

# 第一章 绪论

徐 辰

[cxu@dase.ecnu.edu.cn](mailto:cxu@dase.ecnu.edu.cn)

華東師範大學



# 大纲

2

## □ 分布式系统

- 什么是分布式系统

- 分布式系统的类型

## □ 从数据管理角度看分布式系统

## □ 分布式计算系统

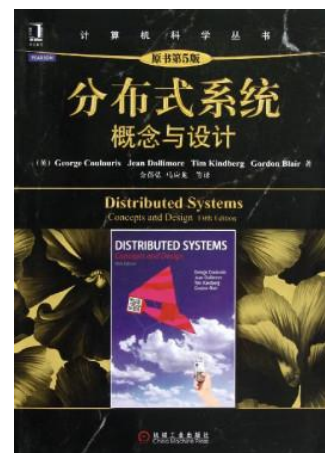
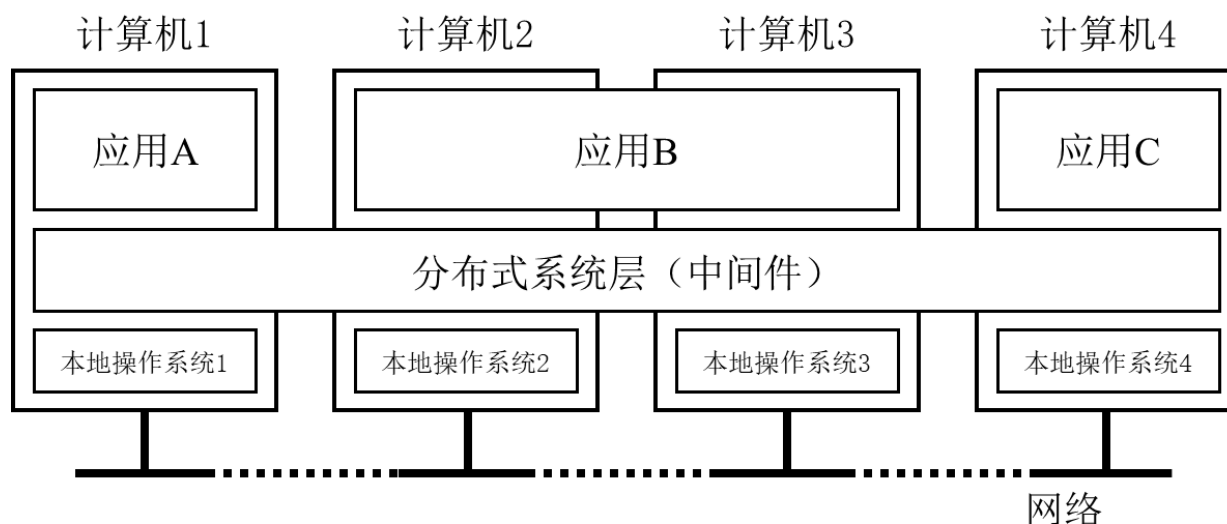
## □ 课程组织结构



