

Advanced Lab 3

Memory, LC3 State Machine and Pipelining

Andrew Wilder

Topics

- Memory
- LC3 State Machine
 - Summary of State Machine
 - Instruction Overview
- Pipelining
 - Throughput
 - Hazards

Memory

Question: The LC3 register file is composed of an array of 16-bit registers, made of D flip-flops. Would this work for memory?

Memory

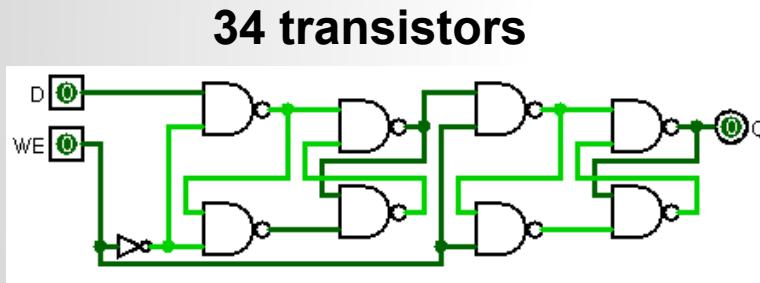
Question: The LC3 register file is composed of an array of 16-bit registers, made of D flip-flops. Would this work for memory?

No:

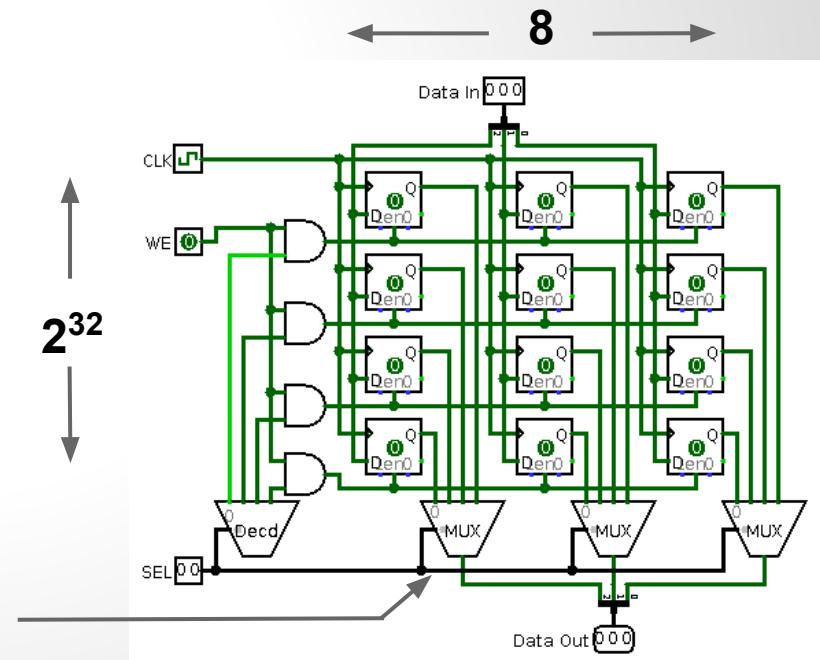
- Too many resources
- Too much power
- Too complex to manufacture

Memory

Consider a 4GB RAM chip. What if this was composed of a D flip-flop array?

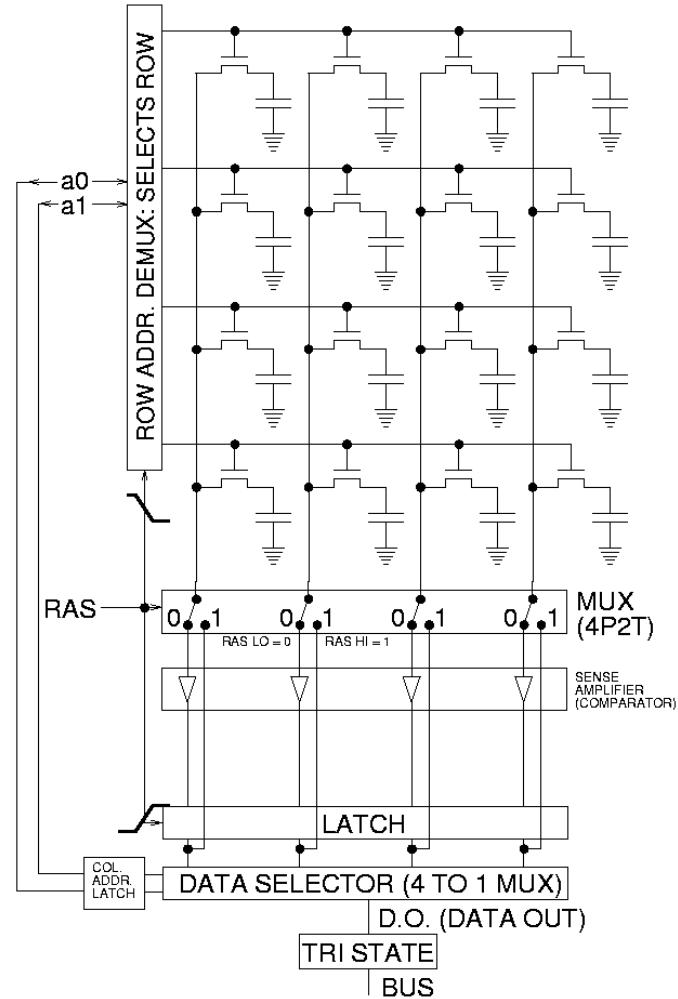


Billions of AND gates and OR gates



Memory

Solution: One transistor and capacitor per bit, constructed in rows and columns. Easier to manufacture, lower power consumption.



Memory

Much like the homework, DRAM is manufactured in smaller pieces on the order of kilobytes, which are combined as abstractions to create larger memories; however, its design takes multiple clock cycles to access.

LC3 State Machine

The LC3 State Machine has very specific standards outlined in appendix C of the textbook. In the following slides, we will cover the execution of each instruction on the datapath.

LC3 State Machine

Inputs:

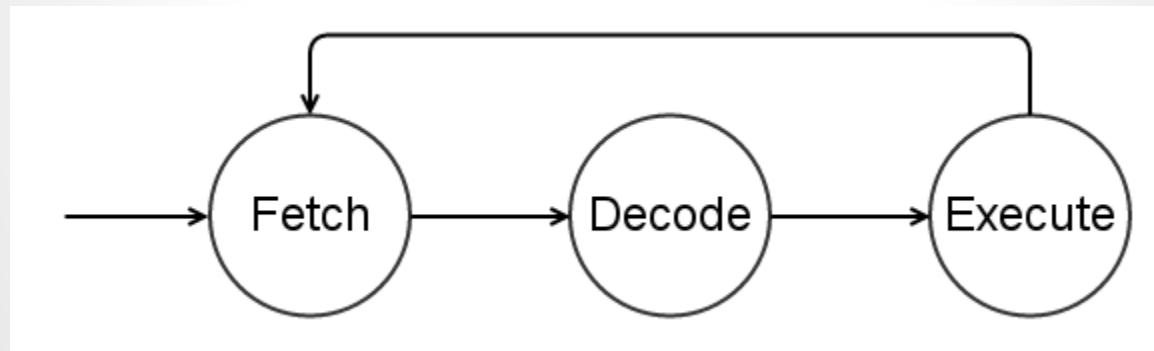
- S0 - S5 (current state bits)
- IR[15:12] (opcode)
- IR[11:9] (hardcoded CC)
- NZP (variable CC)
- IR[5] (imm5 ADD, AND, NOT)
- IR[11] (JSR vs JSRR)

Outputs:

- LD PC,IR,REG,MAR,MDR,CC (6)
- Tri-state buffers (4)
- 5 Multiplexers (7 bits)
- Mem EN/WE (2 bits)
- DR, SR1, SR2 (9 bits)
- ALUK (2 bits)

LC3 State Machine

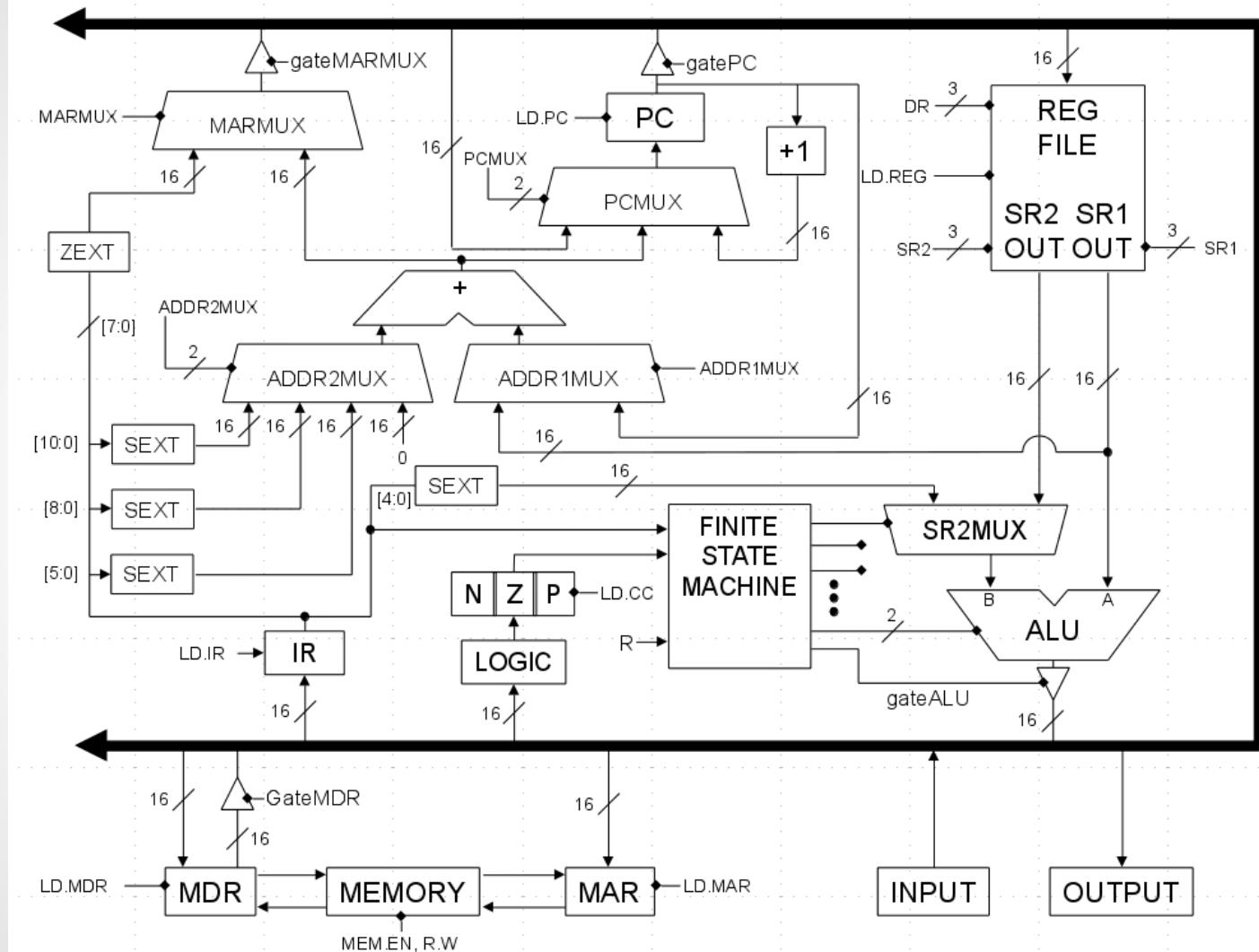
Basic sequence of events: Fetch instruction, decode instruction, execute instruction, repeat.



Fetch

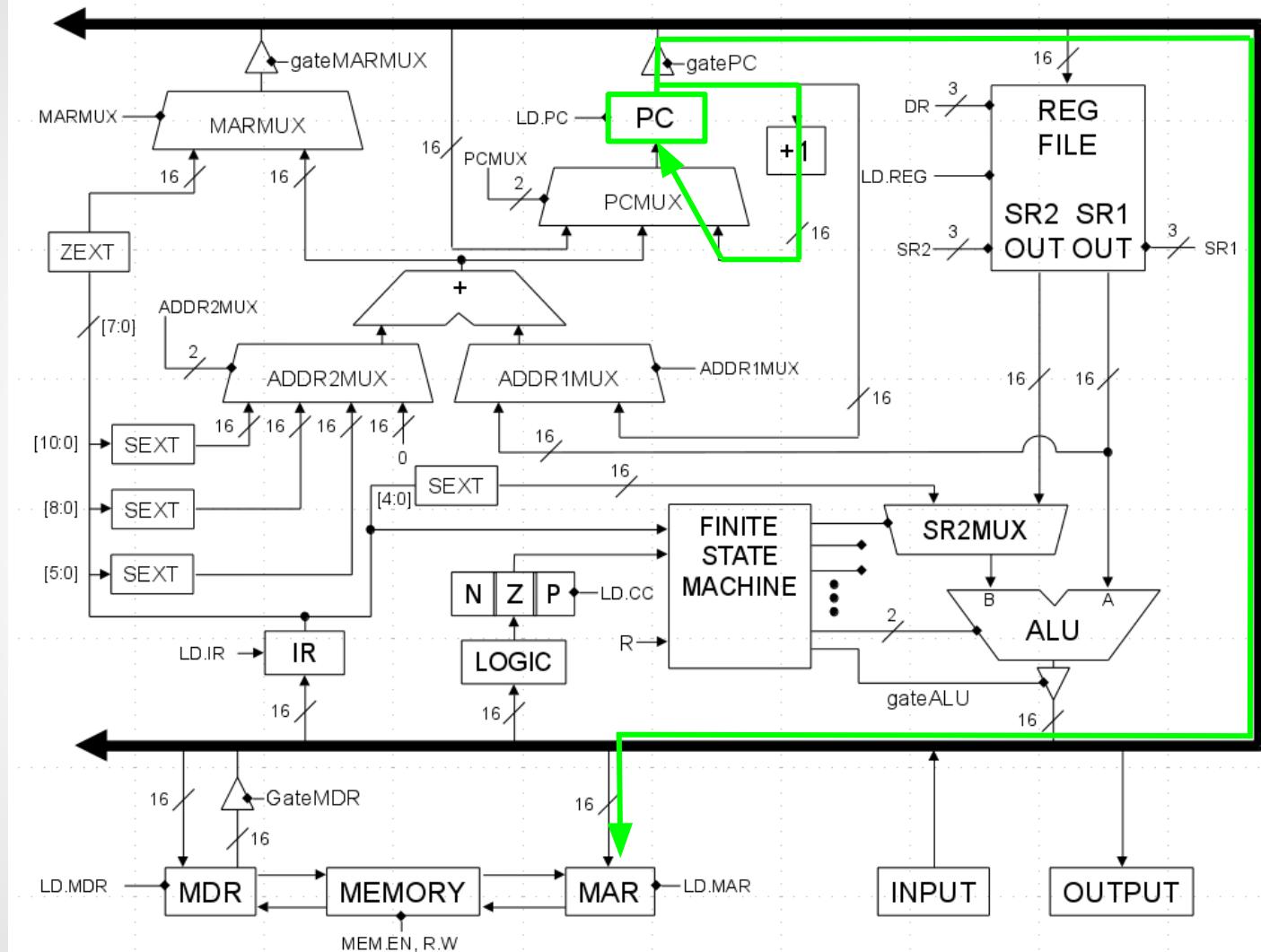
The purpose of fetch is to obtain the next instruction to be executed. This involves going to the memory location pointed to by PC, retrieving the value there, and storing it in IR.

Fetch



Fetch

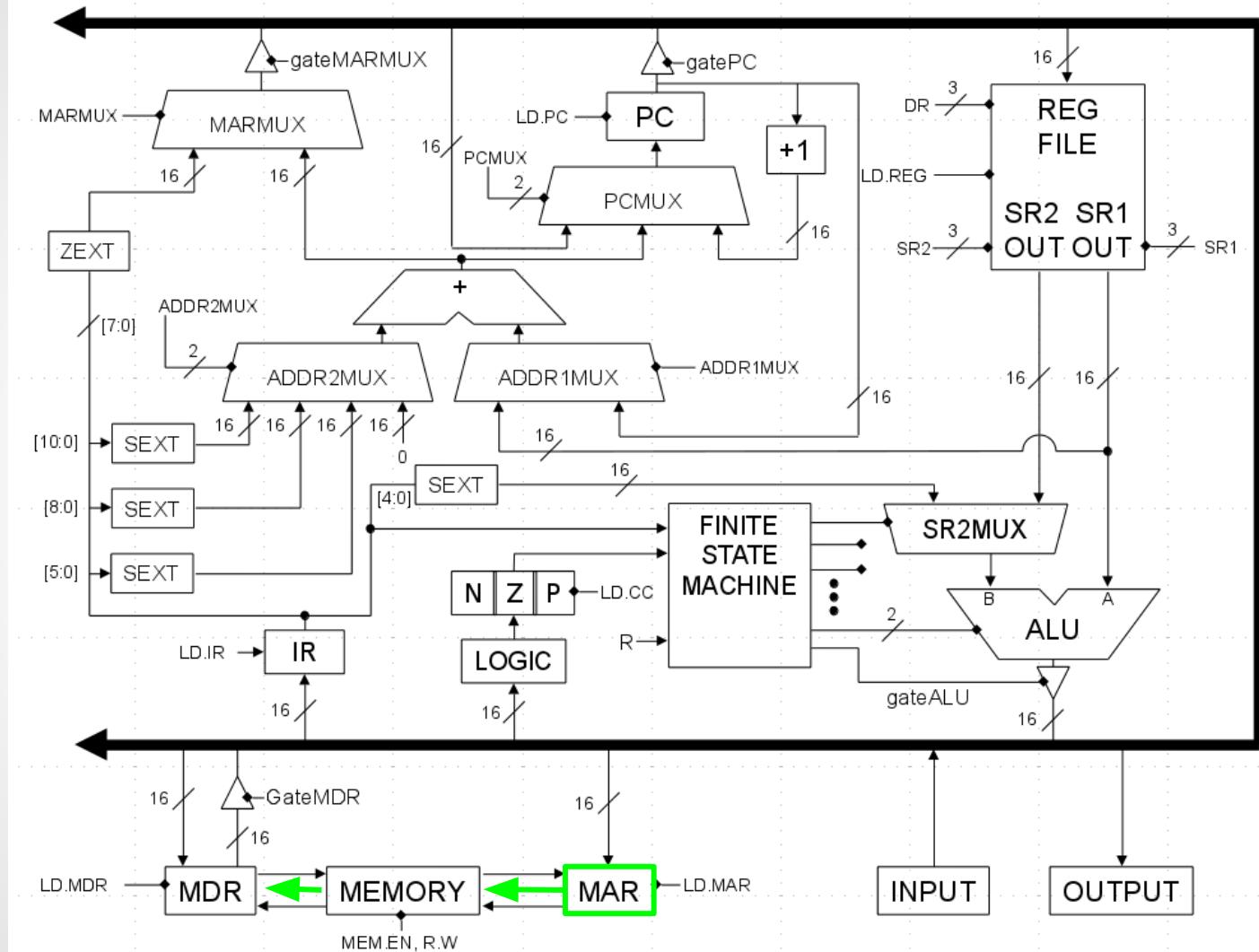
CC1:
gatePC
LD.MAR
PCMUX=PC+1
LD.PC



Fetch

CC1:
gatePC
LD.MAR
PCMUX=PC+1
LD.PC

CC2:
MEM.EN
LD.MDR



Fetch

CC1:

gatePC

LD.MAR

PCMUX=PC+1

LD.PC

CC2:

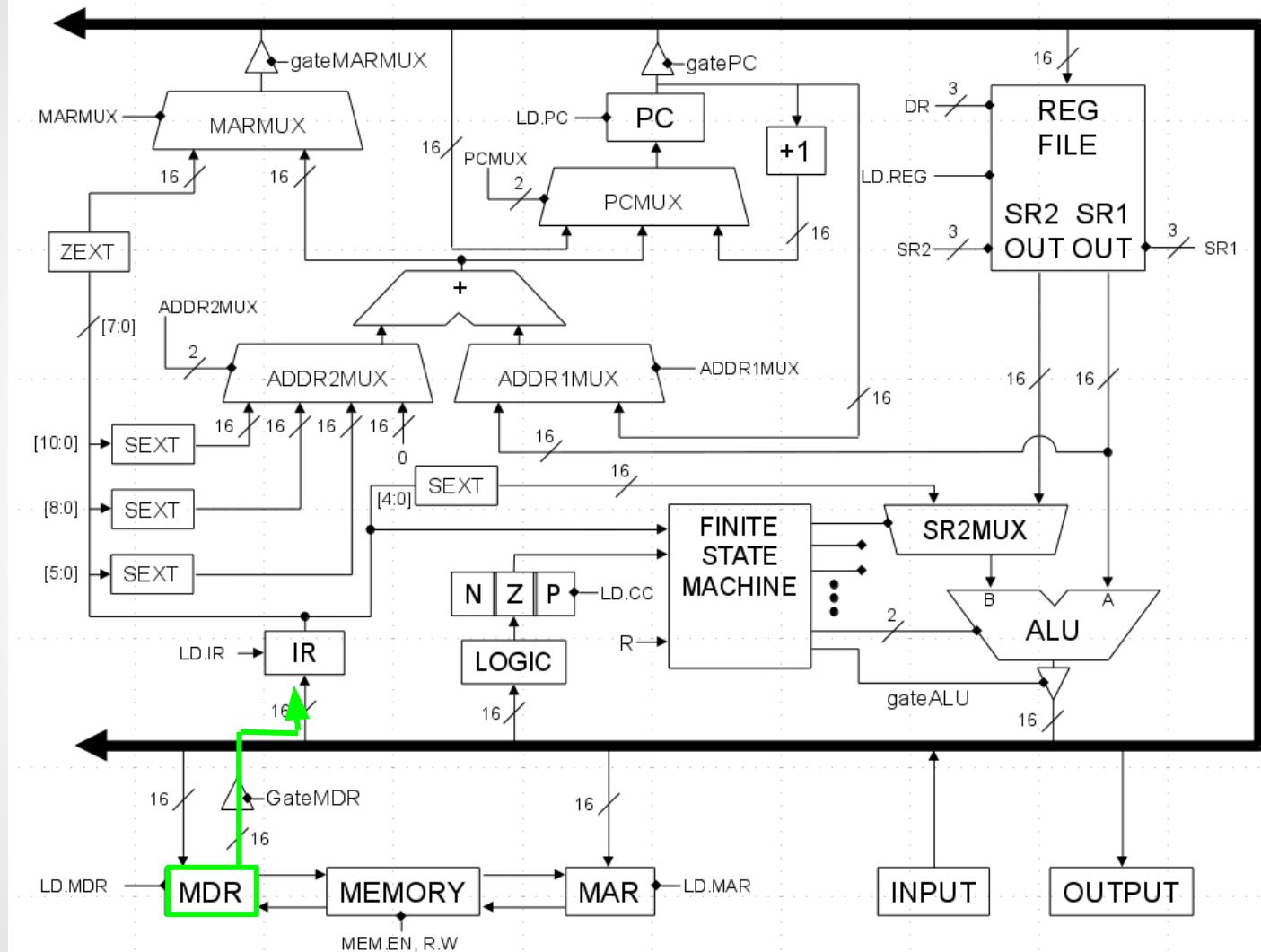
MEM.EN

LD.MDR

CC3:

