



云起之时，开源有道： 新一代云计算课程开放资源建设

王伟

2018年6月9日

Bio

- **王伟**，同济大学计算机科学与技术系副教授，博士生导师
- CCF高级会员，CCF教育工作委员会委员、CCF大数据专委会通讯委员、CCF体系结构专委会委员
- 同济大学开源社区发起人、开源社理事
- 研究方向：可信云服务、开源教育技术、大规模在线学习系统
- E-mail: wwang@tongji.edu.cn



Outline

- 背景
- 课程与教材介绍
- 开放资源建设
- 总结与展望

Emerging trends

Information technology

Big **D**ata

Cloud computing

Blockchain

Artificial intelligence

Organization & Community



Education technology

MOOC + MOOBench

教育信息科学与技术

人工智能/**CS4All教育**

知识经济

Education × IT = New Learning Science

Information source

Learning space

Learning process

Academic

Industry

Open source
community

Pedagogy

概念与基础

云计算概述

分布式原理

云计算架构

原理与技术

虚拟化技术

分布式存储

云计算网络

云计算安全

开发与运维

云原生应用

云操作系统

云端软件

云计算运维

应用与案例

桌面云

软件开发云

大数据与
人工智能

学习环境与师生交互

学习资源与认知工具

学习行为与学习评价

教育大数据

课内复习

课外思考

动手实践

论文研习

知识生产与知识进化

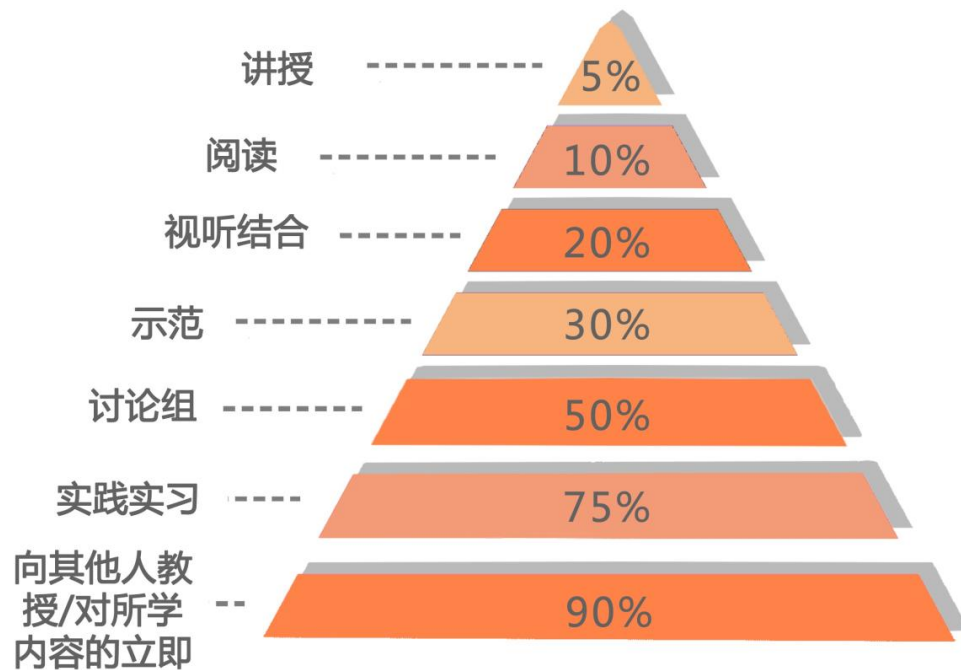
教学内容、方法与资源

学习科学与过程

《云计算原理与实践》建设理念

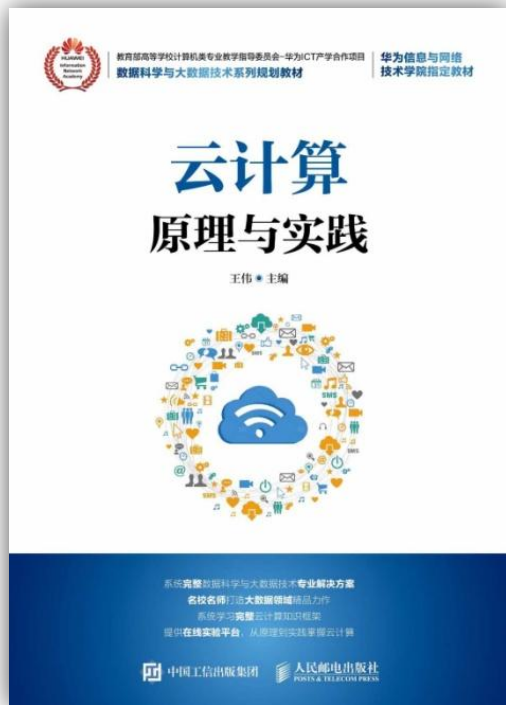


开放融合



动手实践

《云计算原理与实践》教材介绍



- 完整而系统的云计算知识框架
- 丰富的辅导材料
- 贯穿全书的开源软件实践
- 配套在线实训平台：KFCoding
- 微信公众号交流互动平台



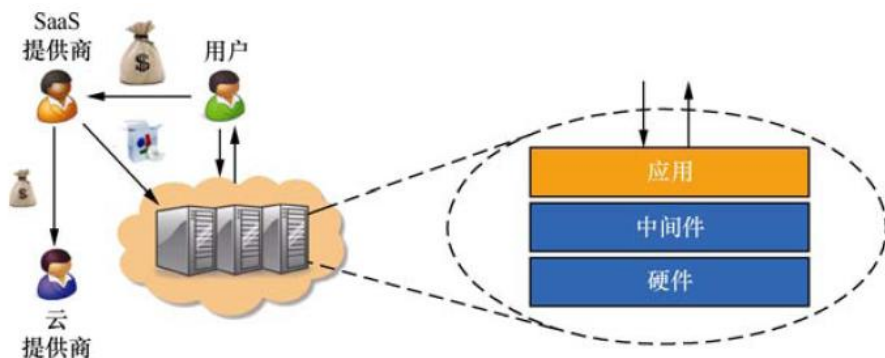
特色1：完整的云计算知识框架

- 三大认识角度
 - 商业模式、计算范式、实现方式
- 四个关键技术
 - 计算、存储、网络、安全
- 四种开发运维维度
 - 云原生应用、云操作系统、云端软件、云运维
- 三大应用场景
 - 桌面云、开发云、大数据与人工智能