# 分布式系统的定义

3

- 《分布式系统：概念与设计》认为：分布式系统是**若干独立计算机的集合**，这些计算机对于用户来说**就像是一个单机的系统**



# 大纲

4

## □ 分布式系统

- 什么是分布式系统

- 分布式系统的类型

## □ 从数据管理角度看分布式系统

## □ 分布式计算系统

## □ 课程组织结构



# 分布式系统的类型

5

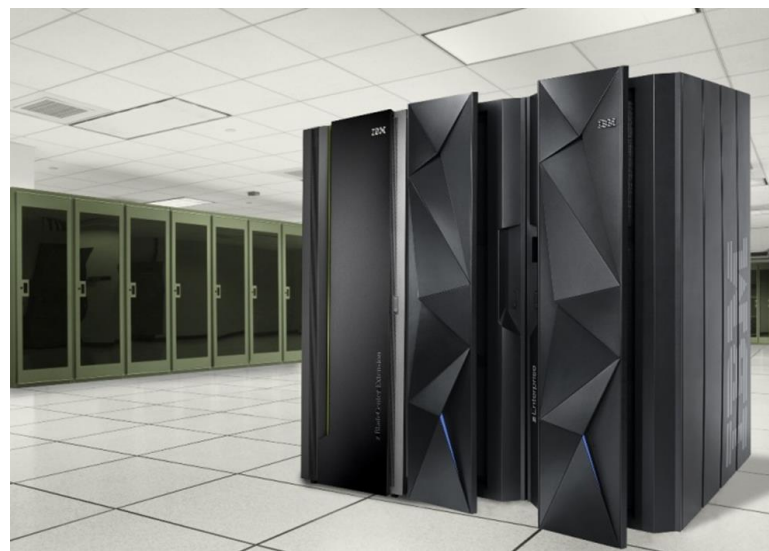
## □ 基于计算机构建的分布式系统

### ✚ 分布式计算系统

- 集群计算
- 网格计算
- 云计算

### ✚ 分布式信息系统

- 事务处理
- 企业信息集成



# 分布式系统的类型

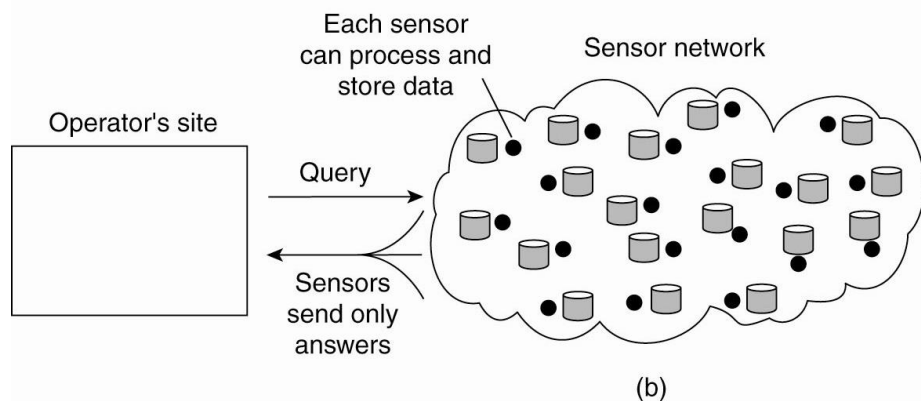
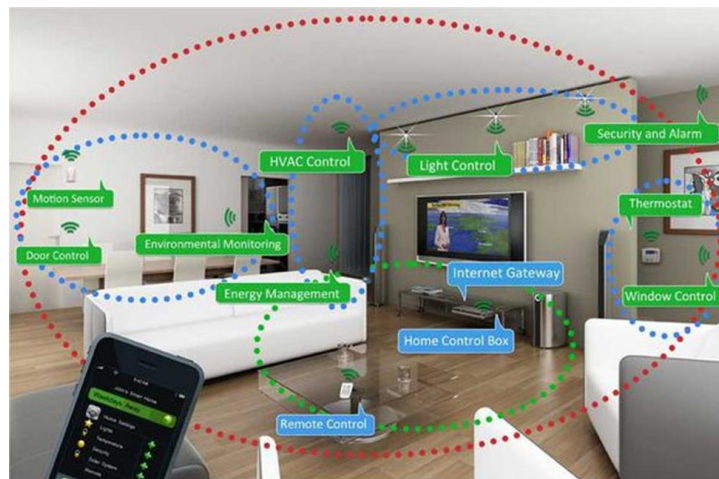
6

## □ 基于微型设备构建的分布式系统

智慧家庭

电子健康

传感网络



# 大纲

7

- 分布式系统
- 从数据管理角度看分布式系统
  - ✚ 数据管理系统的发展回顾
  - ✚ 大数据背景下数据管理面临的挑战
  - ✚ 面向数据管理的分布式系统
- 分布式计算系统
- 课程组织结构



# 数据管理系统发展历史

8

- 第一代：层次、网状数据库系统
  - ✚ 数据的存储与访问从应用程序中分离
- 第二代：关系数据库系统
  - ✚ OLTP
- 第三代：数据仓库系统
  - ✚ OLAP
- 第四代：大数据管理系统
  - ✚ 分布式、可扩展、高可用

杜小勇, 卢卫, 张峰. 大数据管理系统的历史、现状与未来. 软件学报, 2019, 30(1): 127-141





# 大纲

9

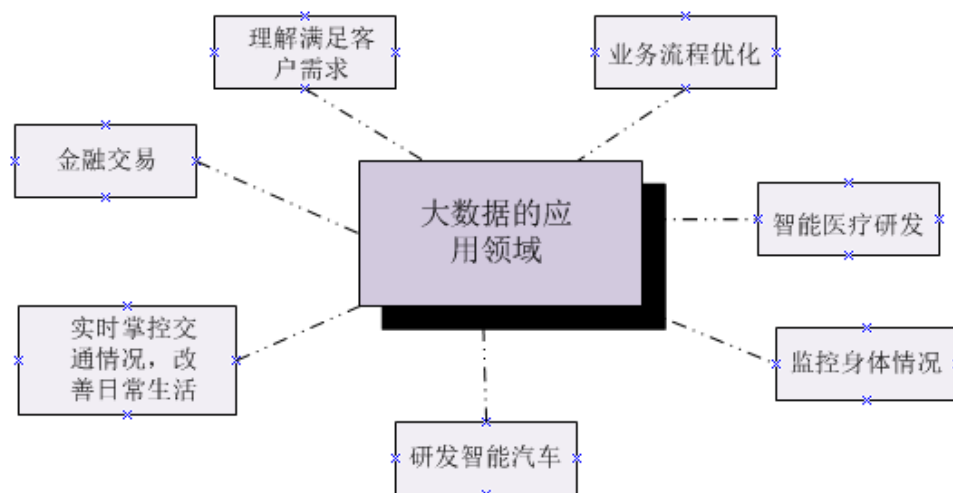
- 分布式系统
- 从数据管理角度看分布式系统
  - ✚ 数据管理系统的发展回顾
  - ✚ 大数据背景下数据管理面临的挑战
  - ✚ 面向数据管理的分布式系统
- 分布式计算系统
- 课程组织结构