gateMDR

LD.IR

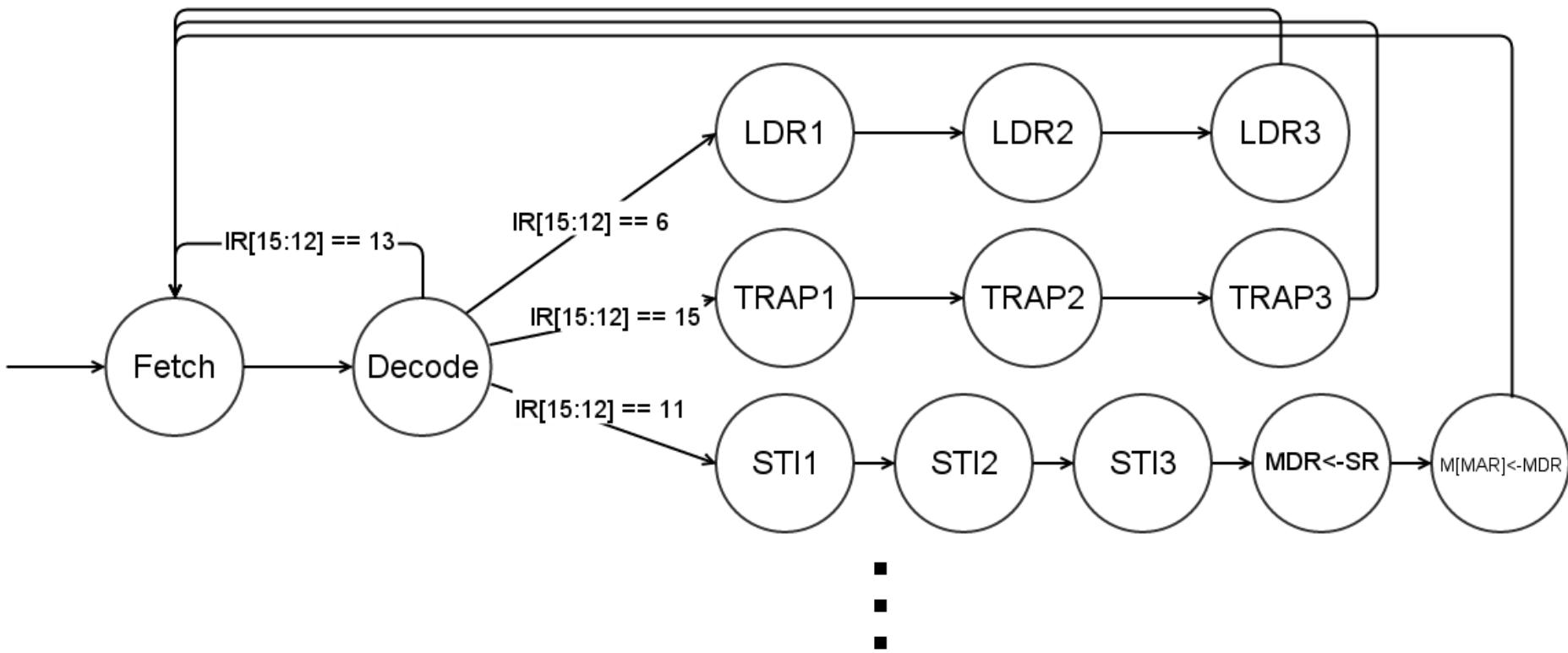


Decode

In the decode stage, the opcode bits are read, and determine the sequence of states to execute next.

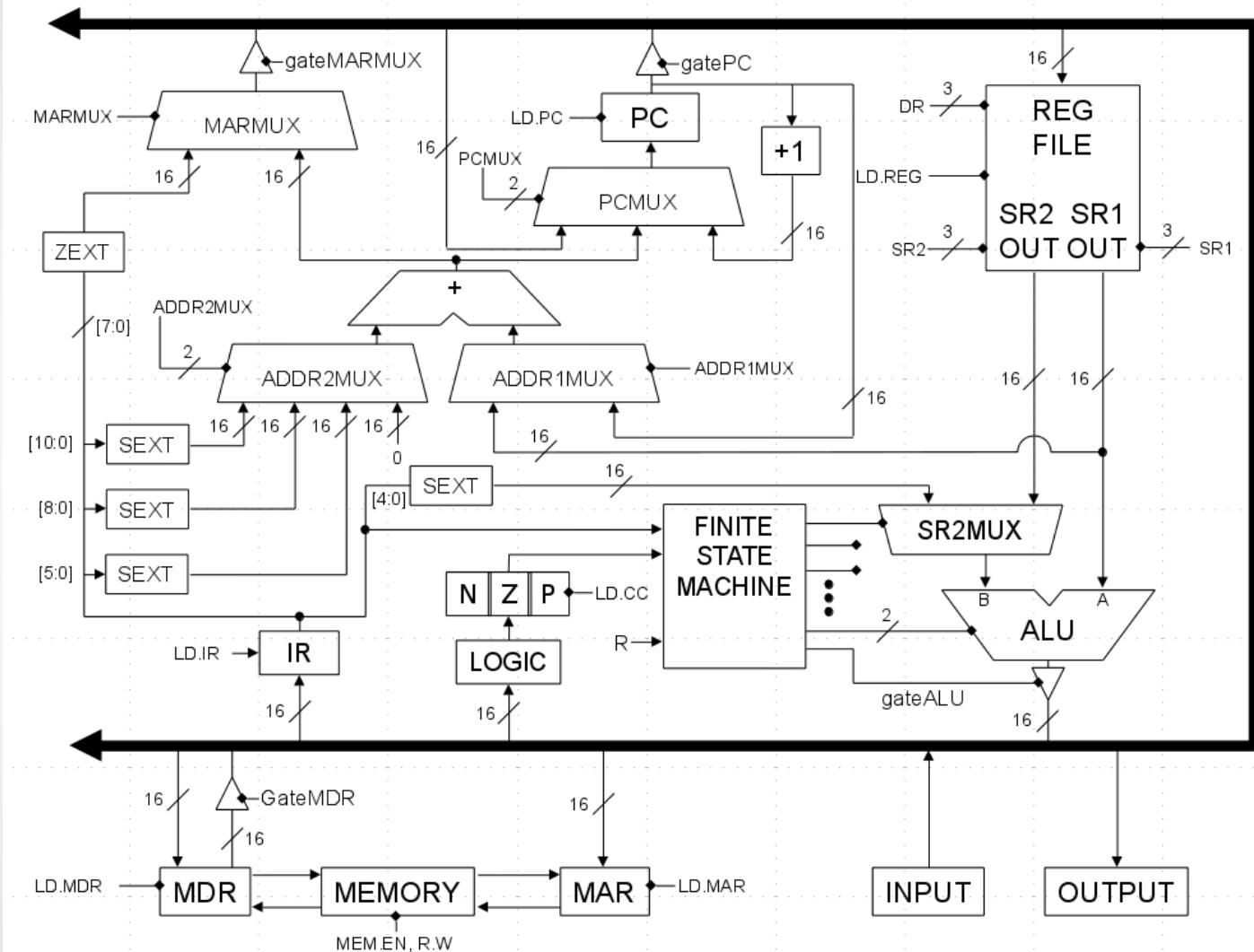
Note: The RTI instruction will be omitted from these slides. Appendix C in the textbook has more information on interrupt support, and interrupts will be covered in more detail in CS 2200.

Decode



Decode

No signals,
because the
decode events take
place within the
state machine.



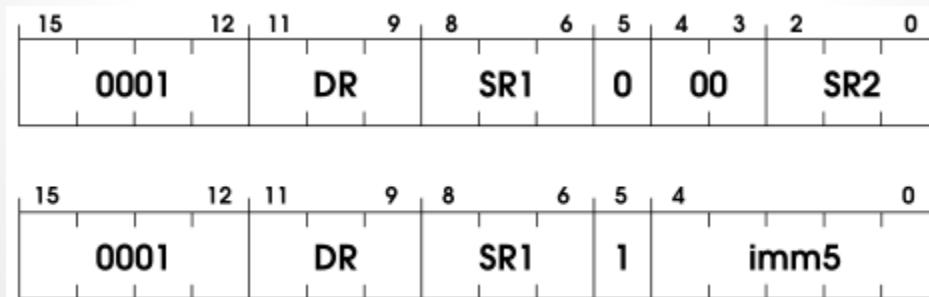
Execute

Each instruction has its own series of states with specific outputs enabled to control the flow of data in the LC3. The 3 general types of instructions are:

- Operations
 - ADD, AND, NOT
- Data transfer
 - LD, LDR, LDI, ST, STR, STI, LEA
- Program control
 - JMP/RET, JSR/JSRR, BR, TRAP

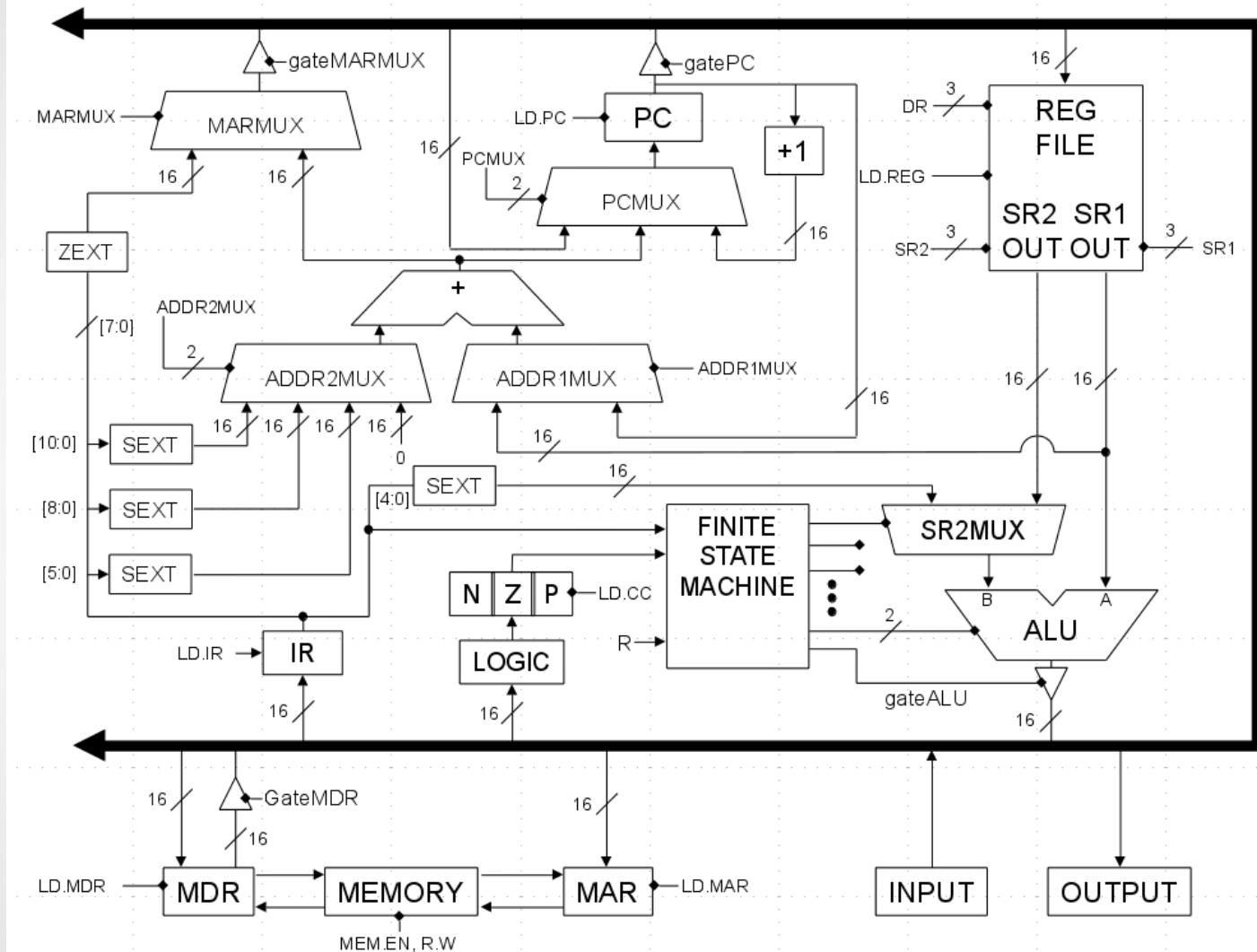
ADD

The ADD instruction adds the contents of a register and a second operand, and stores the result in a register. The second operand may be another register, or an immediate 5-bit value.



$$DR \leftarrow SR1 + (IR[5] == 0 ? SR2 : imm5)$$

ADD (SR2)



ADD (SR2)

CC1:

SR1 = IR[8:6]

SR2 = IR[2:0]

SR2MUX = SR2

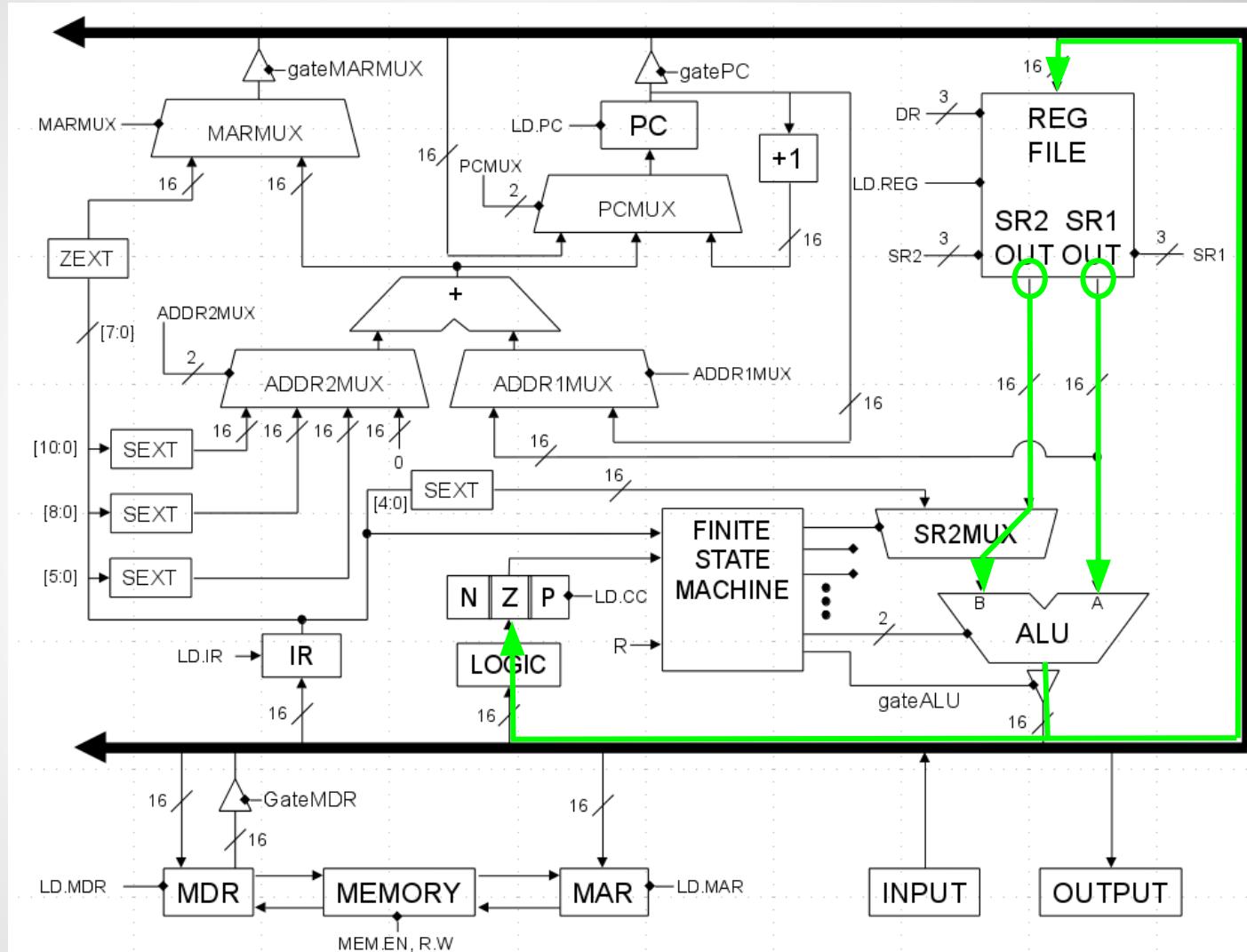
ALUK = ADD

gateALU

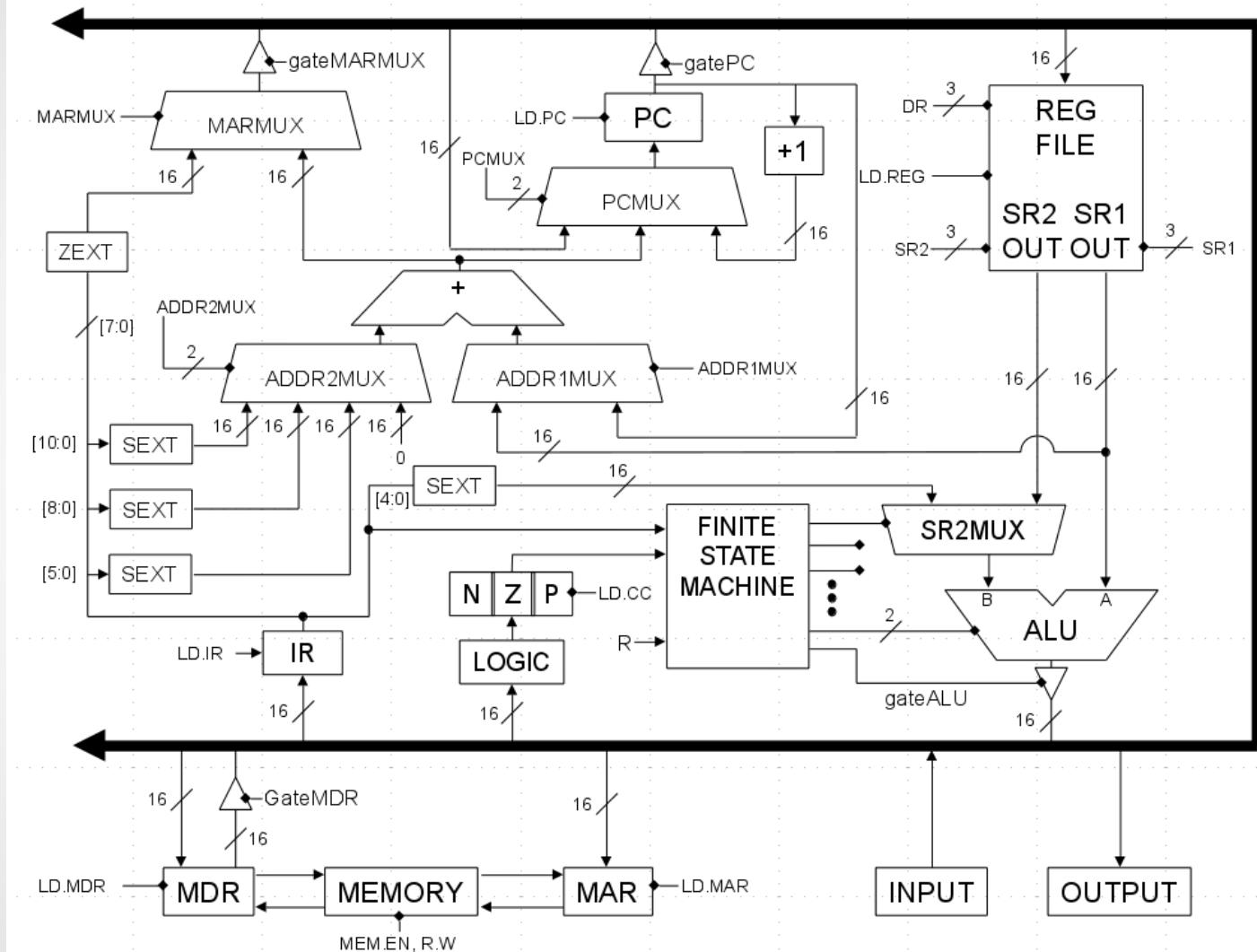
LD.REG

LD.CC

DR = IR[11:9]



ADD (imm5)



ADD (imm5)

CC1:

SR1 = IR[8:6]

SR2MUX = SEXT

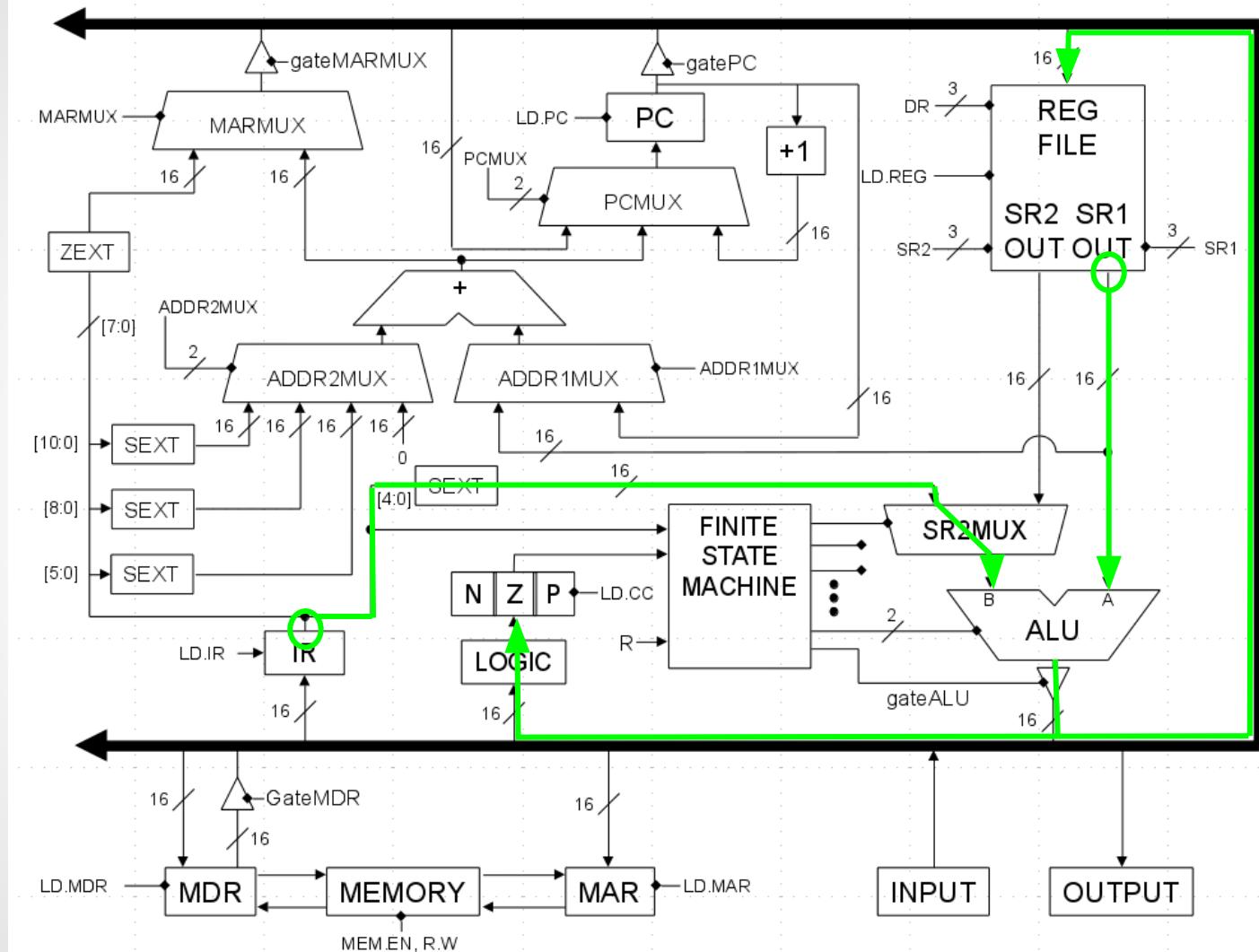
ALUK = ADD

gateALU

LD.REG

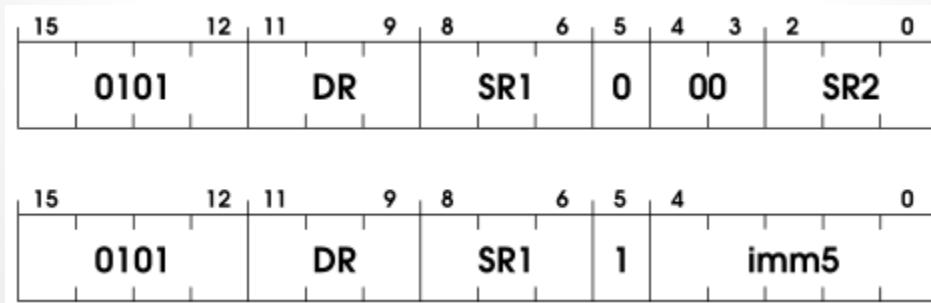
LD.CC

DR = IR[11:9]



