

Introduction to Software Engineering

Jianyong Wang(王建勇)

Department of Computer Science and Technology
Tsinghua University, Beijing, China

群聊：2024秋季软工课程群



该二维码7天内(9月14日前)有效，重新进入将更新

尚未进入《软件工程2024》微信群的同学请尽快扫码或找身边同学邀请入群

叁课堂



王建勇, 计算机系软件所
研究方向: 数据挖掘

清华大学自强科技楼1-812
E-mail: jianyong@tsinghua.edu.cn

授课地点: 建华楼A406
授课时间: 周一第2节 (秋季)
(9:50-12:15)
腾讯会议号: 538-600-6426

裴丹, 计算机系网络所
研究方向: 智能运维

清华大学自强科技楼1-1014
E-mail: peidan@tsinghua.edu.cn

授课地点: 待定
授课时间: 待定 (春季)

李丹, 计算机系网络所
研究方向: 数据中心网络, 网
络智能, 可信任互联网
清华大学 FIT楼 4-104
E-mail: lidan@tsinghua.edu.cn

授课地点: 待定
授课时间: 待定 (春季)

About the Instructor

- Jianyong Wang, Professor
 - IEEE Fellow 、 CAAI(中国人工智能学会) Fellow、 BAAI PI (北京智源人工智能研究院智源学者)
- Institute of Software, Department of Computer Science and Technology, Tsinghua University
- Homepage: <http://dbgroup.cs.tsinghua.edu.cn/wangjy/>
 - Email: jianyong@tsinghua.edu.cn
- Main Research area: Data Mining and Knowledge Discovery
 - Medical Data Mining
 - Knowledge Graph
 - Interpretable Machine Learning

My Research Group

- Currently advising 10 Ph.D. students, 1 M.S. student
- Graduated 10 PhDs & 18 Masters (5 became assistant/associate、full professors)



大作业助教/指导教师

■ 教师

- 王建勇
 - jianyong@tsinghua.edu.cn
- 金奕江(平台运维)
 - yjjin@tsinghua.edu.cn

■ 随堂助教

- 徐嘉诚 (xujc20@mails.tsinghua.edu.cn)

■ 大作业助教

- 董博文 (首席, dongbw18@163.com)
- 钱厚德 (ashitemaru.holder@gmail.com)
- 高焕昂 (cc7w@foxmail.com)
- 王博文 (wangbw21@mails.tsinghua.edu.cn)
- 刘铠铭 (lkm20@mails.tsinghua.edu.cn)
- 郑友捷 (zhengyj21@mails.tsinghua.edu.cn)

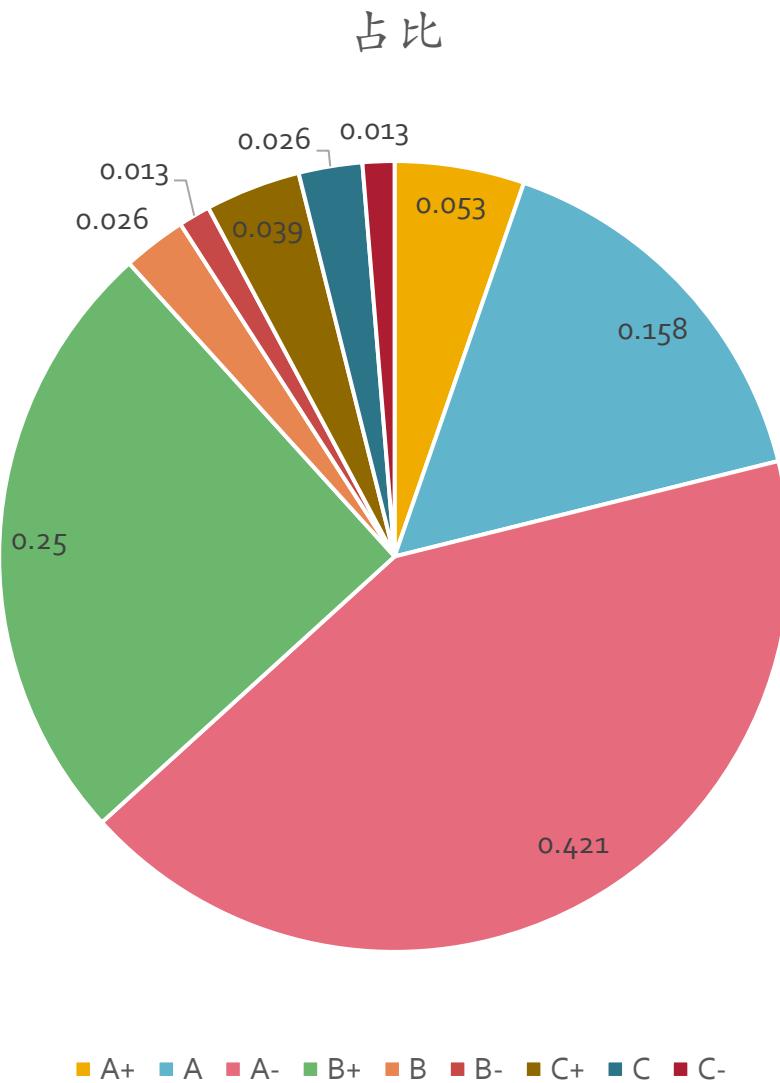
小作业助教

小作业/平台助教	助教
CI/CD 小作业	王博文 (wangbw21@mails.tsinghua.edu.cn)
前端小作业助教	钱厚德 (ashitemaru.holder@gmail.com)
后端小作业助教	高焕昂 (me@c7w.tech)

Grading

- 3 Personal Mini-projects 25%
 - 5% (CI/CD)
 - 8% (后端)
 - 12% (前端)
- 1 Team Project 60%
- 1 Team Project Demo 5%
- In-class quizzes 10%
 - 1%*8 (小测验)
 - 1%*2 (业界分享签到)

2022秋季软工课的成绩分布情况

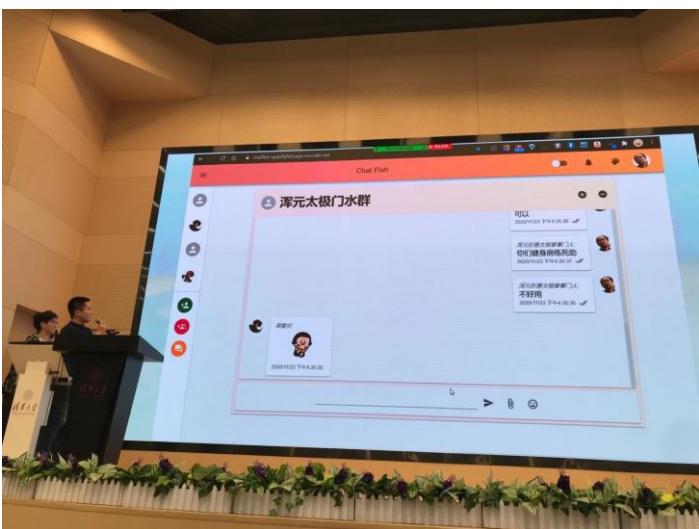


A+/A/A-: 占比63.2%

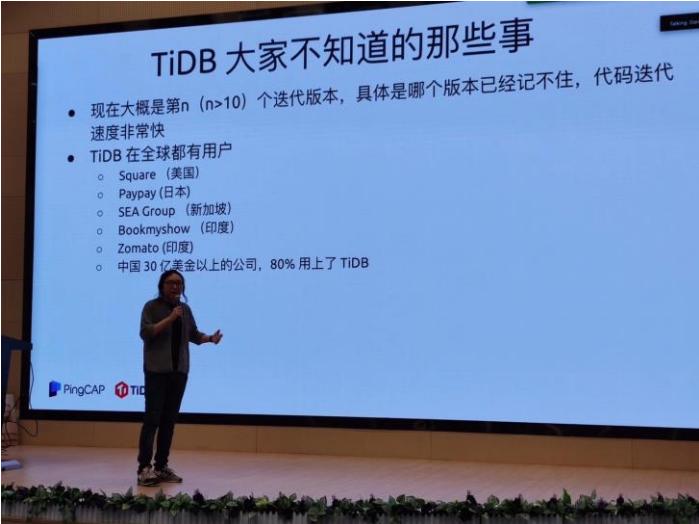
大作业概况

编号	题目	来源	助教
1	智能体比赛	计算机系学生 科协	郑友捷
2	即时通讯系统	字节跳动（复用）	董博文、钱厚德、高焕昂、王博文、刘铠铭

大作业展示



业界分享报告



教学日历

周	日期	授课内容	备注
01	09.09	1、课程简介（1学时） 2、SECODER 平台、大作业内容安排和要求、后端开发讲解及实践（2学时）	1、其中王博文负责讲解 SECODER、大作业要求及“即时通讯”大作业，郑友捷负责讲解“智能体”大作业、高焕昂负责讲解后端开发
02	09.21 <small>(周六)</small>	1、前端开发及 CICD 讲解与实践（3学时）	1、中秋调课：（改上 09.16）课； 2、钱厚德和王博文分别负责前端和 CICD 的讲解 3、课后启动大作业分组，提交分组选题。9月 28 日（周六）确定选题分组名单，设置分组帐号，配置分组项目。【金老师】
03	09.23	1、敏捷开发（1学时） 2、软件测试（2学时）	
04	09.30	1、软件测试（1学时） 2、第一次大作业讲解（分组指导，2学时）	
05	10.07	无课（国庆节放假）	
06	10.14	1、软件测试（1学时） 2、第二次大作业讲解（分组指导，2学时）	
07	10.21	1、软件测试（1学时） 2、代码风格与重构（2学时）	
08	10.28	1、软件架构（1学时） 2、第三次大作业讲解（分组指导，2学时）	
09	11.04	1、软件架构（3学时）	
10	11.11	1、软件架构（1学时） 2、第四次大作业讲解（分组指导，2学时）	
11	11.18	1、软件架构（2学时） 2、需求管理（1学时）	
12	11.25	1、需求管理（1学时） 2、第五次大作业讲解（分组指导，2学时）	
13	12.02	1、需求管理（1学时） 2、运维管理（2学时）	
14	12.09	1、工业界特邀报告（题目待定）	1、12.15 提交大作业
15	12.16	1、工业界特邀报告（题目待定）	1、12.20（周五）各小组交付最终系统、提交文档给助教； 2、12.22（周日）教师和大作业助教大作业（含文档评分）打分沟通会
16	12.23	1、大作业集中展示汇报	1、12.29（周日）助教提交大作业成绩给随堂助教

