Problem Out-of-Order Execution

For this problem, Lu Jiaxin wants to schedule the following codes on an out-of-order core.

The processor is a single-issue core with integer register (x0, x1, . . .) and float register (f0, f1, . . .). The processor uses free ROB entries from low index to high index. Instructions can commit one cycle after writeback, and ROB entries can be reused one cycle after commit. Instructions that depend on others can issue one cycle after the instruction it depends on writes back. Loads and stores take two cycles each, floating-point multiplies take three cycles, and floating-point adds take five cycles.

All functional units are fully pipelined. All the reservation stations are large enough.

Fill out the table with the cycles at which instructions enter the ROB, issue to the functional units, write back to the ROB, and commit. Also fill out the new register names for each instruction. If the instruction producing a source register had already committed before the dependent instruction enters the ROB, use the architectural register name.

Remember that instructions must enter the ROB and commit in order. On each cycle, only one instruction can enter the ROB, one can issue, one can write back, and one can commit.

1 Question A

Code A :

```
fmul.d  f0, f0, f1
fadd.d  f2, f2, f0
```

The ROB has two entries. Use r0-r1 for the two ROB entries.

	Time				Instruction			
	Enter ROB	Issue	WB	Commit	OP	Dest	Src1	Src2
I_1	-1	0	3	4	FMUL.D	r0	f0	f1
I_2	0	4	9	10	FADD.D	r1	f2	r0

2 Question B

Code B :

```
fmul.d  f0, f0, f1
fadd.d  f2, f2, f0
fadd.d  f2, f2, f0
```

The ROB has two entries. Use r0-r1 for the two ROB entries.

	Time				Instruction			
	Enter ROB	Issue	WB	Commit	OP	Dest	Src1	Src2
I_1	-1	0	3	4	FMUL.D	r0	f0	f1
I_2	0	4	9	10	FADD.D	r1	f2	r0
I_3	5	10	15	16	FADD.D	r0	r1	f0

3 Question C

Code C :

```
fld     f0, 0(x1)
fld     f1, 8(x1)
fmul.d  f0, f0, f1
fadd.d  f2, f2, f0
fld     f0, 16(x1)
fadd.d  f2, f2, f0
```

The ROB has four entries. Use r0-r3 for the four ROB entries.

	Time				Instruction			
	Enter ROB	Issue	WB	Commit	OP	Dest	Src1	Src2
I_1	-1	0	2	3	FLD	r0	x1	–
I_2	0	1	3	4	FLD	r1	x1	–
I_3	1	4	7	8	FMUL.D	r2	r0	r1
I_4	2	8	13	14	FADD.D	r3	f2	r2
I_5	4	5	8	15	FLD	r0	x1	–
I_6	5	14	19	20	FADD.D	r1	r3	r0

Problem Virtual Memory & Aliasing Problem

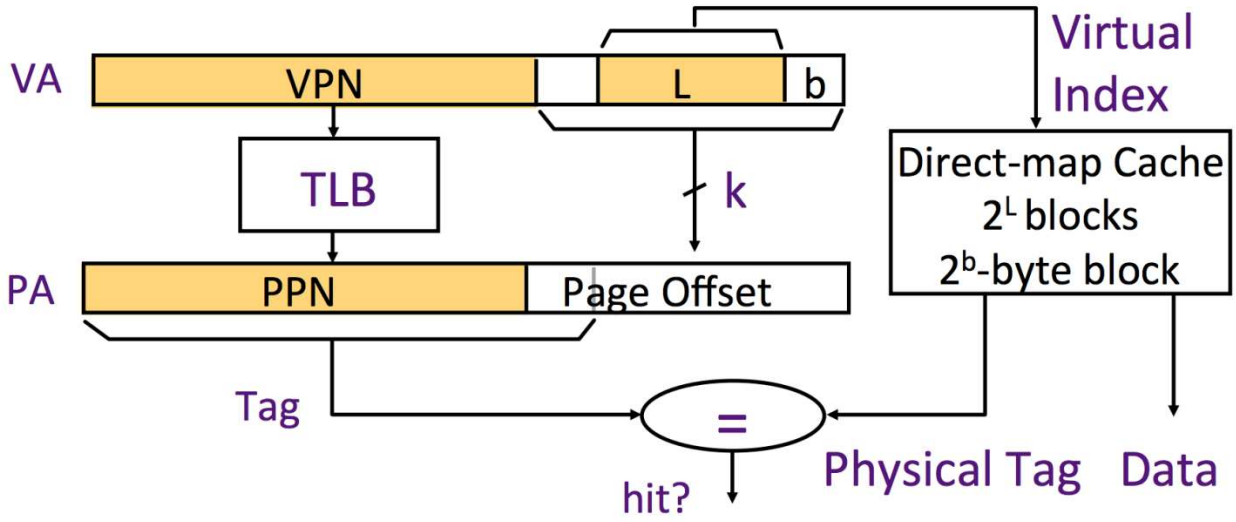
1 Question A

Suppose Lu Jiaxin has a virtual memory system with 2048-byte pages. Assume that physical addresses are 64 bits, each page table entry takes up 64 bits, and each level of the page table takes up a single page total. How many levels of paging would Lu Jiaxin need in order to cover 32 GB of virtual memory?

As the size of each page is 2048 bytes (2^{11} bytes), we need at least 2^{24} pages to cover the 32 GB (2^{35} bytes) virtual memory. Each page can contains 2^8 page addresses, so 3 levels are enough because a 3-level paging can have $(2^8)^3=2^{24}$ pages to store data, which can exactly cover all the virtual memory.