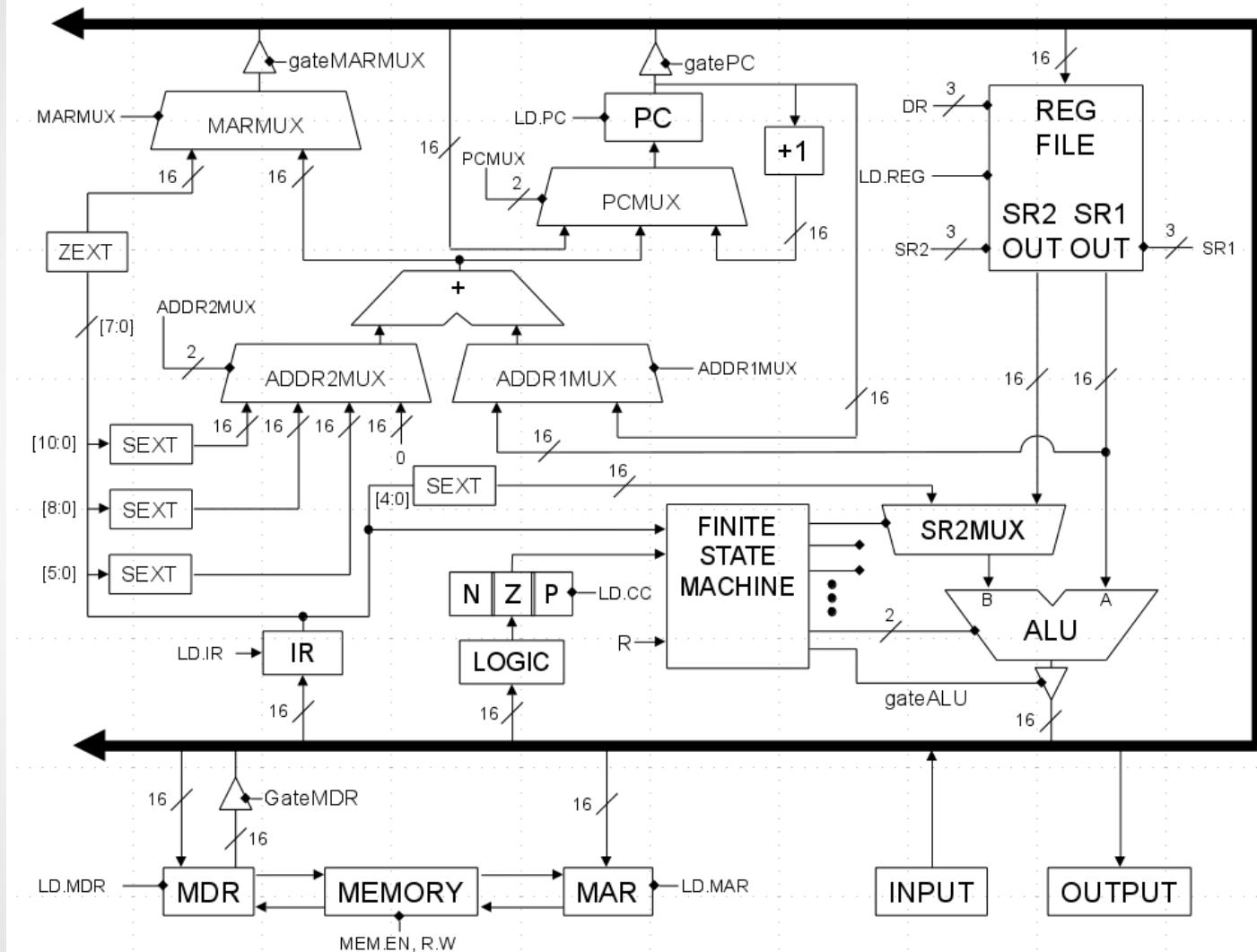
AND

The AND instruction performs a bitwise AND of a register and second operand, and stores the result in a register. AND also has an option for secondary register or immediate value.



$$DR \leftarrow SR1 \& (IR[5] == 0 ? SR2 : imm5)$$

AND (SR2)



AND (SR2)

CC1:

SR1 = IR[8:6]

SR2 = IR[2:0]

SR2MUX = SR2

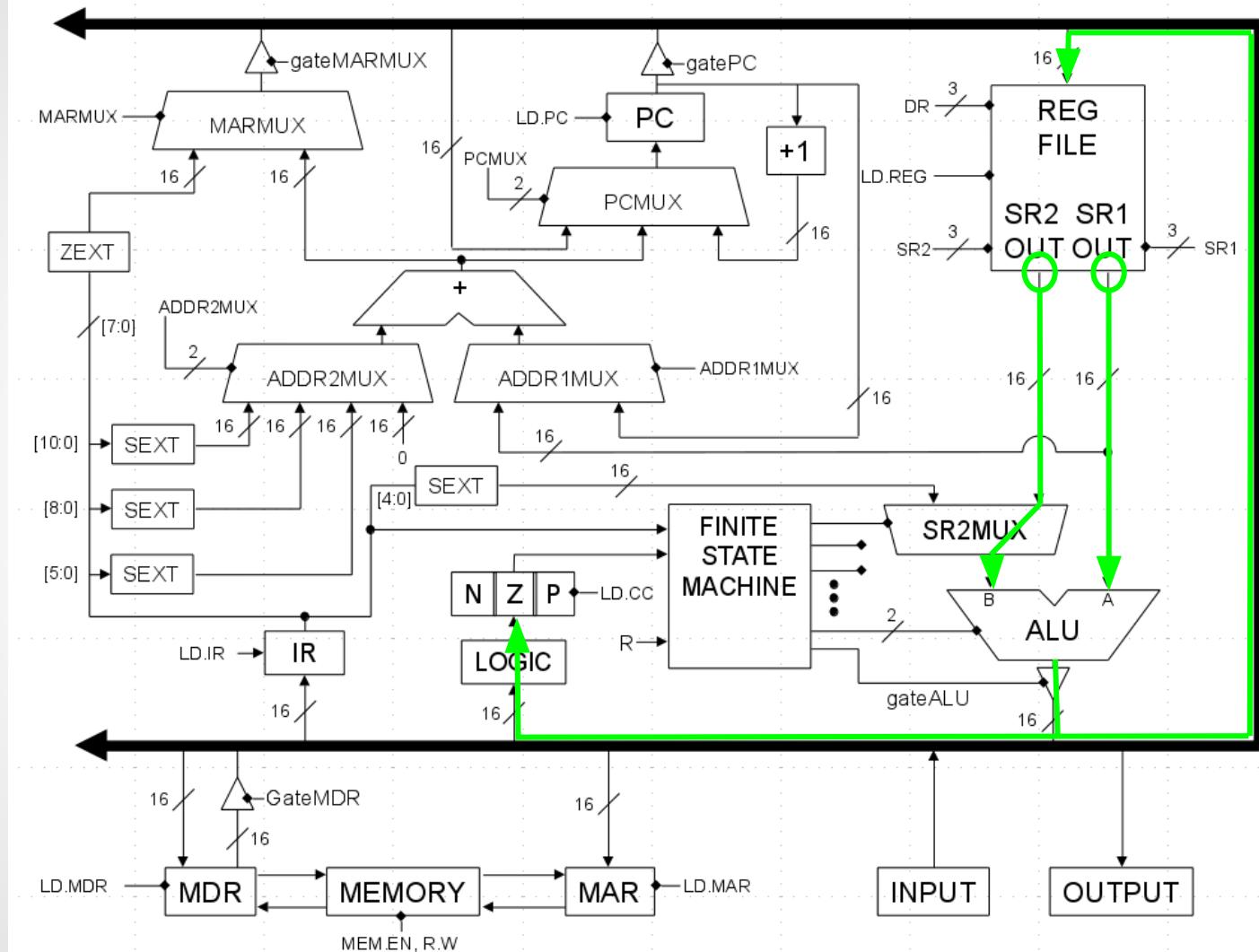
ALUK = AND

gateALU

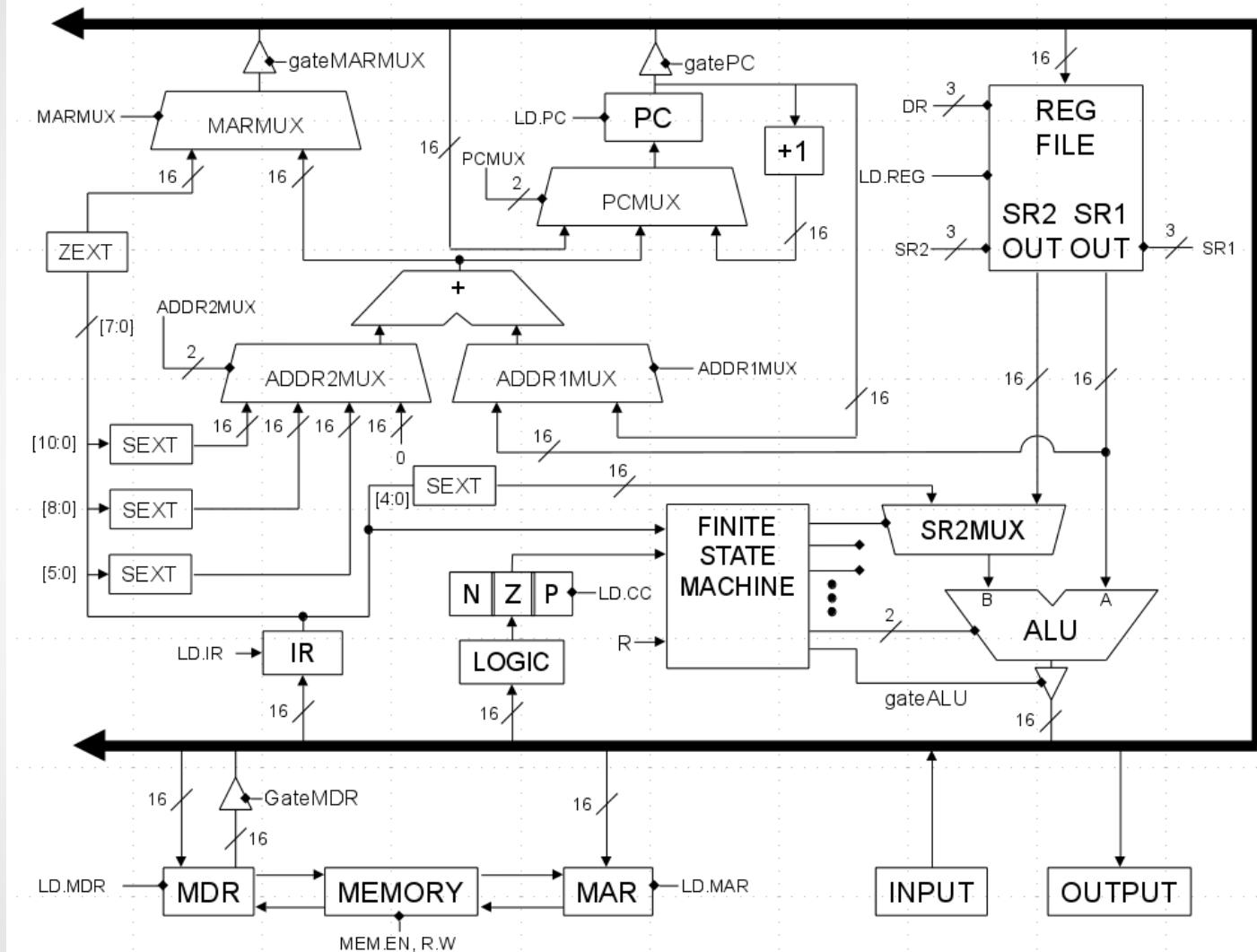
LD.REG

LD.CC

DR = IR[11:9]



AND (imm5)



AND (imm5)

CC1:

SR1 = IR[8:6]

SR2MUX = SEXT

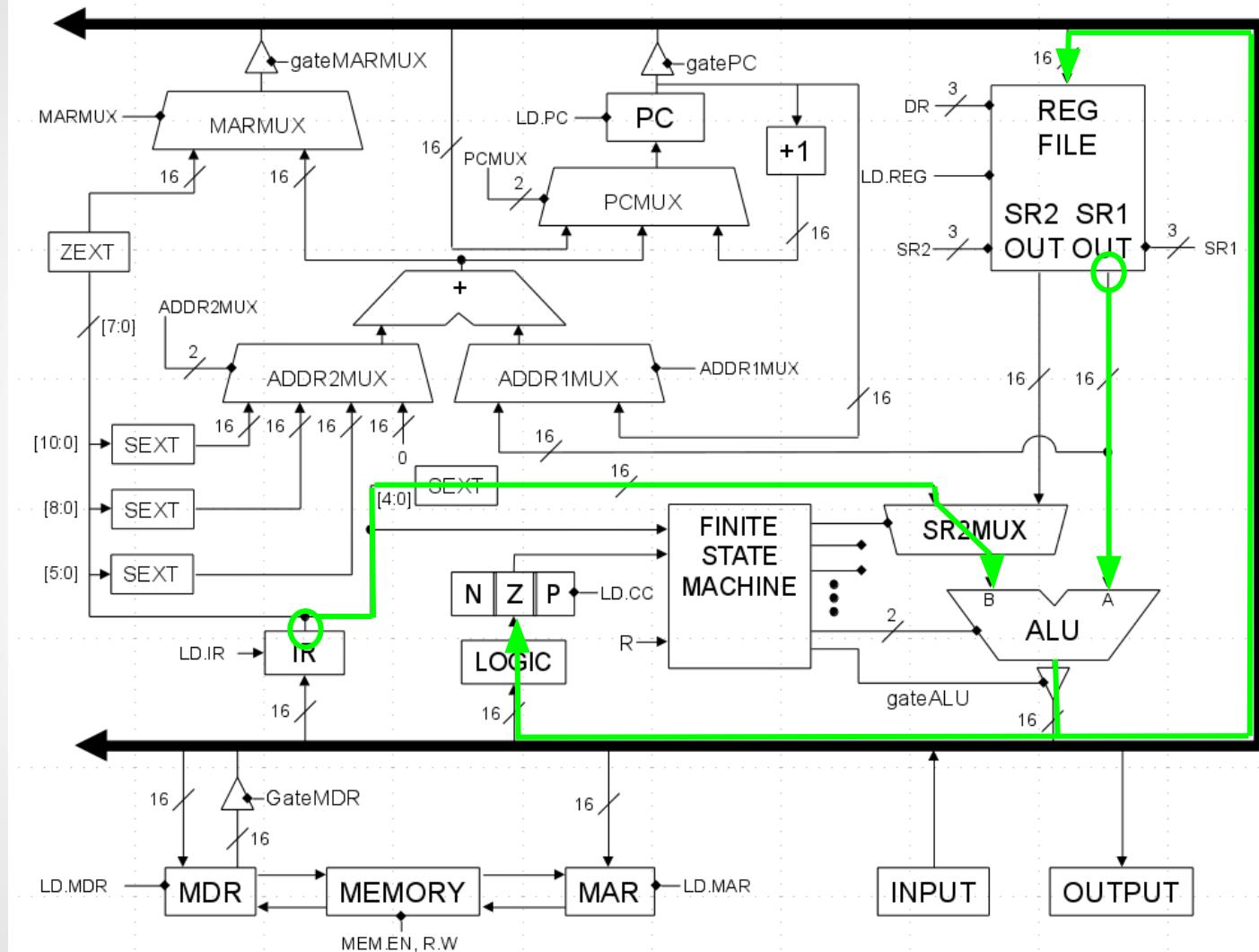
ALUK = AND

gateALU

LD.REG

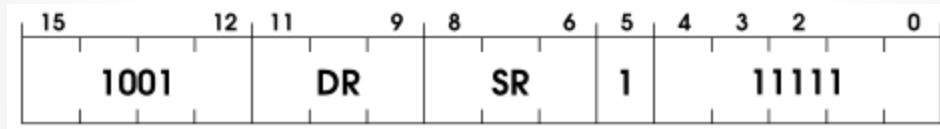
LD.CC

DR = IR[11:9]



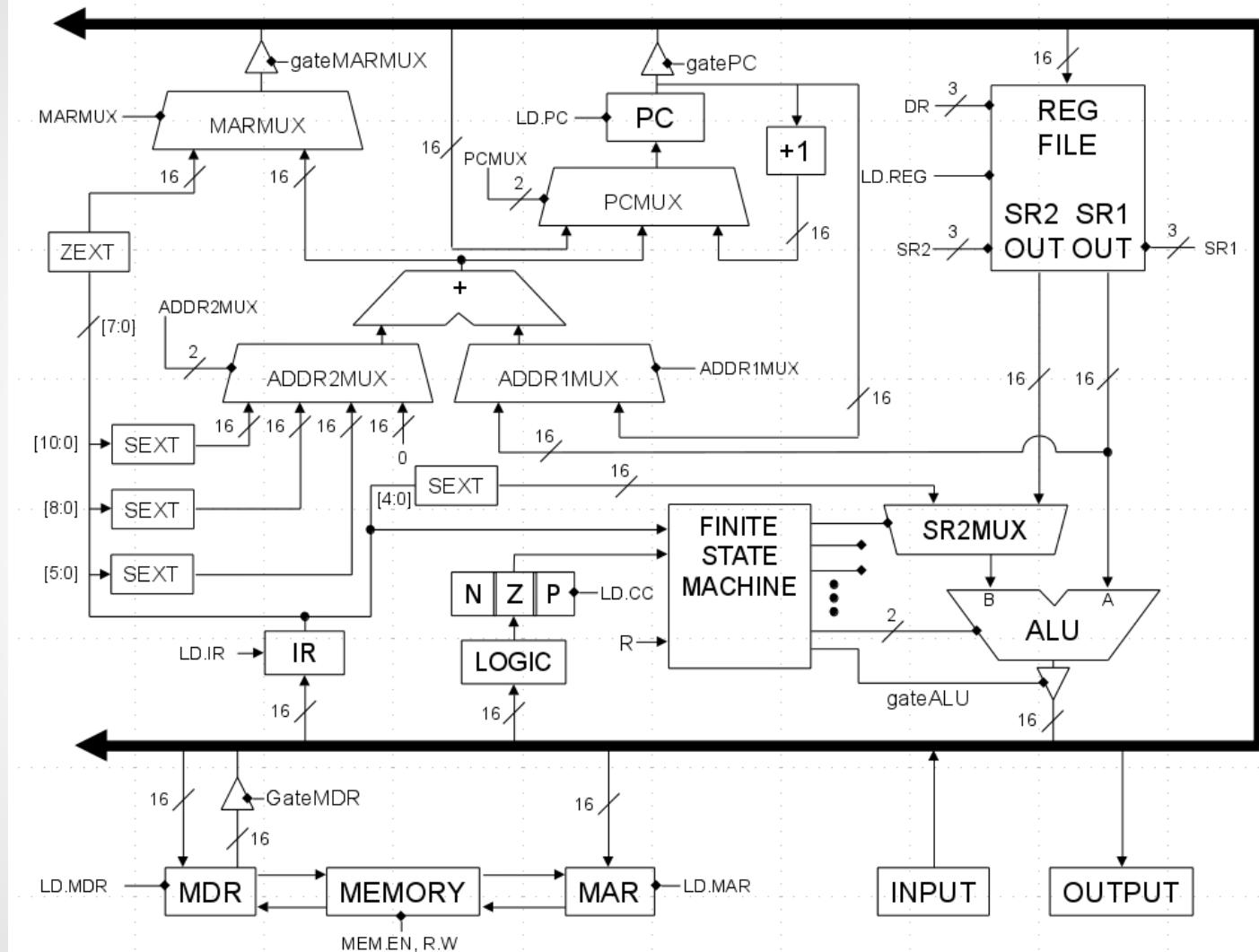
NOT

The NOT instruction performs a bitwise NOT of a register, and stores it in another register.



$$DR \leftarrow \sim SR$$

NOT



NOT

CC1:

SR1 = IR[8:6]

