1. **分布式操作系统中的透明性主要是什么？其中那些透明性容易实现？哪些难实现？**(20分)

|  |  |
| --- | --- |
| 种 类 | 含 义 |
| 位置透明 | 用户不知道资源位于何处 |
| 迁移透明 | 资源可以不改名地随意移动 |
| 复制透明 | 用户不知道有多少个拷贝存在 |
| 并发透明 | 多个用户可以自动的共享资源 |
| 并行透明 | 系统活动可以在用户没有感觉的情况下并行发生 |

难点：

1、创建一个分布式数据库系统可提供有效的存取手段来操纵这些结点上的子数据库。

2、不确定性，控制比较复杂，尤其是在资源管理上要附加许多协调操作—资源属于局部工作站，性能、可靠性对网络的依赖性强，安全保密——基础不好。用户掌握有许软件接口，相应的应用软件较少，需要大力开发。

**设计、实现一个分布式操作系统时，实现透明性的难点是什么？**

所谓分布式透明性就是在编写程序时好像数据没有被分布一样。  
实现透明性的难点  
在于创建一个分布式数据库系统可提供有效的存取手段来操纵这些结点上的子数据库。

2、**分布式互斥中集中式算法、分布式算法、令牌算法。分析一下其中那个算法比较实用，为什么？**(20分)

集中式算法：集中式算法借鉴了集中式互斥算法的思想，在分布式系统中，选出一个进程为协调者 (通过科学的分析制定一套规则) 。协调者对所有的请求进行排队并根据一定的规则授予许可。协调者接受请求以后，检查临界区内的资源是否被其他进程占用。如果是，则它将当前请求进程插入到对应临界资源的请求队列中; 否则，回复一个同意消息给请求进程，通知它可以访问该临界资源。该算法通俗易懂，既能够杜绝死锁、饥饿等现象发生，又能保证资源的互斥访问顺利进行。

但是它也有缺点，由于是集中式管理，所以一旦管理进程出现故障，则整个系统将处于瘫痪状态。因此，管理进程的性能完全决定了算法的效率，应用范围小，难以普及。

分布式算法：分布式算法中运用到广播请求通信，当进程想请求共享资源时，需要首先建立三个变量: 准备进入临界区，实时时间和处理器号，并利用广播通信发送给正在运行的所有进程。该算法的核心思想如下:当进程想进入临界区时,要建立一个包括进入的临界区名字、处理器号和当前时间的消息,并把消息发送给所有其它进程。当进程接收到另一个进程的请求消息时,将分下面三种情况来区别对待:1)若接收者不在临界区中,也不想进入临界区,就向发送者发送OK消息；2)若接收者已经在临界区内就不必回答,而是负责对请求消息排队；3)若接收者要进入临界区,但还没进入,它就会把接收的消息和它发送的消息的时间戳进行对比,取小的那个。如果接收的消息时间戳小,就发OK消息,如果发送的消息时间戳小,那么接收者负责排列请求队列而不发送任何消息。当进程接收到允许消息时,它就进入临界区。从临界区退出时,向队列中的所有进程发送OK消息,并将自己从队列中删除。该算法可以保证访问临界区的互斥性以及无死锁进程、无饥饿进程。但是这种算法有个严重的缺点是算法太复杂并且不健壮,任何一个进程崩溃都会影响到算法的正确性。二是令牌丢失

令牌算法：令牌算法中引入了令牌，所有的进程组成一个环模型，环中每个进程需要知道它的下一个位置的节点的名称。令牌在环上顺序传递，当某个进程拥有令牌时就表明可以访问临界区。当请求进程没有令牌时，算法需要 N 发送任何消息。如果得到令牌的进程不打算进入临界区，它只是简单地将令牌传送给它后面的进程。当每个进程都需要进入临界区时，令牌在环上的传递速度最慢; 相反，当没有进程想要进入临界区时，令牌在环上的传递速度最快。

（该算法的正确性是显而易见的，但是也存在一些问题,比如说,当令牌丢失时,需要重新生成。可是如何检测令牌丢失又是一个困难的问题。还有,如果环中的一个进程崩溃,那么环的连贯性就遭到破坏,算法也就会出现麻烦。）

基于令牌的算法比非基于令牌的算法的时间复杂性和消息复杂性小。不会发生饥饿现象，不需要关心当前谁在临界区中，是通过竞争的方式进入临界区。

综上所述，基于令牌的算法在排除了令牌丢失和进程故障等问题之后，在今后的分布式系统中，能有更好的应用。

1. **分布式操作系统中，用于文件和目录命名的几种常见方法，如果你实现一个分布式操作系统，你会用其中哪一种？为什么？**(20分)

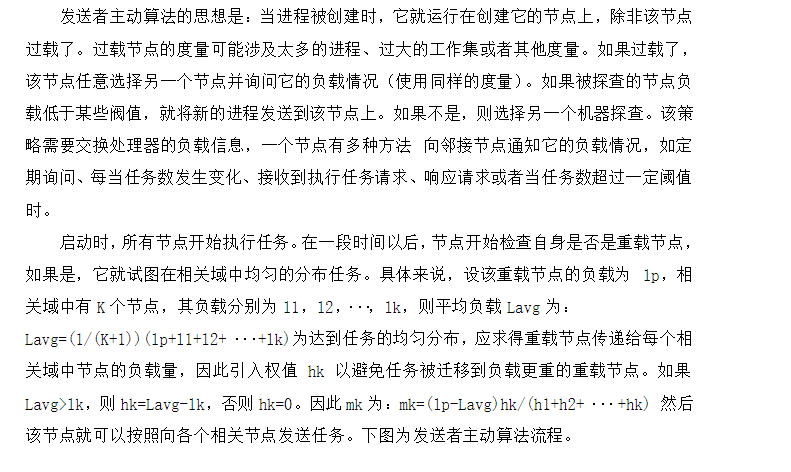
1）机器+路径

2）安装远程文件系统到本地文件分层结构

3）一个在所有机器上看上去都一样的单个名字空间

1. **分布式操作系统中，为什么需要负载平衡？给出一种负载平衡算法？**(20分) **分布式操作系统中的负载是什么？负载平衡算法的组成，影响动态平衡算法效率的因素有哪些？**

P19



1. **在计算机网络环境下，假定有20台已联网微机，其中2台是服务器，其他计算机为客户端机器，规定该分布式文件系统中的目录，文件都有3个副本，论述在该环境下，如果设计实现一个分布式文件系统，给出文件和目录的更新算法？并说明如何实现并发控制**。(20分)

分布式文件系统主要关键技术： 全局名字空间、缓存一致性、安全性、可用性、可扩展性

拷贝算法，使用时，指定一个服务器为主服务器，其他所有服务器为从服务器，当要更新一个复制文件时，我们就将该改变发送至主服务器上，在本地完成修改，然后向各从服务器发出命令，命令他们也完成修改。这样可以在任何一个（主或者从）服务器上进行读操作。这种方法简单，但是有个问题，当主服务器停机时，所有的更细将不能进行。

1. **论述远程过程调用RPC与本地调用的区别？论述RPC调用语义，系统实现中的问题。 什么是rpc,rpc与本地调用的区别，简述调用语义** P10

所谓过程调用，就是将控制从一个过程 A 传递到另一个过程 B, 返回时过程 B 将控制进程交给过程 A。目前大多数系统中, 调用者和被调用者都在给定主机系统中的一个进程中, 它们是在生成可执行文件时由链接器连接起来的, 这类过程调用称为本地过程调用。  
远程过程调用(RPC)指的是由本地系统上的进程激活远程系统上的进程, 我们将此称为过程调用是因为它对程序员来说表现为常规过程调用。处理远程过程调用的进程有两个, 一个是本地客户进程, 一个是远程服务器进程。对本地进程来说, 远程过程调用表现这对客户进程的控制, 然后由客户进程生成一个消息, 通过网络系统调用发往远程服务器。网络信息中包括过程调用所需要的参数, 远程服务器接到消息后调用相应过程, 然后将结果通过网络发回客户进程, 再由客户进程将结果返回给调用进程。因此, 远程系统调用对调用者表现为本地过程调用, 但实际上是调用了远程系统上的过程

