中期答辩记录

苗福涛

至强本身能达到32的加速比吗？？

范：LLC的命中率应该没有这么高，达不到70%，郑亚松之前有一个报告，和这个结果不一样，L2、L3的效果都不很好。

尤：给我感觉高性能应用的数据reuse会比较高，但是高通量应用不应该这么多。你的datasize是否足够大？测的时候分coldcache和warmcache，测试之前要都flush掉，load一些不相关的数据，因为开始测的时候已经有一部分数据在里面了。

答：用户数有关系，几千到上万的级别，关于命中率我回去在确认一下。

范：这种数据有没有和其它的资料产出的结果做一下比较？我记着wordcount terasort的8B的访存没有这么多。

答：和李文明师兄的结果对比过，大体上是差不多，回去再验证一下。

张：XEON的terasort的值有下降但是tile反而有上升，这个在论文中要解释一下

答：这个是和terasort算法相关的。

范：tile-gx在哪儿用的？

张：就是咱们实验室的那那一台。

范：cache命中率，和郑亚松论文的比比，印象中epfl的一篇论文比一下，他这个里面就很低，评估程序是自己摘出来的，那根跟原来的wordcount比较过吗？

答：和hadoop中的wordcount比的话没有太大的意义，它不是在一个核上跑的，本身的框架就占用了很多指令。

范：怎么确定你提取的算法的正确性，行为特征是否和源程序一致呢，你的提取过程要使人信服。

尤：要验证你的结果，应该写出来。

答：我在论文里在补充一下这个说明。

**范：服务类和交互类，还可以在往里面丰富一下，总的论文工作量也够了，但是真正测评的时候1个测试程序比较单薄，有2到3个比较好。**

**答：这些应用行为都是类似的，基本的框架都一样的，所以只选取了一个。**

**范：服务类的还有哪些应该体现出来，可以后续再完善。**

**答：在第三章有说明。**

**尤：你说行为相似，但是怎么证明它行为是相似的，可能代表性不够，图中只画一条线，一条线不能代表同一类，比如第一类数据处理类，也是行为相似，但图中的结果反而不一样。**

**答：数据处理类不太一样，是一些独立的算法，规模比较小，相互独立，数据服务类的虽然是一条曲线但是包含了整个应用中多个关键的workload，比如接受请求，查询数据库，没有办法分开去处理，是一个有机的整体。**

**范：能不能把后面的两个整体的程序的几个功能区拆出来分析一下他的行为，光看一个结果太单一。**

**答：需要在分析一下程序，不过就需要重新定义指标了，这三类的目标是不同，如果单独拆分就没有这个效果了。**

**范：达到这个量时，每一部分对于硬件指标的影响，比如miss rate，每一部分的贡献程度不一样，最终导致这个总量，提取的并不难，但是可以让你的结果更可信。**

尤：实现的runtime是通用的还是捆绑在benchmark里，每一类都不一样

答：数据处理类是通用的，其它的都是绑定在一起的。

张：3.3本章小结的时候列一个表格，三种类型可以一目了然。

尤：你的产出是一个benchmark

答：第一部分的分类分析也算

尤：分析主要依赖于资料，然后选择了几个出来，benchmark的功能是什么，跑完之后有什么分数吗？会产生什么结论

答：主要是看硬件的指标，需要用其它的工具来测试

尤：别人怎么去用？搞个仪表会有数字显示，是不是应该有个分数确定，但是没有包含这些工具，还需要把结果进行量化。

答：还需要依赖其它的工具。

张：这三种分类的依据是什么，为什么这样分合理，会不会已经有其它前人的成果已经覆盖到了，一定要把理由讲清楚。交代一下这个benchmark怎么用，打出一个分数还是需要配合什么工具。

唐：benchmark的指标是什么？

苗：就是三类应用的目标作为指标，比如单位时间的数据量或者在最大在线数等。

唐：实时交互的在线数，是怎么界定的，每个用户的行为都不一样

答：在benchmark来说，记录的是活动状态下的数目

王国江

尤：要列一下技术的情况，怎样做的cache blocking，再展示数据的提升，说一下开题报告里面的内容。

唐：autotune是怎么做的，是通过试不同的参数达到最优的吗》

答：使用脚本生成循环的代码，通过include加入原程序的代码。

尤：你这里面用了unroll吗？

答：因为用SIMD要用， 因为autotune也是需要在内层展开的时候需要用到

范：图的坐标单位要写上

唐：origin实在CPU上测得是吧

答：对

尤：double的还是thread，没有算数据传输的部分？

答：double，没有算

尤：有没有CPU和GPU并行计算的部分

答：没有，CPU和GPU都是串行的

尤：你这个测的是每一个kernel的优化数据，没有跑整个程序吗

答：目前干动力框架是可以跑的，可以达到4.4倍的加速比

尤：你的比较是和纯CPU比

答：这个是数据传输之后达到的加速比，数据规模比较小，网格点如果变化需要修改代码，相对于GPU来说比较小，所以体现不出来太好的加速比，传输的时间和计算的时间比例不太好。