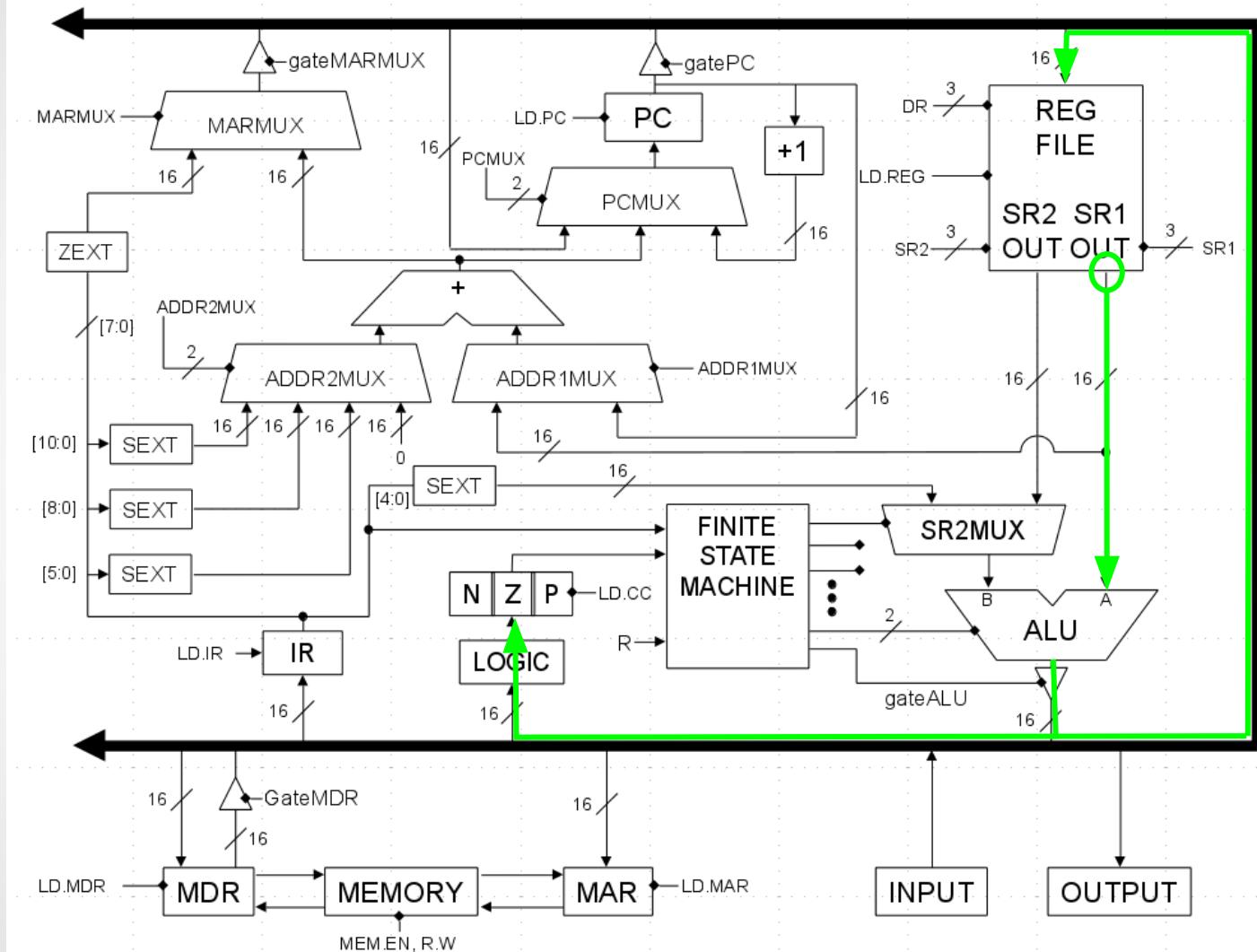
ALUK = NOT

gateALU

LD.REG

LD.CC

DR = IR[11:9]



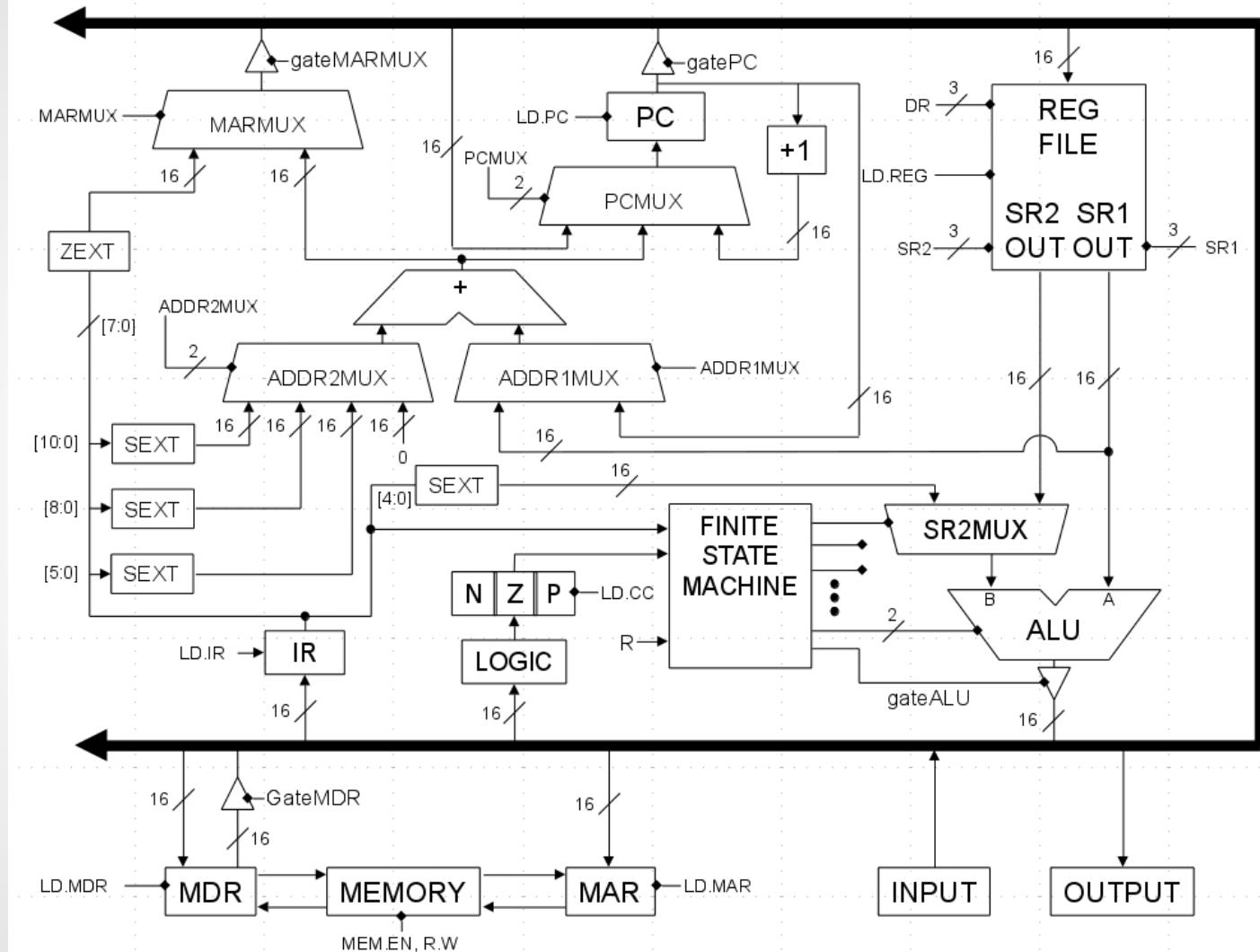
LD

The LD instruction loads a value from memory, nearby the instruction itself, into a register.



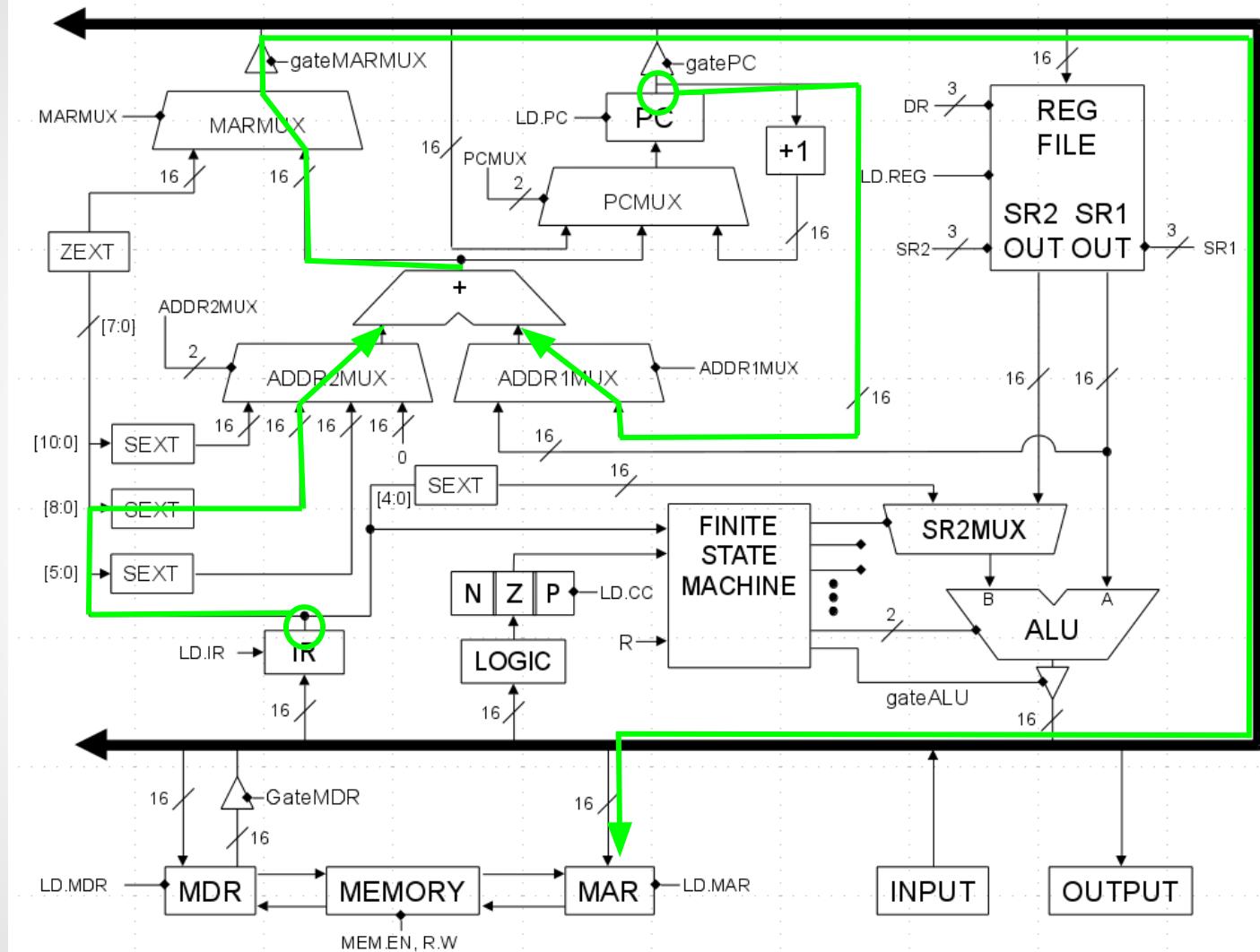
$$DR \leftarrow \text{mem}[PC + PCoffset9]$$

LD



LD

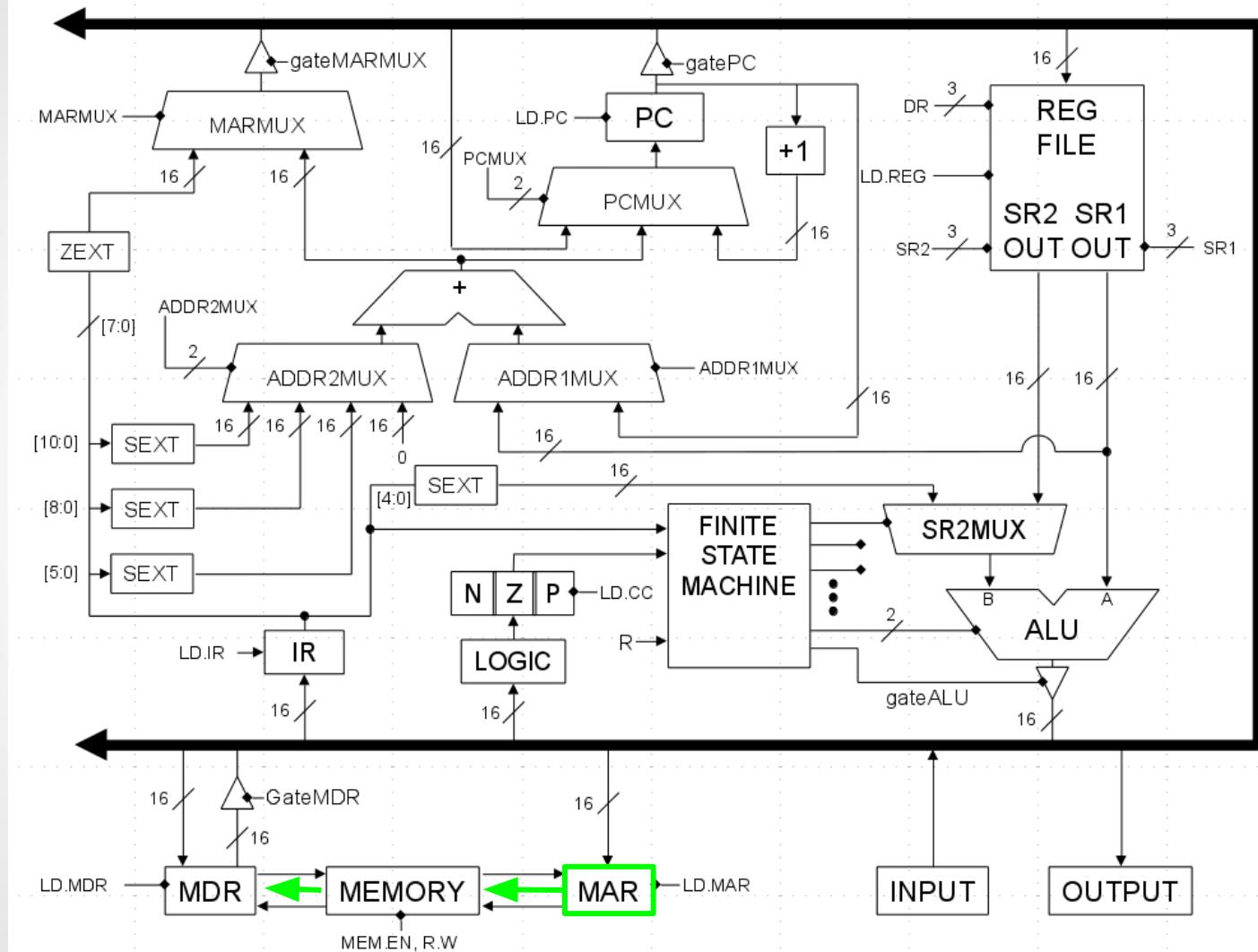
CC1:
ADDR1MUX = PC
ADDR2MUX =
SEXT[8:0]
MARMUX = ADDER
gateMARMUX
LD.MAR



LD

CC1:
ADDR1MUX = PC
ADDR2MUX =
SEXT[8:0]
MARMUX = ADDER
gateMARMUX
LD.MAR

CC2:
MEM.EN
LD.MDR

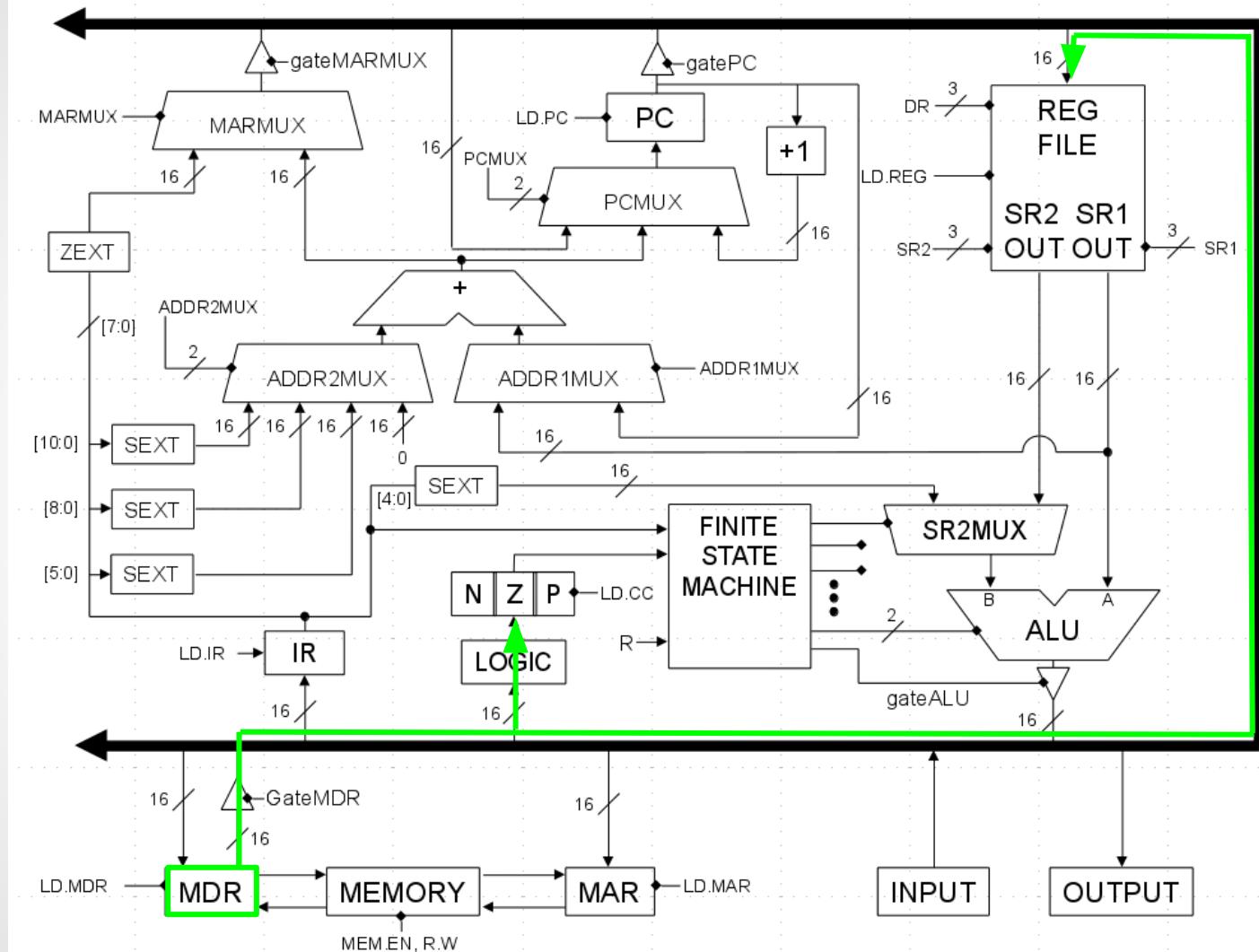


LD

CC1:
ADDR1MUX = PC
ADDR2MUX =
 $\text{SEXT}[8:0]$
MARMUX = ADDER
gateMARMUX
LD.MAR

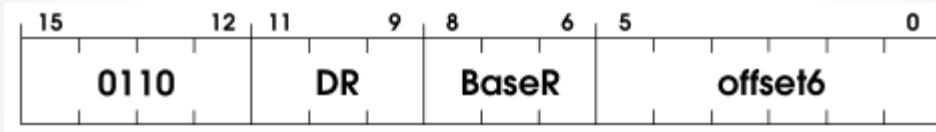
CC2:
MEM.EN
LD.MDR

CC3:
gateMDR
 $\text{DR} = \text{IR}[11:9]$
LD.REG
LD.CC



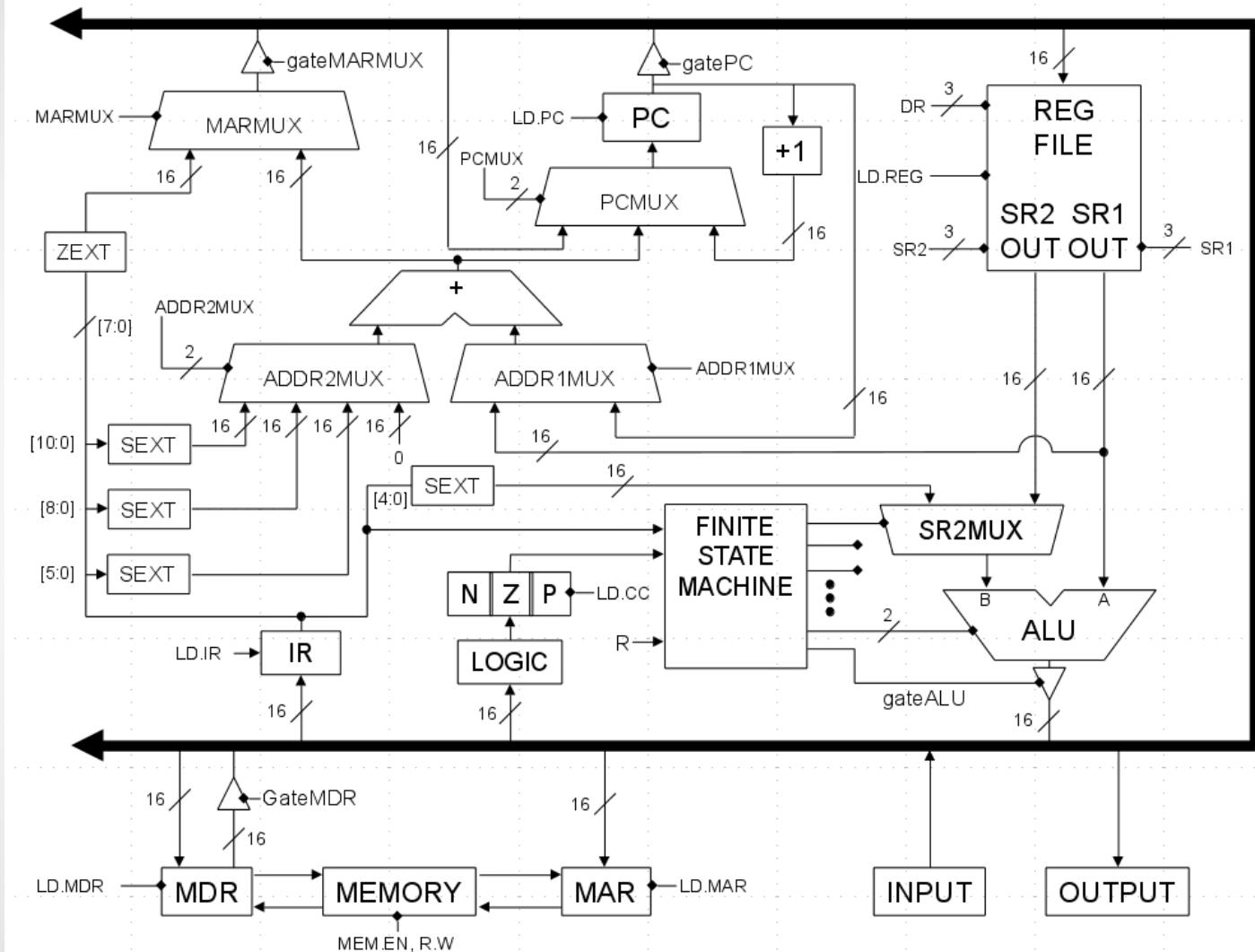
LDR

The LDR instruction loads a value from memory, pointed to by a base register and offset from that location, into a register.