三大认识角度之一：商业模式

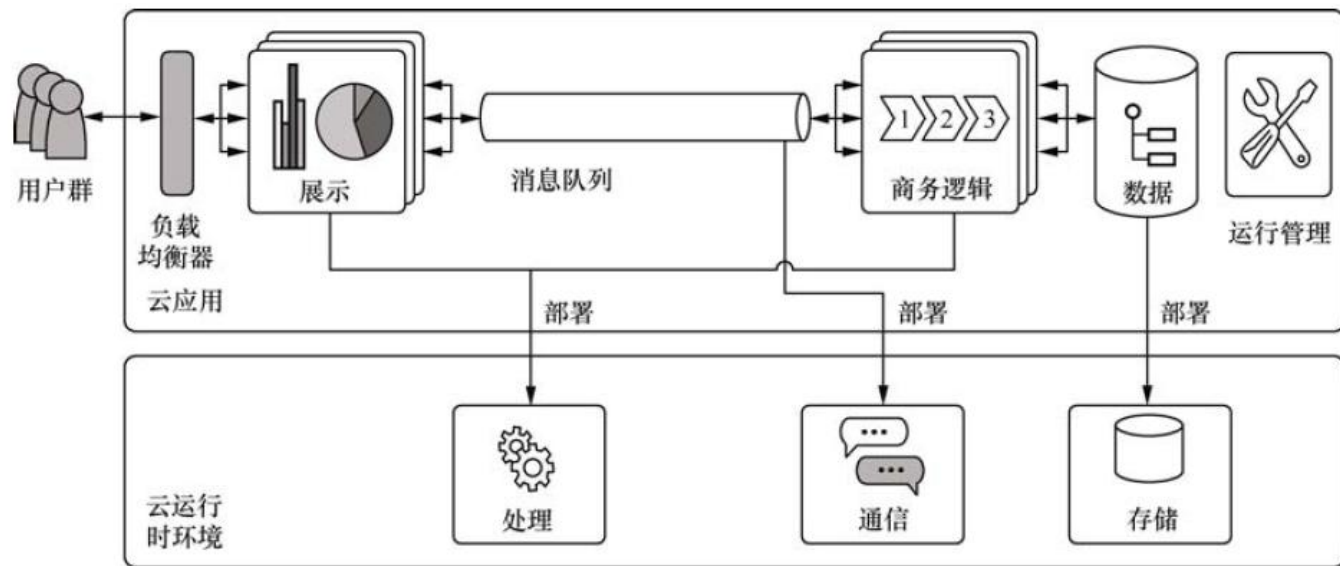


超市模式、电厂模式和云计算模式

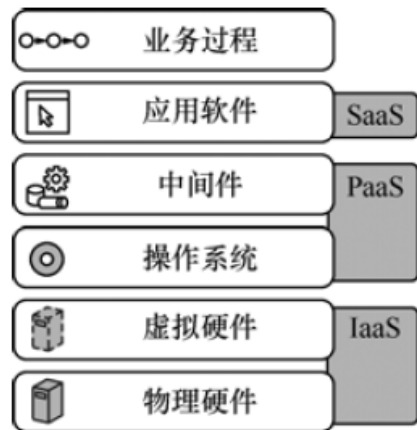


1. 基础架构即服务 (IaaS)
2. 平台即服务 (PaaS)
3. 软件即服务 (SaaS)
4. 一切皆服务 (XaaS)

三大认识角度之二：计算范式

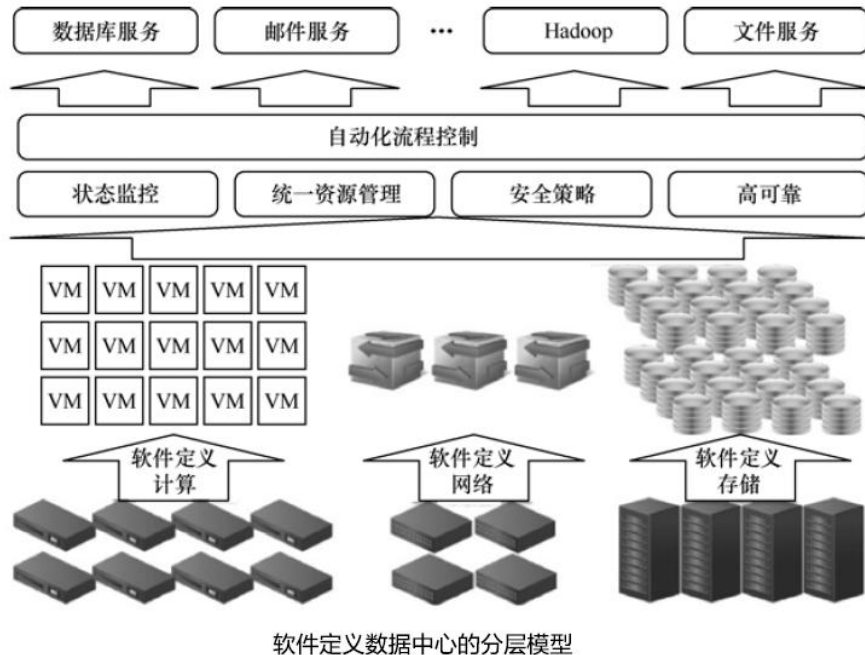
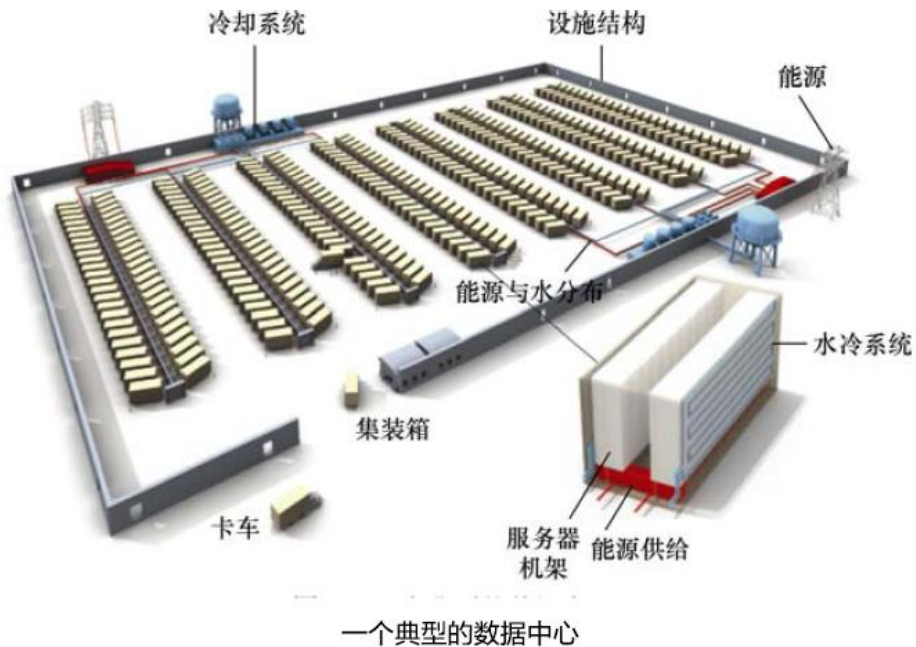


