1. [12 points] Suppose you have two alternatives to enhance a kernel program execution, one is using multiple cores to parallelize 50% of the execution, the other is improving clock rates by 1.4 times which can improve the program on both serial and parallel sequences. Answer the following questions:

(1) if you can only choose one of the enhancement ways metioned above and you have 4 cores available, which one is more favorable? Other costs need not be considered.

(2) Now if you are allowed to use both of the enhancement ways mentioned above, is it possible to improve the overall speedup by 2.4 times? How many cores do you need to achieve this?

(3) Now, in addition to the two enhancement ways (i.e. 4 cores and 1.4 times of clock rate improving), you are provided with a new compilation technique which can improve the parallelization fraction of the program. Then what percentage of parallelization would the compilation technique achieve in order to gain an overall speedup of 4?

2. [20 points] RISC-V is a new instruction set architecture (ISA) that was orginally

designed to support computer architecture research and education, but which becomes

a standard open architecture for industry implementations. The base RISC-V ISA has

a little-endian memory system. LOAD instruction copy data from memory to register

rd in the following format:



where the memory access address is generated by adding register rs1 to the sign-extended（符号扩展） 12-bit offset.

(1) For best performance , the effective address for all loads and stores should be naturally aligned for each data type. What is the memory address feature for a short integer variable?

**(2) What is the maximum byte offset of a base address for LOAD instruction?**

**-211~211-1**

**(3) The funct3 field selects the type of operation. How many instructions the RISC-V is able to support at least? 23✖27？**

(4) Respectively, what will be in the destination register rd loaded from the address …to … an integer? 偏移的地址需要写成扩展符号位的吗？

3.[12 points] Consider the DLX pipeline implementation with all necessary hazard detection and to remove stalls wherever possible. For the C program statement: A=B+E-D, the following sequence of DLX instructions can be executed (assume all variables are in memory):

LW R2, 0(R1) ; load B

LW R3, 4(R1) ; load E

ADD R4,R2,R3; add B and E

LW R5, 8(R1); load D

SUB R6,R4,R5; sub D

SW R6, 12(R1); store A

(1) Please draw a cycle diagram of the code sequence, and indicate forwarding in the cycle diagram.

(2) How to remove pipeline stalls occurred to the code execution? And describe types of the stalls.

4.[12 points] The following code is executed in WinMIPS64 simulator with forwarding and branch target buffer enabled.

daddi r3,r0,1000

daddi r1,r0,4

ld r2,100(r3)

loop: daddi r1,r1,-1

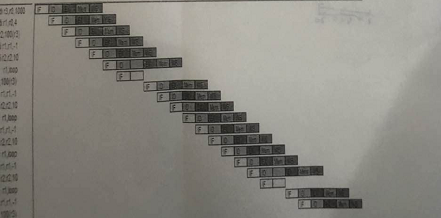
daddi r2,r2,10

bnez r1,loop

sd r2,100(r3)

halt

According to the cycle diagram given by WinMIPS64, answer the following questions:



(1) How many stalls in the execution? Analyse the causes of stalls.

**(2) Calculate the average CPI of the code execution. √（解决）**

5. [13 points]Many processors support out-of-order execution of non-dependent in structions.

(1) How do you understand **out-of-order execution**?

保留站中被阻塞，

(2) What structure serves as “Rest areas for dependent instructions” in an out-of-order execution processor? 保留栈

(3) What kind of hazard is particularly caused by the out-of-order execution?

6.[16 points] There is a sequence of 44-bit bytes addresses 0x59AB34, 0x59CB20, 0x59CB24, 0x59AB24, 0x18CB36, 0x59CB24, 0x59AB34, 0x18CB30,0x59CB2A accessing a cache with 32K bytes of data size in 2-way set associative organization, each cache block having 32 bytes.

(1) Describe **cache location** where each streaming data placed.

(2) For the above reference addresses, how many cache misses occur? Considering three categories, classify each miss occurred.

(3) What is the size of tags for the cache?

7.[7 points] Assume a program with high spatial and temporal locality running on an architecture with typical TLB, page table, and cache. Rank the possibility of the following cases:

**a) TLB miss, page table fault, cache miss**

**b) TLB miss, page table hit, cache hit**

c) TLB hit, page table miss, cache miss

Please rank the possibility in descend and explain the reasons.

8.[8 points] Given a CPU with L1,L2 and L3 cache memory, where the miss rate of I-cache(i.e. L1 instruction cache) is 2%, D-cache (i.e. L1 data cache) miss rate is 4%, and miss penalty for L1 cache is 100 cycles.

(1) What is the average CPI for an application with 20% load/store instructions at the base CPI of 100 cycles?

(2) Why is L1 cache generally separated as instrsuction cache and data cache?