Validity of the single processor approach to achieving large scale computing capabilities

--读书报告

该篇文献是著名的企业家也是计算机科学家Gene M.Amdahl所写的。在这篇文章中，他提出了并行计算中最简单也是最有名的公式--阿姆达尔定律。

在文章中，开篇就提到了因为单处理器的性能提升已经达到了一个瓶颈，很多科学家和工程师认为单处理器的性能已经达到了它所能达到的极限，所以提出了多核处理器的概念，认为只有通过允许多核之间相互协作的方式才能在处理器性能的提升上取得进一步的重大进展。

在之后的研究中，作者发现数据管理是计算负荷中比较重要的一部分，而且它在生产运行中已执行指令的40%，这个比例几乎一直保持不变。在特定的环境中，这个值可能会减少两倍。这样的话，即使是在单核的处理器中，指令的处理效率依然会很高。但是阻碍高速计算的唯一问题不只是数据管理这一部分，还有一部分原因是现实的物理问题往往具有相当大的复杂性，在这些问题中，问题的处理边界往往不规则。在文中在解释这个阻碍因素时，作者举了一个实际的物理例子，在初次涉及时，我不太明白这个例子的意义，因为其中专业名词太多，我也没有进行认真的推演。但是之后通过老师的讲解和我自己再反复的回看，我发现在这个例子中，由于时间步长通常保持恒定，因此对于某些效果而言，较快的传播速率就意味着与更远的点进行交互，最后，在随后遍历网格的过程中就会带来数据的处理问题。这样的例子证明了边界的不规则导致的处理器性能损失问题。

所以为了解决数据管理工作和问题复杂性所影响的问题，作者比较了三种涉及近似相等数量硬件的计算机。之后通过数据证明在并行计算中多核处理器理论上能够提高多少倍的速度，这就是著名的阿姆达尔定律。利用阿姆达尔定律，我们可以计算出改进计算机某一部分而能获得的性能增益。阿姆达尔定律表明，使用某种快速执行模式获得的性能改进受限于可使用此种快速执行方式的时间比例。

在使用阿姆达尔定律的过程中，我发现一个我经常容易犯的错误，那就是容易混淆“可升级部分在升级之前所占时间比例”和“升级部分在升级之后所占时间比例”。如果我测量的不是计算中可以应用该升级的时间，而是测试应用升级之后的时间，那么最后得到的结果显然就是错误的。

之后我为了进一步了解阿姆达尔定律，我搜集了一些资料，我发现阿姆达尔定律不仅仅是简单的一个公式，它还阐述了一个回报递减的规律：如果仅改进一部分计算的性能，在增加改进时，所获得的加速比增量会逐渐减少。我意识到改进一个处理器的性能不仅仅是简单的部分改进，更重要是多个功能之间互相协调，相互权衡，最后换取整体性能的提升。

研读完Gene M.Amdahl的这篇文献，我感受颇深，虽然它是一篇五十多年前的文献，但是它的前瞻性和对计算机体系结构研究的卓越贡献是不可磨灭的。阿姆达尔定律无疑是其中最重要的成果之一，我觉得我们不仅仅应该知道那个著名的公式，更要知道提出背景和推导它的过程，这样对我们自己的基础理论知识的学习是有极大的帮助的。