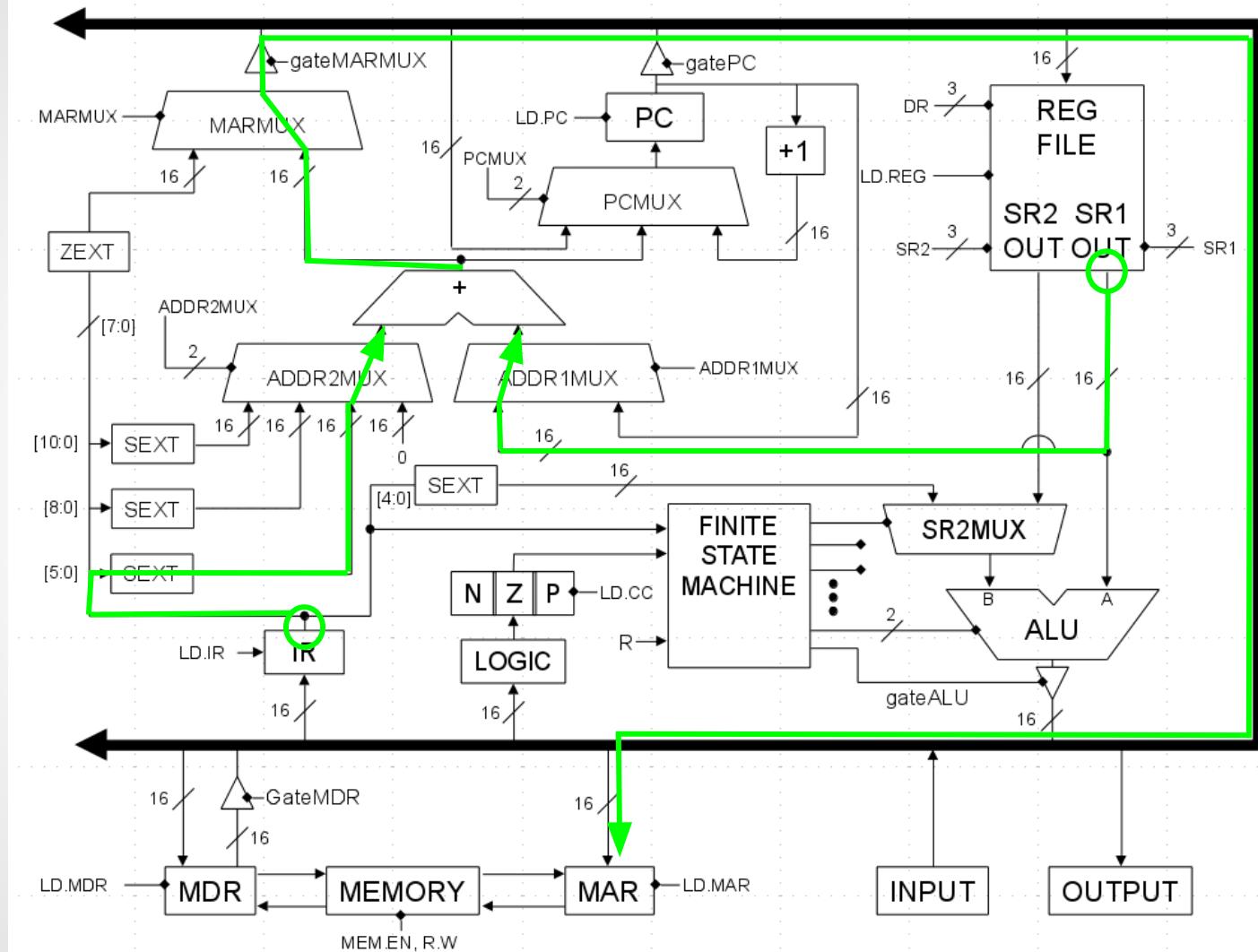
$$DR \leftarrow \text{mem}[\text{BaseR} + \text{offset6}]$$

LDR



LDR

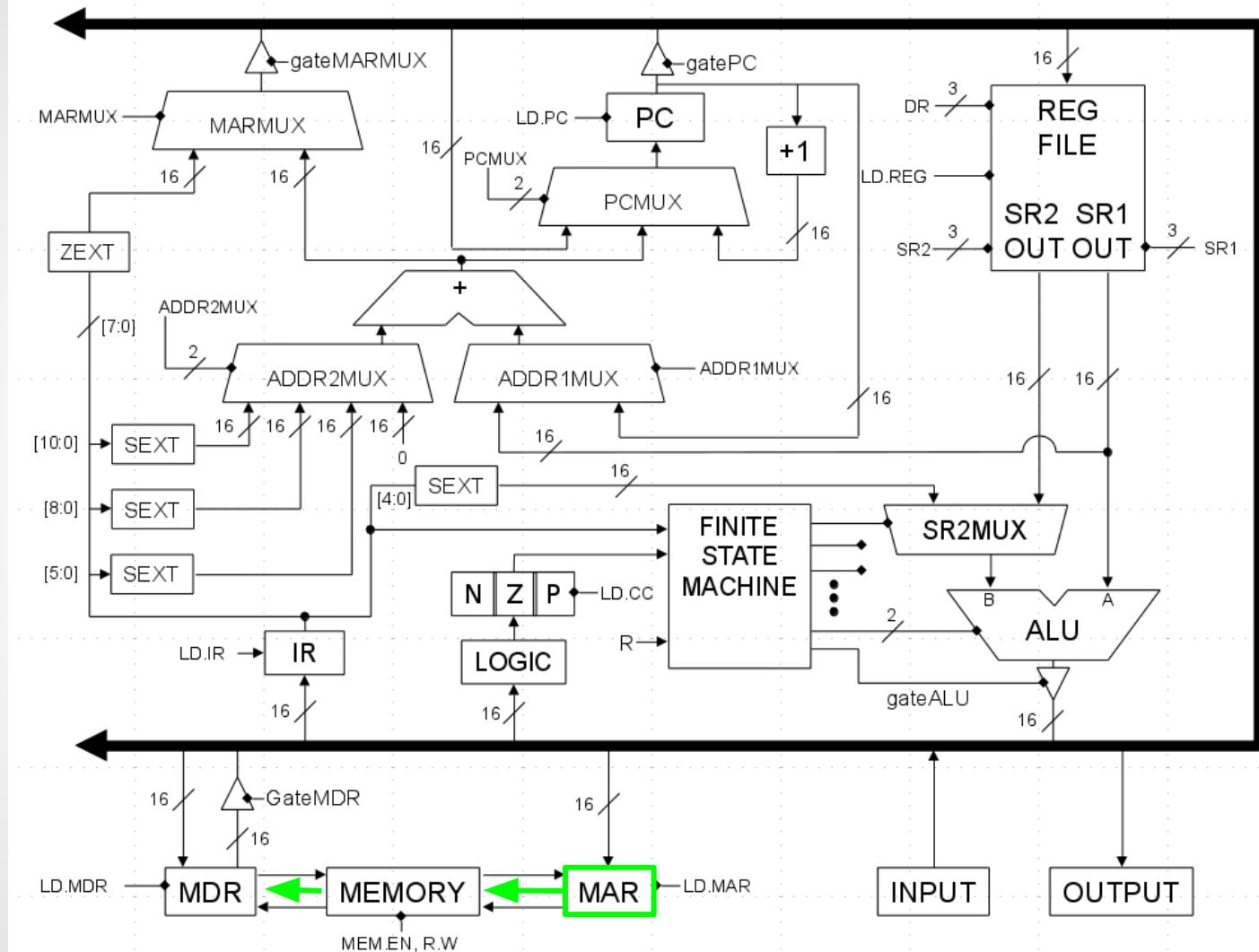
CC1:
ADDR1MUX = BaseR
ADDR2MUX =
SEXT[5:0]
MARMUX = ADDER
gateMARMUX
LD.MAR



LDR

CC1:
ADDR1MUX = BaseR
ADDR2MUX =
SEXT[5:0]
MARMUX = ADDER
gateMARMUX
LD.MAR

CC2:
MEM.EN
LD.MDR



LDR

CC1:
ADDR1MUX = BaseR
ADDR2MUX =

SEXT[5:0]

MARMUX = ADDER

gateMARMUX

LD.MAR

CC2:

MEM.EN

LD.MDR

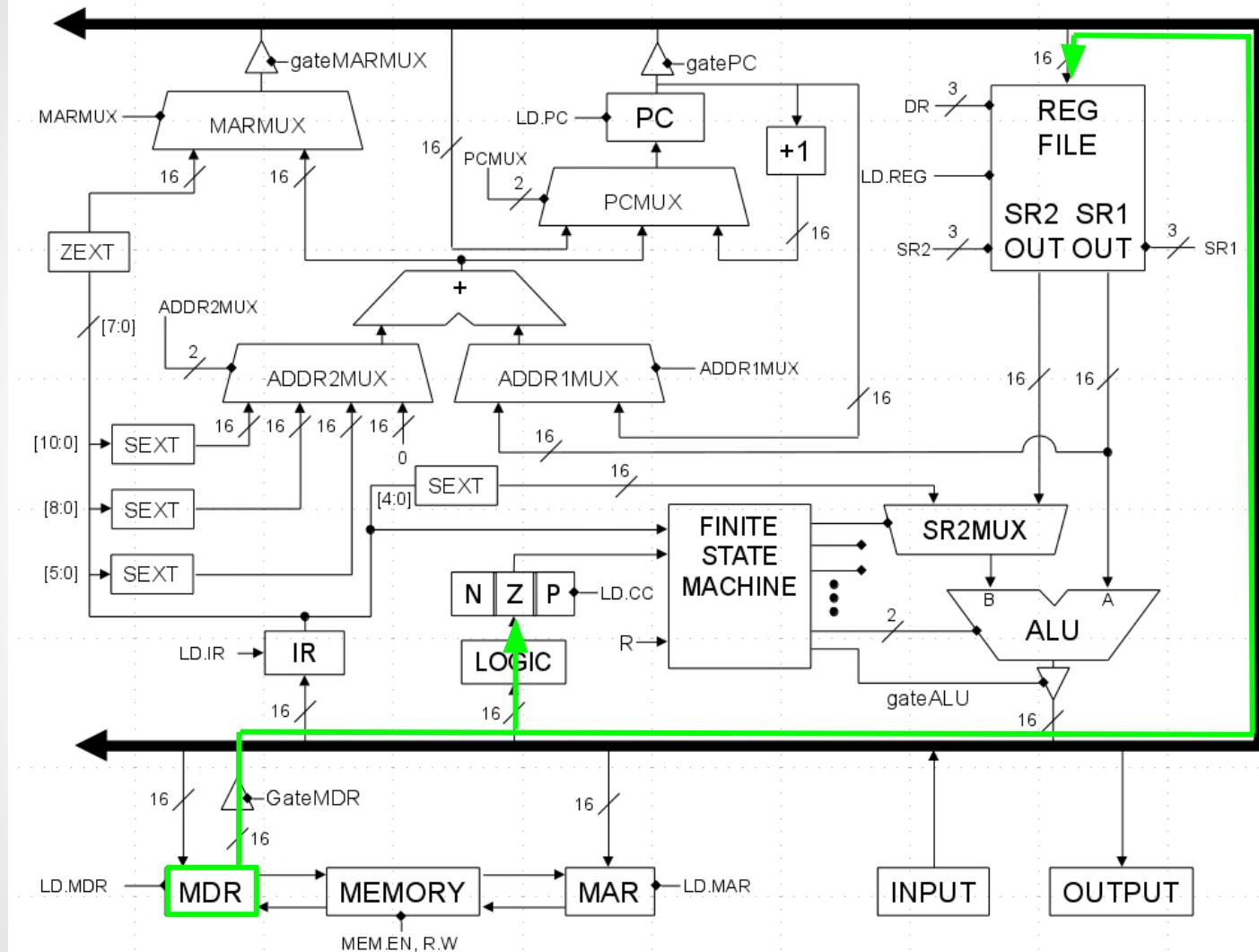
CC3:

gateMDR

DR = IR[11:9]

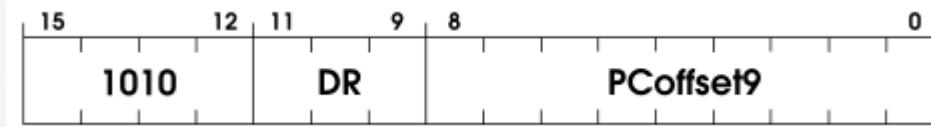
LD.REG

LD.CC



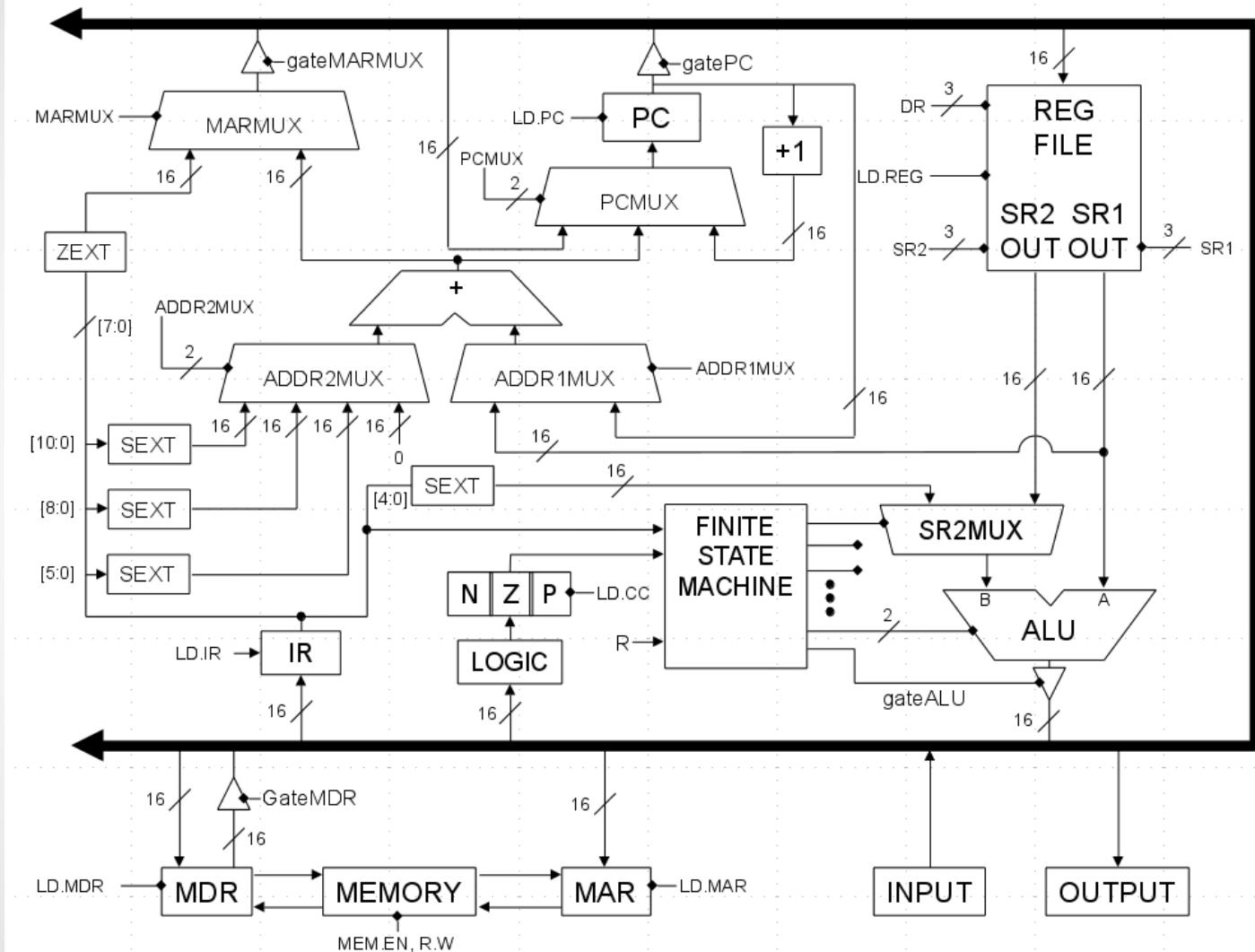
LDI

The LDI instruction takes a value from memory nearby the instruction itself, gets a pointer from that memory location, then loads the contents of the pointed location into a register.



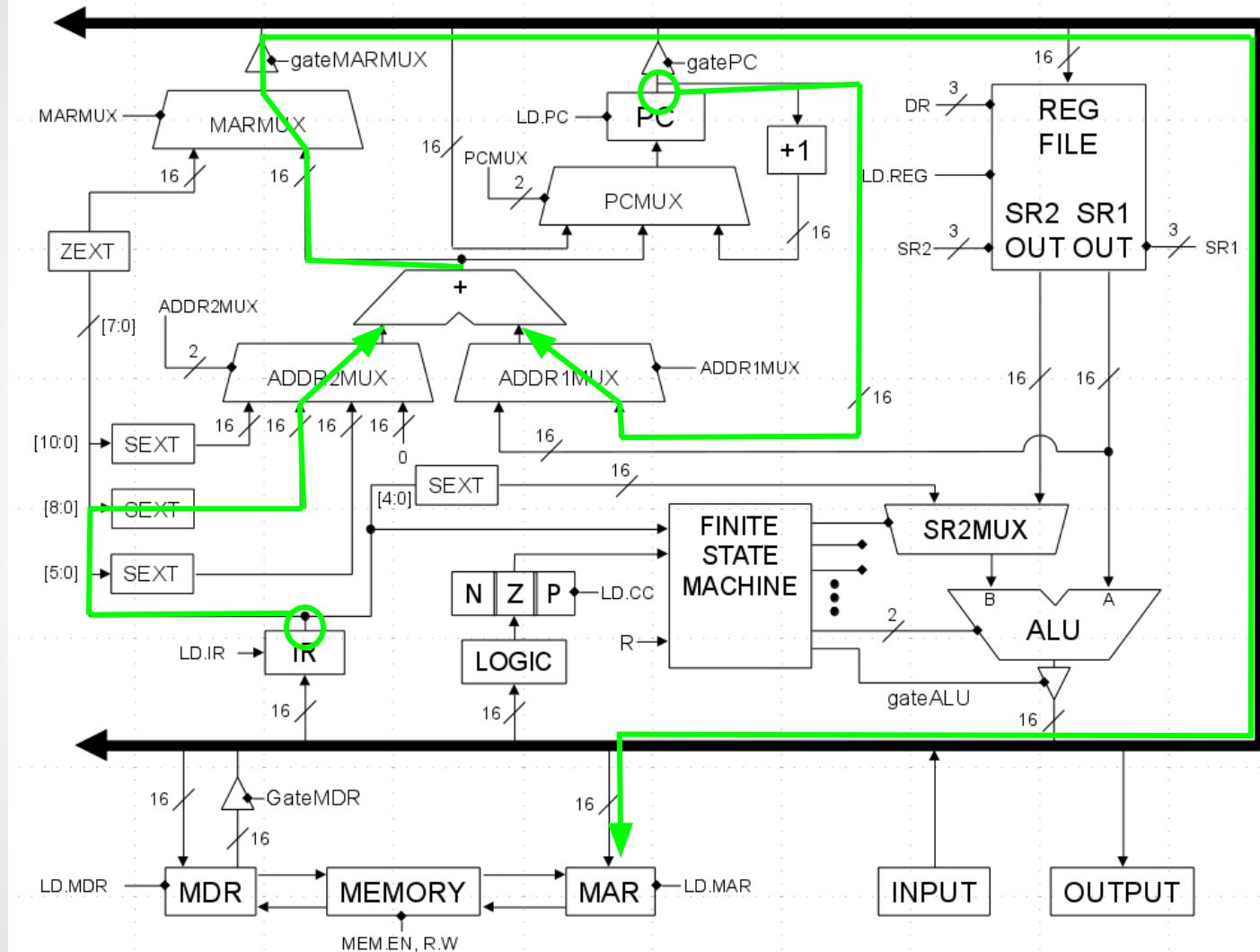
$$DR \leftarrow \text{mem}[\text{mem}[PC + PCoffset9]]$$

LDI



LDI

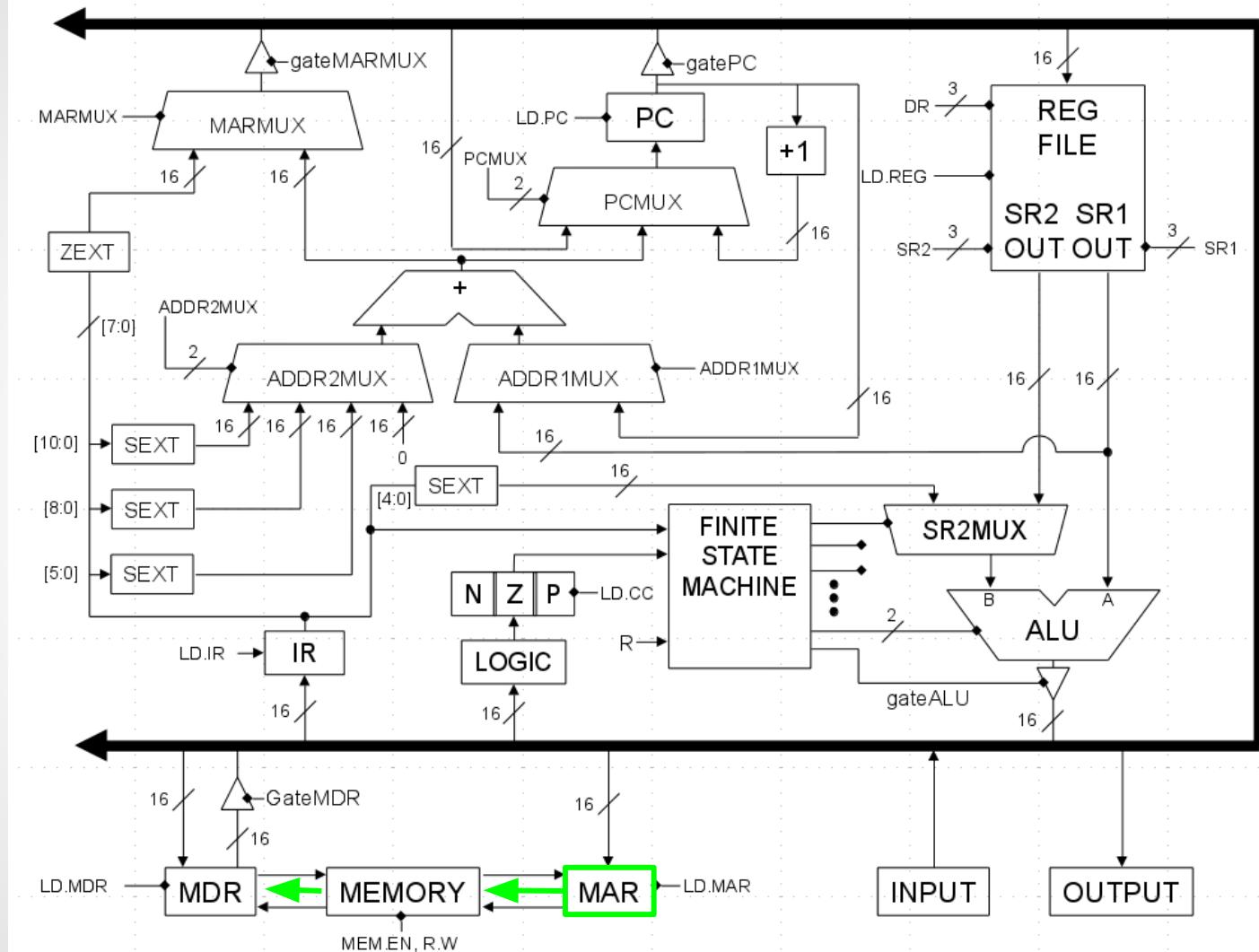
CC1:
ADDR1MUX = PC
ADDR2MUX =
SEXT[8:0]
MARMUX = ADDER
gateMARMUX
LD.MAR