云计算的横向云体逻辑结构

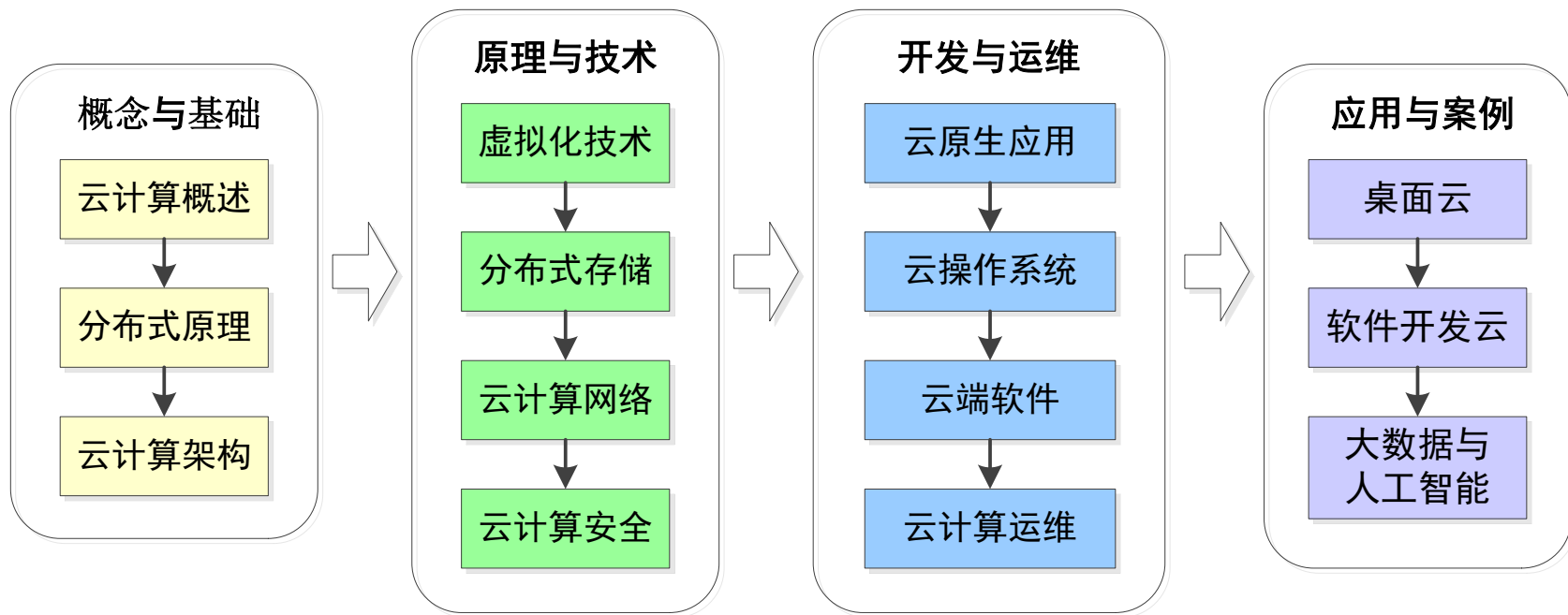


云计算的纵向云栈逻辑结构

三大认识角度之三：实现方式



云计算的知识体系概览



第1章 云计算概述

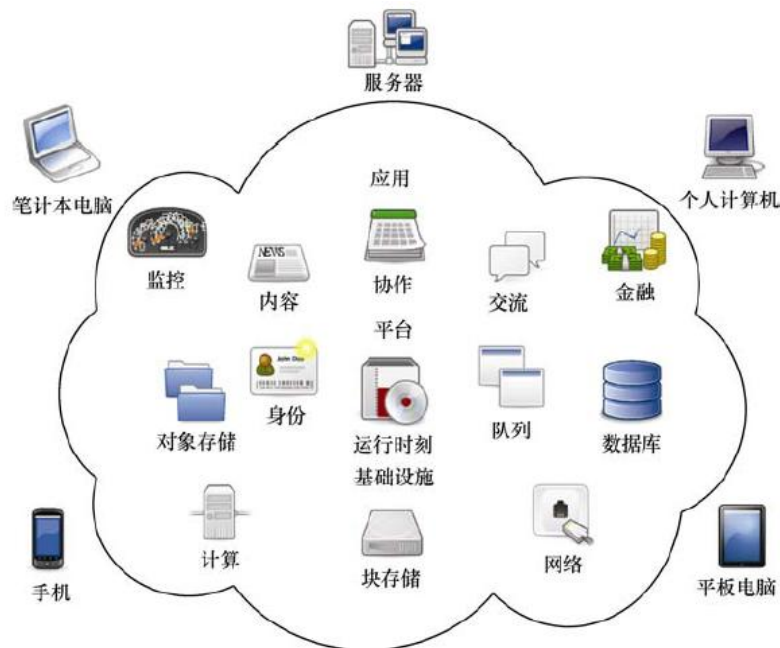
1.1 初识云计算

1.2 云计算的公共特征与分类

1.3 云计算的三元认识论

1.4 云计算的开源方法论

1.5 实践：GitHub



云计算的组成元素 (来源: Wikipedia)

