

试题

《大数据处理》课堂练习题

(习题课: 11月4日)

学号_____ 姓名_____ 成绩_____

注意事项: 1、方式: 模拟考试+习题课

2、时间: 100-120 分钟

3、书写清晰规范, 严禁相互抄袭

一、选择题 (本题共 20 分, 每小题 2 分)

1、云计算体系结构中_____将大量相同类型资源构成同构或接近同构的资源池。

A. 物理资源层 B. 资源池层 C. 管理中间件层 D. SOA 构建层

2、以下_____不属于 Google 的技术。

A. GFS B. MapReduce C. BigTable D. ZooKeeper

3、HDFS 分布式文件系统中数据分块的默认大小是_____。

A. 64KB B. 16MB C. 32MB D. 64MB

4、HDFS 分布式文件系统中每个数据块默认是在_____台 Chunk 服务器上。

A. 1 B. 2 C. 3 D. 4

5、以下哪一项不属于开源系统 Hadoop 中的项目_____。

A. HDFS B. Pig C. GFS D. HBase

6、虚拟化技术中_____特点可以用于实现较好的安全控制。

A. 分区 B. 隔离 C. 封装 D. 相对硬件独立

7、6 个磁盘组成的 RAID5 最多可以容忍_____块磁盘故障而不丢失数据。

A. 1 B. 2 C. 3 D. 4

8、性能介于寄存器和 Flash 之间的存储设备是_____。

A. 磁带 B. PRAM C. 磁盘 D. 光盘

9、介于 PB 和 ZB 之间的容量量化单位是_____。

A. GB B. EB C. MB D. YB

10、_____不属于大数据技术的 4 “V” 主要特点。

A. Volume B. Void C. Variety D. Velocity

二、论述题（本题 45 分）

1、论述云计算技术的四层体系结构并分别说明每层体系结构的功能（12 分）。

2、简述 Google 的分布式文件系统 GFS 中三方架构的主要模块（9 分）。

3、简要说明服务器虚拟化技术的 4 个主要特点，并举例说明各自的应用（16 分）。

4、假设每个磁盘的容量为 10GB，根据 RAID 技术的相关知识填表（8 分）

级别	盘个数	容量 (GB)	存储效率
RAID0	6		
RAID1		40	
RAID5			87.5%
RAID6		40	

四、系统设计与分析题（本题 35 分）

1、背景说明：传统的 B/S 架构中读取数据过程中数据要经过存储服务器→应用服务器→客户端的流程，但是随着用户量的急剧增加和数据的爆炸式增长，这种模式会导致应用服务器出现崩溃或者负载过重的现象，可扩展性很差，因此急需设计一种满足大数据处理和存储需求的分布式系统。

问题说明：现在为了解决这个问题，由你来设计一个系统解决上述问题，可以参考 GFS 和 HDFS 的系统架构，答题时要结合图形和文字说明设计的系统。（20 分）

1、通过学习本课程，在你的知识体系中大数据生态系统应该包括哪几个层次？并谈谈你对大数据发展和应用的趋势，可通过举例的方式来叙述其如何改变我们未来的生活。（15 分）

答案

一、选择题

1-5 B D D C C

6-10 B A B B B

二、论述题（注意：有给出答案即可酌情给分）

1、答案点：云计算技术体系结构分为4层：物理资源层、资源池层、管理中间件层和SOA构建层，如图3所示。物理资源层包括计算机、存储器、网络设施、数据库和软件等；资源池层是将大量相同类型的资源构成同构或接近同构的资源池，如计算资源池、数据资源池等。构建资源池更多是物理资源的集成和管理工作，例如研究在一个标准集装箱的空间如何装下2000个服务器、解决散热和故障节点替换的问题并降低能耗；管理中间件负责对云计算的资源进行管理，并对众多应用任务进行调度，使资源能够高效、安全地为应用提供服务；SOA构建层将云计算能力封装成标准的Web Services服务，并纳入到SOA体系进行管理和使用，包括服务注册、查找、访问和构建服务 workflow 等。管理中间件和资源池层是云计算技术的最关键部分，SOA构建层的功能更多依靠外部设施提供。

《大数据处理》课堂练

二、论述题（注意：有给出答案即可酌情给分）

1、答案点：云计算技术体系结构分为4层：物理资源层、资源池层、管理中间件层和SOA构建层，如图3所示。物理资源层包括计算机、存储器、网络设施、数据库和软件等；资源池层是将大量相同类型的资源构成同构或接近同构的资源池，如计算资源池、数据资源池等。构建资源池更多是物理资源的集成和管理工作，例如研究在一个标准集装箱的空间如何装下2000个服务器、解决散热和故障节点替换的问题并降低能耗；管理中间件负责对云计算的资源进行管理，并对众多应用任务进行调度，使资源能够高效、安全地为应用提供服务；SOA构建层将云计算能力封装成标准的Web Services服务，并纳入到SOA体系进行管理和使用，包括服务注册、查找、访问和构建服务 workflow 等。管理中间件和资源池层是云计算技术的最关键部分，SOA构建层的功能更多依靠外部设施提供。

