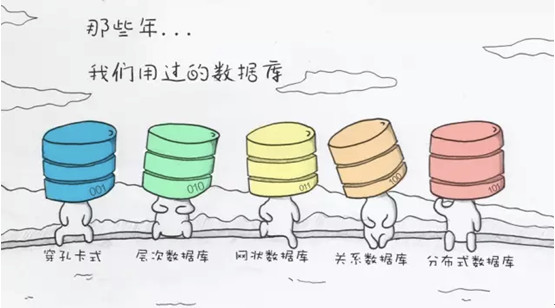
<https://www.csdn.net/article/a/2015-09-02/15827778>

**摘要：** 短短几十年的时间，数据库经历了从远古时代到现代的发展。关系数据库，NoSQL，NewSQL，OLTP，OLAP……满眼的专业词汇你都知道是什么吗？下面是我们精心整理的“数据库发展史”，也许这才是了解数据库发展的正确姿势。



  短短几十年的时间，数据库经历了从远古时代到现代的发展。关系数据库，NoSQL，NewSQL，OLTP，OLAP……满眼的专业词汇你都知道是什么吗?下面是我们精心整理的“数据库发展史”，也许这才是了解数据库发展的正确姿势。

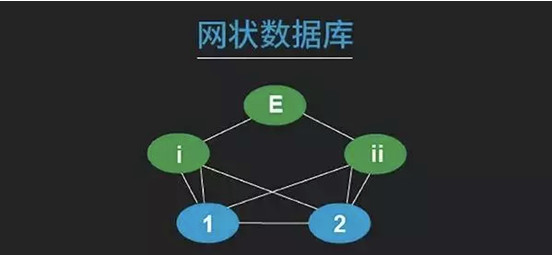




在很久很久很久以前，计算机还还只叫做“分析机”的时候，其实就已经有了“数据库”­——数量巨大的“穿孔卡片”。当然也存在所谓数据管理——也就是对穿孔卡片进行存储和处理，其工作量之大，工作内容之繁琐，消耗的不仅是脑力，也是体力和空间。



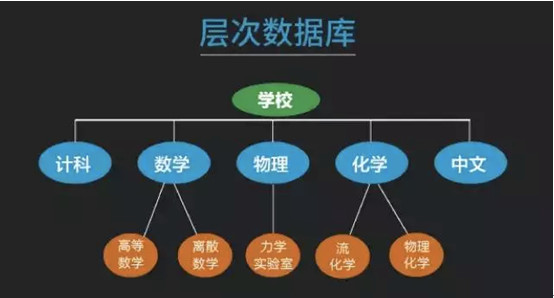
http://articles.csdn.net/uploads/allimg/150902/208_150902093812_1.jpg



网状结构开辟了便捷的数据管理之先河，它表示了数据以及数据之间的联系，但是实体之间往往关系复杂又互相纠缠，所以“网”越来越大，越来越密。关系复杂，操作复杂，所以一段时间内都被使用着。

通用电气公司 1961年开发的IDMS集成数据管理系统为网状结构数据库的代表产品，在一时间占领了一壁江山的地位。

http://articles.csdn.net/uploads/allimg/150902/208_150902094041_1.jpg

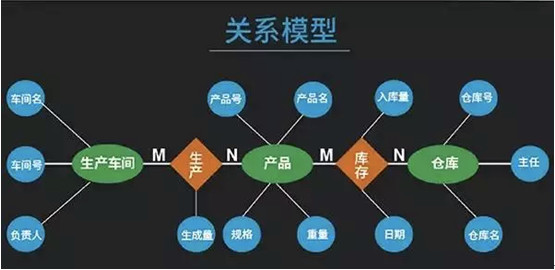


顾名思义，层次数据库是用树形结构表示实体之间联系，层次清晰，但是对于复杂的数据结构模型，树形深度会很大，有很多父子关系，数据会形成冗余。

IBM 1968年 开发了IMS，为代表产品，被较广泛地使用着。(IMS的热销也为IBM后续新型数据库推出埋下障碍，详见后续分析~)