第2章 分布式原理

2.1 分布式计算概述

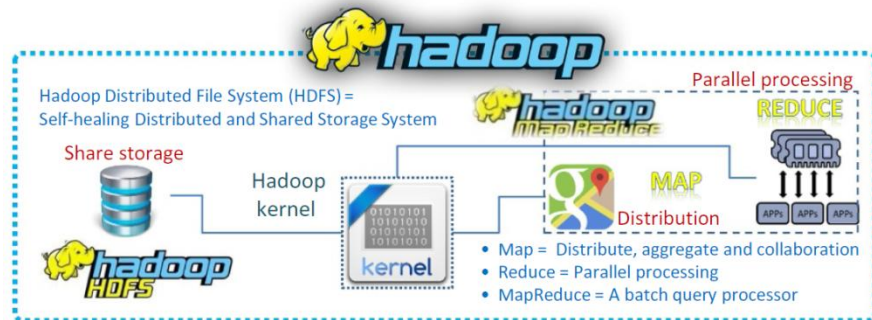
2.2 分布式计算的理论基础

2.3 分布式系统概述

2.4 分布式系统的进阶

2.5 典型的分布式系统

2.6 实践：Apache Hadoop



第3章 云计算架构

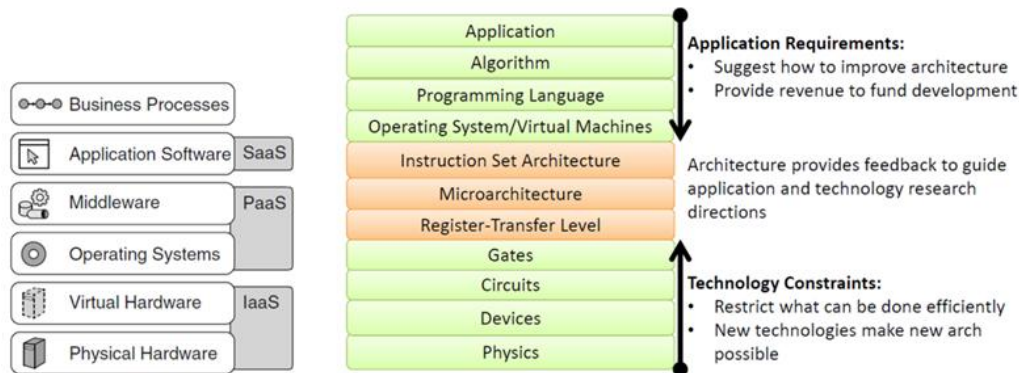
3.1 云计算的本质

3.2 云计算的架构

3.3 云栈和云体

3.4 软件定义的数据中心

3.5 实践：OpenStack



第4章 虚拟化技术

4.1 虚拟化的定义

4.2 服务器虚拟化

4.3 商用虚拟机技术

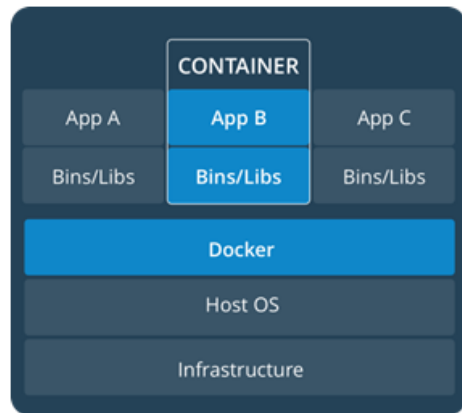
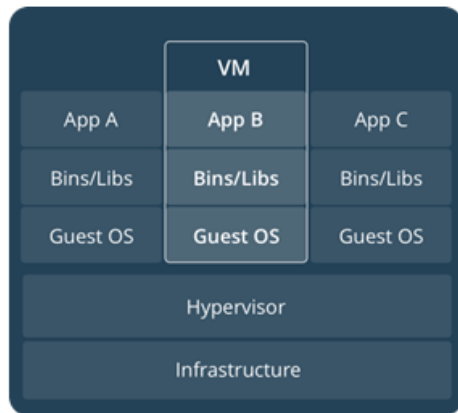
4.4 新型硬件虚拟化