# 大数据的应用

10

- 大数据无处不在，金融、汽车、零售、餐饮、电信、能源、政务、医疗、体育、娱乐等在内的社会各行各业都已经融入了大数据的印迹



# 典型的大数据应用实例

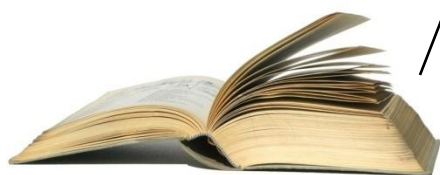
11



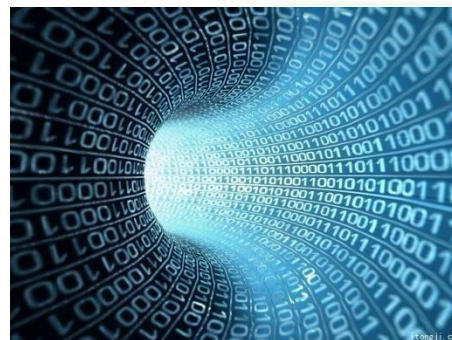
Kevin Spacey



David Fincher



英国同名小说《纸牌屋》



大数据分析

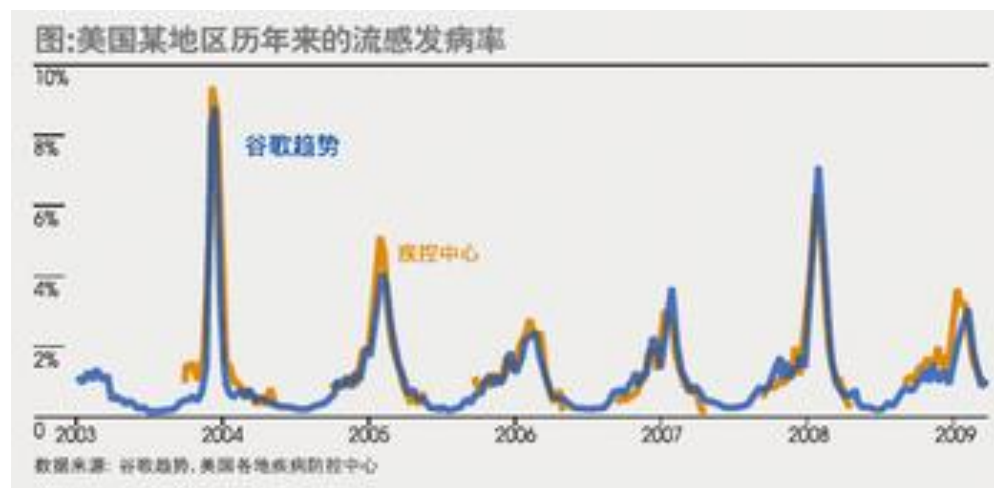


风靡全球的美剧《纸牌屋》

