parallel Query Evaluation System](#), IEEE Transactions on Knowledge and Data Engineering February
- 2014, [Morsel-Driven Parallelism: A NUMA-Aware Query Evaluation Framework for the Many-Core Age](#), SIGMOD

Vectorization vs Compilation

博客

- [Overhead of a Generalized Query Execution Engine](#), from [The Pivotal Engineering Journal](#)

论文

- 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

论文

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

Hash Table

课程

- CMU [Database Systems \(15-445/645\)](#), by Andy Pavlo
 - [Hash Tables](#)

博客

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

Bloom Filter

论文

- 2018, [SuRF: Practical Range Query Filtering with Fast Succinct Tries](#), SIGMOD

事务

隔离级别

博客

- [一致性模型](#), by siddontang
- [Understanding Isolation Levels in a Database Transaction](#)
- [浅析数据库事务的隔离性](#)
- [MySQL 的事务隔离级别和实现原理](#)
- [数据库内核杂谈](#), by 顾仲贤
 - [事务、隔离、并发 \(1\)](#)
 - [事务、隔离、并发 \(2\)](#)
 - [事务、隔离、并发 \(3\)](#)

论文

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

并发控制

课程

- CMU 数据库 [Database Systems \(15-445/645\)](#), by Andy Pavlo
 - [Concurrency Control Theory](#)
 - [Two-Phase Locking Concurrency Control](#)
 - [Timestamp Ordering Concurrency Control](#)
 - [Multi-Version Concurrency Control](#)
- CMU 数据库进阶 [Advanced Database Systems \(15-721\)](#), by Andy Pavlo
 - [Multi-Version Concurrency Control \(Design Decisions\)](#)
 - [Multi-Version Concurrency Control \(Protocols\)](#)
 - [Multi-Version Concurrency Control \(Garbage Collection\)](#)

论文

- 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

网络

课程

- CMU 数据库进阶 [Advanced Database Systems \(15–721\)](#), by Andy Pavlo
 - [Networking Protocols](#)

论文

- 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

存储

NoSQL 系统

书籍

- [NoSQL 精粹](#)

博客

- [What is NoSQL database?](#)

论文

- 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 管理

课程

- CMU 数据库 [Database Systems \(15–445/645\)](#), by Andy Pavlo
 - [Buffer Pools](#)

论文

- 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

磁盘 IO

博客

- [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](#)

论文

- 2016, [Design Tradeoffs of Data Access Methods](#), SIGMOD
- 2016, [Designing Access Methods: The RUM Conjecture](#), EDBT

B+ 树

博客

- [B树、B+树索引算法原理（上）](#) by [codedump](#)
- [B树、B+树索引算法原理（下）](#)
- [B-link Tree: 一种 B+Tree 的并发优化](#)
- [B+ Tree 的并发优化](#)

课程

- CMU Database Systems (15–445/645), by [Andy Pavlo](#)
 - [Trees Indexes I](#)
 - [Trees Indexes II](#)
- CMU Advanced Database Systems (15–721), by [Andy Pavlo](#)
 - [OLTP Indexes \(B+Tree Data Structures\)](#)

论文

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

项目

- [boltdb](#)
- [sqlite](#)

LSM Tree

博客

- [LSM-Tree 设计](#)
- [浅析 LSM Tree \(WiscKey, Bourbon、Rum\)](#)
- [LevelDB 源码剖析](#)

论文

- 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](#)

项目

- [leveldb](#)
- [rocksdb](#)
- [goleveldb](#)
- [pebble](#)
- [badger](#)

数据分区

博客

- [TiDB Internal \(I\) – Data Storage](#)
- [Partitioning Behavior of DynamoDB](#), by Parth Modi

论文

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

复制/一致性

博客

- [Tick or Tock? Keeping Time and Order in Distributed Databases](#), by Liu Tang
- [Raft 算法](#)

论文

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

基准测试

博客

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

Papers:

- 2010, [Benchmarking Cloud Serving Systems with YCSB](#), SOCC

HTAP

博客

- [What is HTAP?](#) , single store
- [HTAP: HYBRID TRANSACTIONAL AND ANALYTICAL PROCESSING](#)
- [Making An HTAP Database Reality](#)

论文

- [TiDB: A Raft-based HTAP Database](#)
- [Hybrid Transactional/Analytical Processing: A Survey](#)

其他

- 数据库排名 DB-ranking: <https://db-engines.com/en/ranking>
- CNCF-Database 全景图: <https://landscape.cncf.io/card-mode?category=database>
- dbdb.io: 各类数据库大汇总: <https://dbdb.io/>
- 数据库社区: [墨天轮](#)