《分布式系统》课程建设与教学实践

陈榕

上海交通大学

2021.12

《分布式系统》课程背景

关键支撑作用

分布式系统

- 分布式存储系统
- 分布式文件系统
- 分布式计算系统
- 分布式任务调度
- ...

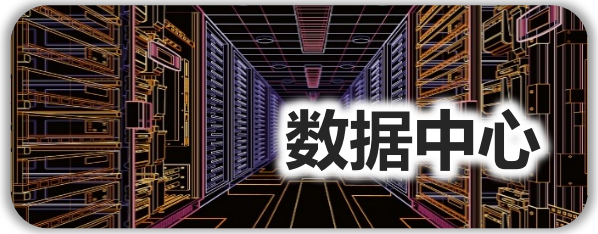
- 性能
- 弹性
- 容错
- 安全
- ...

《分布式系统》课程背景

大数据

云计算

移动计算



数据中心

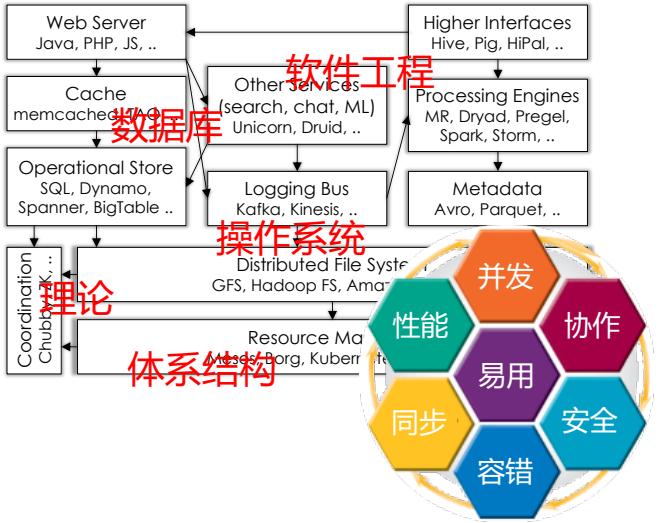
关键支撑作用

分布式系统

1. 基本概念
2. 设计原则
3. 关键方法
4. 核心技术
5. 系统工具

典型案例

《分布式系统》课程挑战



知识体系：多、散、变、活

应用

硬件

分布式系统

繁荣时期

课程教材

- 覆盖面差异大
- 缺乏公认体系
- 内容相对陈旧

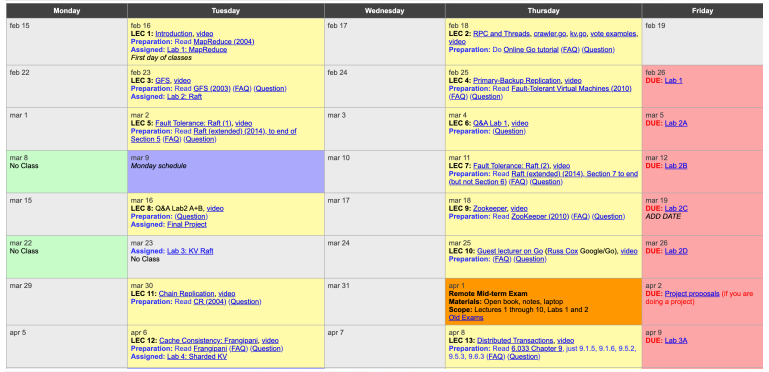
大部分国外课程并未指定教材，课程内容体系也有较大差异

《分布式系统》课程借鉴

国内外大学《分布式系统》课程调研

- 围绕“主题”安排课程内容
- 配套论文研读、实验项目等
- 无指定教材（仅参考读物）
- 分布式系统领域专家主讲

MIT: 6.824



Monday	Tuesday	Wednesday	Thursday	Friday
feb 15	feb 16 **LEC 1:** Introduction, video **Preparation:** Read MapReduce (2004) **Assigned:** Lab 1: MapReduce *First day of classes*	feb 17	feb 18 **LEC 2:** RPC and Threads, crawler.go, kv.go, vote examples, video **Preparation:** Do Online Go tutorial (FAQ) (Question)	feb 19
feb 22	feb 23 **LEC 3:** GFS, video **Preparation:** Read GFS (2003) (FAQ) (Question) **Assigned:** Lab 2: Raft	feb 24	feb 25 **LEC 4:** Primary-Backup Replication, video **Preparation:** Read Fault-Tolerant Virtual Machines (2010) (FAQ) (Question)	feb 26 **DUE:** Lab 1
mar 1	mar 2 **LEC 5:** Fault Tolerance: Raft (1), video **Preparation:** Read Raft (extended) (2014), to end of Section 5 (FAQ) (Question)	mar 3	mar 4 **LEC 6:** Q&A Lab 1, video **Preparation:** (Question)	mar 5 **DUE:** Lab 2A
mar 8 No Class	mar 9 *Monday schedule*	mar 10	mar 11 **LEC 7:** Fault Tolerance: Raft (2), video **Preparation:** Read Raft (extended) (2014), Section 7 to end (but not Section 8) (FAQ) (Question)	mar 12 **DUE:** Lab 2B
mar 15	mar 16 **LEC 8:** Q&A Lab2 A+B, video **Preparation:** (Question) **Assigned:** Final Project	mar 17	mar 18 **LEC 9:** Zookeeper, video **Preparation:** Read ZooKeeper (2010) (FAQ) (Question)	mar 19 **DUE:** Lab 2C *ADD DATE*