4.5 实践：Xen虚拟化技术

4.6 实践：KVM虚拟化技术

4.7 轻量级虚拟化

4.8 实践：Docker容器



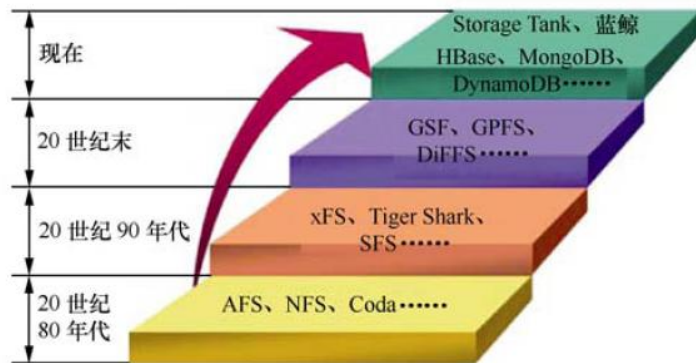
第5章 分布式存储

5.1 分布式存储的基础

5.2 文件存储

5.3 从单机存储系统到分布式存储系统

5.4 实践：分布式存储系统Ceph



第6章 云计算网络

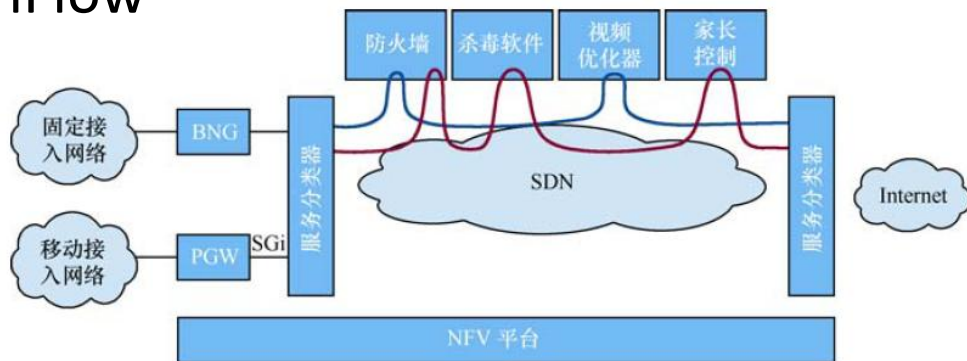
6.1 网络基本概念

6.2 数据中心网络：云计算的骨架

6.3 网络虚拟化

6.4 租户网络管理

6.5 实践：用Mininet搭建OpenFlow



第7章 云计算安全

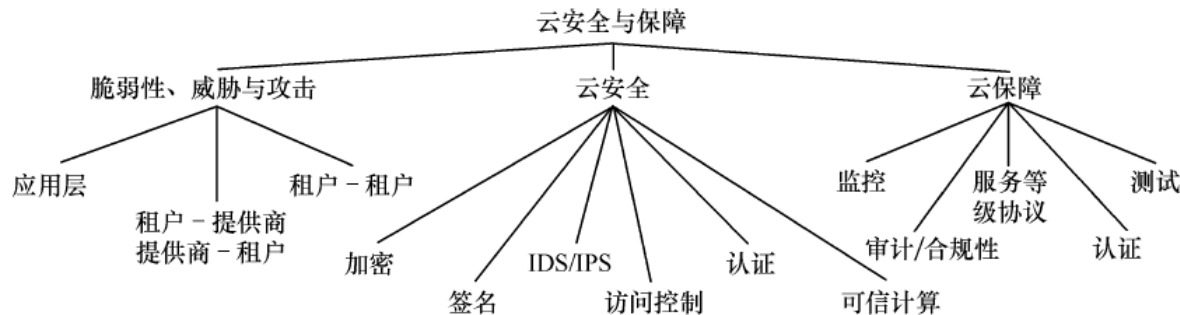
7.1 云安全概述

7.2 虚拟机安全

7.3 云存储安全

7.4 云数据安全

7.5 实践：全同态加密算法



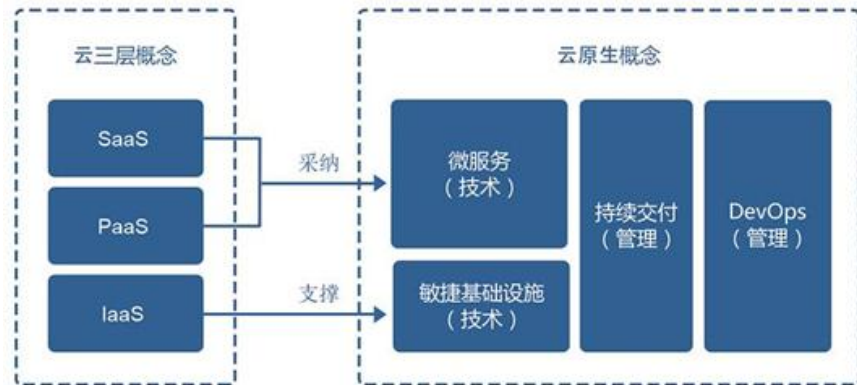
第8章 云原生应用的开发

8.1 云原生的相关概念

8.2 云原生应用开发实践的12要素

8.3 云原生应用开发

8.4 实践：基于Node.js的云原生应用开发



第9章 云计算操作系统

9.1 计算机软件与操作系统

9.2 UNIX类操作系统的发展

9.3 云操作系统概述

9.4 云计算编程模型与环境

9.5 云操作系统的资源调度

9.6 实践：Mesos



第10章 云端软件

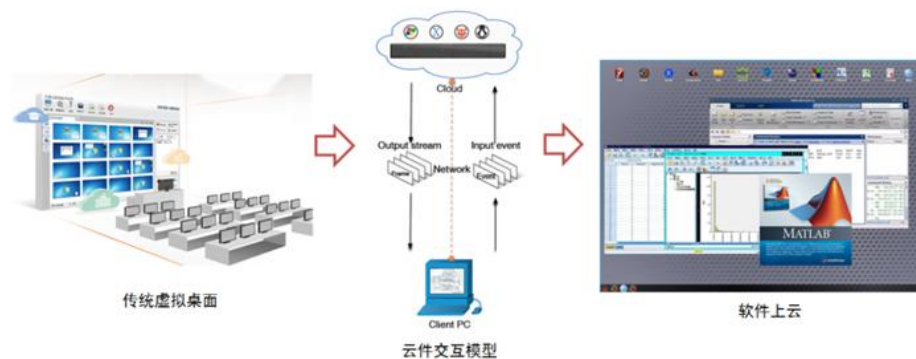
10.1 从软件到云件

10.2 云件系统架构与运行原理

10.3 云件的开发模式

10.4 云件在大规模在线实训平台中的应用

10.5 实践：云件应用开发实例



第11章 云计算运维

11.1 云服务环境的监控

11.2 云监控解决方案

11.3 智能运维

11.4 实例：智能运维在大视频运维中的应用



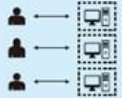
第12章 桌面云

12.1 桌面云概述

12.2 桌面云架构与关键技术

12.3 桌面云典型应用案例

12.4 实践：基于OpenStack的桌面云

	会话	虚拟桌面池	个人虚拟机
良好 ★			
较好 ★★			
最好 ★★★			
个性化	★★	★★	★★★
应用程序兼容性	★★	★★★	★★★★
易于管理	★★★★	★★	★
成本效益	★★★★	★★	★

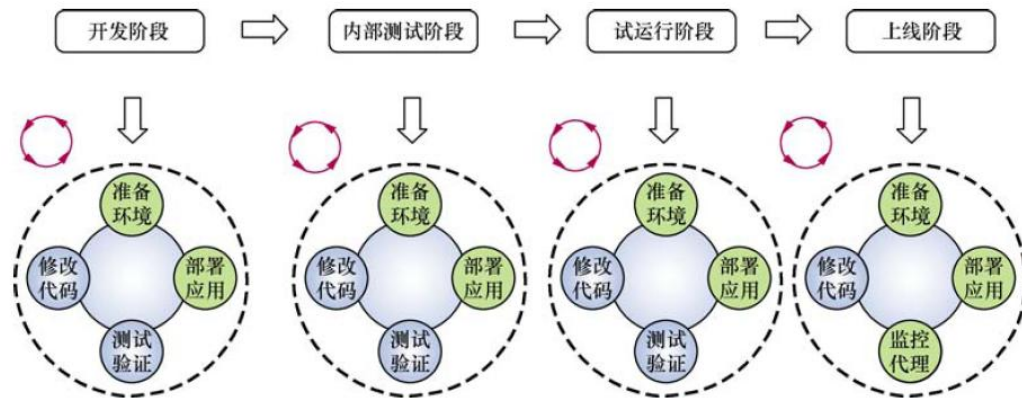
第13章 软件开发云

13.1 软件开发云的概念

13.2 华为软件开发云服务

13.3 DevCloud技术方案

13.4 实践：DevCloud实战



第14章 大数据与人工智能

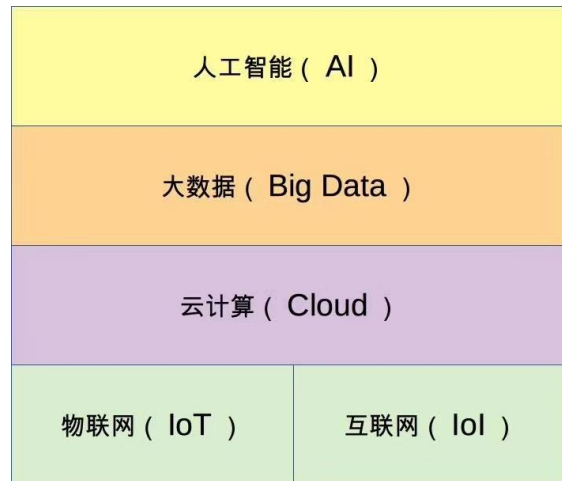
14.1 初识大数据

14.2 初识人工智能

14.3 云计算、大数据与人工智能的关系

14.3.1 云计算与大数据的融合

14.3.2 云计算与人工智能的融合



特色2：丰富的辅导材料

- **课内复习：**帮助读者复习本章的一些基本核心概念，从章节内容中基本上就可以找到对应答案；
- **课外思考：**围绕本章的核心内容的问题，一方面引发读者的思考，另一方面需要读者查找一些课外资料才能找到答案，甚至包括一些开放问题；
- **动手实践：**围绕本章内容，结合对应的开源软件和工具，开展动手实践的活动，这也是本书的特色之一，动手实践已经成为新工科背景之下课程内容的必要组成；
- **论文研习：**阅读所推荐的66篇学术论文，深度调研与本章内容相关的话题，培养读者学术论文阅读与理解的能力，从中也可以找到很多云计算领域的最新前沿内容。

