一、假定处理器运行频率为700MHZ，最大向量长度为64，载入/存储单元的启动开销为15个时钟周期，乘法单元为8个时钟周期，加法/减法单元为5个时钟周期。在该处理器上进行如下运算，将两个包含单精度复数值的向量相乘：

```
for (i=0;i<300;i++) {
        c_re[i] = a_re[i] * b_re[i] – a_im[i] * b_im[i];
        c_im[i] = a_re[i] * b_im[i] + a_im[i] * b_re[i];
}
```

（1）这个内核的运算密度为多少（注：运算密度指运行程序时执行的**浮点运算数**除以主存储器中**访问的字节数**）？

（2）将此循环转换为使用条带挖掘（Strip Mining）的VMIPS汇编代码。

（3）假定采用链接和单一存储器流水线，需要多少次钟鸣？每个复数结果值需要多少个时钟周期（包括启动开销在内）？

（4）如果向量序列被链接在一起，每个复数结果值需要多少个时钟周期（包含开销）？

（5）现在假定处理器有三条存储器流水线和链接。如果该循环的访问过程中没有组冲突，每个结果需要多少个时钟周期？

二、假定一个虚设GPU具有以下特性：

- 时钟频率为1.5GHz
- 包含16个SIMD处理器，每个处理器包含16个单精度浮点单元
- 片外存储器带宽为100GB/s

（1）不考虑存储器带宽，假定所有存储器延迟可以隐藏，则这一GPU的峰值单精度浮点吞吐量为多少GFLOP/s？

（2）在给定存储器带宽限制下，这一吞吐量是否可持续？

三、假定有一种包含10个SIMD处理器的GPU体系结构。每条SIMD指令的宽度为32，每个SIMD处理器包含8个车道，用于执行单精度运算和载入/存储指令，也就是说，每个非分岔SIMD指令每4个时钟周期可以生成32个结果。假定内核的分岔分支将导致平均80%的线程为活动的。假定在所执行的全部SIMD指令中，70%为单精度运算，20%为载入/存储。由于并不包含所有存储器延迟，所以假定

SIMD 指令平均发射率为 0.85。假定 GPU 的时钟速度为 1.5GHz。

（1）计算这个内核在这个 GPU 上的吞吐量，单位为 GFLOP/s。

（2）对于以下改进中的**每一项**，吞吐量的加速比为多少？

① 将单精度车道数增大至 16。

② 将 SIMD 处理器数增大至 15（假定这一改变不会影响所有其他性能度量，代码会扩展到增加的处理器上）。

③ 添加缓存可以有效地将存储器延迟缩减 40%，这样会将指令发射率增加至 0.95.