项目准备

迭代

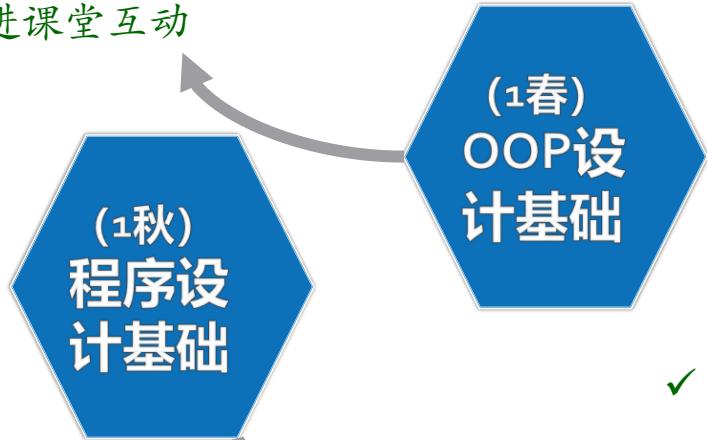
开发

产品交付

清华大学计算机系软件系列课程

1. Software development process
2. Working with non-technical customers
3. Testing
4. Legacy code and change

- ✓ 培养OO思维，拓展专业视野
- ✓ “小教员+助教”立体辅导
“自主学习”结合“帮扶带”
- ✓ 平时实践75%作为基础
上机考试25%体现差异
- ✓ 引入报告环节
增进课堂互动



- ✓ 课号合并，2夏留白；平行课堂，统一标准
围绕程序设计训练，各有侧重
- ✓ 语言：QT + Python + Java
- ✓ 模块：图形、网络通讯、多线程、
Web、RegEx、...

- ✓ 探索个性化培养与考核
针对实际，落实因材施教

- ✓ 立足C/C++，回归基础
- ✓ 强化测试意识、调试能力
- ✓ 尽早培养、规范编程风格
- ✓ 多设习题课，促进朋辈学习

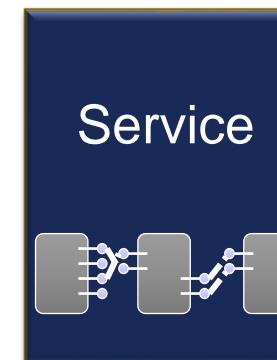
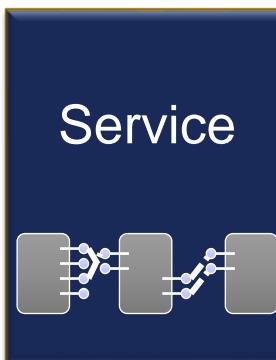
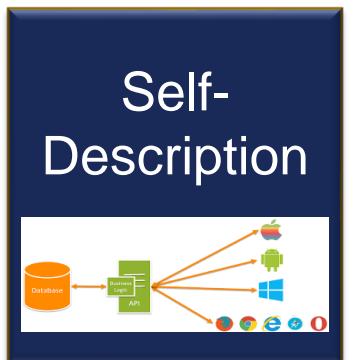
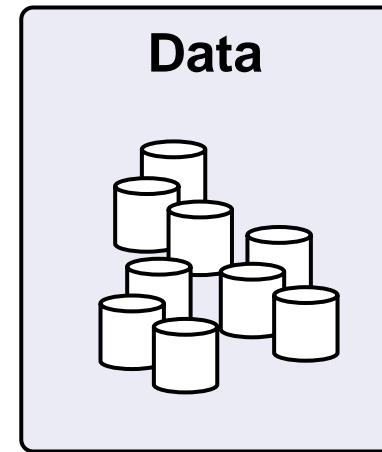
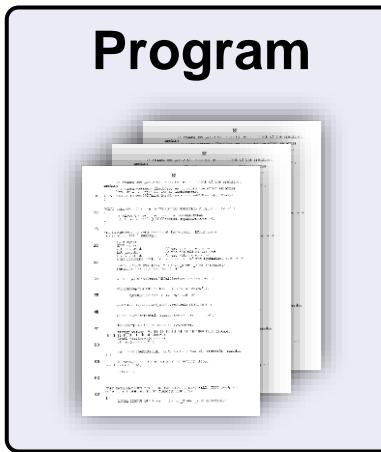
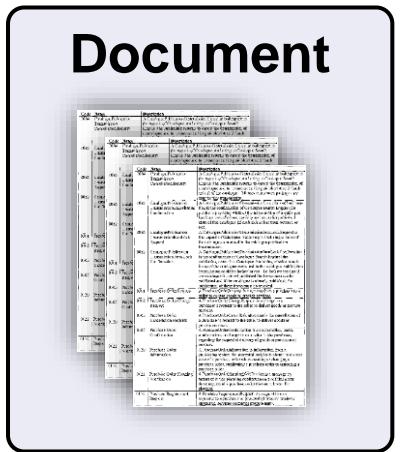
- ✓ 改为2春、3秋同开，
降低3秋整体课时
- ✓ 精选内容
重新分配学时
- ✓ 强化过程性
训练与考核
- ✓ 调整挑战性
提高参与度
- ✓ 增加随堂小测验
- ✓ 增加个人小作业