LDI

CC1:
ADDR1MUX = PC
ADDR2MUX =
SEXT[8:0]
MARMUX = ADDER
gateMARMUX
LD.MAR

CC2:
MEM.EN
LD.MDR



LDI

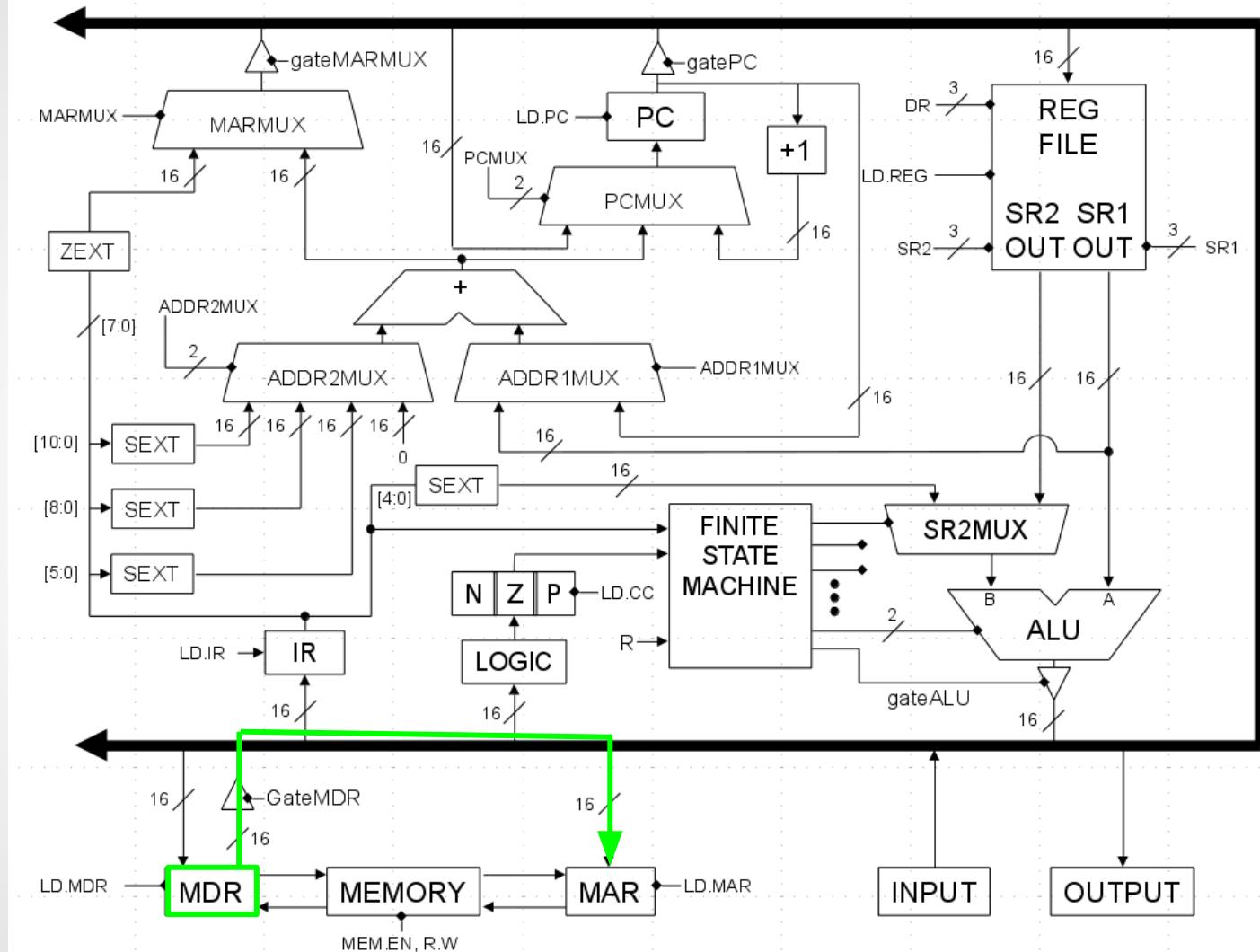
CC1:
ADDR1MUX = PC

ADDR2MUX =
SEXT[8:0]

MARMUX = ADDER
gateMARMUX
LD.MAR

CC2:
MEM.EN
LD.MDR

CC3:
gateMDR
LD.MAR



LDI

CC1:
ADDR1MUX = PC

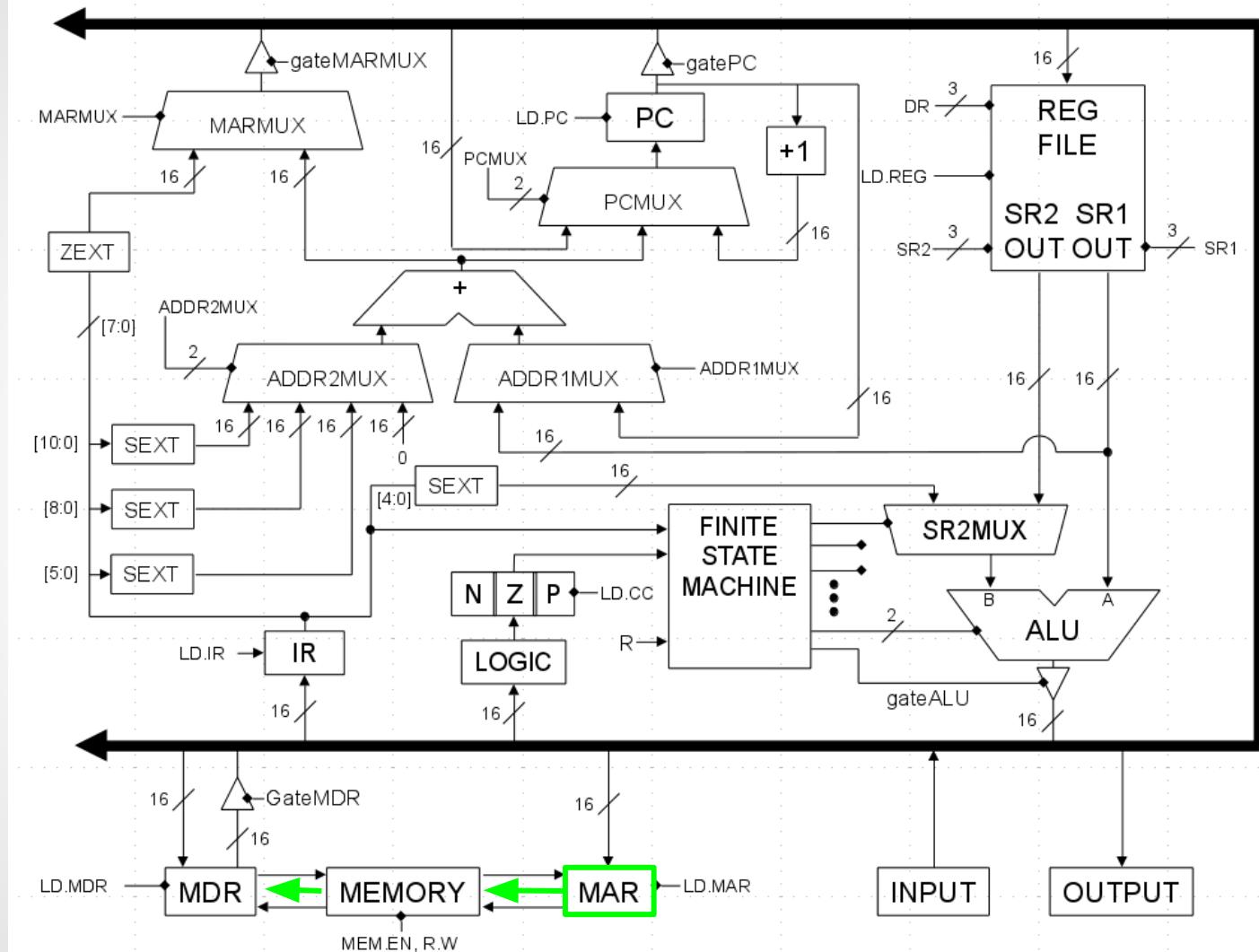
ADDR2MUX =
SEXT[8:0]

MARMUX = ADDER
gateMARMUX
LD.MAR

CC2:
MEM.EN
LD.MDR

CC3:
gateMDR
LD.MAR

CC4:
MEM.EN
LD.MDR



LDI

CC1:

[...]

CC2:

MEM.EN
LD.MDR

CC3:

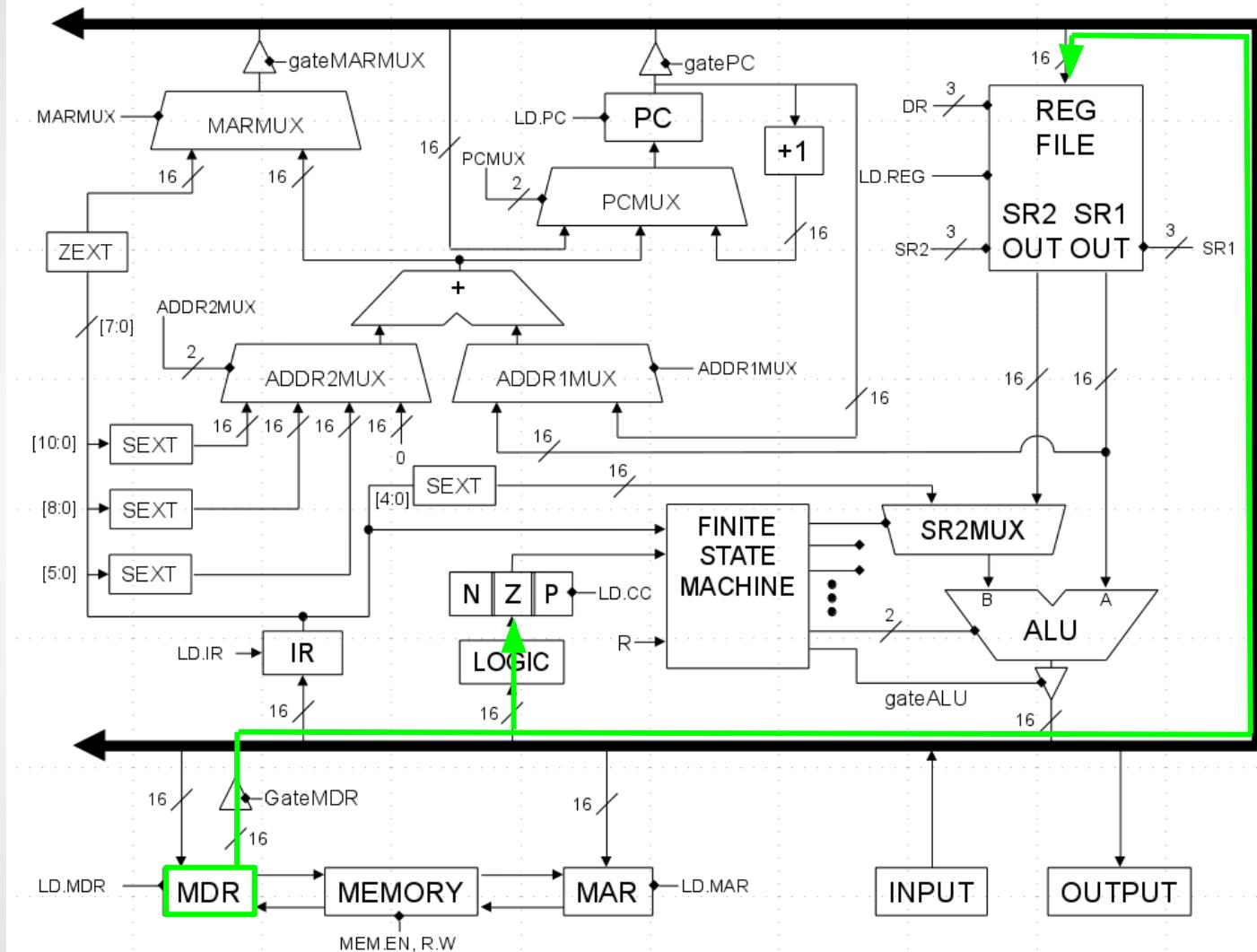
gateMDR
LD.MAR

CC4:

MEM.EN
LD.MDR

CC5:

gateMDR
DR = IR[11:9]
LD.REG
LD.CC

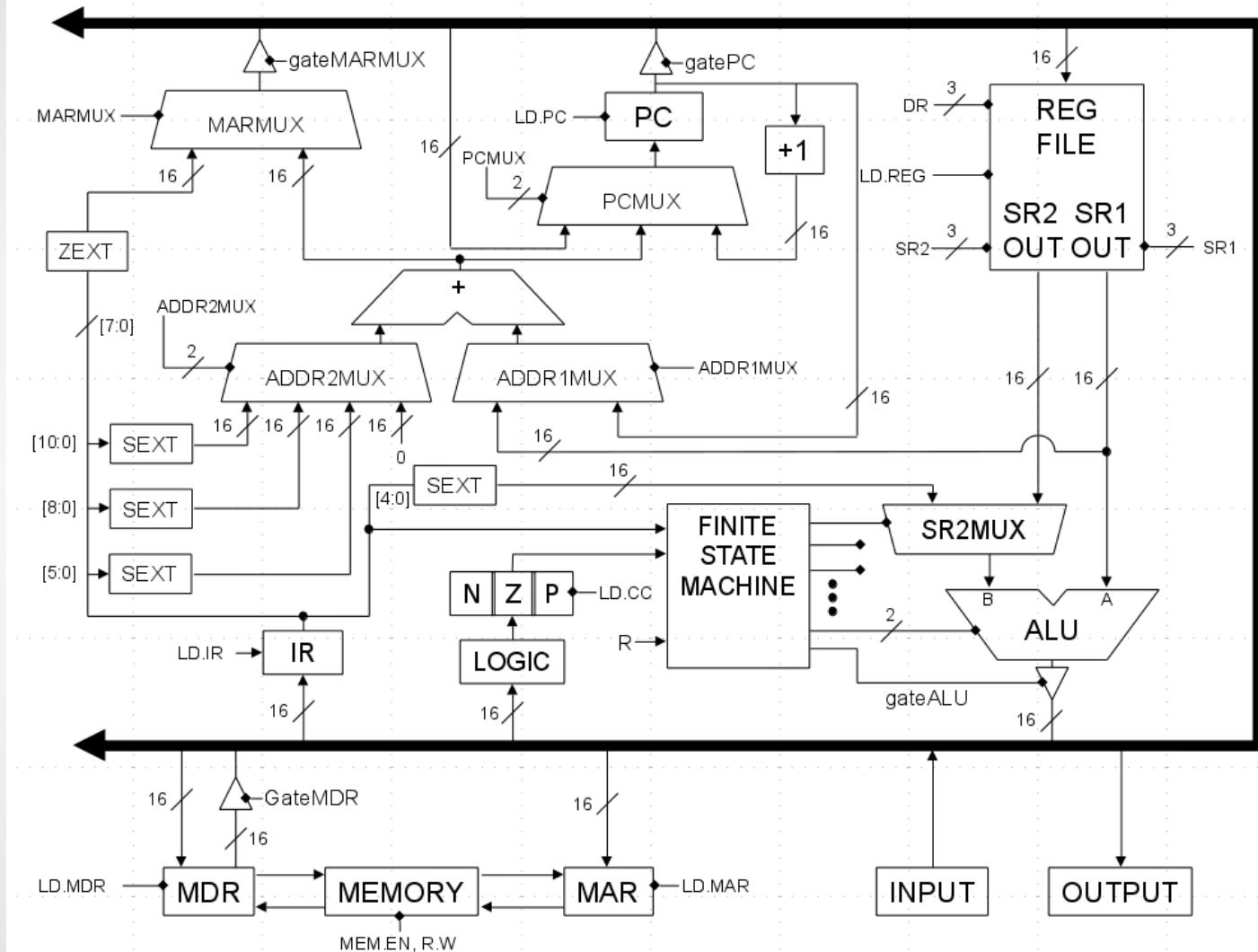


ST

The ST instruction stores a value from a register into memory, at a location near the instruction itself.

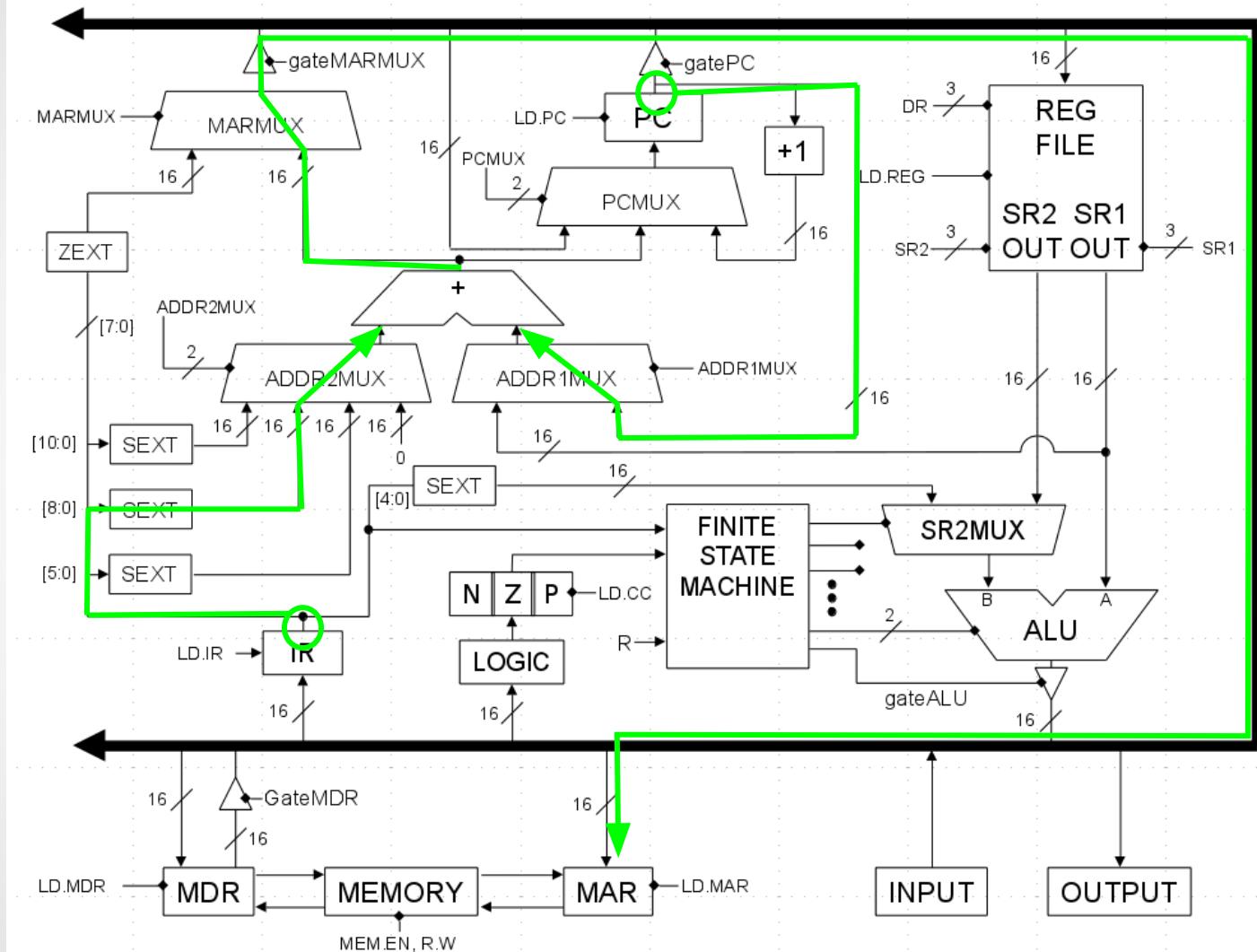

$$\text{mem}[\text{PC} + \text{PCoffset9}] \leftarrow \text{SR}$$

ST



ST

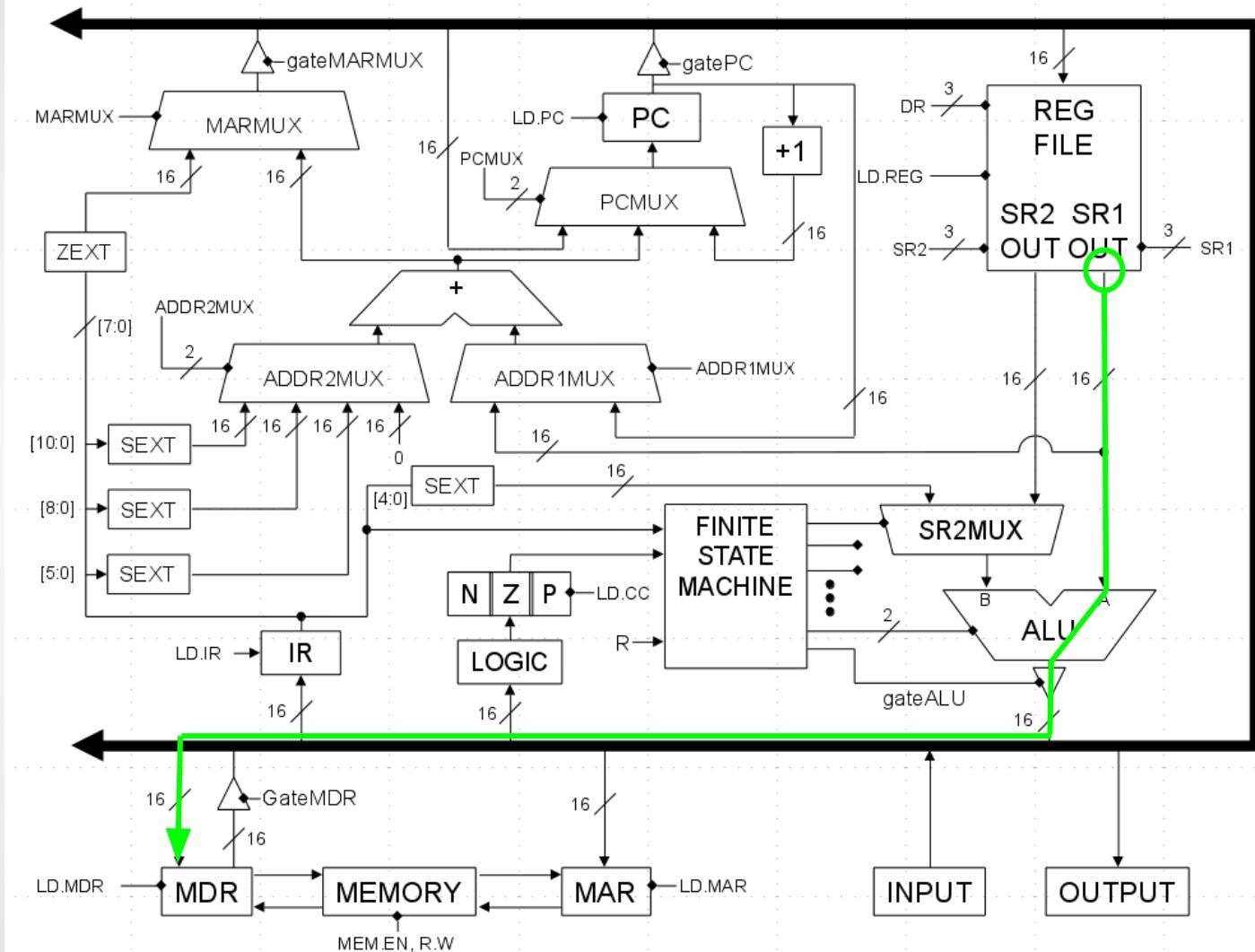
CC1:
ADDR1MUX = PC
ADDR2MUX =
SEXT[8:0]
MARMUX = ADDER
gateMARMUX
LD.MAR