mar 22 No Class	mar 23 **Assigned:** Lab 3: KV Raft No Class	mar 24	mar 25 **LEC 10:** Guest lecturer on Go (Russ Cox GoogleGo), video **Preparation:** (FAQ) (Question)	mar 26 **DUE:** Lab 2D
mar 29	mar 30 **LEC 11:** Chain Replication, video **Preparation:** Read CR (2004) (Question)	mar 31	apr 1 **Remote Mid-term Exam** **Materials:** Open book, notes, laptop **Scope:** Lectures 1 through 10, Labs 1 and 2 Old Exams	apr 2 **DUE:** Project proposals (if you are doing a project)
apr 5	apr 6 **LEC 12:** Cache Consistency: Frangipani, video **Preparation:** Read Frangipani (FAQ) (Question) **Assigned:** Lab 4: Sharded KV	apr 7	apr 8 **LEC 13:** Distributed Transactions, video **Preparation:** Read 6.033 Chapter 9, just 9.1.5, 9.1.6, 9.5.2, 9.5.3, 9.6.3 (FAQ) (Question)	apr 9 **DUE:** Lab 3A

16主题 | 16篇论文 | 4个实验

主讲：Frans Kaashoek

Stanford: CS244b

Week of	Monday	Wednesday
Apr 6 - Apr 10	**Introduction** **Handout:** FLP	Zookeeper
Apr 13 - Apr 17	Two-phase Commit	Paxos Mode Simple
Apr 20 - Apr 24	Raft	HARP
Apr 27 - May 1	Practical Byzantine Fault Tolerance	Honey Badger
May 4 - May 8	HotStuff	Coral
May 11 - May 15	Dynamo	COPS
May 18 - May 22	GFS	Aurora

15主题 | 15篇论文 | 自选项目

主讲：Jinyang Li

NYU: Distributed Systems

Date	Lecture	Lecture Preparation
9/4	Introduction, RPC and threads ds-intro.pdf ds-rpcthreads.pdf	[Online Go Tutorial]
9/11	primary/backup replication ds-pb.pdf ds-gfs.pdf	[Google File System] [Viewstamp Replication]
9/18	linearizability linearizability.pdf	[Linearizability]
9/25	consensus: paxos paxos.pdf	[Paxos]
10/2	consensus: raft ds-raft.pdf	[Raft]
10/9	causal consistency ds-causal.pdf	[Lamport Clocks] [Bayou]

11主题 | 21篇论文 | 5个实验

主讲：David Mazières

《分布式系统》课程借鉴

国外大学《分布式系统》课程分析

优点

- 适合于分布式系统特点(技术点分散、不系统、变化快)
- 强调问题解决能力、系统设计能力，能紧跟技术发展前沿

问题

- 缺少展示技术/理论的发展历程、对于不同设计权衡的思辨
- 欠缺基础知识/背景介绍，对系统软件开发能力有较高要求

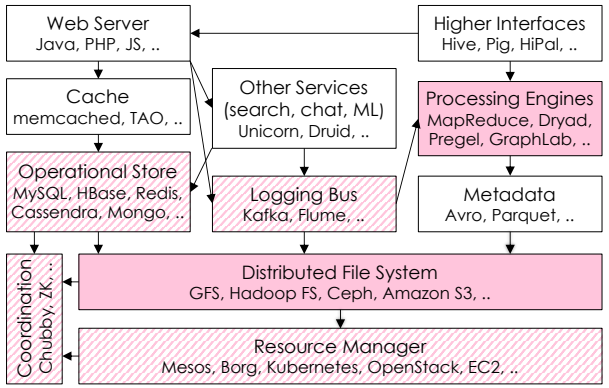
课程建设案例（上海交通大学/软件学院）

课程对象

- 软件工程专业研究生(各方向：SE | Theory | AI | System ...)

课程目标

- 掌握分布式系统的核心问题、共性技术、(S|H需求/特征)设计权衡
- 了解领域研究与产业前沿、历史变迁、未来趋势
- 重点培养学生(全体研究生)
  1. 解决问题的系统化手段和方法
  2. 能力：选型、组合、扩展、优化等



分布式系统软件栈

课程建设案例（上海交通大学/软件学院）

课程负责人

- 陈榕（软件学院）
- 研究方向：计算机系统、分布式系统
- 主要职责：课程建设、课程主讲等

在OSDI/SOSP/NSDI上发表多项分布式系统研究工作

领导研发分布式数据库、图计算、键值存储等系统

课程信息

- 开设时间：2017～2021（春）
- 课程信息：专业基础课、48学时（3学分）
- 选课人数：49、74、75、69、70（全体研究生）

课程建设案例（上海交通大学/软件学院）

建设思路

- “经典”与“前沿”相结合 + “指定”与“自选”相结合

建设特色

- 以“主题”为中心，“新(2010s)旧(1990s)”对比
  - 先“旧”：核心问题、基础概念
  - 再“新”：(软硬件需求/环境)变化 >> 技术变迁和演化(权衡)
- 半自选团队SHOW：前沿论文(同主题、5年内)的讲解+演示
  - 团队任务、现学现用(变化>>权衡)、发挥自身特色和优势
  - 演示：围绕论文自选内容(如、应用/技术/实验 | 复现/开发/分析 ..)

课程建设案例（上海交通大学/软件学院）

课程内容

48课时

= 16周

x3课时

———

1. 旧

2. 新

3. SHOW

课程建设案例（上海交通大学/软件学院）

课程内容

48课时

= 16周

x3课时

——

1. 旧

2. 新

3. SHOW

基础

1. 分布式系统简介
2. 顺序一致性
3. 最终一致性
4. 原子性/日志
5. 分布式事务处理
6. 分布式提交协议
7. 分布式共识算法

系统

8. 分布式文件系统
9. 分布式键值存储
10. 数据并行计算系统
11. 分布式图计算系统
12. 分布式数据划分
13. 分布式任务调度
14. 分布式容错方法
15. 分布式系统安全(可选)
16. 新网络硬件优化(可选)

课程建设案例（上海交通大学/软件学院）

课程内容

48课时

= 16周

x3课时

——

1. 旧

2. 新

3. SHOW

基础

1. 分布式系统简介
2. 顺序一致性
3. 最终一致性
4. 原子性/日志
5. 分布式事务处理
6. 分布式提交协议
7. 分布式共识算法

存储

系统

8. 分布式文件系统
9. 分布式键值存储
10. 数据并行计算系统
11. 分布式图计算系统
12. 分布式数据划分
13. 分布式任务调度
14. 分布式容错方法
15. 分布式系统安全(可选)
16. 新网络硬件优化(可选)

课程建设案例（上海交通大学/软件学院）

课程内容

48课时

= 16周

x3课时

——

1. 旧

2. 新

3. SHOW

基础

1. 分布式系统简介
2. 顺序一致性
3. 最终一致性
4. 原子性/日志
5. 分布式事务处理
6. 分布式提交协议
7. 分布式共识算法

计算

系统

8. 分布式文件系统
9. 分布式键值存储
10. 数据并行计算系统
11. 分布式图计算系统
12. 分布式数据划分
13. 分布式任务调度
14. 分布式容错方法
15. 分布式系统安全(可选)
16. 新网络硬件优化(可选)

课程建设案例（上海交通大学/软件学院）

课程内容

48课时

= 16周

x3课时

——

1. 旧

2. 新

3. SHOW

基础

1. 分布式系统简介
2. 顺序一致性
3. 最终一致性
4. 原子性/日志
5. 分布式事务处理
6. 分布式提交协议
7. 分布式共识算法

系统

8. 分布式文件系统
9. 分布式键值存储
10. 数据并行计算系统
11. 分布式图计算系统
12. 分布式数据划分
13. 分布式任务调度
14. 分布式容错方法
15. 分布式系统安全(可选)
16. 新网络硬件优化(可选)

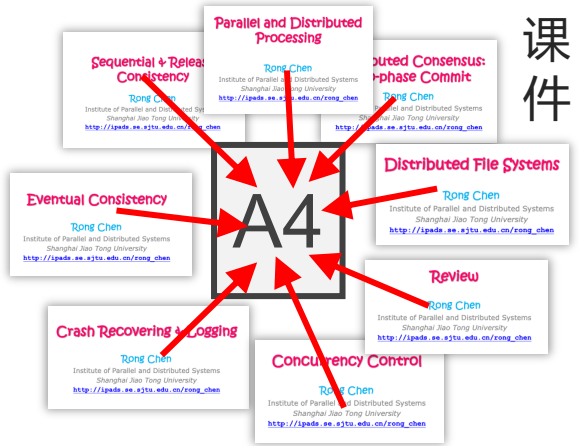
高级

课程建设案例（上海交通大学/软件学院）

考核方式

- 书面考核(50%)：问题分析、方法设计、优化丨权衡、...（半开卷）
- 团队SHOW(40%)：论文讲解、代码/功能/实验演示（评分：教师 + 学生）
- 平时成绩(10%)：论文预习问答、团队SHOW提问

半开卷：一张A4

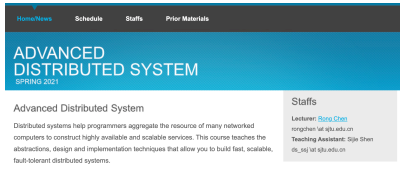


课件

新工具辅助

二维码 (问卷/投票)

课程主页



上交CANVAS系统

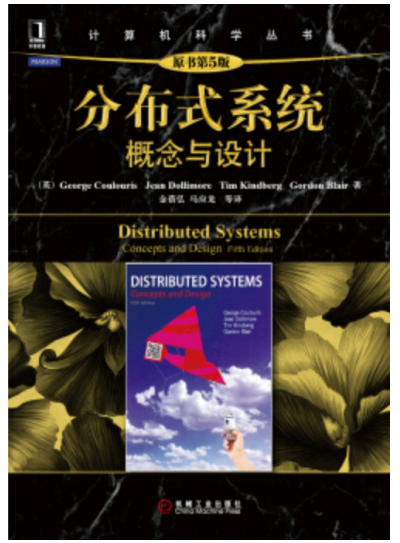
课程建设案例（上海交通大学/软件学院）

课程材料

- 课件/PPT（公开）
- 无指定教材(提供自学参考读物)