Objectives

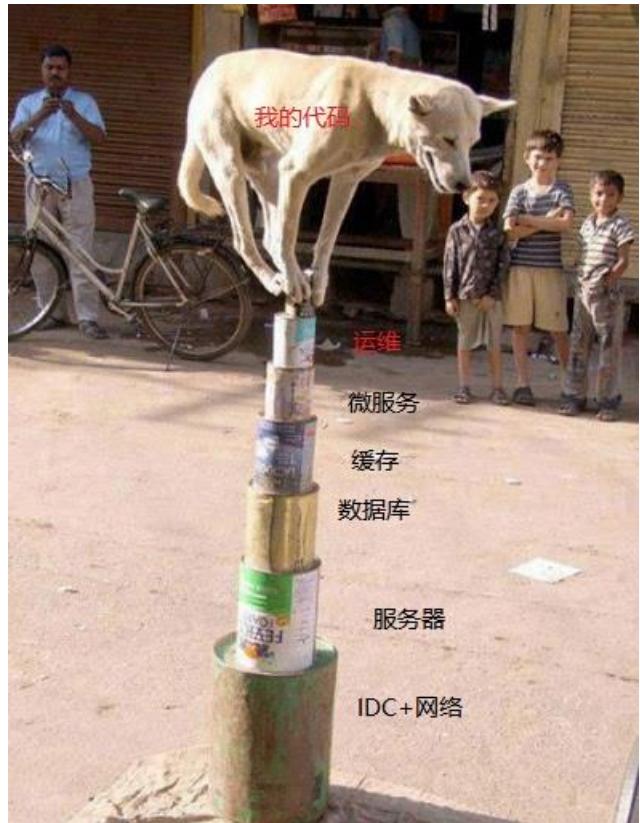
#1: To Understand Software As Systems

#2: To Build Software As Engineers

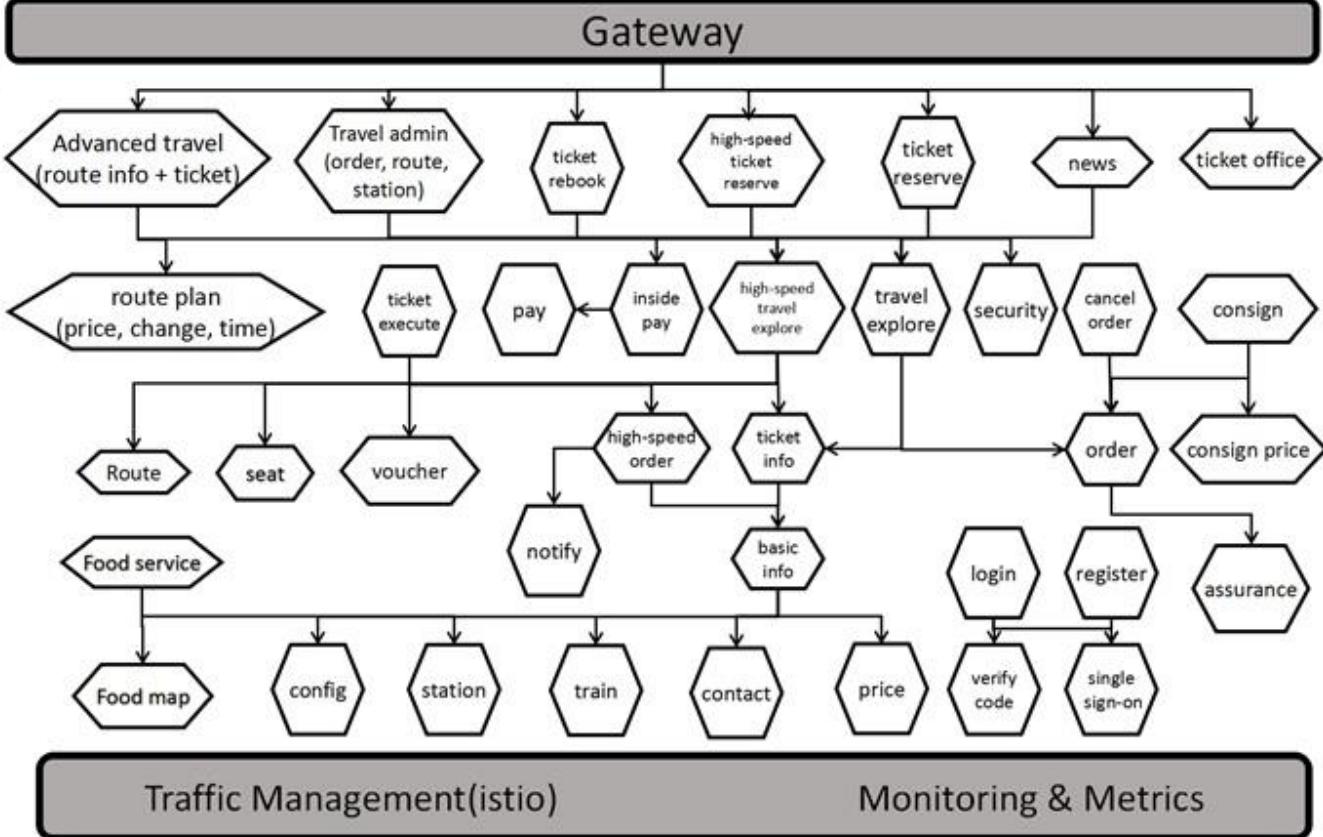
What is Software?

