第一章

1. 软件危机的表现和原因

第二章 软件体系结构建模

1. “4+1”模型

第三章 软件体系结构风格

1. 客户/服务器风格

2. 浏览器/服务器风格

（两者逻辑特点 优点 缺点）

第四章

1. 高并发度对网站性能的影响(不考)

第五章

1. 分布式系统的概念、优点和缺点（中间件）

2. CORBA技术标准（对象请求代理如何连接需求端和服务端、名录服务过程和原理）

第六章

1. Web Service的基本层次结构（概念、层次结构、XML、WSDL、SOAP）

2. SOAP

（消息格式、协议绑定、服务交换协议）

3. WSDL

（三个属性、六个元素、元素与属性的映射关系、四种消息交换的方式-异同点-描述角度）

XFire—不涉及

**第七章**

1. **分布式系统的特征（与SOA的关系、产生的动机）。**

**产生动机：**(1) 信息交换；(2) 资源共享；(3) 通过重复提高可靠性；(4) 通过并行化提高性能；(5) 通过专门化简化设计；(6) 问题本身的特点决定。

**与SOA的关系：**SOA是面向服务的架构，将业务逻辑提炼成共用服务，对外提供服务，也不一定就用webservice通信，也可以是socket、EJB等等。目标就是建成一个服务组件的“超市”，供别人自行选用。而分布式，主要还是从部署的角度，将应用按照访问压力进行归类，主要目标是充分利用服务器的资源，避免资源分配不均，譬如流行的三层架构等。

在分布式的环境中，将各种功能都以服务的形式(SOA)提供给最终用户或者其他服务。如今，企业级应用的开发都采用面向服务的体系架构来满足灵活多变，可重用性高的需求。

1. **面向服务体系架构（哪些参与者、三种角色之间的互操作方式-对应关系）**

看PPT。

1. **SOA基本特征（看PPT）**

**第八章**

1. 现有IT系统的主要问题

（云计算提出原因）

1. 云计算架构

（三种不同的架构-每个架构针对什么样的用户、特征）

第九章

1. 大数据的概念

（了解为主、四个维度-四个V）

第十章

1. 大数据计算体系

2. 数据存储系统

非结构化数据、结构化数据；两种存储系统的区别；非关系型数据库是否分区？

文件存储-两种HDFS/GFS两者异同

CAP理论

三个概念-最终一致性。。。

四类数据库比较

类型

1. 数据处理系统

4. 数据应用系统

第十一章

1. Hadoop总体架构

组成结构P4、P5

1. HDFS文件系统

主从节点直接的功能部署P10、存储结构好处P12、命名空间的管理P14/P16、第二名称节点作用、容错回复机制P23、机架感知

1. 分布式存储架构（HBase）

P28-32（region、store、HFile）

P33 HBase表设计特征、表索引、二次索引P45

1. Hadoop资源管理与作业调度P7

P47 三个组件（功能、服务、业务流程）

Zookeeper 业务流程P52、P54

Oozie P55 提供工作流的位置、节点概念

YARN P59 部署方式

第十二章 MapReduce计算模型

1. MapReduce计算架构

批量化计算的设计模式

分治法P6

四大组件P7 和物理节点之间的映射

JobTracker、TaskTracker和Oozie 的区别

P8 主要任务

P10-11 分片概念、与Block的相互映射

Map数目（了解）

P19 三个工作阶段

第十三章 交互式处理

一次写入、多次读取

1. Dremel数据模型与存储结构

数据逻辑结构和实际存储结构

P4 逻辑上 嵌套式结构

P6 物理存储结构 （行和列存储的存储效率，列更好）

P9 还原嵌套结构（R值和D值）

1. 并行查询
2. Drill

整体架构、运算过程与Dremel相似

第十四章 流计算

1. 流计算模型

P5 两者模型的区别（延时、吞吐率）

P8 （计算）

2. Storm计算架构

3. Storm工作机制

映射关系

（Spark中的流计算不涉及）

第十五章

1. 内存计算模型

为解决什么问题

1. 分布式缓存体系

缓存服务器概念、非永久存储媒介

P6-9 内存技术（数据压缩、列存储结构、分区、指插入差异数据）

P11 Memchache工作流程、内存管理

1. 内存数据库

P19 整体结构

P20-21 两个架构的比较、优缺点（全内存架构和读写分离架构）

1. MemCloud计算架构（不涉及）

第十六章 设计模式

（不涉及）