多处理机系统

科普：

各行各业对计算机的计算能力一直有着强大的需求，而更强的计算能力意味着需要更多的CPU周期

更多的CPU周期，说明单位时间内，CPU的周期更多，信号传输的次数也就更多

单位时间内更多的CPU周期，说明CPU的主频越大。周期=频率的倒数(反比关系)。想象一下单位时间，周期被挤压，频率变快

那么频率与信号传输的波长有什么关系呢？毕竟信号就是一种各种波的叠加

按照爱因斯坦的相对论，电子信号的速度不可能超过光速，这个速度在真空中大约是30cm/ns

在铜线或光纤中电子信号的传播速度小于光速，约为20cm/ns

那么10GHz(频率)的时钟，信号的传送距离总共不超过2cm

- 为什么呢？
- 距离也就是波长
- 波长=波速x周期=波速x频率的倒数
- 波长=20cmx10的9次方/s x 1/(10x10的9次方Hz)=2cm/s
- hz是一个频率的单位，物体在一秒钟之内振动一次，它的频率就是1hz；1s=1/Hz

对于100GHz的计算机，整个传送路径长度最多为2mm，而一台1THz(1000GHz)的计算机，传送距离就不足100微米了。传送距离变短了，要求硬件的制造工艺越来越高。基本的散热问题越加凸显，因为计算机越小越难散热

曲线提高计算能力的处理方式是大规模使用并行计算机，那么出现的问题就是机器之间的通信

电子(或光学)部件之前的所有通信，归根结底是在它们之间发送消息(具有良好定义的位串 bit string)。其差别在于所涉及的时间范围、距离范围和逻辑组织

多处理机

共享储存器多处理机：让两个或更多的CPU全部共享访问一个公用的RAM

多处理机操作系统，是通常的操作系统。它们处理系统调用，进行存储管理，提供文件系统并管理I/O设备。进程同步，资源管理以及调度

多处理器硬件

UMA(Uniform Memory Access，统一存储器访问)多处理器

- 基于总线的UMA多处理机体系结构
  - 最简单的多处理机是基于单总线的，两个或更多的CPU以及一个或多个存储器模块都使用同一个总线进行通信
  - 存在的问题，当一个CPU需要读一个存储器字时，它需要首先检查总线忙否。如果总线空闲，该CPU把所需字的地址放到总线上，发出若干控制信号，然后等待存储器把所需的字放到总线上。当某个CPU需要读写存储器时，如果总线忙，CPU只是等待，直到总线空闲。那么受到总线带宽的限制，多数CPU在大部分时间里是空闲的
  - 解决方案是为每个CPU添加一个高速缓存(cache)，这个高速缓存可以位于CPU芯片的内部、CPU附近、在处理器板上或所有这三种方式的组合。许多操作可以从本地高速缓存上得到满足，总线流量就大大减少了
  - 高速缓存一致性协议：当CPU试图在一个或多个远程高速缓存中写入一个字时，总线硬件检测到写，会把这个信号放到总线上通知所有其他的高速缓存。其他高速缓存中为干净的副本，则丢弃该副本，如果其他高速缓存中有脏的副本，则脏的副本必须写回到存储器或者发送到写者CPU的高速缓存上
- 使用交叉开关的UMA多处理机
  - 为了避免单总线对于多处理机数量的限制，就需要使用新的互联网络(交叉开关的方式)，水平线(进线)和垂直线(出线)的每个相交位置上是一个交叉点
  - 交叉开关最好的一个特性是，它是一个非阻塞网络(但对于两个CPU同时试图访问同一个模块的时候，还是会出现内存的竞争)
- 使用多级交换网络的UMA多处理机
  - 基于简单2x2开关的多处理机设计，有两个输入和两个输出

NUMA(Nonuniform Memory Access，非一致存储器访问)

- 访问本地存储器模块快于访问远程存储器模块
- 具有三种关键特性
  - 具有对所有CPU都可见的单个地址空间
  - 通过LOAD和STORE指令访问远程存储器
  - 访问远程存储器慢于访问本地存储器

多核芯片

- 针对晶体管数量级的增加，如何使用晶体管？
  - 给芯片添加数兆字节的高速缓存(缓存也是主要由晶体管构成的)
  - 将两个或者多个完整的CPU，通常称为核(core)，放到同一个芯片上
- CPU可以共享高速缓存或者不共享，但是它们都共享内存。考虑到内个内存字都有唯一的值，这些内存是一致的。当某个CPU修改了该字，所有其他高速缓存中的该字都会被自动地并且原子性地删除来确保一致性(窥探，snooping)
- 片上系统(SoC)，与所有核都是对等的对称多个芯片不同，片上系统含有一个或者多个主CPU，但是同时还包含若干个专用核，例如视频与音频解码器、加密芯片、网络接口等

众核芯片

- 众核芯片是指包含几十、几百甚至成千上万个核心的多核处理器
- 为了保证缓存一致性的问题，存在一致性壁垒
- 唯一已证明可适用于众核的编程模型是采用消息传递和分布式内存实现的
- 图像处理单元(GPU)是当今最为常见的众核。拥有专有内存和成千上万个微小核的处理器。与通用处理器相比，GPU在运算单元的电路上预留了更多的晶体管，而在缓存和控制逻辑上则更少

异构多核

- 把一个GPU和一些通用处理器核封装在一起

多处理机操作系统类型

- 每个CPU有自己的操作系统
  - 静态地把存储器划分成和CPU一样多的各个部分，为每个CPU提供其私有存储器以及操作系统的各个私有副本
  - 所有的机器共享一套磁盘及其他的I/O设备，还允许灵活地共享存储器
  - 存在的问题
    - 进程系统调动是调用本机的操作系统表中的数据结构
    - 每个操作系统都有自己的表，进程间没有共享
    - 没有共享物理页面
    - 会出现高速缓存不一致的情况
- 主从多处理器
  - 操作系统的一个副本机器数据表都在CPU1上，而其他CPU的系统调用都重定向到CPU1上
  - 如果有剩余的CPU时间，还可以在CPU1上运行用户进程
- 对称多处理机
  - 在存储器中有操作系统的一个副本，但任何CPU都可以运行它
  - 这个模型动态平衡进程和存储器，因为它只有一套操作系统数据表。消除了主CPU的瓶颈

第 3 部分