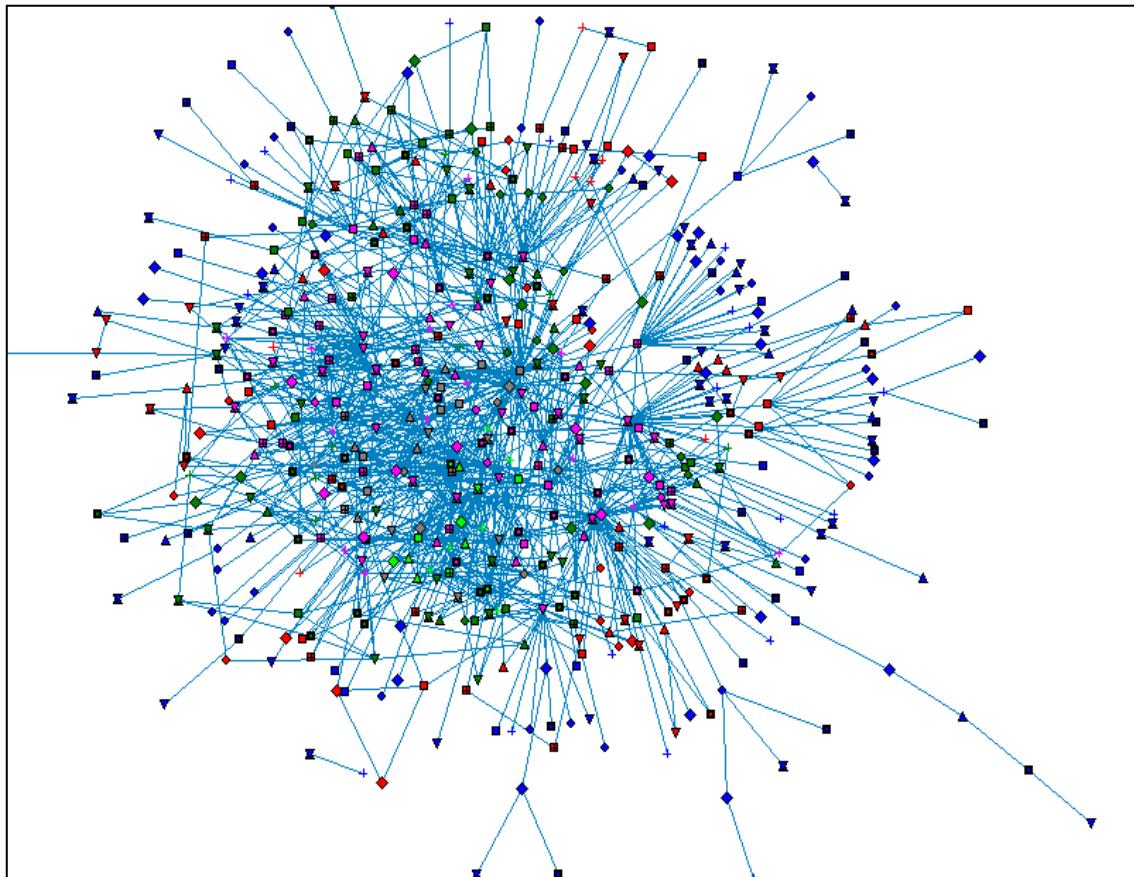


Code vs System

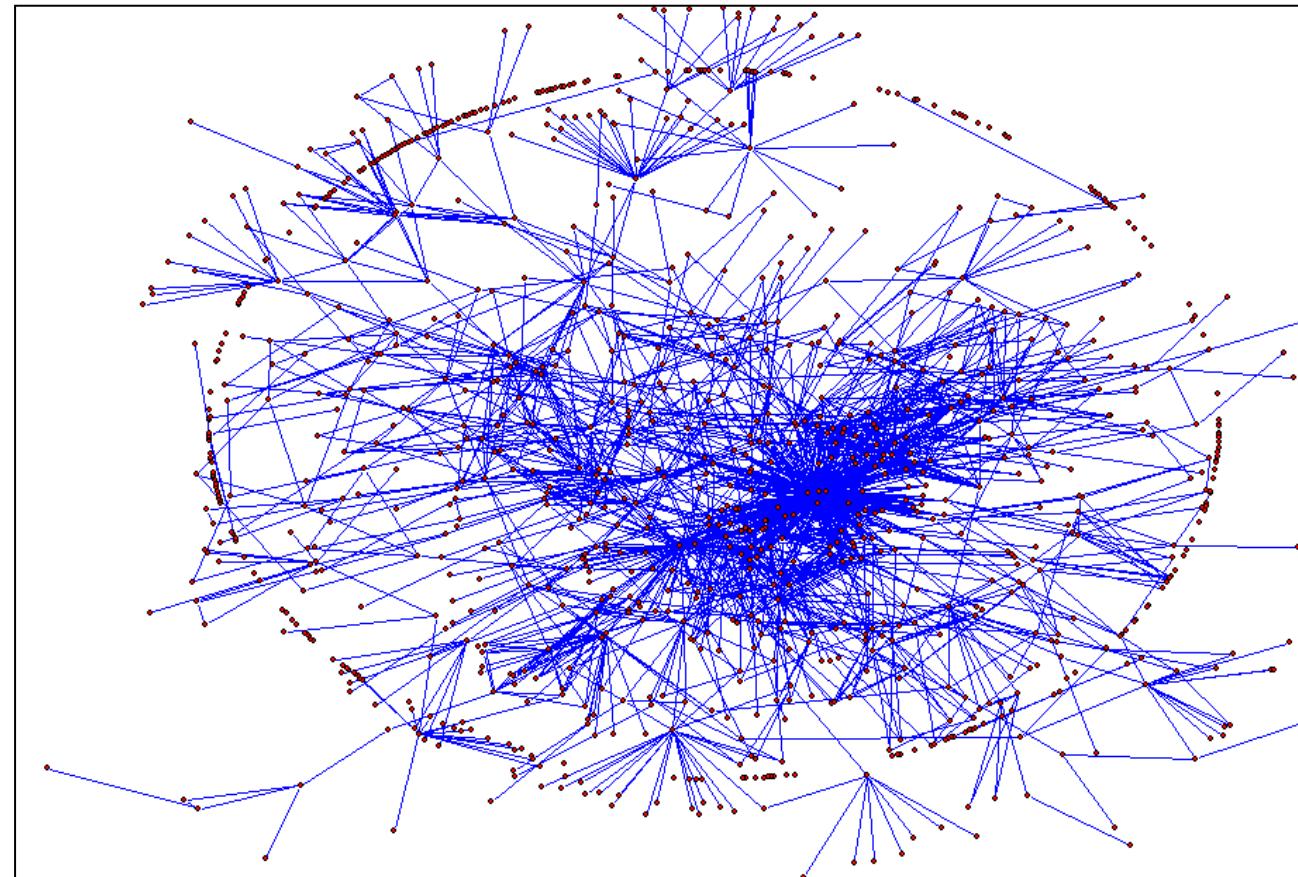


Service Discovery (k8s)
Service Registry
Load Balance

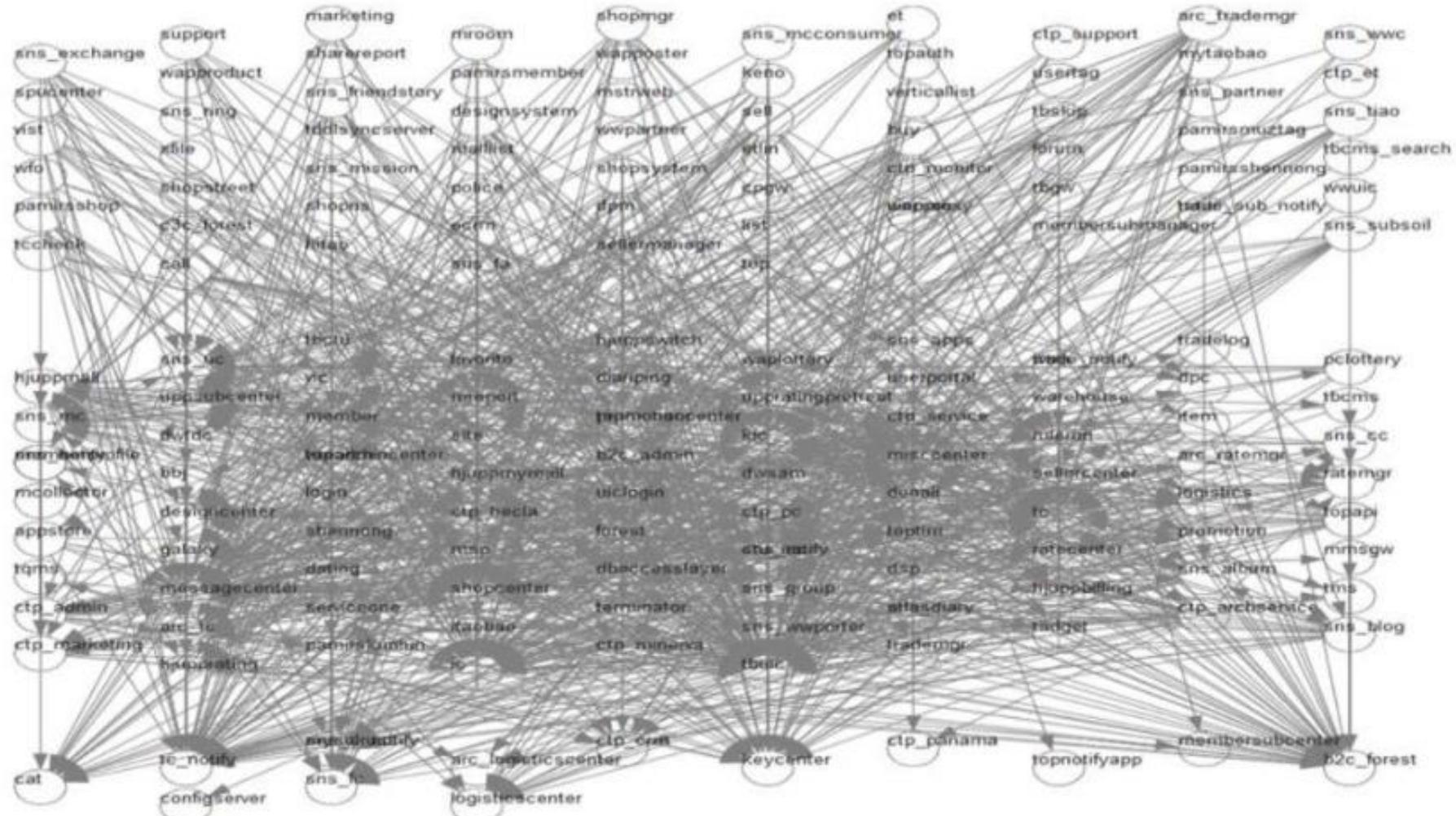




Linux内核有630个函数，存在1841个函数调用。图中将函数表示为节点，调用关系表示为边



应用服务器支撑软件Tomcat，含1019个类，类与类之间有2109个继承或聚合关系。图中将类表示为节点，关系表示为边

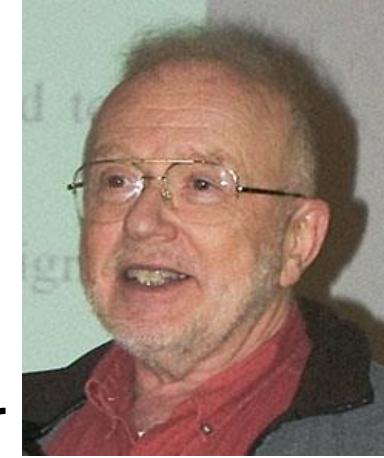


2012 淘宝核心链路应用拓扑图

What is SW engineering (vs. programming)?

“Multi-person development of multi-version programs”

“...combining separately written programs and making them suitable for use by people who had not written them...These topics usually received little or no attention in traditional programming courses”



David Lorge
Parnas



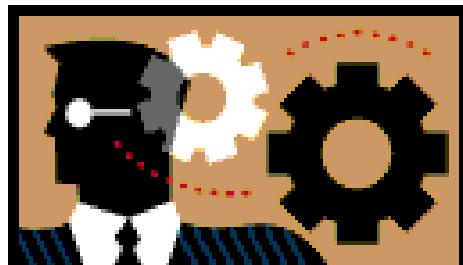
-Fred
Brooks, Jr.
(1999 Turing
Award
winner)

Software Engineering

Software Engineering is the establishment & use of sound **engineering principles** in order to obtain economically software that is reliable & work efficiently on real machines.

--- Fritz Bauer, 1969, NATO

“The application of a **systematic, disciplined, quantifiable** approach to the **development, operation, and maintenance** of software; that is, the application of engineering to software.”

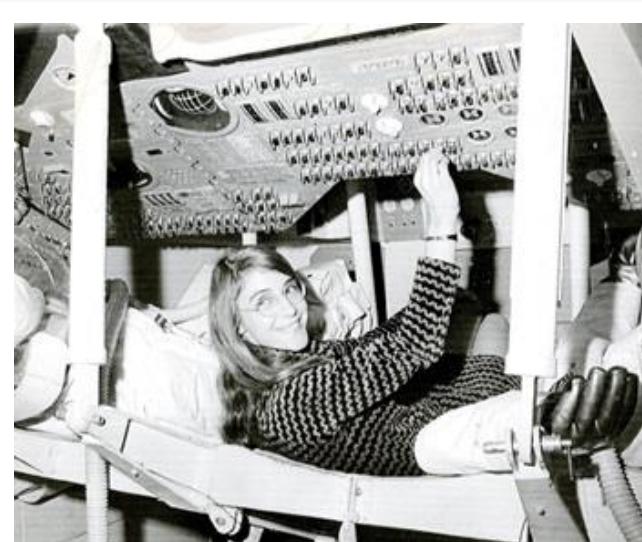


--- “IEEE Standard Glossary of Software Engineering Terminology”, IEEE, Piscataway, NJ std 610.12-1990, 1990.

Margaret Hamilton coined term “Software Engineering”



Margaret Hamilton, lead software engineer of the Apollo Project, stands next to a huge stack of code written by her and her team, in 1969.



Margaret Hamilton in an Apollo Command Module



Apollo 8
software team
at MIT, c.1965

- When was the term software engineering first introduced?
 - The term ‘software engineering’ was suggested at conferences organized by NATO in 1968 to discuss the ‘software crisis’. The software crisis was the name given to the difficulties encountered in developing large, complex systems in the 1960s.

What is *missing* from our student's technical skills?

1. Software development process
2. Working with non-technical customers
3. Testing
4. Legacy codes
5. And many others, such as software architecture, requirement engineering, operations

Software Process Spectrum

- Waterfall
- Incremental
- Prototyping
- Spiral

Process Models



Ad Hoc development

Light-weight Process

- Agile

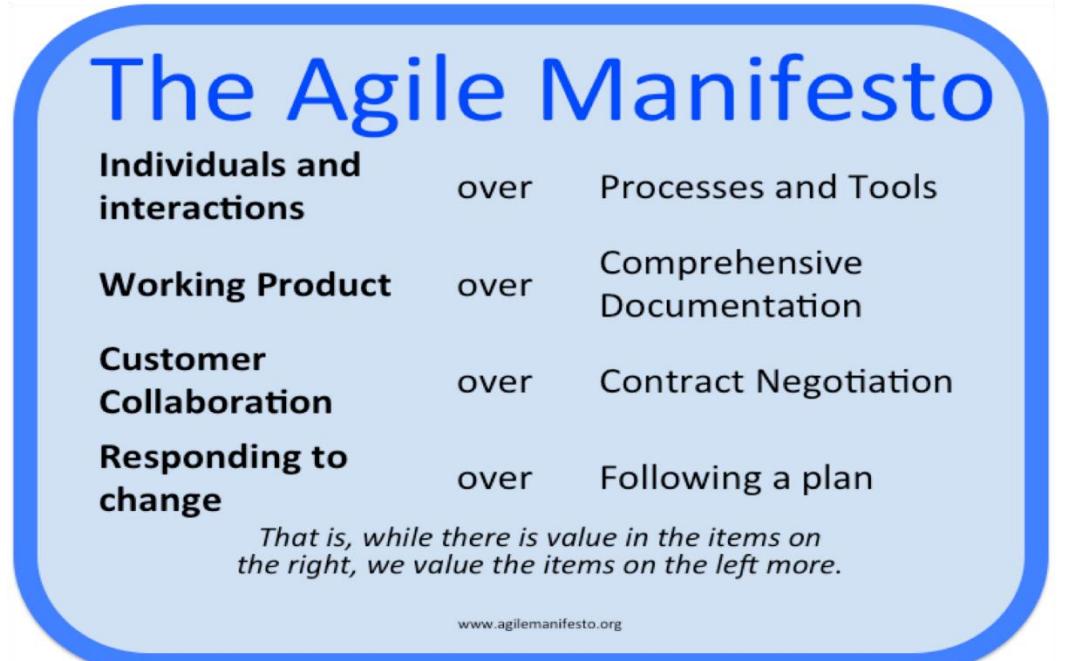


- CMM
- SPICE

Heavy-weight Process



Agile Software Development



Changing requirements

*And the users exclaimed with a laugh and a taunt:
“It’s just what we asked for, but not what we want.”*
—Anonymous