ST

CC1:
ADDR1MUX = PC
ADDR2MUX =
SEXT[8:0]
MARMUX = ADDER
gateMARMUX
LD.MAR

CC2:
SR1 = IR[11:9]
ALUK = PASSA
gateALU
LD.MDR

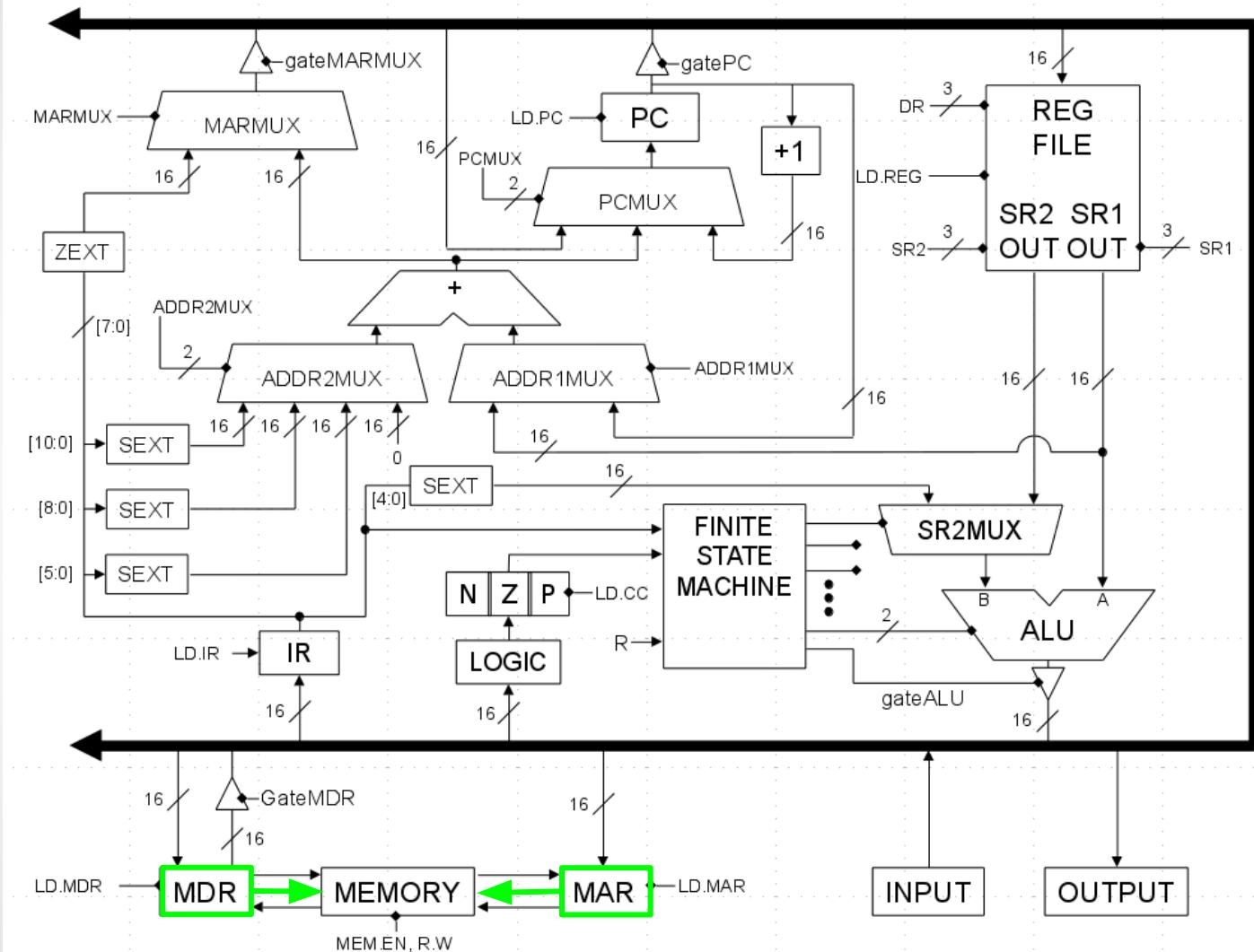


ST

CC1:
ADDR1MUX = PC
ADDR2MUX =
 $\text{SEXT}[8:0]$
MARMUX = ADDER
gateMARMUX
LD.MAR

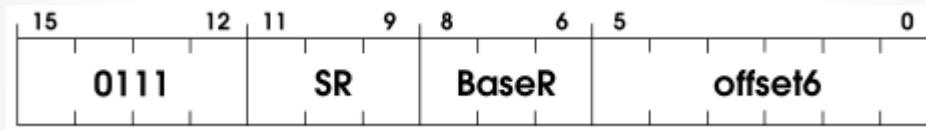
CC2:
SR1 = IR[11:9]
ALUK = PASSA
gateALU
LD.MDR

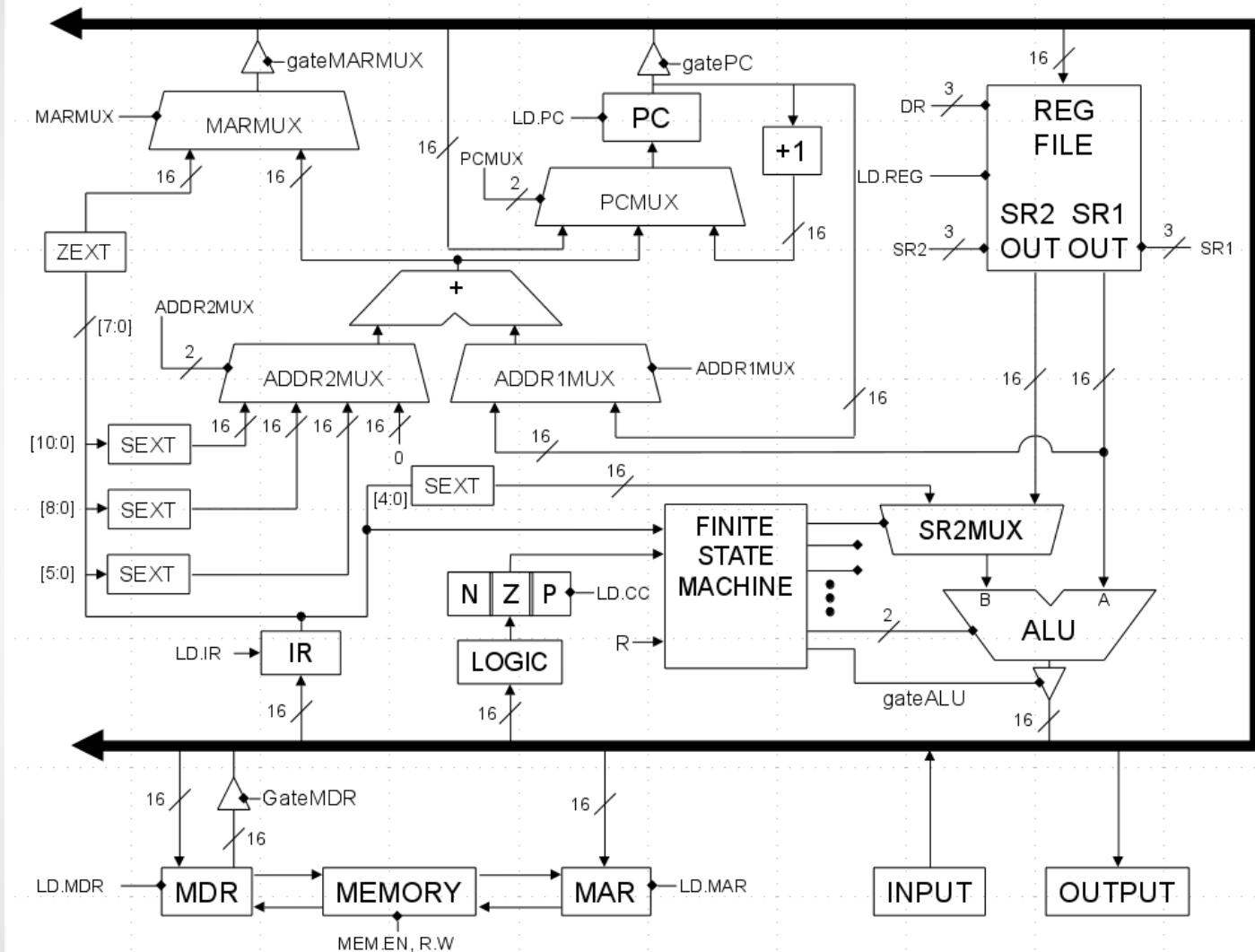
CC3:
MEM.EN
R.W



STR

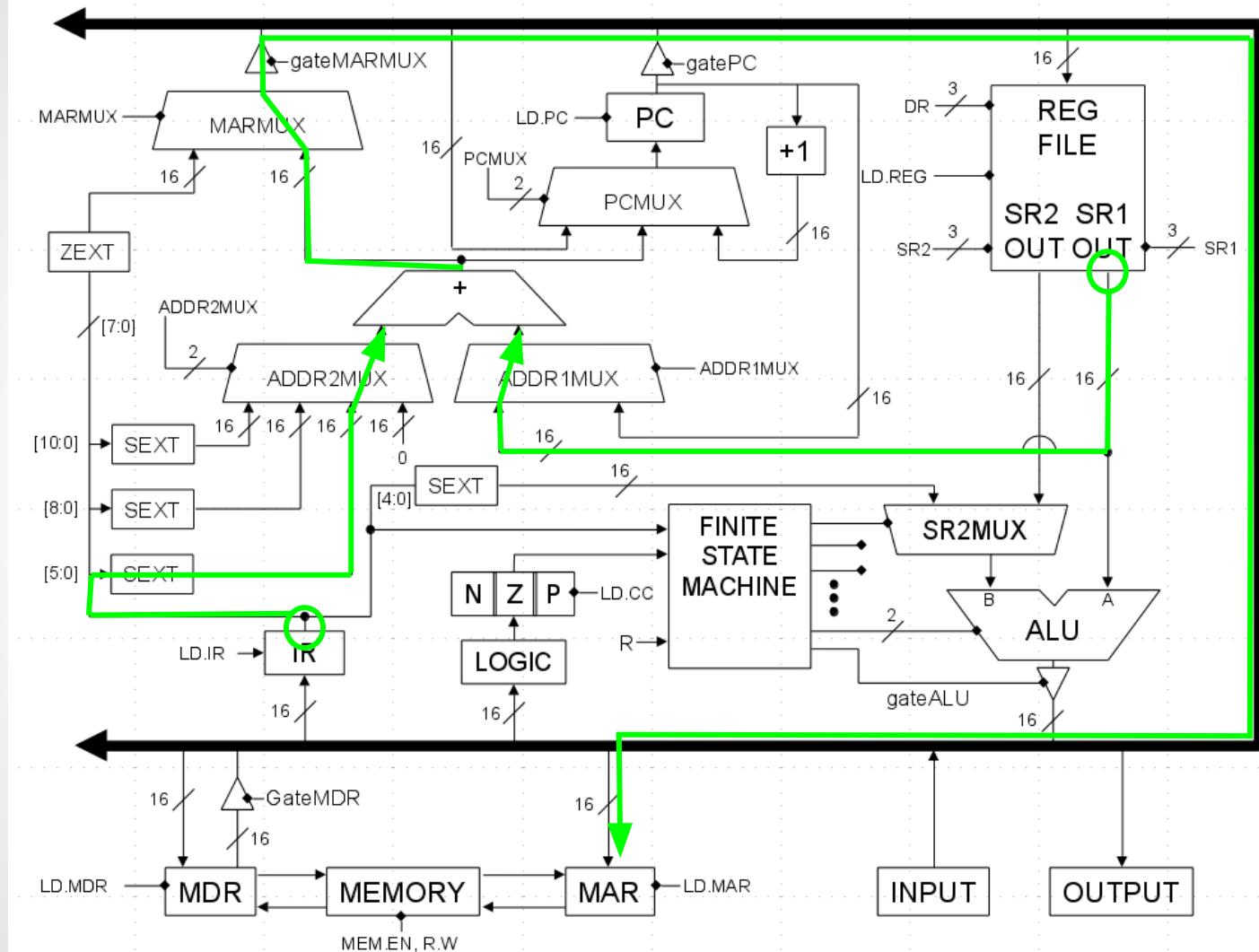
The STR instruction takes a location pointed to by a base register and offset from that location, and stores the value from a source register at that location.


$$\text{mem}[\text{BaseR} + \text{offset6}] \leftarrow \text{SR}$$

STR

STR

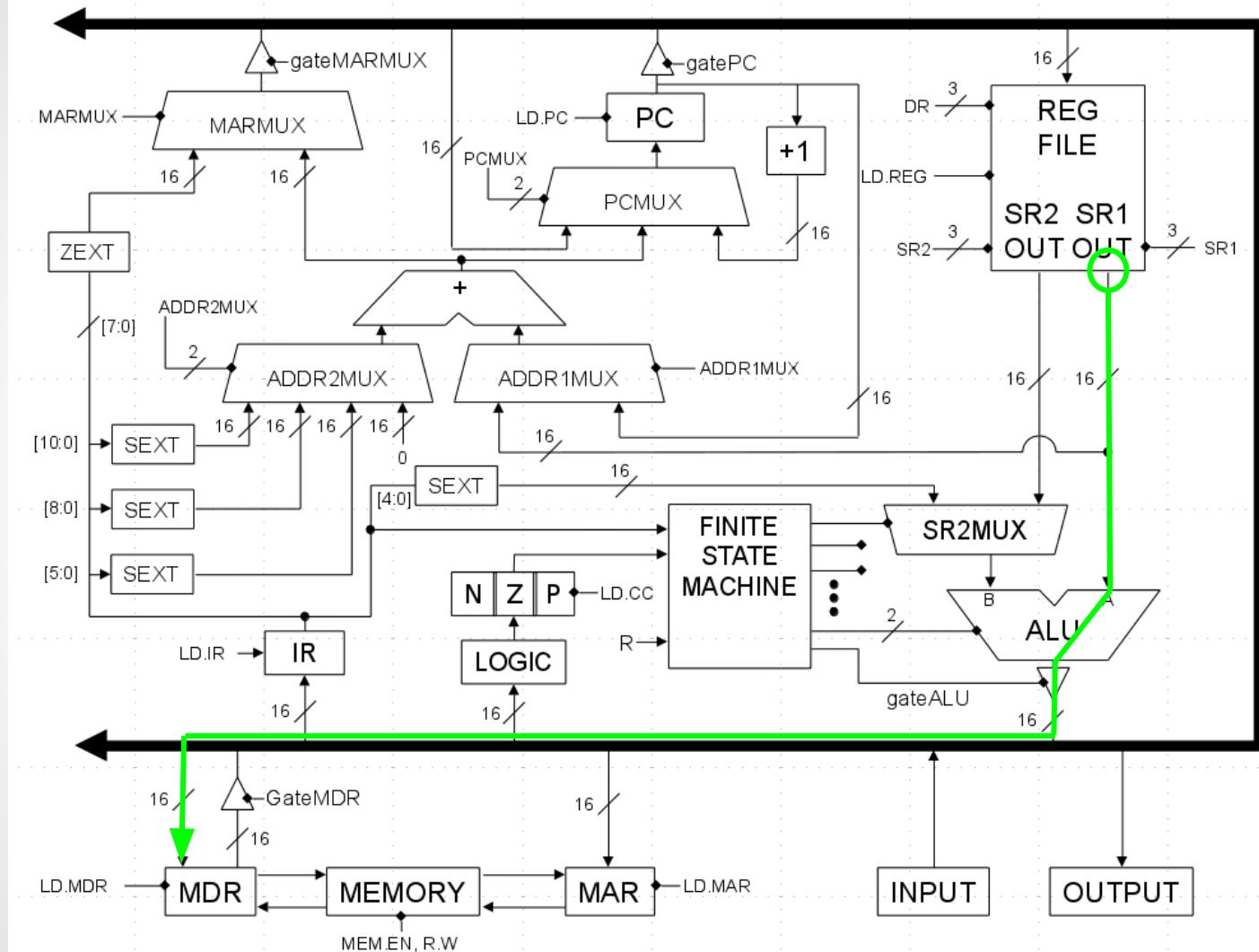
CC1:
ADDR1MUX = BaseR
ADDR2MUX =
SEXT[5:0]
MARMUX = ADDER
gateMARMUX
LD.MAR



STR

CC1:
ADDR1MUX = BaseR
ADDR2MUX =
SEXT[5:0]
MARMUX = ADDER
gateMARMUX
LD.MAR

CC2:
SR1 = IR[11:9]
ALUK = PASSA
gateALU
LD.MDR

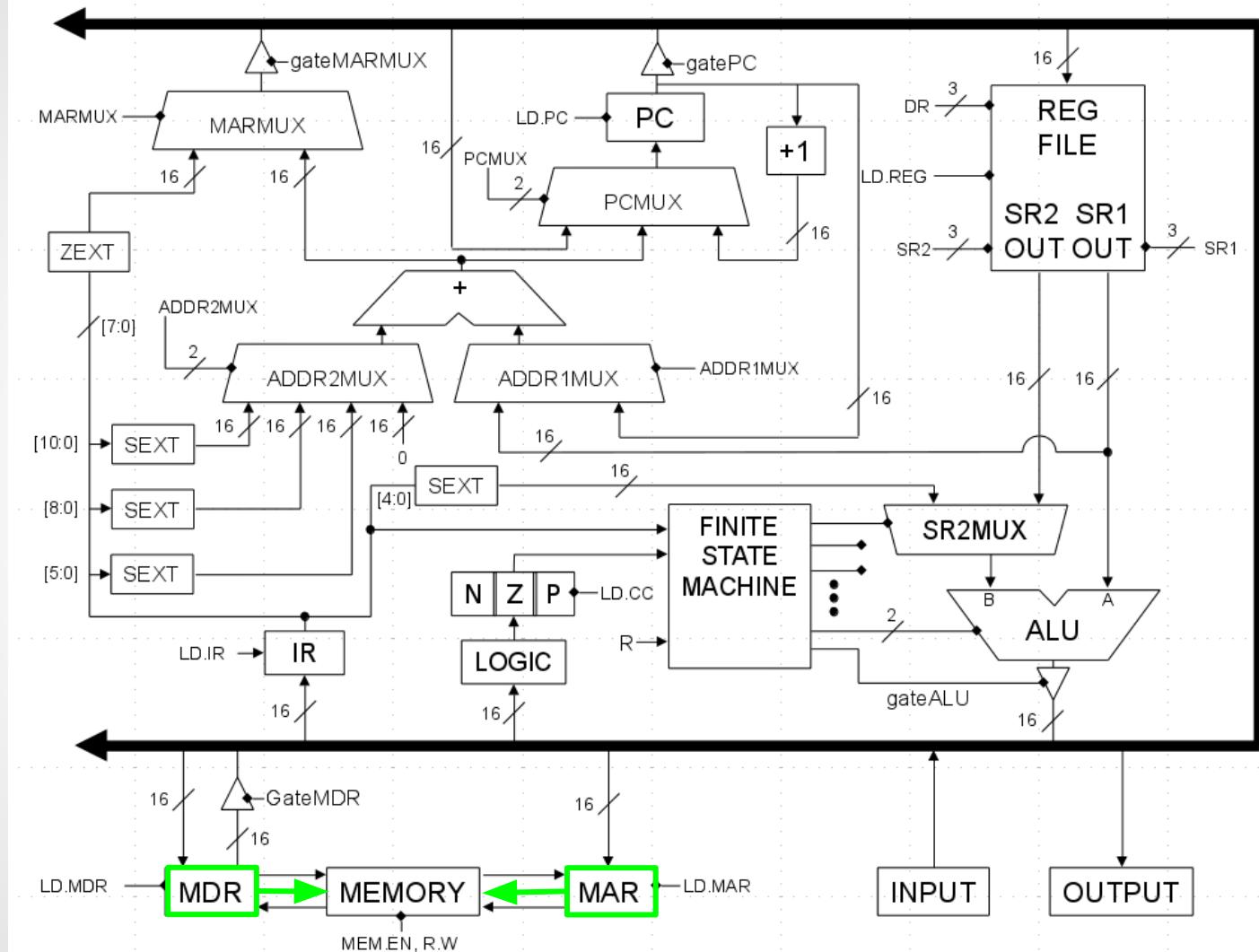


STR

CC1:
ADDR1MUX = BaseR
ADDR2MUX =
SEXT[5:0]
MARMUX = ADDER
gateMARMUX
LD.MAR

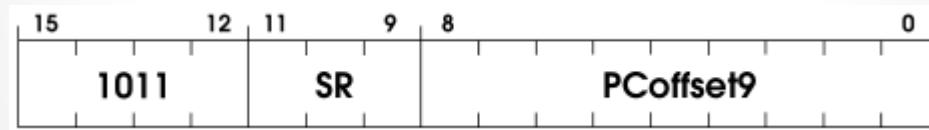
CC2:
SR1 = IR[11:9]
ALUK = PASSA
gateALU
LD.MDR

