

第1--3章测试

go back

Name：黄智超 Class：2018大数据2班

一、单选题（Num：15，total 45.0 points）

1	以下有关Paxos算法描述正确的是（）。	(3.0分)
A、 acceptor对收到的决议进行投票，leaner只能观察投票结果但没有投票权		
B、 proposer负责提出投票决议，leaner对决议进行投票		
C、 client是提议的产生者，会将提议交给指定的proposer；proposer负责提出投票决议		
D、 acceptor和leaner共同对收到的决议进行投票		
Right Answer： A My Answer： A		
Answer analysis：		
2	有关NTP描述错误的是（）。	(3.0分)
A、 NTP是常见的分布式时间同步机制		
B、 在分布式NoSQL环境中，也需要通过NTP实现时间同步		
C、 在分布式系统中，各节点必须要与标准时间精确同步		
D、 NTP可以用于内网或公网环境		
Right Answer： C My Answer： C		
3	以下有关Paxos算法描述正确的是（）。	(3.0分)
A、 在提议发起阶段，accepter决定是否接受提议，并向proposer发送accepted		
B、 Paxos算法可以分为prepare、discribe、accept、release四个阶段		
C、 在提议发起阶段，accepter如果发现收到的提议编号是最新的，则会拒绝该提议		
D、 在提议发起阶段，proposer会将提议和编号向半数以上的accepter发送prepare请求		
Right Answer： D My Answer： D		
Answer analysis：		
4	以下属于文档式数据库的是（）。	(3.0分)
A、 MongoDB		
B、 HBase		

C、 Redis
D、 MySQL
Right Answer： A My Answer： A
Answer analysis： 5 非关系型数据库一般采用（）方式对数据进行存储、组织、管理和处理。 (3.0分)
A、 纵向扩展 B、 横向扩展 C、 以上都不对 D、 升级硬件配置
Right Answer： B My Answer： B
6 有关NoSQL描述正确的是（）。 (3.0分)
A、 列式数据库不可以通过数据字典方式实现数据压缩 B、 列式数据库执行insert或update操作一般通过追加方式实现 C、 在键值数据库中，条件查询是其强项，不支持回滚操作 D、 文档数据库只能根据键创建索引
Right Answer： B My Answer： B
7 有关文档式存储模式描述正确的是（）。 (3.0分)
A、 JSON只支持字符串类型的数据 B、 文档式存储模式中的值一般为半结构化内容，可以通过JSON方式组织 C、 文档式存储模式无法预先定义结构，且文档不能嵌套 D、 文档式存储模式的嵌套结构比传统型关系型数据库的join查询开销大
Right Answer： B My Answer： B
8 有关HDFS架构描述错误的是（）。 (3.0分)
A、 HBase使用HDFS作为底层存储系统 B、 在HDFS架构中，Namenode负责存储文件的元数据 C、 在HDFS架构中，Datanode负责其存储的数据块的读写，并决定是否对数据存储状况进行调整 D、 在HDFS架构中，Namenode对Datanode中存储的数据块进行定期检查，标记异常Datanode
Right Answer： C My Answer： C
9 以下有关Hadoop组件，描述正确的是（）。

(3.0分)	
<div><div>A、 MapReduce负责对集群中的CPU、内存等资源进行管理</div><div>B、 Hive可以将HDFS文件映射为传统的二维表，支持将SQL语句转换为MapReduce过程</div><div>C、 HDFS文件系统是存储基础，负责对大数据文件和存储集群进行管理，实现数据快速检索</div><div>D、 Spark通过Zookeeper向分布式集群提交分布式处理任务</div></div>	
<div>Right Answer： B My Answer： B</div>	
<div>10 有关布隆过滤器描述正确的是（）。</div> <div>(3.0分)</div>	
<div><div>A、 增大二进制向量的大小，可以降低布隆过滤器的误报率</div><div>B、 布隆过滤器利用一个不定长的二进制向量和哈希函数构成</div><div>C、 布隆过滤器中的二进制位对应值为1，则说明该元素肯定不存在于集合中</div><div>D、 布隆过滤器中元素与二进制位的映射关系是一对一</div></div>	
<div>Right Answer： A My Answer： A</div>	
<div>11 （）指存储设备通过电缆直接连到服务器。</div> <div>(3.0分)</div>	
<div><div>A、 NAS</div><div>B、 DAS</div><div>C、 SAN</div><div>D、 以上都不对</div></div>	
<div>Right Answer： B My Answer： B</div>	
<div>12 以下有关NoSQL描述正确的是（）。</div> <div>(3.0分)</div>	
<div><div>A、 NoSQL数据模型需要预先定义表结构，但表没有固定的列名和值域</div><div>B、 NoSQL是分布式非关系型数据库的统称，有统一的数据模式</div><div>C、 与关系型数据库相比，NoSQL数据库在通用性、分布式部署和大数据检索方面具有优势</div><div>D、 NoSQL数据库不能替代关系型数据库，需要根据业务场景和需求进行灵活运用</div></div>	
<div>Right Answer： D My Answer： D</div>	
<div>13 有关键值对存储模式描述错误的是（）。</div> <div>(3.0分)</div>	
<div><div>A、 数据表中每一行只有行键和数值</div><div>B、 键值模式可以对值进行排序，实现数据快速定位</div><div>C、 键值模式的表是无结构的，值可以是任何类型</div><div>D、 键值模式的NoSQL通常不支持对值建立索引</div></div>	