Software’s strength is its evolvability—
its ability to change and adapt—
but that’s a poor fit for “big design up front” (BDUF)
or “top down” approaches



Expensive Bugs



World

Boeing engineers blame cheap Indian software for 737 Max problems

Boeing engineers say 737 Max software outsourced from India's HCL Technologies & Cyient. HCL denies link to troubles of 737 Max, two of which crashed.

The Max planes were involved in two deadly crashes

The need for testing



产品还没测试直接投入生产时



测试的时候一切OK, 真正上线的时候...



The need for testing

阿丽亚娜火箭失事

<https://www.bilibili.com/video/BV127411L7ob?from=search&seid=16865301561105523756>

<https://zhuanlan.zhihu.com/p/114486843>

1996年6月4日。“阿丽亚娜-5”运载火箭进行首次测试发射，结果是失败的。在火箭发射后的第37秒就被迫自爆。经过检查后发现，原因竟然是因为控制火箭飞行的软件发生了故障。资料显示由于64位浮点运算错误的变为16位整数运算，造成程序崩溃后处理器发生数据溢出，读取感测角度水平数据运算时错误地读入垂直数据，导致火箭在高速下进行了90度水平滚转而崩溃，触发了自毁装置。造成了非常大的损失。

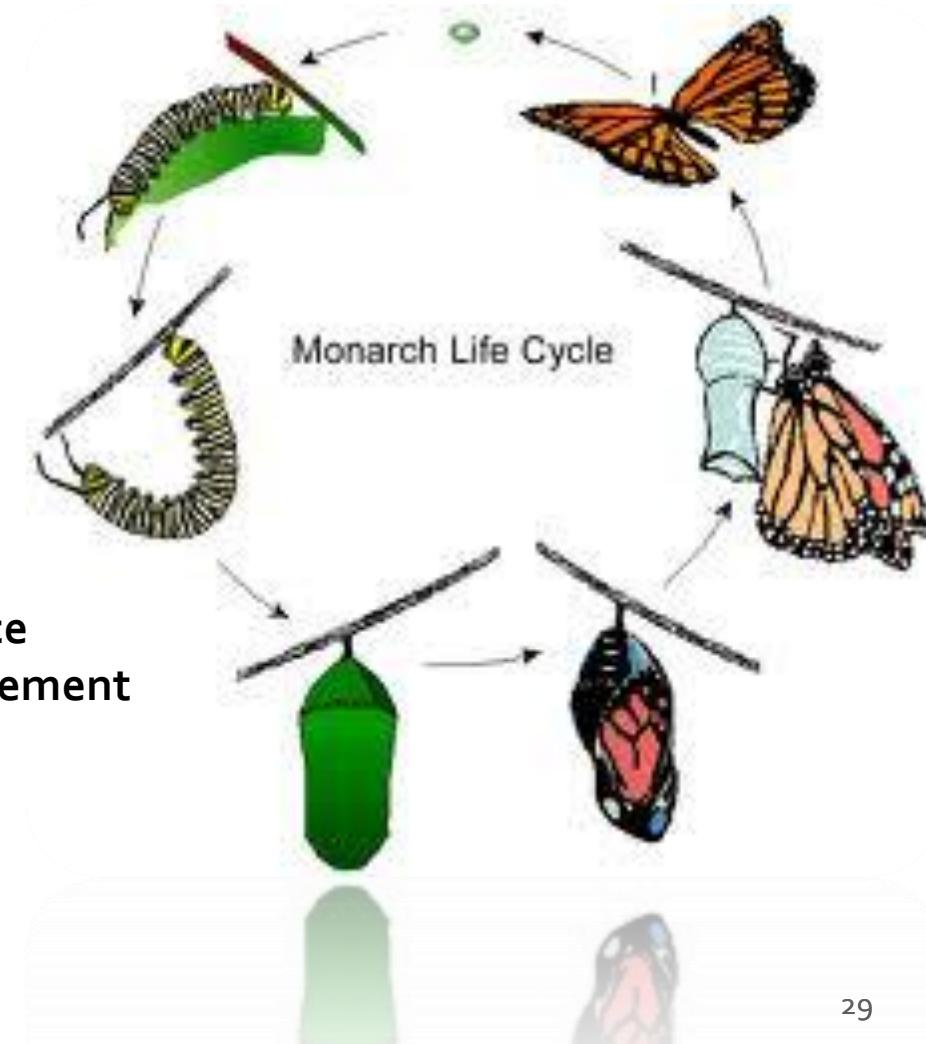


Software Life cycle

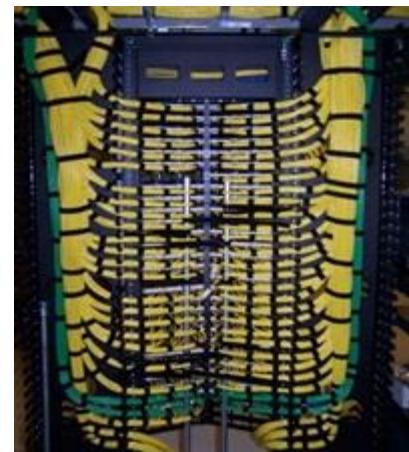
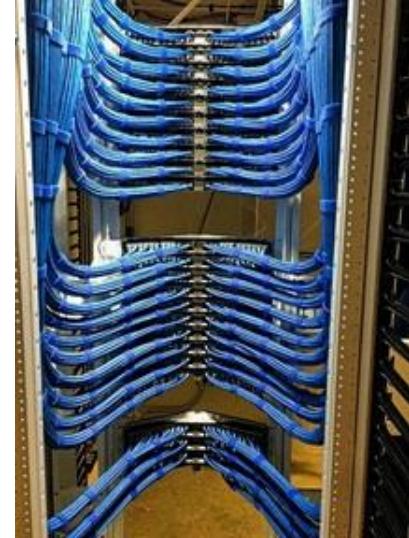
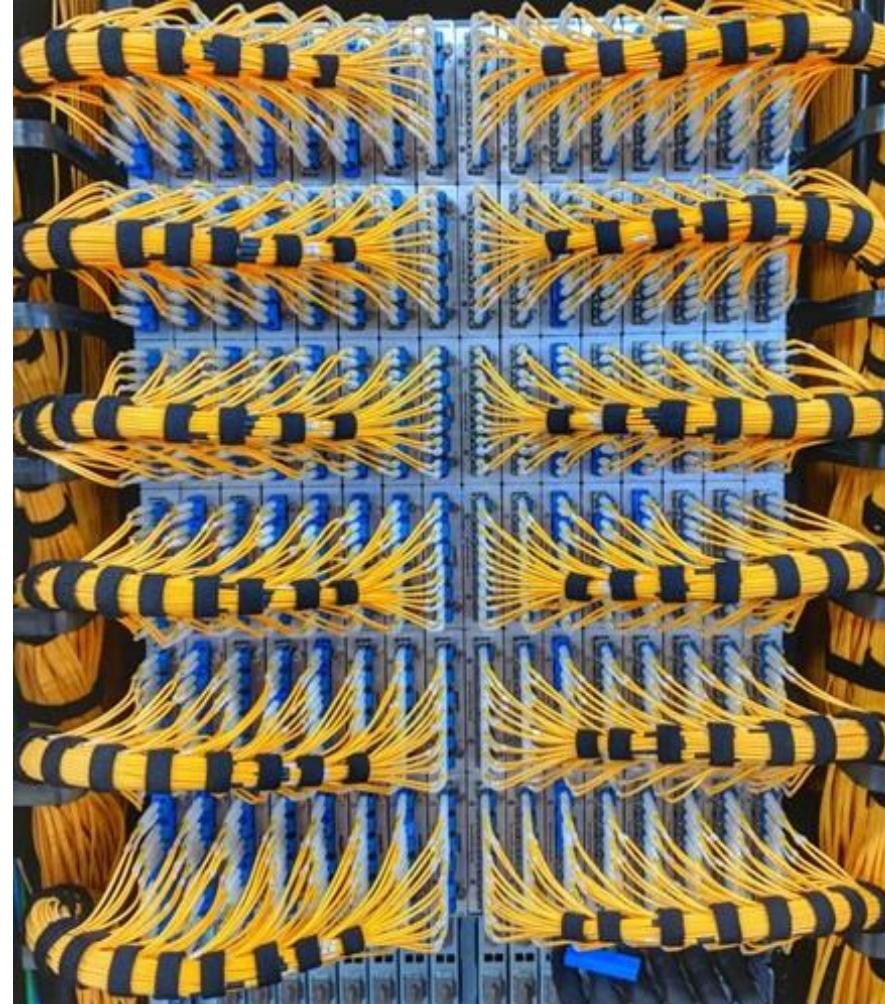
- Requirements
- Specification (Analysis)
- Design
- Implementation
- Integration
- **Maintenance**
- Retirement

“60/60 rule”:

- 60% of software cost is maintenance
- 60% of maintenance cost is enhancement



Legacy SW vs. Beautiful SW



Legacy Code



偶然间看到自己多年前写的代码



当我试图清理几行所谓的旧代码的时候



Software Change

- Software upgrade



Introduce
new feature



Improve
performance



Fix bugs

- Configuration change

- e.g., traffic switching for load balancing reasons

- Occurs frequently

- 10K+ per day in Baidu

Impact of Erroneous Software Upgrades

2012.10, Google

Google Apps Incident Report

Gmail Partial Outage - December 10, 2012

Prepared for Google Apps customers

The following is the...