尤：要说清楚是纯CPU的优化可以达到2倍的优化，加上GPU效果不太好由于数据规模的问题

唐：干动力框架是一个大程序吗

答：是一个大程序的一部分，占用的时间可以达到89%，所以优化的必要

范：怎么用SIMD指令加速的

答：用的是insentric。

唐：这些的origin是用的是最好的吗，

尤：应该在优化的时候都是O3选项，

答：前面的大概能达到10倍左右的效果

范：GPU对比CPU的多少倍？

答：大概是100倍左右。CPU用的是XEON5670

范：事实上只用了一个核

答：是的，原先的程序是串行的。基准就是一个处理器核作为基准。

唐：比如8个小函数，他们平均的计算时间大概是多少

答：大概是500-600毫秒，模型的需求的迭代次数比较高，每次迭代的计算量比较少。

尤：为什么没有试一个CPU跑一个MPI，一个multiple process跑多个thread

答：没有，根据我读的资料显示，？？？？？

尤：最终想看的是一个整体的效果，而不是单个的情况。在CPU上做cacheblocking有加速的，新旧程序在天河1上跑了吗？

答：用的是32个处理器核来测的，测得总时间也不是很稳定，会有20%的偏差。

唐：跑了多长时间

答：16到20分钟的情况，迭代次数没有特别长，所以不能再天河上跑。

尤：还是需要在天河上跑一个整体的，规模小可以多测几组。

答：理论上整体的程序有20%的提高。

唐：你在天河上测是独占分区吗

答：不是，所以不稳定，网络通讯是不稳定的，

唐：可以再所里的平台上试一试独立的

答：移植难度比较大，

尤：迭代次数能不能设置的长一些，testcase规模要大一些，提交一个作业，放着去做就行。

答：目前还有一些工作要做，所以没有把工作的重点放在大规模的测试上。

唐：在GPU上的优化数据都出来了，方法是原先有的方法，应该将工作再深入一点，比如GPU的峰值有多少差距，是不是还有挖掘的空间，理性的东西比较少，有优化的方式就用，而不是分析什么样的优化是有效的。出于什么原因用了这个方法。

答：用的都是比较经典的方法

张：论文还要在推敲，怎么优化的，为什么这么优化再添加一下。

范：为什么纵轴的单位是MFLOPS

答：kernel的数据输入是固定的，是用操作数处于时间。

唐：其实直接就是时间。既然用了这个单位，就更有必要对比一下峰值性能，挖潜，对后面的人的工作也有指导意义。

答：峰值性能在5%到10%左右

尤：要从程序里分析为什么这样，比如像数据重用率比较低。

唐：应该是优化前就有一个预估，然后在进行实现，然后分析。

答：理论分析还是比较缺乏。

唐：应该用一些vtune这样的硬件计数器分析一下，从系统结构的角度来考虑，因为机器的什么特点采用了这种优化的方法。

唐：使用GPU加速的结果还是和单个CPU带比较的，还是有问题，不考虑数据传输4.4%，那如果用多个CPU可能达到更好的效果

尤：虽然openmpi没有达到比较好的效果，可以自己写一些线程程序把他分配到多个CPU上来做，可能用GPU是不成功的，所以你可以放弃GPU的加速，

唐：因为kernel比较小，所以放在GPU上跑移动开销很大，方法上是有问题的。

尤：现在既然有这个结论了，可以把主要的方向转过来，做到CPU上来。

张：应该加一个一般优化干动力框架的时候需要有对比，要对比前人的优化方法，和你的优化技术有什么区别，这样你就能说明白你为什么这么优化了，思路就比较清楚了。

尤：优化方法没有从算法入手，优化的并不是算法本身，你应该从程序结构入手来阐述。

倪文显

唐：图上的1500MHz是DDR的数据速率，不是频率。

张：一个卡就一个片吗，实验是多少，应该在表中把预期指标和实测的结果做一个说明。感觉图中的有误导性，一个DPU芯片的性能就是别人的4、5倍这是不现实的

唐：图上的DDR1250后面不是MHZ的单位，你这个DDR不能跑到1666了吗

答：DDR控制器不支持，比较不稳定。

唐：你的ppt排版还可以更好，观众会有一些迷惑。

唐：GPU也是有转码功能了不如NVIDIA，可以比较一下，单位时间转多少路，一张卡200W，1080p 30针，能转多少路，单芯片和单芯片比，单卡和单卡比。

唐：sandy bridge 4核已经不是新的东西，可以用一些intel主流cpu的志强2670，来对比一下。一开始你讲标准的可以简化一些，DPU-m本身要介绍一下，DPU的图，更趋向于软件架构，还需要说一下DPU的硬件图，比如有多个VPU这件事，后面显得比较突兀，至少应该用示意图表示一下。循环等待的图里面应该把原来简单设计的是什么样子的描述一下，用一张图或者简短的文字说明一下，有什么缺陷，再去看你的方法，然后有什么优势，要有对比。比如之前的设计应该会等待多长时间，那么新的改进的设计的效果是能减少多少时间，一定要突出改进的效果。

异步去做的时候两者都用了多长时间，能隐藏多上时间的同步延迟。

答：这些数据都有，回去我把它放在论文和ppt里。

张：在答辩的时候不能假设下面的听众都明白背景和原有的问题。

唐：要把整个解决问题的逻辑展现清楚，那块比较慢，怎么解决。工作量都在，但是要在论文中和答辩中体现出来。

张：应该又定量的东西。