网状和层次模型被使用着，直到——

http://articles.csdn.net/uploads/allimg/150902/208_150902094234_1.jpg



以上两种数据库解决了数据集中和共享问题，但是在数据独立性和抽象级别上仍由很大欠缺。用户在对数据存取时，让需要明确数据的存储结构，指出存储路径。1970年，IBM研究员E.F.Codd提出关系模型的概念，关系模型横空出世。

数据存放在一种称为二维表的存储结构中，整个数据库是由若干个相互关联的二维表组成。

关系数据库解决了冗余问题，简化更新。在当前被广泛应用于运营商金融电力制造物流等行业。其代表产品均为大名鼎鼎的：Oracle、DB2(IBM)、MS SQLServer(MS)、Informix、MySQL(开源，支持几乎所有操作系统，使用最多)等。

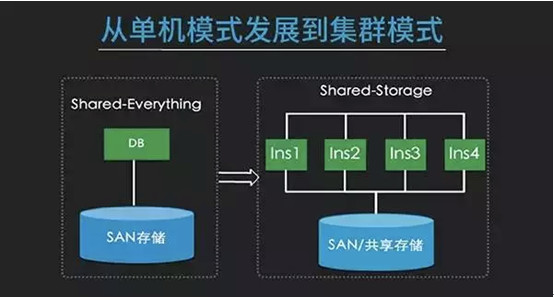
话说IBM提出关系数据库的时候就启动了大名鼎鼎的研究项目：System R，论证一个全功能关系数据库的可行性，但是直到1985年才作为正式产品推向市场(因为IBM的官僚体系并且之前层次数据库为IBM重点产品，导致参与人员不积极);而此时早已发现商机并已推出关系数据库的Oracle创始人Ellision已成千万富翁。





当网络访问量越来越大，数据库无法满足业务需求时，解决的方法有两个：拼数量和拼装备。

http://articles.csdn.net/uploads/allimg/150902/208_150902094633_1.jpg



随着业务增加，为了能够提供较高并发的用户访问，数据库架构从单机模式发展到了集群模式，处理节点增多，处理节点共享存储。这样的集群架构能支持高并发，但是共享存储侧存在扩展瓶颈。

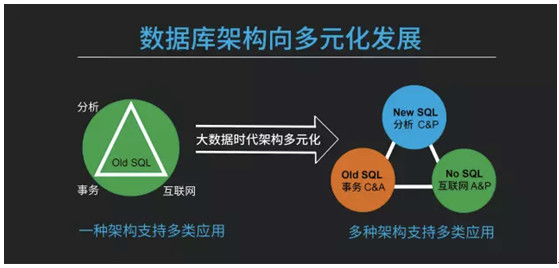
这样的架构现在在金融、物流运营商等传统行业依旧支持各式各样的信息系统。

http://articles.csdn.net/uploads/allimg/150902/208_150902094754_1.jpg

传统关系数据库是处理永久稳定的数据，维护数据的完整性一致性，但是对于实时没有响应，常规数据库存储完全基于磁盘，而内存的读写速度高于磁盘几个数量级。因此对于运营商的实时账务系统，用户查询话费等对实时要求很高的数据处理就转变成用内存数据库中。

Oracle TimesTen、IBM SolidDB、SQLite、Redis为内存关系数据库的代表产品。

话说传统关系数据库经历了几十年的风平浪静，单一的系统架构支撑着各种场景的业务。但随着业务的发展，企业系统的数据特别是互联网上的业务，面临着大数据量、高并发的趋势。根据特点，应用也大致分为三种，分别是分析型、事务型和互联网型。

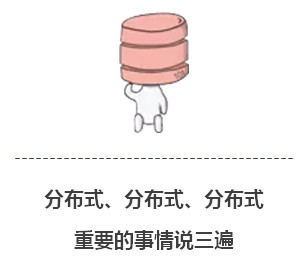


分析型：即OLAP，强调的是高IO和规模化数据处理，如银行信用记录分析、大数据分析型业务，需要New SQL数据库。

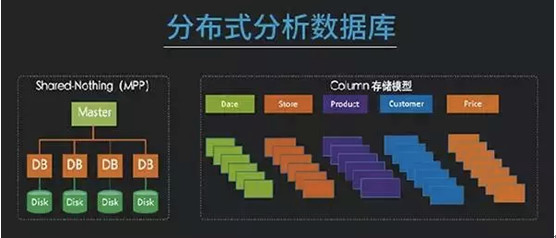
事务型：即OLTP，强调数据库访问的的高并发和响应时间，如运营商实时话费查询和银行交易流水记录，需要Old SQL数据库。