Right Answer: B My Answer: B	
14	<p>Paxos算法描述正确的是（）。</p> <p>(3.0分)</p>
<p>A、 Paxos算法是为了解决分布式系统时节点状态更新问题</p> <p>B、 Paxos算法是基于消息的一致性算法</p> <p>C、 Paxos算法是为了解决分布式系统的如何数据分片问题</p> <p>D、 Paxos算法是为了解决分布式系统的数据存储策略问题</p>	
Right Answer: B My Answer: B	
Answer analysis:	
15	<p>以下数据库管理系统，满足CAP原理中的一致性和可用性，放弃分区容错性的是（）。</p> <p>(3.0分)</p>
<p>A、 Cassandra</p> <p>B、 MongoDB</p> <p>C、 HBase</p> <p>D、 MySQL</p>	
Right Answer: D My Answer: D	

二、多选题（Num: 5, total 25.0 points）

1	<p>在分布式集群中，经常存在的问题有（）。</p> <p>(5.0分)</p>
<p>A、 局部网络故障有暂时的网络拥塞或网络设备故障</p> <p>B、 节点故障有节点的暂时故障</p> <p>C、 局部网络故障和少量节点故障</p> <p>D、 节点存储数据的永久损失</p>	
Right Answer: ABCD My Answer: ABCD	
2	<p>以下数据库管理系统，满足CAP原理中的一致性和分区容错性，放弃可用性的是（）。</p> <p>(5.0分)</p>
<p>A、 Cassandra</p> <p>B、 MongoDB</p> <p>C、 HBase</p> <p>D、 MySQL</p>	
Right Answer: BC My Answer: BC	
3	<p>以下说法正确的是（）。</p> <p>(5.0分)</p>

A、 基本可用是BASE理论的核心	
B、 一次写入多次读取是分布式系统中重要指标	
C、 一致性是传统关系型数据库的优势，体现在满足ACID4个方面。	
D、 在分布式NOSQL系统中，最终一致性要求是为了更好的实现伸缩性和可靠性、可用性	
Right Answer： CD My Answer： BCD	
4 NoSQL的特点有（ ）。 (5.0分)	
A、 NoSQL主要针对海量数据的管理	
B、 NoSQL数据库一般具有统一的数据模型	
C、 HBase属于非关系数据库，MySQL是关系数据库	
D、 NoSQL软件一般是收费的	
Right Answer： AC My Answer： AC	
5 有关数据采集描述，正确的是（ ）。 (5.0分)	
A、 互联网是大数据采集的源头，例如记录用户访问页面的情况。	
B、 大数据采集的过程就是把原始数据加载到分布式的大数据管理系统的过程。	
C、 离线采集是指直接监控数据源的变化，数据分发服务主动查看数据变化并获取数据。	
D、 在线采集是指直接监控数据源的变化，数据分发服务主动查看数据变化并获取数据。	
Right Answer： ABD My Answer： ABD	

三、判断题（Num： 10， total 30.0 points）

1 NTP可以解决分布式系统中各节点时间不一致问题。 (3.0分)	
Right Answer： <input checked="" type="checkbox"/> My Answer： <input checked="" type="checkbox"/>	
2 HDFS采用“一次写入多次读取”文件模型，具有数据分块存储、分块多副本的机制。 (3.0分)	
Right Answer： <input checked="" type="checkbox"/> My Answer： <input checked="" type="checkbox"/>	
3 在分布式系统中，NoSQL负责数据管理和简单查询，分布式处理由MapReduce等模块负责。 (3.0分)	
Right Answer： <input checked="" type="checkbox"/> My Answer： <input checked="" type="checkbox"/>	
4 在分布式系统中，如果满足一致性和分区容错性，则必须容忍系统响应迟疑。 (3.0分)	

Right Answer: ☒ My Answer: ☒

5 NewSQL是关系数据库的延伸，具有关系数据库的特性，但不具有NoSQL的特点。
(3.0分)

Right Answer: ☐ My Answer: ☐

6 在HDFS分布式文件系统中，采用机架感知的存储策略实现数据多副本。
(3.0分)

Right Answer: ☒ My Answer: ☒

7 分布式处理需要解决的问题包括数据分割，任务调度、监控和管理，数据管理。
(3.0分)

Right Answer: ☐ My Answer: ☐

8 NoSQL数据库通过降低系统的可用性，提高可扩展性和可通用性。
(3.0分)

Right Answer: ☐ My Answer: ☐

9 在分布式系统中，用户查询相关数据时，必须了解整个集群的拓扑结构和存储状态。
(3.0分)

Right Answer: ☐ My Answer: ☐

10 Datanode可以完成EditLog与FsImage的合并操作，减小editLog文件大小。
(3.0分)

Right Answer: ☐ My Answer: ☐

Answer analysis: