

第六章作业

1.

- (1)
 - 理想状态下应该是线性加速比，实际情况往往是亚线性加速比
 - 在大多数并行系统中，在数据规模确定的情况下，程序的加速比随节点数的增加而增加，但是大多数机群系统的节点间是共享物理传输介质的，这就使得许多并行程序的加速比在节点数目超过某一个值之后会随着节点数的增加而减少
- (2)
 - 超线性加速比
 - 最通常的情况，并行执行时所访问的数据都驻留在每个处理器的cache中，而顺序执行时必须访问存储器系统中较慢的部分，因为此时数据无法全都驻留在单个cache中。
 - 第二种超线性加速比情况来自于算法方面，比如当所要查找的元素已发现因而终止搜索时。当搜索以并行方式进行时，能以不同的次序有效地进行，意味着总的被搜索数据量将少于顺序搜索的情况。因此，并行执行将完成较少的工作。(存在性搜索)

2.

- (1)
 - 流水线并行：
 1. $RESULT1 = R1 \bowtie R2$
 2. $RESULT2 = RESULT1 \bowtie R3$
- (2)
 - 不适合流水线并行，在操作进行中数据逐渐减少，计算负载不均匀

3.

- 都分为三个阶段：partition、build和probe阶段。partition阶段将R表的数据划分到不同的分区上，build阶段在各个分区上构建哈希表，probe阶段先将S表中的元组映射到对应的分区上，再进行连接操作。
 1. 不阻塞分区
 - Build阶段不会被Partitioning阶段阻塞。因为Build阶段只关心hash表的产生，对于它输入的数据是否是完整的并不关心。此种方法只扫描一次输入关系并即时生成输出，通过另外的线程去读取输入产生的hash表等结果，从而不必要等到所有的线程都结束扫描再进入下一个阶段
 2. 阻塞分区
 - 多次扫描输入数据，以生成分区。多步骤传递数据：

1. 扫描R并计算每个哈希键的元组数量的直方图在某个偏移量处的基数
2. 使用此直方图通过计算前缀和来确定输出偏移量
3. 再次扫描R并根据散列键进行分区

4.

- (1)
 - **模式结构**
 - 应用程序接触——**全局外模式**
 - 全局外模式（通过映像匹配）——**全局概念模式**
 - 对全局概念模式中全局关系的逻辑划分——**数据分片**
 - 以一定冗余度将子关系分片分配到多个节点上——**数据分配**
 - 分配到各个节点的数据分片——**局部概念模式**
 - 局部概念模式在节点内部的存储模式——**局部内模式**
 - **映像**
 - 全局外模式——全局概念模式
 - 全局概念模式——分片模式
 - 分片模式——分配模式
 - 分配模式——局部概念模式
 - 局部概念模式——局部内模式
 - **数据独立性**
 - 逻辑数据独立性
 - 分片透明性
 - 分配位置透明性
 - 局部数据模型透明性
 - 物理数据透明性
- (2)
 - 事务可分为两类
 - 局部事务（local transaction）——访问、更新局部分布式系统中的某个数据库就能完成的事务
 - 全局事务（global transaction）——需要访问、更新分布式数据库系统中的多个数据库才能完成的事务
 - 在分布式事务系统中，每个节点都有各自独立的事务管理器（transaction manager），该事务管理器能保证节点内的局部事务的ACID特性，而事务协调器（transaction coordinator）则负责协调节点间不同事务的处理。

5.

- 多数协议的读写规则：
 - 设数据有 V 个副本，则读操作至少读 V_r 个副本，写操作至少写 V_w 个副本，且满足：
 1. $V_r + V_w > V$ ，以保证读操作一定能读到最新数据
 2. $V_w > V/2$ ，以保证消除写冲突

6.

- SpannerDB属于**半关系型模型**，其数据模型采用了**目录层级结构设计**，把相关联的表放在同样的位置，利用数据的局部性提高性能，解决由于物理分布引起的效率问题，包括事务的执行效率，包括了目录和表数据结构。目录是包含公共前缀的连续key的集合。一个目录是数据放置的基本单元；表和关系数据库表类似，具备行、列和版本值。每个表有包含一个或多个列的主键，主键形成了行的名称，每个表都定义了从主键列到非主键列的映射。Spanner的模型属于半关系型模型，数据库根据客户端分割为一个表或者多个表的层次结构，类似于传统关系数据库中一对多关系。客户端应用使用 INTERLEAVE IN 语句在数据库模式中声明这个层次结构。该层次结构的上层的表是一个目录表，目录表中每行都具有键K，和子孙表中的所有以K开始(以字典顺序排序)的行一起，构成了一个目录。
- 异：
 - Spanner的tablet不一定是一个行空间内按照词典顺序连续的分区，可以是行空间内的多个分区
 - Spanner以目录为单位，构建的是层次结构的半关系模型
 - 时间字段不再是Key的一个组成部分，而是把时间戳单独作为一个独立项，从而构成了一个多版本的数据库
- 同：
 - 两者都是Google设计的分布式数据库系统，都具有高度的可扩展性和高性能