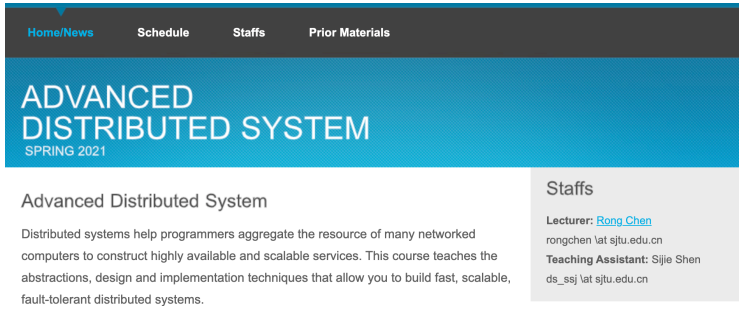
  《分布式系统：概念与设计（原书第5版）》

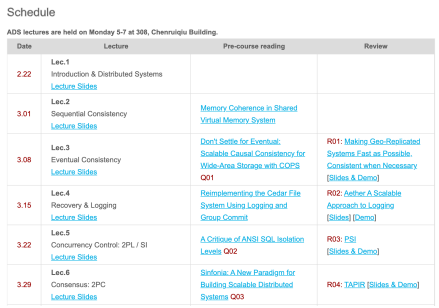
  George Coulouris, Jean Dollimore, Time Kindberg, Gordon Blair

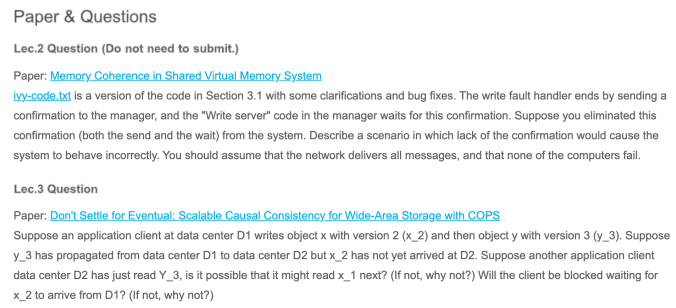


- 课程主页：https://ipads.se.sjtu.edu.cn/courses/ads/index.shtml

  历年资料：课件、论文、团队SHOW







课程建设案例（上海交通大学/软件学院）

教学效果

学生评教结果[1]

- 2021年春：评教档次：A1（全校排名：130/1047、学院排名：10/164）
- 2019年春：评教档次：A1（全校排名：312/1354、学院排名：28/175）

最高档次

评教内容：

教学规范 | 教学内容 | 教学方法 | 师生互动 | 教学效果 | 教学资源

建议在之后教学中改进

后续科研基础

如：2018春学生谢夏婷 | 基于RDMA的分布式图数据动态迁移（ATC 2019）

2017春学生王思源 | 基于RDMA+GPU的分布式图查询优化（ATC 2018）

1. 2018-2019学年起可查询学生评教结果，2020年因疫情未进行评教

小节

分布式系统成为云计算、大数据、人工智能等新兴领域的关键支撑

《分布式系统》是软件工程一级学科研究生核心课程之一

课程建设应充分考虑课程特点、学生情况、培养目标

特色化道路（上海交通大学软件学院）

- “经典”与“前沿”相结合：以“主题”为中心，采用“新旧”对比
- “指定”与“自选”相结合：半自选团队SHOW(运用所学、发挥特长)

五年课程建设初步成型，内容/方法/工具等仍需保持与时俱进

衷心感谢 敬请不吝指正