- An update to Google's load balancing software
- Poor performance to Gmail for 18 minutes

Issue

For a brief period of time, users experienced issues with Gmail. The number of users receiving errors increased over time.

Actions and Root Cause

Background: The load balancing software routes users' requests to Google data centers around the world for processing and serving content, such as search results and email.

Between 8:45 AM PT and 9:13 AM PT, a routine update to Google's load balancing software was rolled out to production. A bug in the software update caused it to incorrectly interpret a portion of Google data centers as being unavailable. The Google load balancers have a failsafe mechanism to prevent this type of failure from causing Google-wide service degradation, and they continued to route user traffic. As a result, most Google services, such as Google Search, Maps, and AdWords, were unaffected. However, some services, including Gmail, that require specific data center information to efficiently route users' requests, experienced a partial outage.

2014.11, Microsoft Azure

Update on Azure Storage Service Interruption

WEDNESDAY, NOVEMBER 19, 2014



JASON ZANDER
CVP, Microsoft Azure Team

- A performance update to Azure Storage
- Reduced capacity across services utilizing Azure Storage

Wednesday, November 19,

As part of a performance update to Azure Storage, an issue was discovered that resulted in reduced capacity across services utilizing Azure Storage, including Virtual Machines, Visual Studio Online, Websites, Search and other Microsoft services. Prior to applying the performance update, it had been tested over several weeks in a subset of our customer-facing storage service for Azure Tables. We typically call this "flighting," as we work to identify issues before we broadly deploy any updates. The flighting test demonstrated a notable performance improvement and we proceeded to deploy the update across the storage service. During the rollout we discovered an issue that resulted in storage blob front ends going into an infinite loop, which had gone undetected during flighting. The net result was an inability for the front ends to take on further traffic, which in turn caused other services built on top to experience issues.

Impact of Erroneous Configuration Changes

2014.1. Dropbox

Outage post-mortem

Akhil Gupta | January 12, 2014 |

- Planned maintenance to upgrade the OS on some machines
- Dropbox service been down for three hours

We use thousands of machines for redundancy. We have one base file master and two replica machines for redundancy. We also have incremental data backups and store them in a separate environment.

On Friday at 5:30 PM PT, we had a planned maintenance scheduled to upgrade the OS on some of our machines. During this process, the upgrade script checks to make sure there is no active data on the machine before installing the new OS.

A subtle bug in the script caused the command to reinstall a small number of active machines. Unfortunately, some master-replica pairs were impacted which resulted in the site going down.

2014.6, Facebook

Facebook outage caused by software system update

20 June 2014 | By Hollie Luxford



[in Share](#) [Tweet 0](#) [Like 0](#)

Social networking site Facebook suffered a worldwide outage yesterday after an issue while updating the configuration of one of its software systems.

The worldwide outage began at 1:50 AM PT on June 20.

Facebook

- Update the configuration of the software systems
- Failed Facebook for 31 minutes

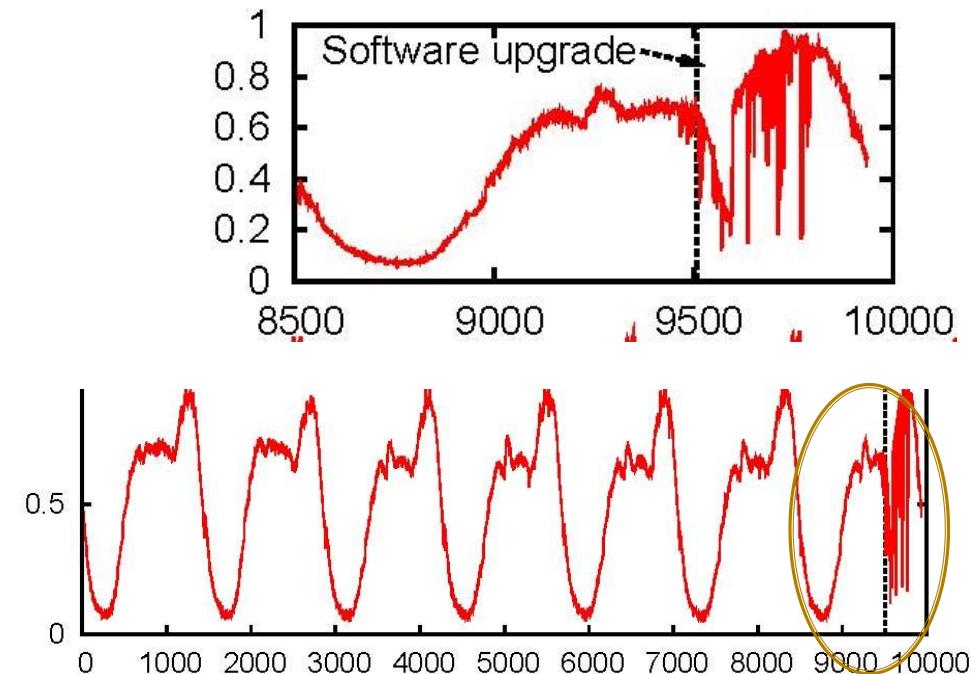
Case Study: An Erroneous Software Upgrade in Advertising

Methodology

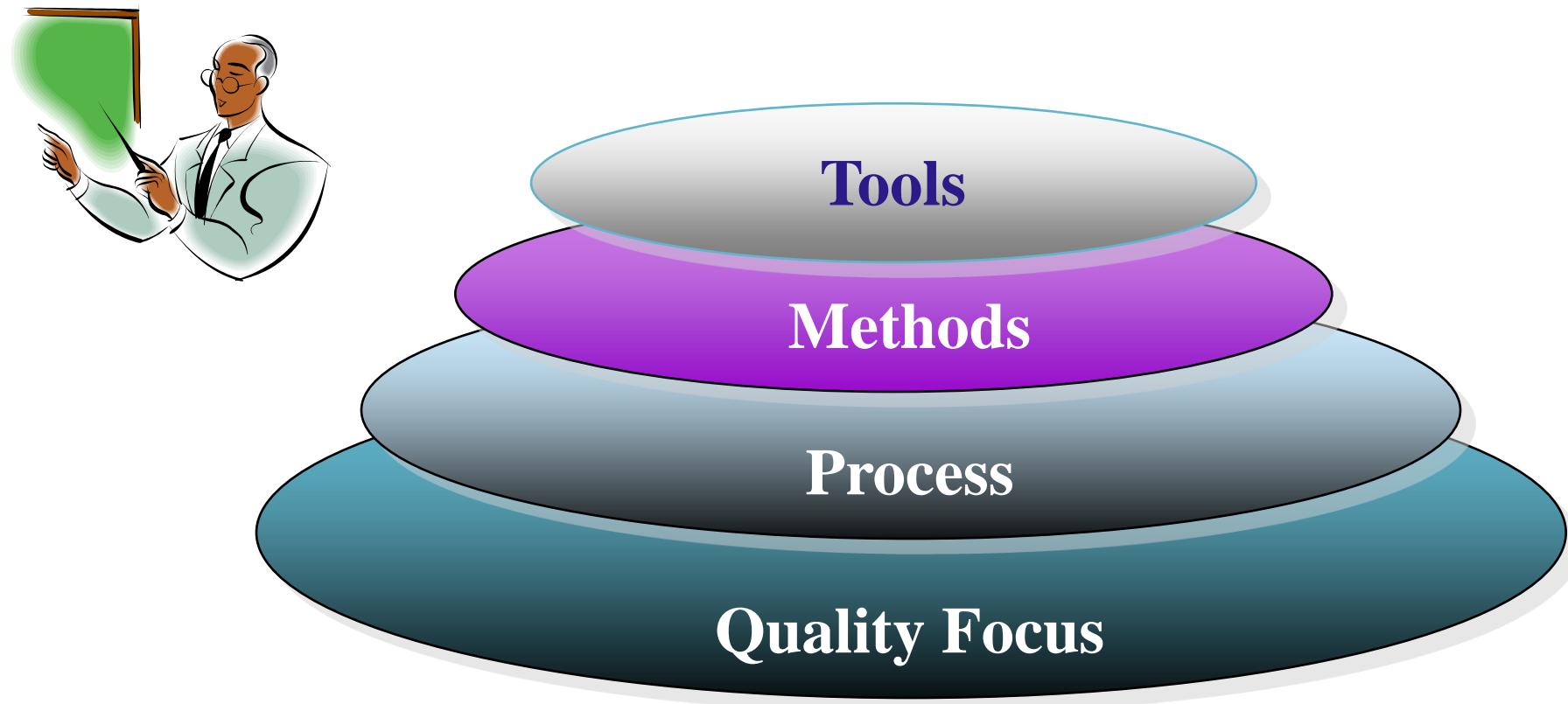
- A fraction of software changes
- Not deliver the results to the operators
- The operators assess software changes independently

Manual troubleshooting:

- 1.5 hours



Software Engineering: A Layered Technology



SE Body of Knowledge

- 15 SWEBOK key areas[IEEE]

Software Foundations

- Computing Foundations
- Mathematical Foundations
- Engineering Foundations

Software Management

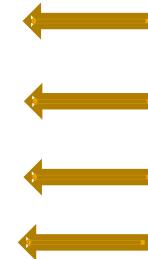
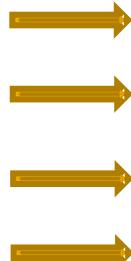
- Software Maintenance
- Configuration Management
- Engineering Management
- Engineering Process

Software Development

- Software Requirements
- Software Design
- Software Construction
- Software Testing

Software Professional Practices

- Engineering Economics
- Software Quality
- Engineering Methods
- Professional Practices



What Makes a Great Software Engineer?