# 典型的大数据应用实例

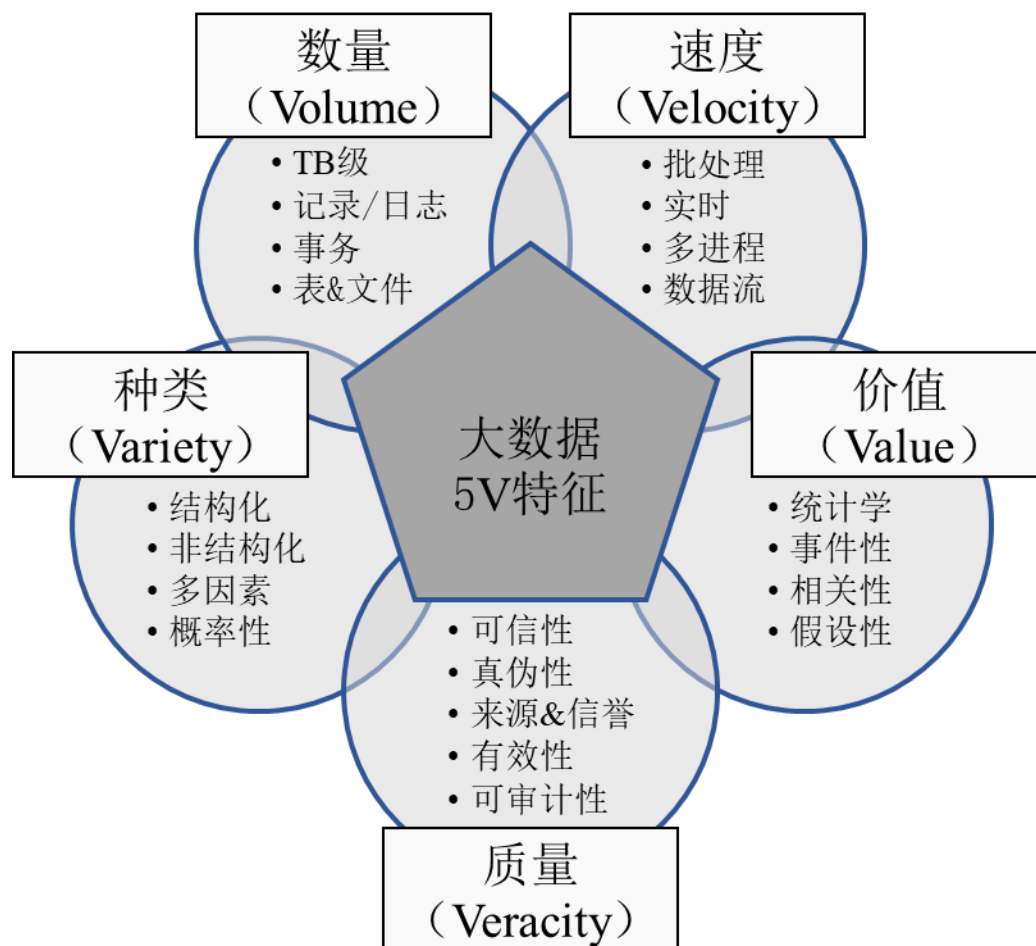
12

## □ 通过跟踪搜索词相关数据来判断全美地区的流感情况



# 大数据5V

13

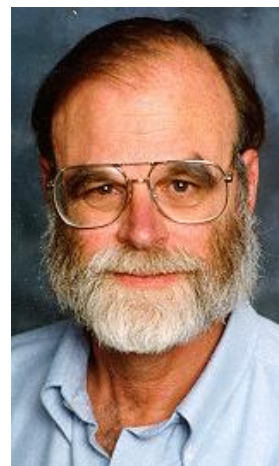
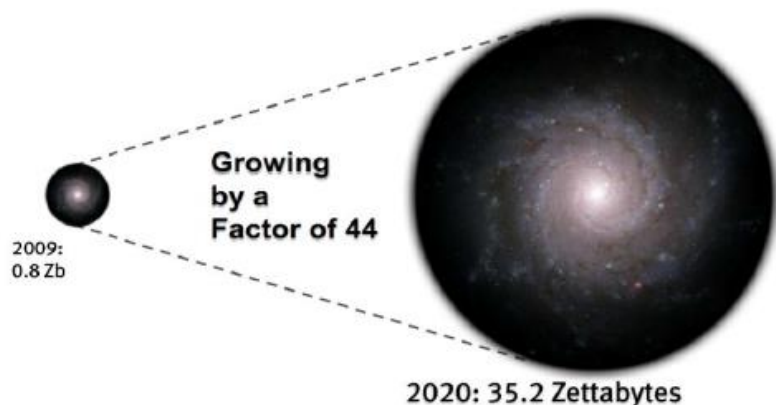


# 数量大 (Volume)

14

## □ 数据量大:

- ✚ 包括采集、存储和计算的量都很大
- ✚ 相对于技术水平和应用水平而言
- ✚ Jim Gray: 大数据“摩尔定律”
  - 每18个月新增的数据量等于有史以来数量之和