2 Question B

What is the aliasing problem? Describe it in your words. Let's consider a feasible solution for the aliasing problem, Virtual Indexed & Physical Tagged Cache.



To avoid the aliasing problem, what condition should meet among L, k, b? Why? (L is cache index. k is page offset. b is block offset)

Aliasing problem is a situation where two virtual addresses share the same physical address so that some unexpected consequences will occur when the data is modified from some aliased name.

$L+b \leqslant k$ because if $L+b>k$, the virtual index will include some bits in VPN, which will still cause an aliasing problem when two VPNs are different but share the same PPN.

3 Question C

You are asked to design a virtually indexed, physically tagged cache. A page is 4096 bytes. The cache must have 256 lines of 64 bytes each. What associativity must the cache have to not worry about aliases? (Hint : cache size is bigger than page size, so the cache should be set- associative)

As the size of the page is 4096 (2^{12}) bytes, the page offset should have a length of 12 bits. So let the length of the virtual index be 6 and the length of the cache block offset be 6, and the remained bits in physical address be the tag. Clearly there are 64(2^6) different indexes here while we have 256(2^8) lines in cache. Hence, the cache can be 4 ways set-associative.

Problem Branch Prediction

For the following question, we are interested in the performance of branches when executing the following code. For each part, assume that N = 4 and the array A has the values {1, 7, 2, 5}.

C code	RISC-V Assembly

`for (int i = 0; i < N; i++) {`
`  int b = A[i];`
`  if (b >= 5)`
`  c += b;`
`  if (b < 4)`
`  c -= b;`
`}`
	`  li x1, A`
	`  li x4, N`
	`loop:`
	`  lw x2, 0(x1)`
	`  li x5, 5`
	`  blt x2, x5, skip1`
	`  add x3, x3, x2`
	`skip1:`
	`  li x5, 4`
	`  bge x2, x5, skip2`
	`  sub x3, x3, x2`
	`skip2:`
	`  addi x1, x1, 4`
	`  addi x4, x4, -1`
	`  bnez x4, loop`

1 Question A

Fill out the table to show what the predictions will be if the branch predictor is a BHT indexed by PC with two-bit counters. If the most- significant bit of a counter is 1, the predictor predicts taken. Otherwise it predicts not taken. The counters are initialized to weakly not-taken (01). The Counter column in the table shows the state of the counter before the branch is executed. Assume that the BHT is large enough that no aliasing of instruction addresses will occur.

	Instruction	Counter	Prediction	Actual
i = 0	blt x2 x5, skip 1	01	not taken	taken
	bge x2, x5, skip2	01	not taken	not taken
	bnez x4, loop	01	not taken	taken
i = 1	blt x2 x5, skip 1	10	taken	not taken
	bge x2, x5, skip2	00	not taken	taken
	bnez x4, loop	10	taken	taken
i = 2	blt x2 x5, skip 1	01	not taken	taken
	bge x2, x5, skip2	01	not taken	not taken
	bnez x4, loop	11	taken	taken
i = 3	blt x2 x5, skip 1	10	taken	not taken
	bge x2, x5, skip2	00	not taken	taken
	bnez x4, loop	11	taken	not taken

What is the prediction accuracy for each branch? What is the prediction accuracy overall?

Branch	Accuracy
blt	0 %
bge	50 %
bnez	50 %
overall	33.33 %

2 Question B

Now assume we change the branch predictor to a BHT indexed by PC and a single bit of global history. Assume the global history is initialized to 0, the counters are initialized to 01 (weakly not-taken), and there is no aliasing. The Counter columns in the table show the state of the counters before the branch is executed. Fill out the table with the predictions.

	Instruction	Global History	Counter 0	Counter 1	Prediction	Actual
i = 0	blt x2 x5, skip 1	0	01	01	not taken	taken
	bge x2, x5, skip2	1	01	01	not taken	not taken
	bnez x4, loop	0	01	01	not taken	taken
i = 1	blt x2 x5, skip 1	1	10	01	not taken	not taken
	bge x2, x5, skip2	0	01	00	not taken	taken
	bnez x4, loop	1	10	01	not taken	taken
i = 2	blt x2 x5, skip 1	1	10	00	not taken	taken
	bge x2, x5, skip2	1	10	00	not taken	not taken
	bnez x4, loop	0	10	10	taken	taken
i = 3	blt x2 x5, skip 1	1	10	01	not taken	not taken
	bge x2, x5, skip2	0	10	00	taken	taken
	bnez x4, loop	1	11	10	taken	not taken

What is the prediction accuracy for each branch? What is the prediction accuracy overall?

Branch	Accuracy
blt	50 %
bge	75 %
bnez	25 %
overall	50 %

3 Question C

If you run this code with a large array containing uniformly randomly distributed values, which branch do you expect to get the most benefit from global history and why?

The bge instruction, which determine b < 4, will get the most benefit because it follows the blt instruction, which determine b >= 5, as we know that whether a value is bigger than or equal to 5 if strongly associated with whether it is less than 4. When a value is bigger than or equal to 5, we can assert that it is bigger than 4, and when a value is less than 5, we can predict that it is less than 4 as there are only one counterexample: 4.

4 Question D

Explain the motivation for using both BHT and BTB branch-prediction structures in the same implementation.

BTB can predict the branch result early in IF stage. When BTB fails, which means the branch instruction is not in BTB, we can apply BHT in ID stage so that we can still have a prediction to avoid being stall. So we use both BHT and BTB to better the branch prediction.