P. Li, A. Ko, J. Zhu, Microsoft & Univ. of Washington, Proc. ICSE 2015

59 interviews with recognized SE experts at Microsoft

- Always improving
- Passionate
- Technically open-minded
- Data-driven

Personal Characteristics Decision Making

- Knows people & organization
- Updates mental models when learn new skills/facts/context
- Consider situation at multiple levels when making judgments
- Can reason about complex & intertwining ideas

- Helps others understand by tailoring explanations to them
- Creates shared success possibly via personal compromises
- Creates “safe haven” where others can learn from mistakes
- Gives honest feedback

Team Skills Technical Skills

- Elegant solutions
- Creative thinking when faced with limitations of current solutions
- Anticipates technical needs based on prior experience

Our goal is for you to learn
disciplined,
customer-focused,
agile software engineering

prepare you to function as member
of professional software engineering
community

Course Outline

Project
Practice, Practice, Practice



Towards Good Code

Version Control

Coding Style

xUnit

Towards Good System

Requirements

Design

Testing

Project

Process

Software Engineering: Practice, Practice and Practice!



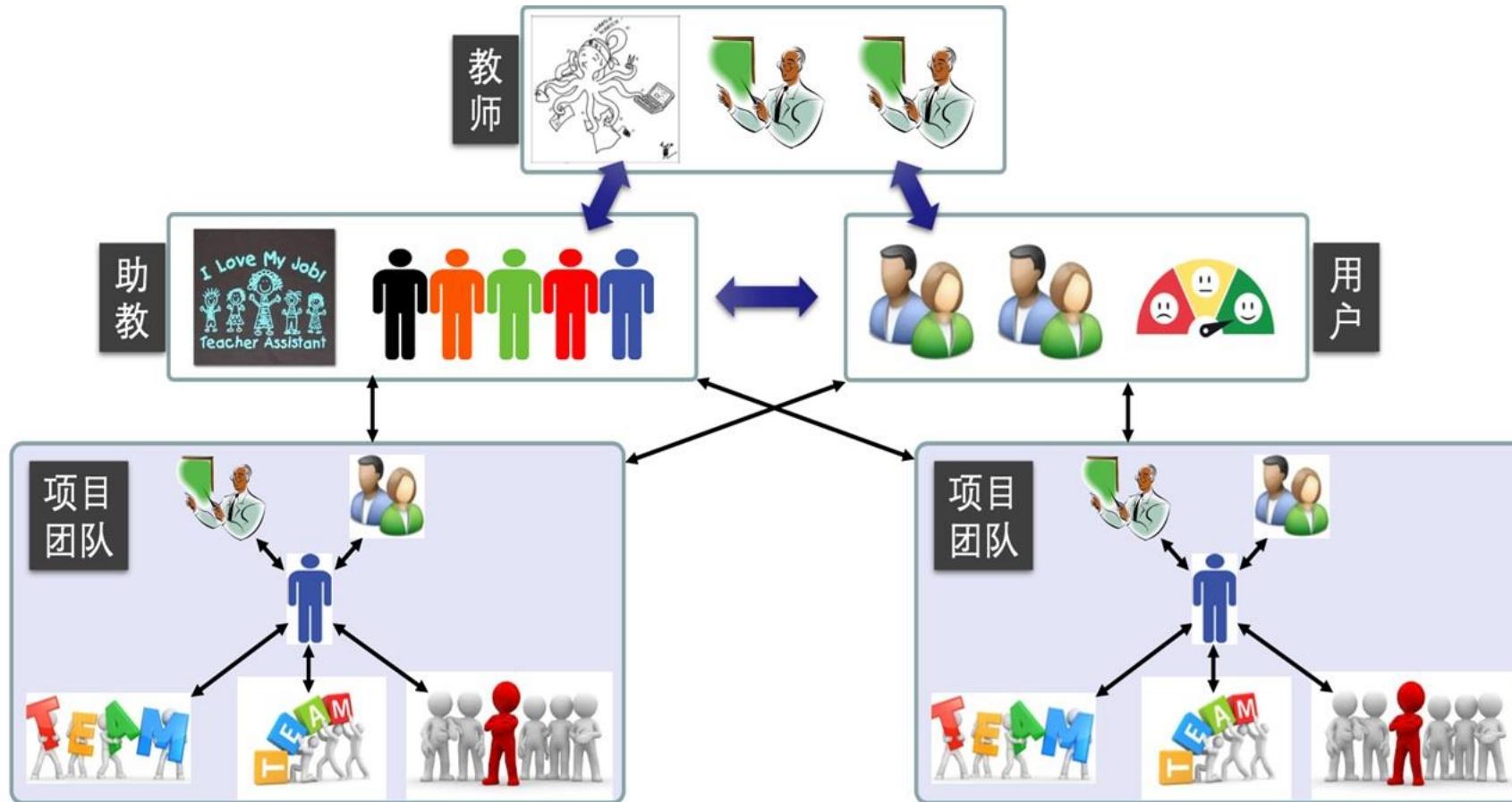
Team Project

If you want to be incrementally better: Be competitive.

If you want to be exponentially better: Be cooperative.



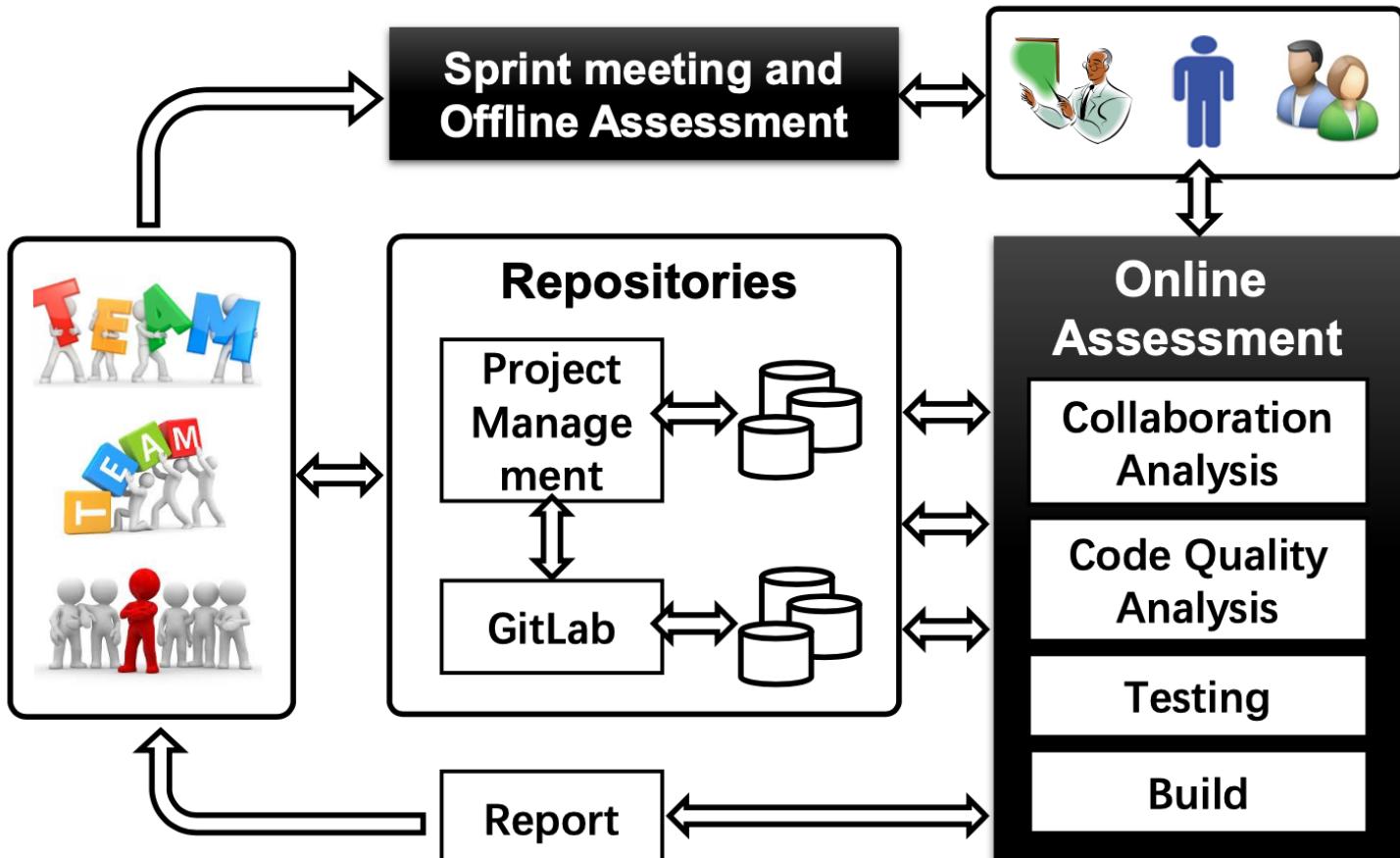
Team work!