# 种类多 (Variety)

15

## □ 结构化

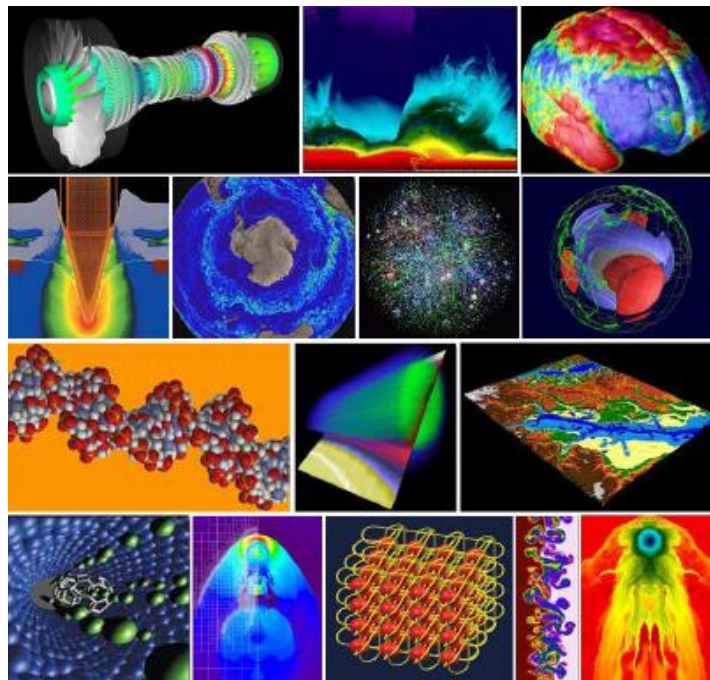
+ 日志

## □ 非结构化

+ 图片

+ 音频

+ 视频





# 速度快（Velocity）

16

## □ 大数据的产生速度非常快

- 双十一：2017年11月11日凌晨5分22秒交易数量为25.6万笔/秒，给底层系统带来的峰值为4200万次/秒





# 价值高 (Value)

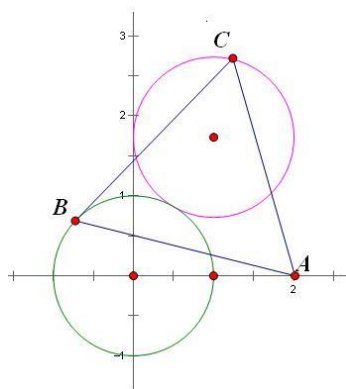
17

## □ 大数据蕴含了非常高的潜在价值

- 利用机器学习、数据挖掘或人工智能等方法，可以从数据中挖掘出对未来趋势或模式的预测分析有价值的信息、甚至发现新规律、新知识



实验归纳



模型推演



仿真模拟



Jim Gray第四范式

数据密集型科学发现

# 质量低（Veracity）

18

□ 大数据通常都是自动采集的，天然地具有噪音

✚ 数据准确性

✚ 可信赖度

基础数据	60345 浏览PV	6489 浏览UV	1452 外部分享	88 导出PV	暂无 平均停留时间	暂无 平均访问深度
转化数据	366456 商品曝光次数	2504 商品浏览UV	799 下单笔数	570 付款订单	¥ 101456.20 付款金额	22.76% 曝光购买付款率



# 为数据管理系统带来的机遇与挑战

19

- **数据量大**：要求大数据管理系统能够存储和管理TB甚至PB级别的数据
- **数据种类多**：要求大数据管理系统能够在同一平台处理多种类型的数据
- **产生速度快**：要求大数据管理系统能够在数据产生的过程中实时处理数据



# 为数据管理系统带来的机遇与挑战

20

- **潜在价值高**：要求大数据管理系统能够支持使用机器学习、数据挖掘或人工智能等方法进行分析处理
- **数据质量低**：要求数据管理系统要考虑适应支持更复杂的数据治理、数据分析和机器学习技术



# 大纲

21

- 分布式系统
- 从数据管理角度看分布式系统
  - ✚ 数据管理系统的发展回顾
  - ✚ 大数据背景下数据管理面临的挑战
  - ✚ 面向数据管理的分布式系统
- 分布式计算系统
- 课程组织结构



# 数据管理系统的设计理念

22

## □ 关系数据库时代：one-size-fits-all

- ✚ 所有有关数据管理的问题都应该交由关系数据库来解决
- ✚ 希望“一劳永逸”



**Michael Stonebraker, Ugur Çetintemel:**

"One Size Fits All": An Idea Whose Time Has Come and Gone (Abstract).  
ICDE 2005

# 数据管理系统的设计理念

23

## □ 大数据时代：one-size-does-not-fit-all

- 许多应用特别是互联网应用，比如社交网络、知识图谱、搜索引擎、“秒杀”营销等带来的数据管理问题，使用传统的关系数据库已经无法满足应用处理的要求
- 从“一劳永逸”走向“**分类定制**”：解耦数据库系统的各个模块，并从模型、可靠性、可伸缩性、性能等方面对各个模块进行重新设计