互联网型：以web2.0的产生而带来的大量类似SNS的数据，数据之间关系较弱，以文档、图片、视频等为代表，需要No SQL数据库。

显然，针对三种不同的应用，需要更强劲的“新三样”数据库。



http://articles.csdn.net/uploads/allimg/150902/208_150902095048_1.jpg



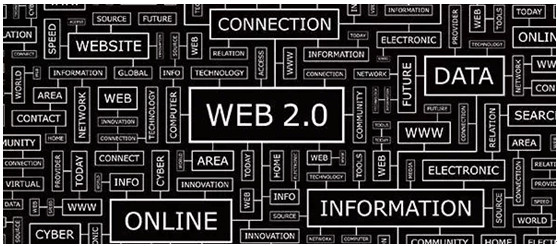
为了满足企业经营者的决策支持和经营分析需要对大量的生产数据进行多维环境下的特定查询和报表分析，传统的集中式数据库架构对处理大量数据的计算和存储能力显得很苍白。

此时出现了以SQL为访问接口面向OLAP业务场景的新型关系数据库：如MPP架构的数据库产品，其能够进行并行分析处理能力，高速数据装载以及数据压缩，在线扩展能力等。

另外列式数据库的出现在批量数据处理和即席查询业务中极大降低了I/O开销，广泛应用于数据分析领域。

Sybase IQ、 Infobright、Vertica等为分布式分析数据库代表。

http://articles.csdn.net/uploads/allimg/150902/208_150902095406_1.jpg

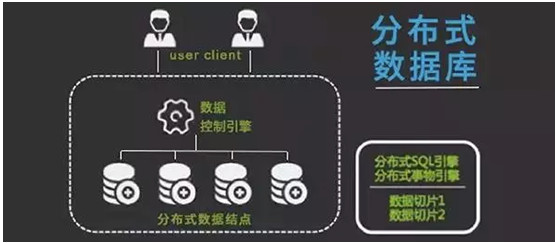


Web2.0的发展产生了超大规模和高并发的SNS类型的纯动态网站。传统数据库扩展性差，大数据下IO压力大，表结构更改困难。互联网时代的高并发需要非关系型、分布式数据存储，不保证数据的ACID特性，NoSQL 非关系型数据库在批量数据处理和即席查询业务中极大降低了I/O开销，独有的数据存储方式提供单表数十万的QPS。采用无共享架构、结构简单、弹性可扩展、读写吞吐量大。

阿里云OTS为分布式非关系数据库代表产品。

OTS为开放结构化数据服务，构建于阿里云飞天分布式系统的NoSQL数据库服务，提供海量结构化数据的存储和实时访问。OTS以实例和表的形式组织数据，通过数据分片和负载均衡技术，实现规模上的无缝扩展。

http://articles.csdn.net/uploads/allimg/150902/208_150902095506_1.jpg



New SQL和Nosql数据库在DT时代的众多业务场景中得以应用，但是其对MVCC模型的支持和事务处理中对数据的强一致性保证较弱。传统关系型数据库在面向银行交易，运营商业务办理等典型OLTP的业务无法大显身手，基于关系数据库的单实例和集群架构在交易型业务中的扩展能力受限。

分布式数据库的产生，能够提供水平扩展、突破单节点数据库服务器的IO能力限制的方案，解决数据库扩展性能，保证高效的持续的在线交易业务。

阿里云DRDS、数梦工场DTBase为分布式事务数据库代表产品。

DTBase提供分布式数据库引擎，提供对数据的智能分库分表，支持分布式SQL优化，分布式DDL和分布式DML。针对业务量的变化可平滑扩缩容，独有的高效数据迁移，扩缩容对应用几乎无影响。