今年项目管理
变化：(1)每组
3人；(2)不再设
用户代表

项目评价

线上+线下；团队+个人



- ◆ **Combined Assessment**
 - ◆ **Online automatic** evaluation based on the DevOps platform
 - ◆ **Offline manual** evaluation based on regular Sprint meetings and in person discussion
- ◆ **Continuous assessment**
 - ◆ Timely feedback to continuous deliveries
- ◆ **Personalized assessment**
 - ◆ For each team and individual students.

持续监控、反馈与评价

教学日历

周	日期	授课内容	备注
01	09.09	1、课程简介（1学时） 2、SECODER 平台、大作业内容安排和要求、后端开发讲解及实践（2学时）	1、其中王博文负责讲解 SECODER、大作业要求及“即时通讯”大作业，郑友捷负责讲解“智能体”大作业、高焕昂负责讲解后端开发
02	09.21 <small>（周六）</small>	1、前端开发及 CICD 讲解与实践（3学时）	1、中秋调课：（改上 09. 16）课； 2、钱厚德和王博文分别负责前端和 CICD 的讲解 3、课后启动大作业分组，提交分组选题。9月 28 日（周六）确定选题分组名单，设置分组帐号，配置分组项目。【金老师】
03	09.23	1、敏捷开发（1学时） 2、软件测试（2学时）	
04	09.30	1、软件测试（1学时） 2、第一次大作业讲解（分组指导，2学时）	
05	10.07	无课（国庆节放假）	
06	10.14	1、软件测试（1学时） 2、第二次大作业讲解（分组指导，2学时）	
07	10.21	1、软件测试（1学时） 2、代码风格与重构（2学时）	
08	10.28	1、软件架构（1学时） 2、第三次大作业讲解（分组指导，2学时）	
09	11.04	1、软件架构（3学时）	
10	11.11	1、软件架构（1学时） 2、第四次大作业讲解（分组指导，2学时）	
11	11.18	1、软件架构（2学时） 2、需求管理（1学时）	
12	11.25	1、需求管理（1学时） 2、第五次大作业讲解（分组指导，2学时）	
13	12.02	1、需求管理（1学时） 2、运维管理（2学时）	
14	12.09	1、工业界特邀报告（题目待定）	1、12. 15 提交大作业
15	12.16	1、工业界特邀报告（题目待定）	1、12. 20（周五）各小组交付最终系统、提交文档给助教； 2、12. 22（周日）教师和大作业助教大作业（含文档评分）打分沟通会
16	12.23	1、大作业集中展示汇报	1、12. 29（周日）助教提交大作业成绩给随堂助教

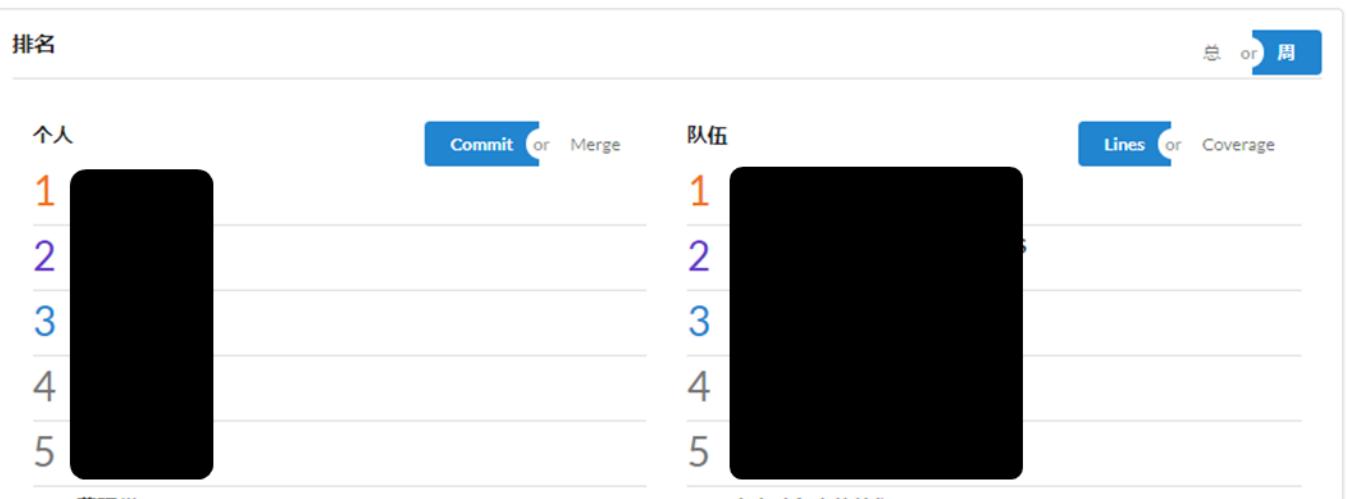
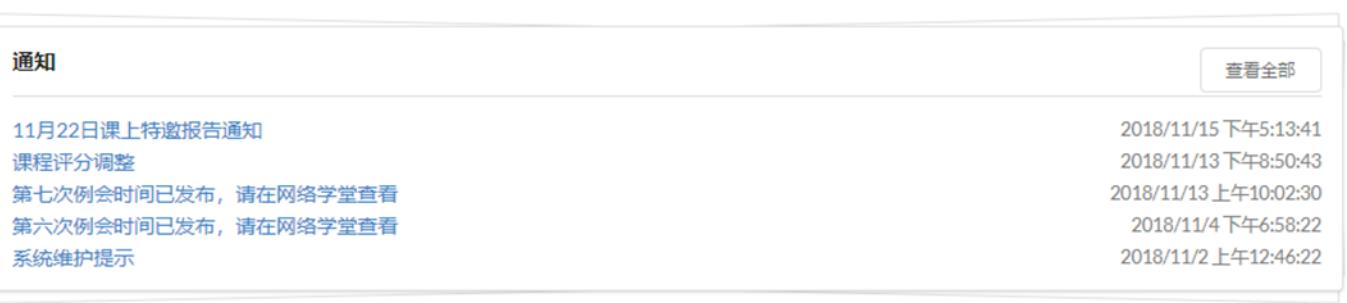
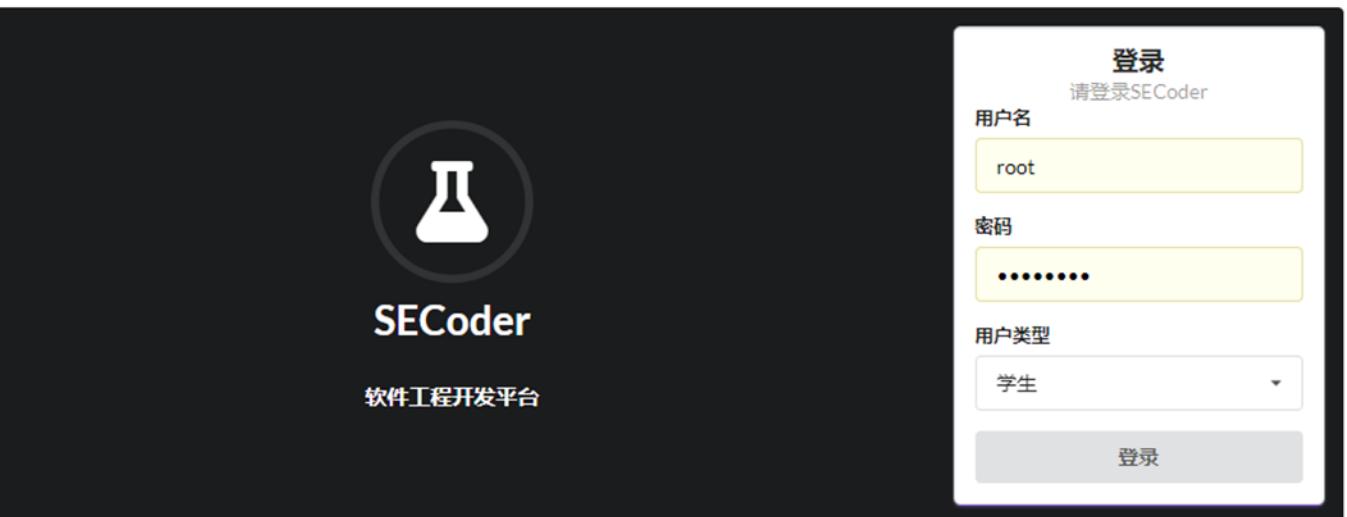
项目准备

迭代

开发

产品交付

LET'S GO!



Thank you!