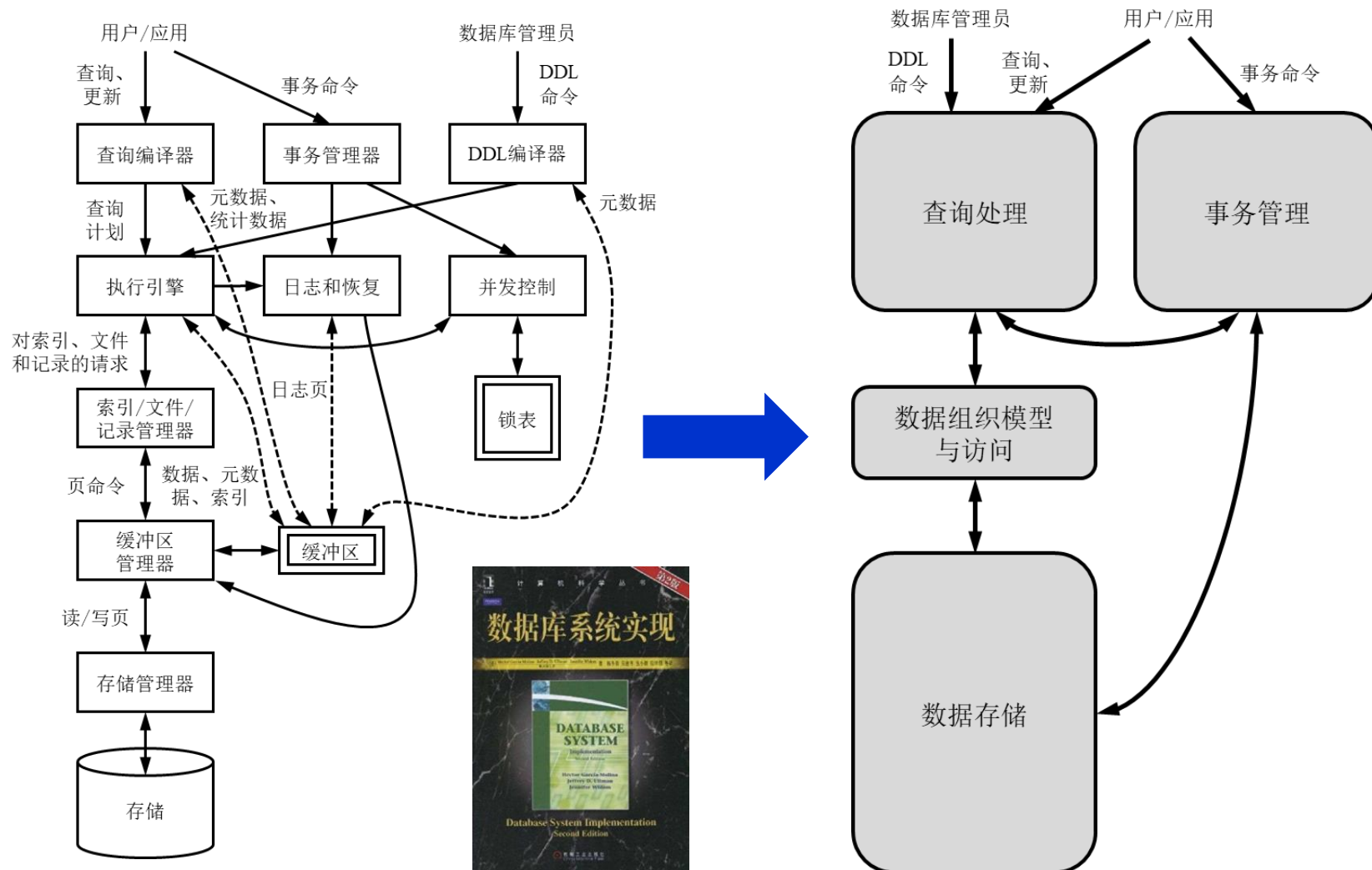
**Michael Stonebraker**, Ugur Çetintemel:

"One Size Fits All": An Idea Whose Time Has Come and Gone (Abstract).  
ICDE 2005



# 关系数据库系统的架构及模块

24





# 从集中式走向分布式

25

- 大数据的5V特征决定了运用分布式系统技术研发大数据管理系统是必然选择

集中式	分布式	
数据存储	分布式存储系统	HDFS、Kafka、Alluxio等
数据组织模型与访问	分布式NoSQL系统	BigTable、Dynamo、Cassandra、Neo4j等
查询处理	分布式计算系统	MapReduce、Spark、Flink等
事务管理	分布式NewSQL系统	Spanner、OceanBase、TiDB等

# 大纲

26

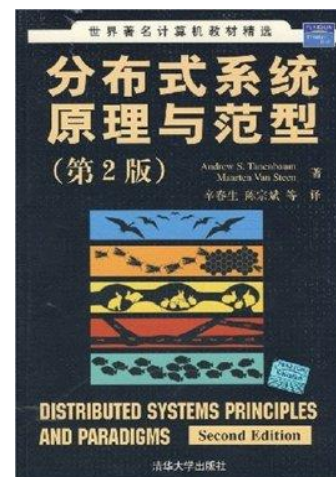
- 分布式系统
- 从数据管理角度看分布式系统
- 分布式计算系统
  - ✚ 什么是分布式计算系统
  - ✚ 系统生态圈
- 课程组织结构