调用语义：

（1）、last-of-many 对执行一个远程过程调用而言，被调用的过程可能执行若干次，但规定其最后一次执行的结果作为返回结果

（2）、at-most-once 若调用者收到了回复消息，则被调用的过程正确地完成了它的一次（仅仅一次）执行。如果调用者没收到回复消息，或者，如果调用者在获得回复消息之前发生故障，那么，这时的调用效果就看作是根本就没有执行相应的过程

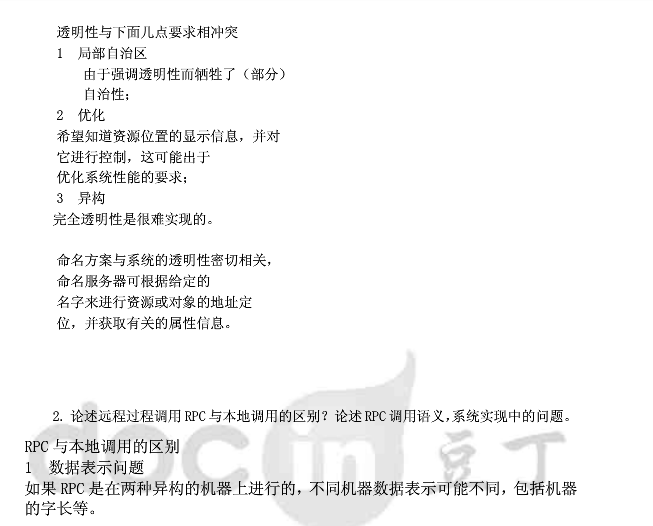
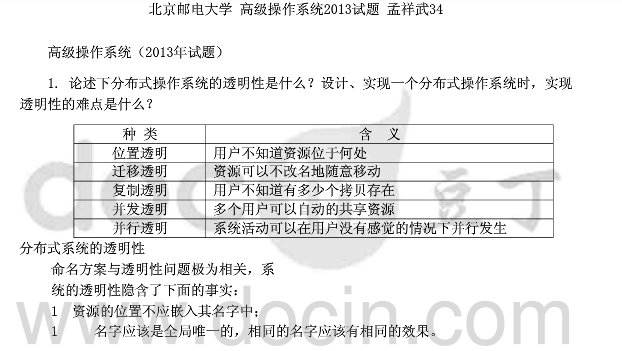
（3）、at-least-once 在场点正常情况下,则远程过程至少执行一次，且回复消息可能返回一次或多次。在场点故障时，就不能保证远程过程是否已被执行或曾返回任何回复消息

（4）、exactly-once若server正常，则远程过程将恰好执行一次，并返回一个调用结果。

同ssend/reveive通信原语有许多变种一样，RPC也有一些不同的形式。例如可以允许异步远程过程调用，因此，调用者和被调用者可以并行执行，调用者负责在稍后某一时刻执行一个所谓的会和（rendezvous）来获取调用结果

问题：实现RPC 时要解决的典型问题包括：（1）参数编组与解组，（2）调用语义或在不同地址空间的参数传递语义，（3）在客户端与服务器之间的控制与数据传输协议，还有（4）绑定或如何发现一个服务提供者，以及如何从客户端连接它。

1. **分布式操作系统中事务的并发控制采用什么机制？给出两种并发控制方法，并比较两种方法的优点和缺点。** P27
2. **假设有5台计算机、10部智能移动终端构成局域网，通过局域网，计算机、5部只能移动终端之间可以相互读写其他机器硬盘上的数据，基于这个环境，设计一个分布式文件系统，给出主要实现技术，说明一下，该分布式文件系统中一个文件多副本的更新过程。**
3. **简述操作系统的发展简史，推动操作系统发展的因素** P6 P4
4. **简述命名方案，命名方案与透明性问题的关系，如果让你设计一个操作系统，你会选哪一个命名方案** P26



1. **简述遗传算法，并解释编码的含义** P18
2. **解锁和时间戳的相同之处和不同之处，为什么操作系统不能用信号量** P28-29 无时钟