示例：第1章和第3章

课内复习

1. 云计算的定义是什么？
2. 云计算的公共特征有哪几个？
3. 云计算按照部署方式和服务类型分别分成哪几类？
4. 如何从三元认识论的角度理解云计算？
5. 云计算作为一种计算范式可以分成哪两种结构？
6. 开源软件、自由软件和免费软件的区别与联系是什么？

课外思考

1. 计算系统是如何演变成今天的云计算的？
2. 如何理解“开源是种方法论”？
3. 开源技术是如何促进云计算发展的？

动手实践

1. Git是目前世界上最流行的开源分布式版本控制系统之一，用于敏捷高效地处理项目。Git是Linux Torvalds 为了帮助管理Linux内核开发而开发的一个开放源码的版本控制软件。Git与常用的版本控制工具CVS、Subversion等不同，它采用了分布式版本库的方式，不必需要服务器端软件支持。

- 任务：在Linux上安装某个版本的Git软件，试着创建和管理一个版本库，并熟悉Git的各种操作。

2. Github是由克里斯·万斯特拉斯（Chris Wanstrath）、海伊特（PJ Hyett）与汤姆·普雷斯顿·沃纳（Tom Preston-Werner）三位开发者在2008年4月创办，主要提供基于Git的版本托管服务。

GitHub目前已经成为最好用的免费开源项目托管站点。

- 任务：在Github的官方网站上注册一个账号，然后通过实际的项目熟悉Github的各种操作。

论文研习

1. 参考本书“论文阅读”部分的论文[1][2][3]，学习如何阅读一篇学术论文。
2. 阅读“论文阅读”部分的论文[5]，深入了解加州大学伯克利分校当年对云计算的一些观点，并和今天云计算的发展现状进行比较。

课内复习

1. 分布式存储的定义是什么？
2. 分布式存储有哪几种类型？
3. SAN和NAS的区别是什么？
4. 比较不同文件系统的特点。

课外思考

1. 是否存在一种文件系统能够应对所有类型的文件存储？为什么？
2. Paxos的原理和机制是什么？

动手实践

1. Ceph从2004年提交了第一行代码，至今为止已经十多年了。这个起源于Sage博士论文，最致力于开发下一代高性能分布式文件系统的项目，现在也成为了开源社区众人皆知的明星项目。随着云计算的发展，Ceph乘上了OpenStack的春风，受到各大厂商的欢迎，成为IaaS三大组件计算、网络、存储之一。

- 任务：通过Ceph的官方网站下载并安装使用最新的软件，进一步了解Ceph的原理。
- 任务：理解并实践CRUSH（Controlled Replication Under Scalable Hashing）算法。

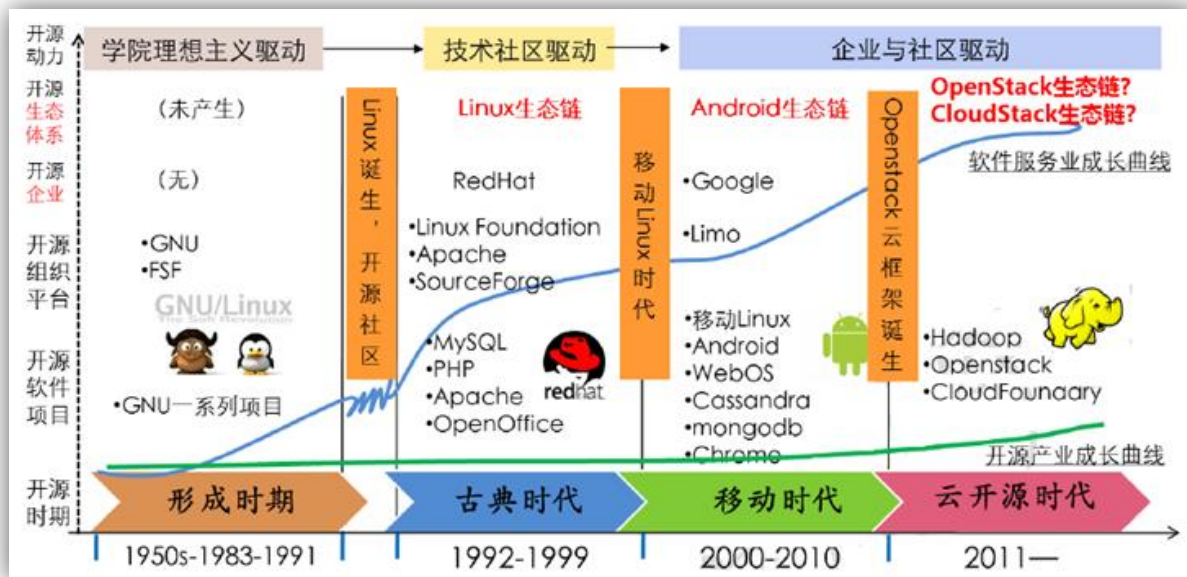
2. Hadoop分布式文件系统（HDFS）是一个高度容错性的系统，适合部署在廉价的机器上。HDFS能提供高吞吐量的数据访问，非常适合大规模数据集上的应用。HDFS是Apache Hadoop Core项目的一部分。

- 任务：通过Hadoop的官方网站下载并安装使用最新的Hadoop软件，进一步了解HDFS的工作原理。

论文研习

1. 阅读“论文阅读”部分的论文[23]，深入理解Google的GFS技术。
2. 阅读“论文阅读”部分的论文[25]，深入理解Key-Value存储技术。
3. 阅读“论文阅读”部分的论文[28]，全面了解云端存储技术。

特色3：开源实践



1. Git
2. Hadoop
3. Spark
4. OpenStack
5. KVM
6. Xen
7. Docker
8. Ceph
9. Mininet
10. OpenDaylight
11. Node.js
16. Kubernetes
17. Mesos
18. TensorFlow

- 今天已经不是开源和闭源之间的选择与博弈，而是大面积的开源与开源项目之间的竞争，这里面包含了**技术、社区、人才**等众多方面的综合因素。所谓，**得开发者得天下**！
- 例如，**Kubernetes社区**战略性的全方位吸引开发者的策略令人叹为观止！



《云计算原理与实践》所涉及到的开源软件

(按在书中出现的先后顺序)

1. GitHub	https://github.com/	16. Kubernetes	https://kubernetes.io/
2. Git	https://git-scm.com/	17. Mesos	http://mesos.apache.org/
3. Hadoop	http://hadoop.apache.org/	18. Rancher	https://rancher.com/rancher/
4. Globus	http://toolkit.globus.org/toolkit/	19. WebRTC	https://webrtc.org/
5. HTCondor	https://research.cs.wisc.edu/htcondor/	20. CloudwareHub	https://github.com/cloudwarel
6. OpenStack	https://www.openstack.org/	21. Ganglia	http://ganglia.info/
7. KVM	https://www.linux-kvm.org/	22. Nagios	https://www.nagios.org/
8. Xen	https://www.xenproject.org/	23. Cacti	https://www.cacti.net/
9. Docker	https://www.docker.com/	24. Zabbix	https://www.zabbix.com/
10. Ceph	https://ceph.com/	25. Ntop	https://www.ntop.org/
11. Mininet	http://mininet.org/	26. cAdvisor	https://github.com/google/cad
12. OpenDaylight	https://www.opendaylight.org/	27. Prometheus	https://prometheus.io/
13. HELib	https://github.com/shaih/HElib/	28. Sysdig	https://sysdig.com/opensource
14. FHE-CODE	https://github.com/rdancer/fhe	29. TensorFlow	https://www.tensorflow.org/
15. Node.js	https://nodejs.org/	30. Spark	http://spark.apache.org/

特色4：实训平台

- 大规模开放在线实训平台（MOOBench，慕班）是MOOC平台的延续与补充，主要是针对理工类课程提供大规模的、开放的、在线的实训环境。
 - Learning at Scale（大规模教育）
 - Learning by doing（新工科的诉求）
 - 开放的实训资源（Open）
 - 独享、快速、高效、灵活
 - 智能化的交互学习
 - 师生交互 $\leftrightarrow$ 机生交互（AI）

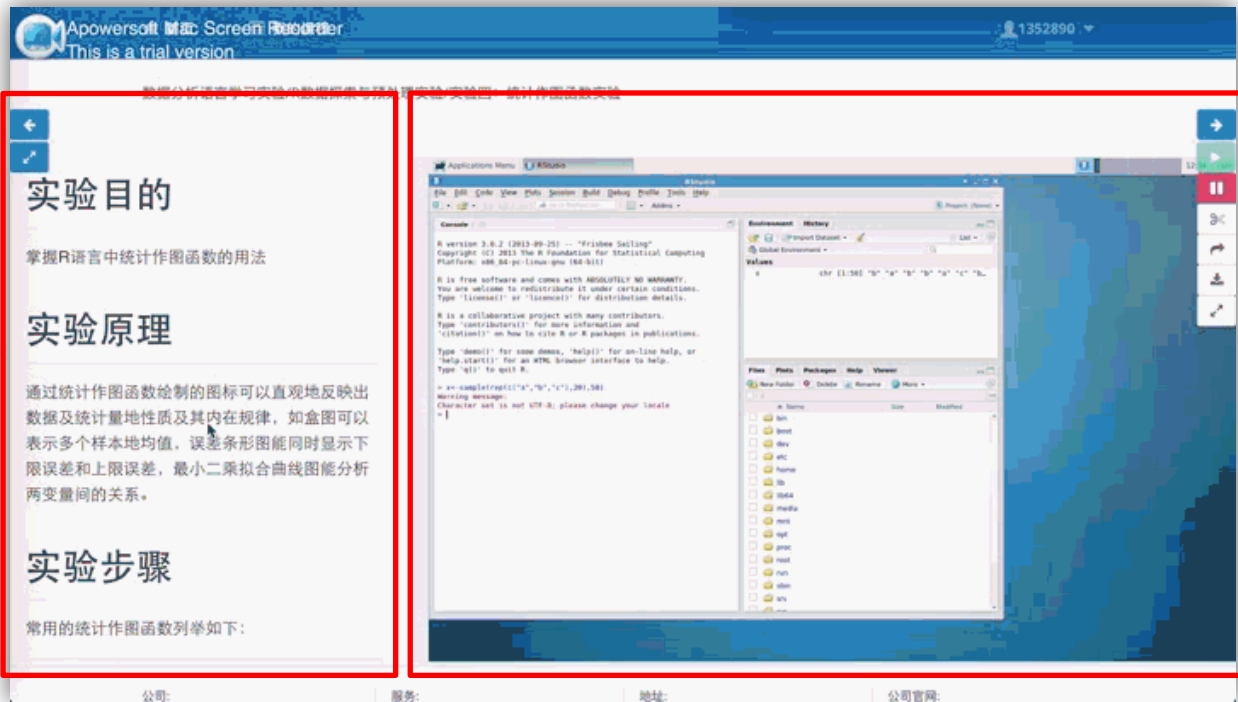


第一代MOOBench示范演示

实验教程

隐藏
启动
暂停
截图
共享
下载
全屏

实验区域



新一代交互式开发者学习社区



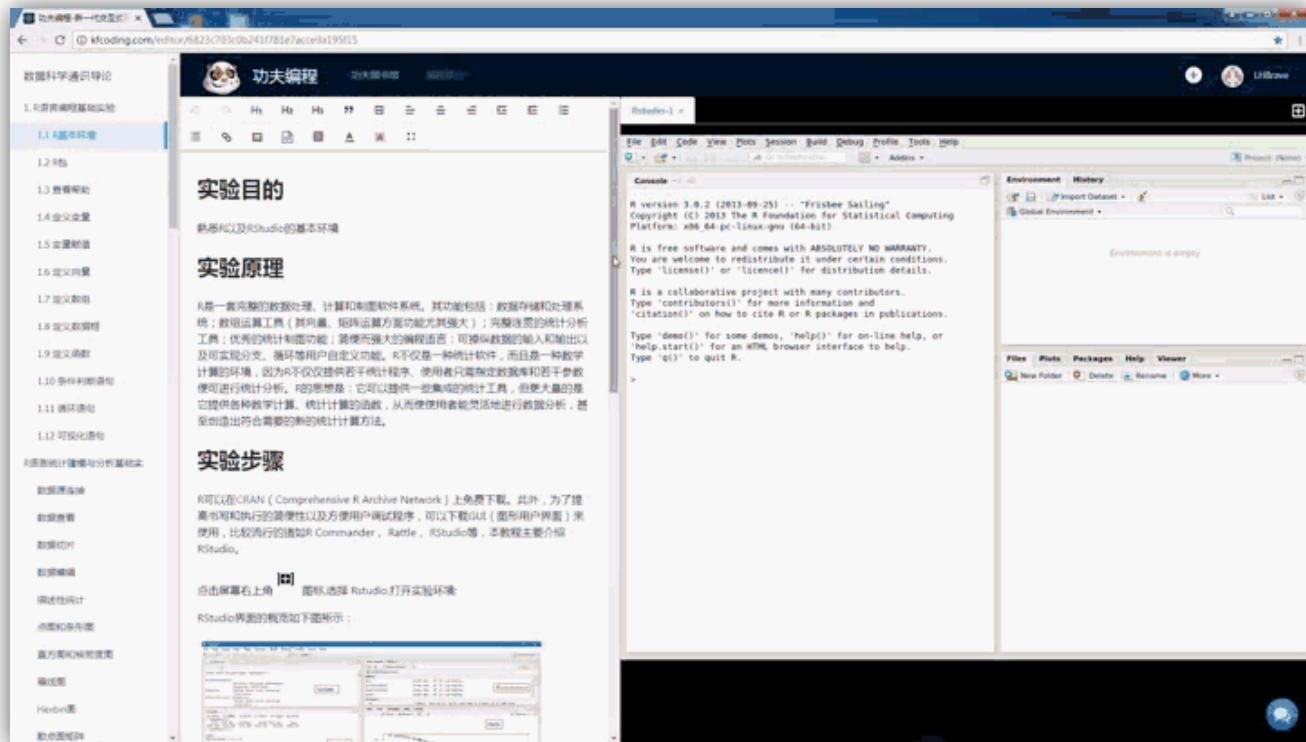
新一代交互式开发者 学习社区

在浏览器中学习各种编程语言与开源技术

Github 登录

www.kfcoding.com

新一代交互式开发者学习平台演示



开放（Open）的内容创作工具

云计算原理与实践

实践01 学习Git

+ 添加章节

章节目录

功夫编程 功夫图书馆 编程擂台

工具栏

请开始你的表演！

文字

表格

表格示例

图片

新一代交互式开发者学习社区

在浏览器中学习各种编程语言与开源技术

Github 登录

node-4 x

Linux工具库

Python环境

Nodejs环境

Nginx服务器

Rstudio(GUI)

实训环境

云件技术

Apache Hadoop & Family

HERISE

hadoop

YARN

hive

TensorFlow

社交

《云计算原理与实践》课程的实践

云计算原理与实验

实验1-Git实训任务

实验2-Hadoop实训任务

实验3-KVM实训任务

实验4-Docker实训任务

实验5-Ceph实训任务

实验6-Mininet实训任务

实验7-Kubernetes实训任务

实验8-Mesos实训任务

实验9-TensorFlow实训任务

 功夫编程

功夫图书馆 编程擂台

 willtongji

现在假设我们有一个helloworld.txt文本文件，里面内容是“Hello world!”，可以放进版本库里。

步骤3

```
git add helloworld.txt
```

把文件添加到仓库。

步骤4

```
git commit -m "a helloworld file"
```

把文件提交到仓库。

-m后是本次提交说明，利于从历史记录找到改动的部分。

commit命令一次可以提交多个文件。

步骤5

接下来编写一下helloworld.txt文件，在里面加入“Hello world again! (回车) I'm Git。”

步骤6

运行 `git status` 。可以查看到我们修改的记录。

node-1 x

```
root@acaeaea16-4ab7-44f5-9e3f-03a353c2a7
08:/# mkdir Git_temp
root@acaeaea16-4ab7-44f5-9e3f-03a353c2a7
08:/# git init
Initialized empty Git repository in /.git/
root@acaeaea16-4ab7-44f5-9e3f-03a353c2a7
08:/# git add helloworld.txt
fatal: pathspec 'helloworld.txt' did not match any files
root@acaeaea16-4ab7-44f5-9e3f-03a353c2a7
08:/#
```

知识
生产



简书

交流故事，沟通想法



开源社区

在线书籍/文档

专家/技术大神

内容
消费

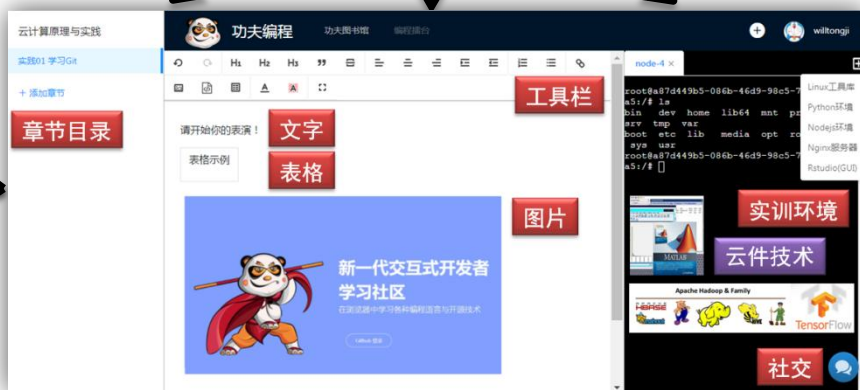
学校



公司



个人



激励
产权



软件
生产



开源
软件

商业
软件



个体/组织机构/行业软件

总结：开放资源建设

- 传统开放资源
 - 课程大纲、课程讲义、课件PPT、习题、MOOC
- 特色开放资源
 - 大规模在线实训平台
 - KFCoding学习社区



总结与展望

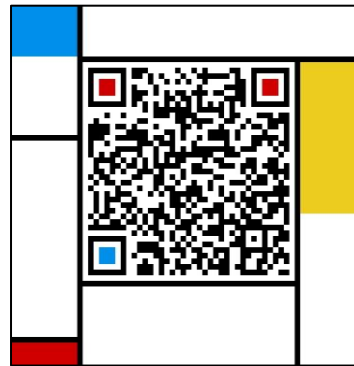
- 计算机类课程面临重大改革，真正做到以学生为本；
- 密切关注教育信息科学与技术：
 - Education × IT = New Learning Science
- 大规模开放在线实训平台将成为大数据、人工智能、新工科背景下的重要的教育信息化场景；
- “人工智能 + 教育大数据”将会在接下来的教育中发挥更加实质性的影响。

Thanks

课程公众号



个人微信



I believe the future is open!