# 分布式计算系统的定义

27

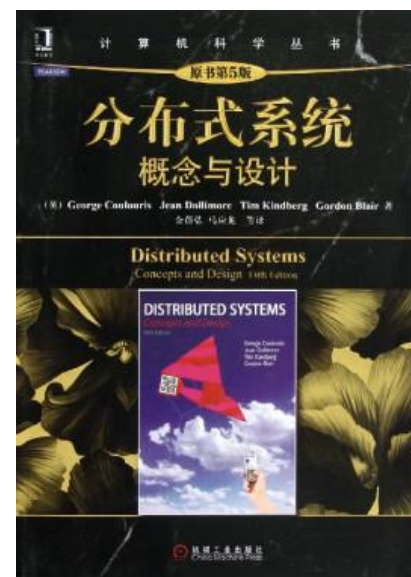
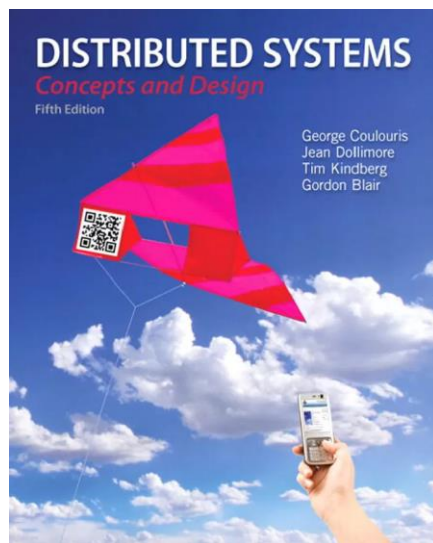
- 《分布式系统：原理与泛型》认为分布式计算系统是用于实现高性能计算的系统，所说的高性能计算主要面向解决科学计算领域的复杂数值计算问题，并不用于进行大规模数据处理



# 分布式计算系统的定义

28

- 《分布式系统：概念与设计》对分布式计算的概念进行了延伸，将能够支持**大规模数据处理**的MapReduce视为能够提供分布式计算服务的系统



# 分布式计算系统的定义

29

- 分布式计算系统是作用于若干独立计算机之上、使得这些计算机能够协同执行计算完成某项应用的软件系统
- 分布式计算系统归根结底是为了解决某些类别的应用问题而设计的分布式系统



# 应用驱动分布式计算系统的发展创新

30



# 应用的类型

31

- **计算密集型**应用：计算机在处理这些应用时CPU的处理能力成为了首要限制性因素
  - ✚ 高性能计算领域的数值运算
  - ✚ 人工智能领域的深度学习模型训练
- **数据密集型**应用：计算机在处理这些应用时I/O带宽（包括磁盘I/O、网络I/O等）成为了首要限制性因素
  - ✚ 大数据领域的数据处理



# 大纲

32

- 分布式系统
- 从数据管理角度看分布式系统
- 分布式计算系统
  - ✚ 什么是分布式计算系统
  - ✚ 系统生态圈
- 课程组织结构





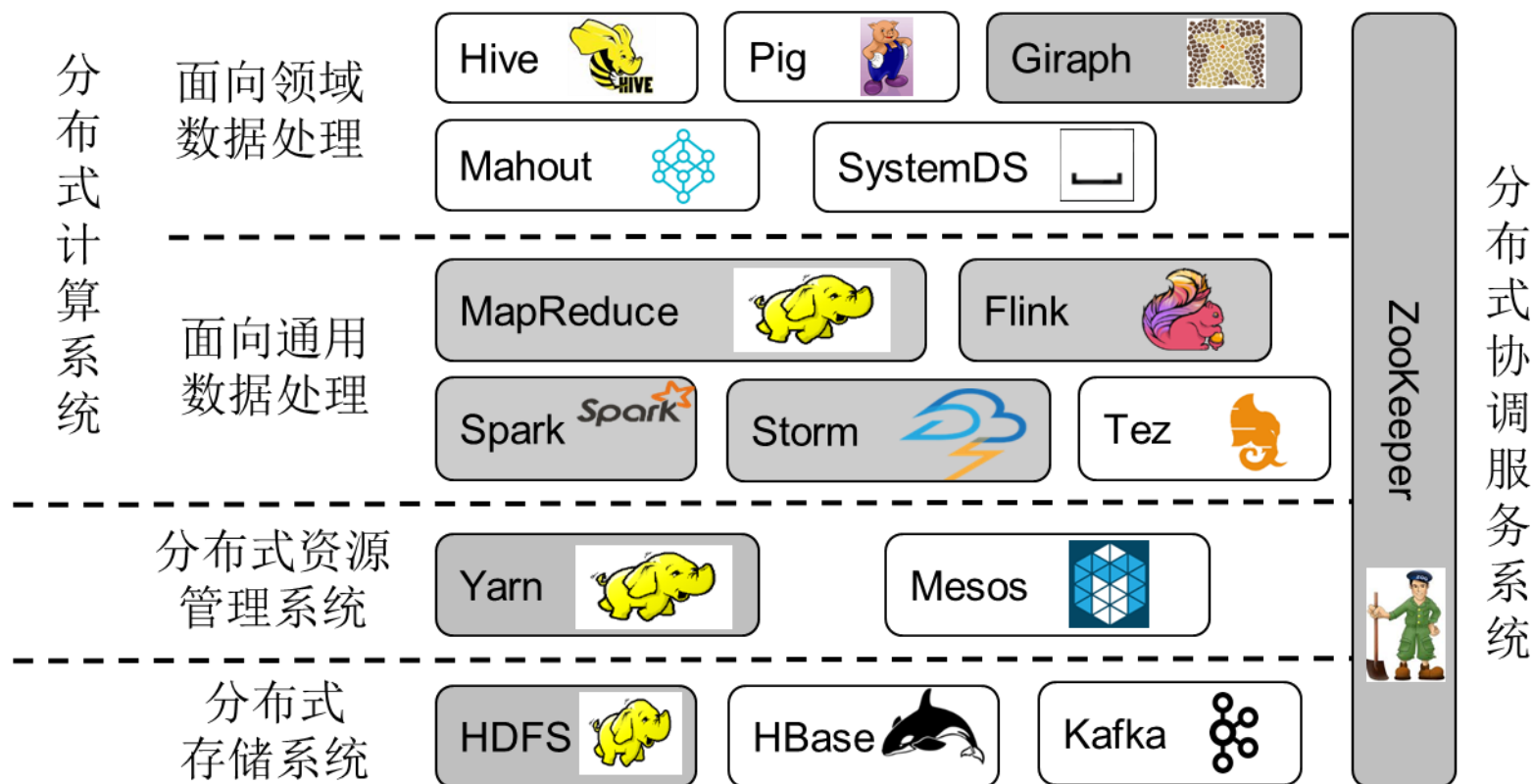
# 面向数据密集型应用的分布式计算系统

33

应用	类型	代表性系统
通用数据处理	批处理系统	MapReduce、Spark、Tez
	流计算系统	Storm、Spark Streaming
	批流融合系统	Flink、Structured Streaming
领域数据处理	图处理系统	Giraph、GraphX
	机器学习系统	Mahout、SystemDS、Mllib
	查询分析系统	Hive、Pig

# 分布式计算系统生态圈

34



# 大纲

35

- 分布式系统
- 从数据管理角度看分布式系统
- 分布式计算系统
- 课程组织结构



## □ 设计思想：

- ✦ 该系统应对的**应用场景**如何？
- ✦ **为什么**需要构造这个系统？
- ✦ 该系统所处理的**数据模型**是怎样的？
- ✦ 该系统采用何种**计算模型**？
- ✦ 如果系统支持迭代计算，那么采用何种**迭代模型**？
- ✦ .....

## □ 体系架构：

- ✚ 作为一个分布式系统，该系统包含哪些部件，即由**哪些进程或线程**构成？
- ✚ 各个部件都扮演着**怎样的角色**？
- ✚ 系统中的各个部件都以**怎样的次序**共同工作才能保障用户应用程序的顺利执行？
- ✚ .....

## □ 工作原理：

- ✚ 系统如何根据用户编写的应用程序生成相应的执行计划，以及是否可以进行优化？
- ✚ 面对数据密集型应用，系统如何减少数据I/O的代价？
  - 这里I/O的代价表现为节点本地读写数据的磁盘I/O代价以及节点之间通过网络进行数据传输的代价，后者在分布式场景下显得尤为突出

✚ .....

## □ 容错机制：

- 早期的并行计算系统部署在高性能服务器甚至是超级计算机当中，通常不需要考虑容错问题
- 商用服务器集群中的节点发生故障是常见的现象，系统如何保证在某些节点有故障的情况下不影响用户应用程序正常执行，即故障对用户透明？
- .....

## □ 编程示例：

- ✚ 使用这些系统的编程接口进行简单编程的一般步骤是怎样的？
- ✚ 使用这些系统进行编程是否有通用的框架？如何提高某些程序的性能？
- ✚ 使用不同系统实现同一个应用所编写的程序有何差异？
- ✚ .....



# 课程内容结构

41

类别			内容	章节
全书总览			绪论	第1章
分布式存储系统			HDFS	第2章
分布式计算系统	通用数据处理系统	批处理系统	MapReduce	第3章
			Spark	第4章
		流计算系统	Storm	第7章
			Spark Streaming	第8章
	批流融合系统	基础	第9章	
		Flink	第10章	
	领域数据处理系统	图处理系统	Giraph	第11章
分布式资源管理系统			Yarn	第5章
分布式协调服务系统			ZooKeeper	第6章

# 本章小结

42

- 分布式系统
- 从数据管理角度看分布式系统
- 分布式计算系统
- 课程组织结构