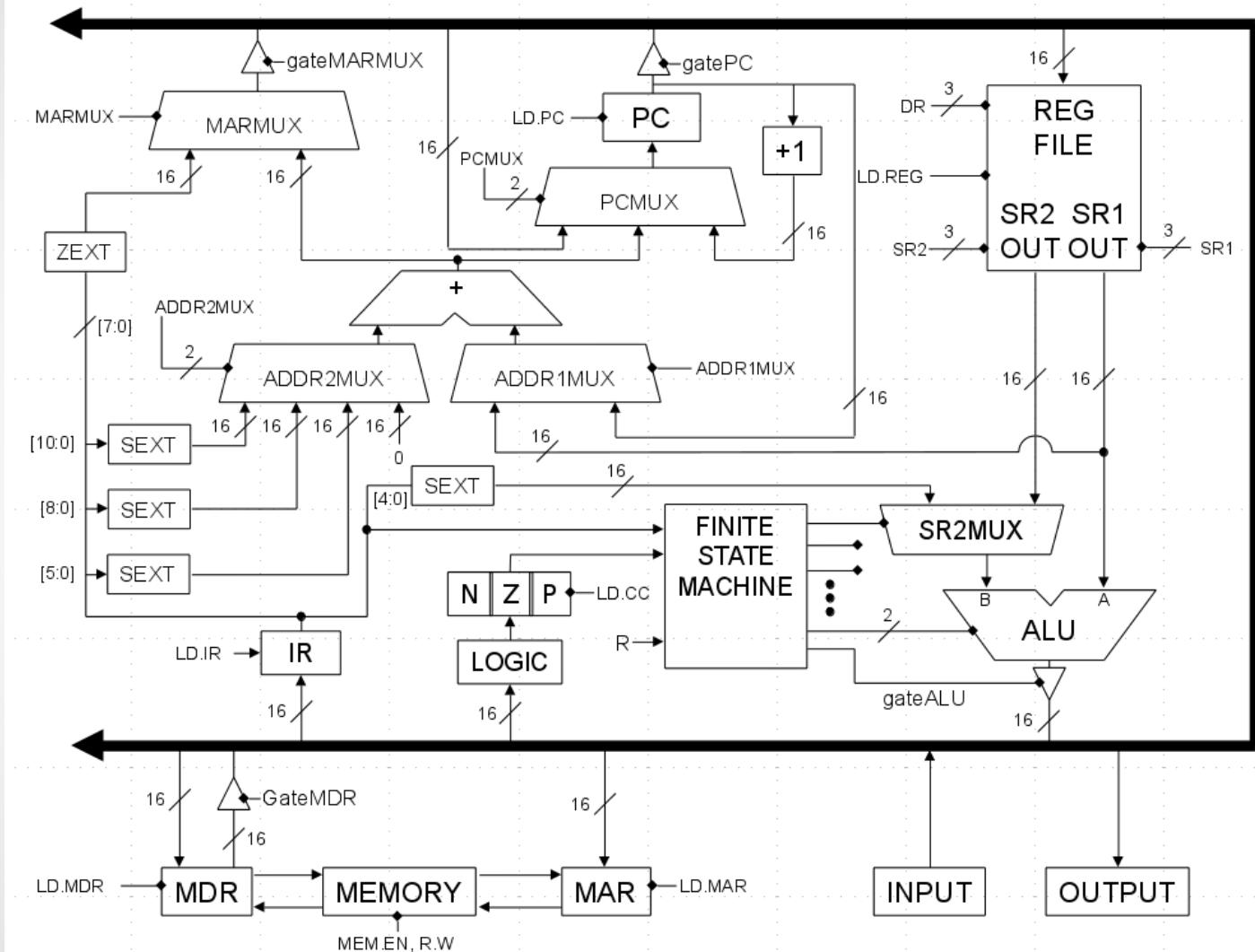
CC3:
MEM.EN
R.W



STI

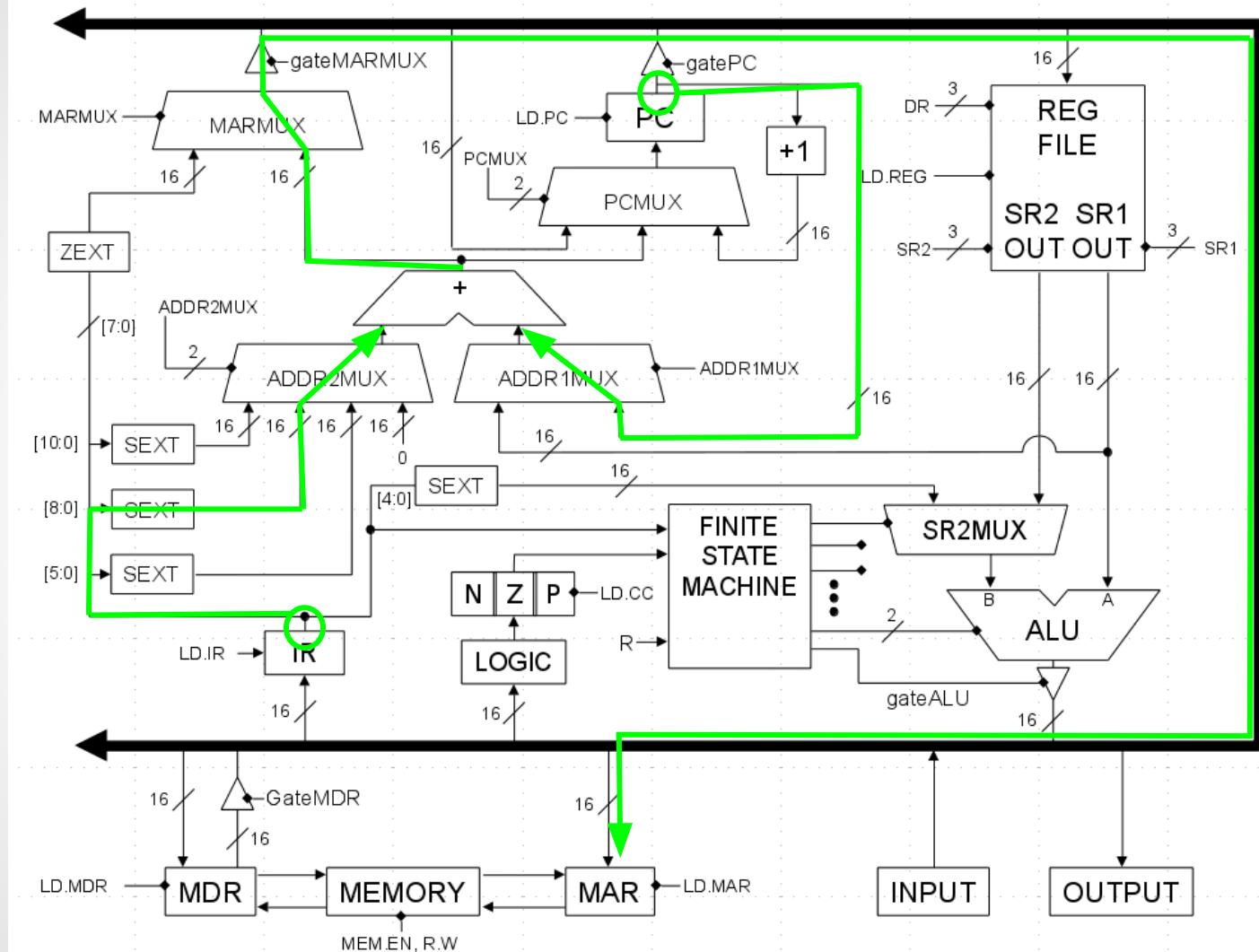
The STI instruction takes a value from memory nearby the instruction itself, gets a pointer to that memory location, then stores the contents of a source register at that location.


$$\text{mem}[\text{mem}[\text{PC} + \text{PCoffset9}]] \leftarrow \text{SR}$$

STI

STI

CC1:
ADDR1MUX = PC
ADDR2MUX =
SEXT[8:0]
MARMUX = ADDER
gateMARMUX
LD.MAR



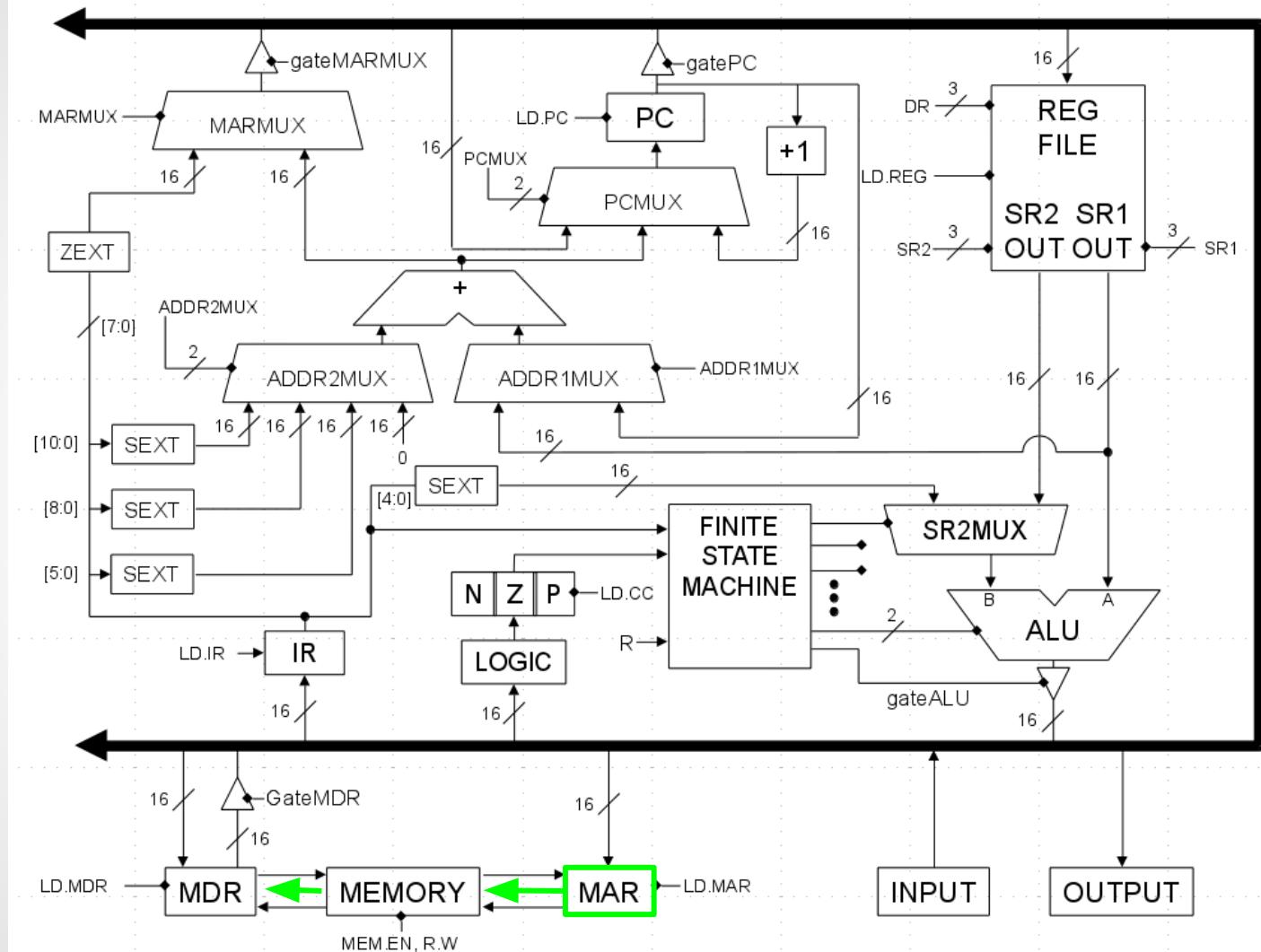
STI

CC1:
ADDR1MUX = PC

ADDR2MUX =
SEXT[8:0]

MARMUX = ADDER
gateMARMUX
LD.MAR

CC2:
MEM.EN
LD.MDR



STI

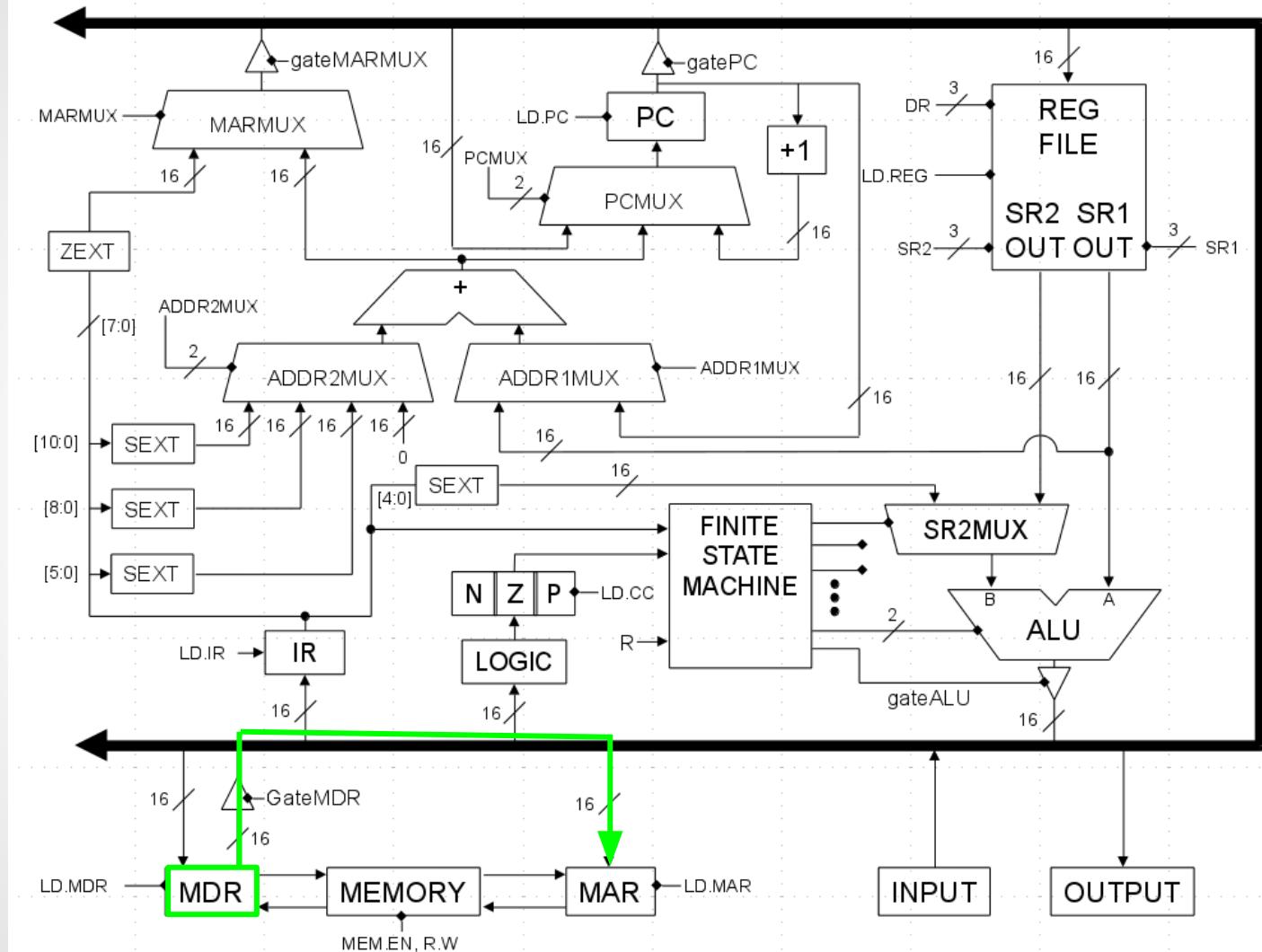
CC1:
ADDR1MUX = PC

ADDR2MUX =
SEXT[8:0]

MARMUX = ADDER
gateMARMUX
LD.MAR

CC2:
MEM.EN
LD.MDR

CC3:
gateMDR
LD.MAR



STI

CC1:

[...]

CC2:

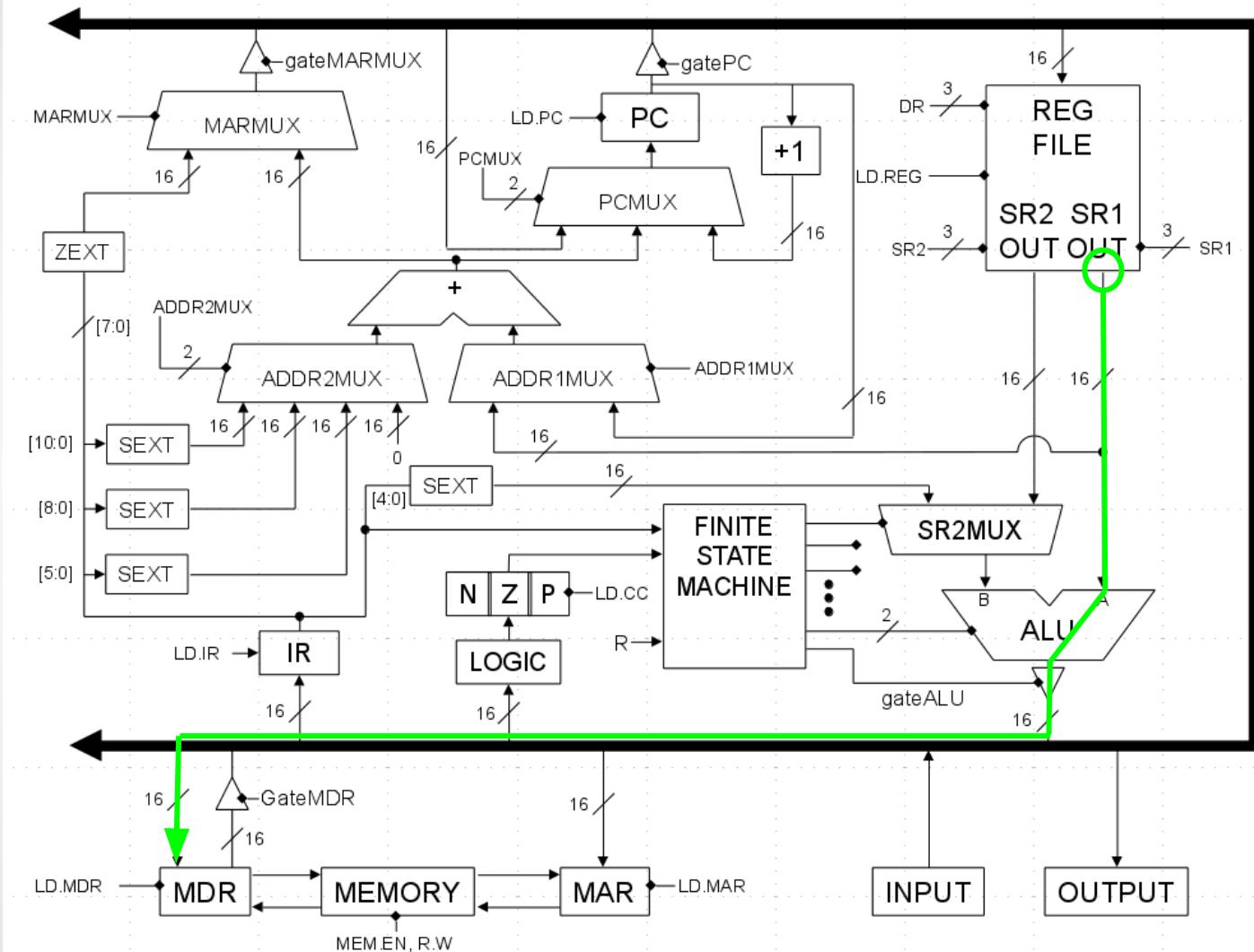
MEM.EN
LD.MDR

CC3:

gateMDR
LD.MAR

CC4:

SR1 = IR[11:9]
ALUK = PASSA
gateALU
LD.MDR



STI

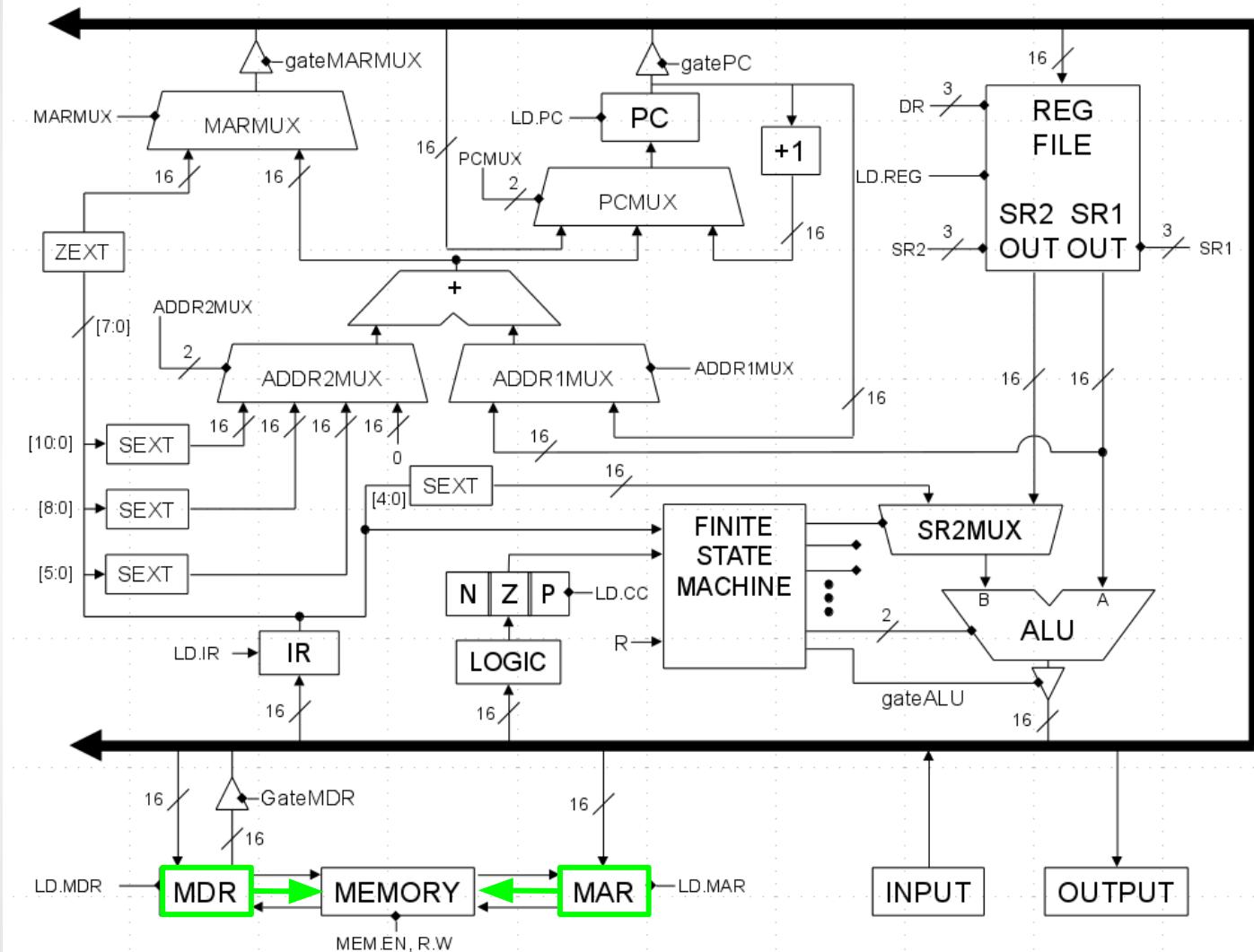
CC1:
[...]

CC2:
[...]

CC3:
gateMDR
LD.MAR

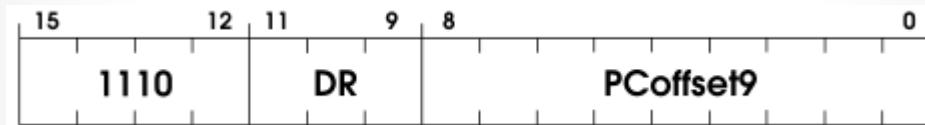
CC4:
SR1 = IR[11:9]
ALUK = PASSA
gateALU
LD.MDR

CC5:
MEM.EN
R.W



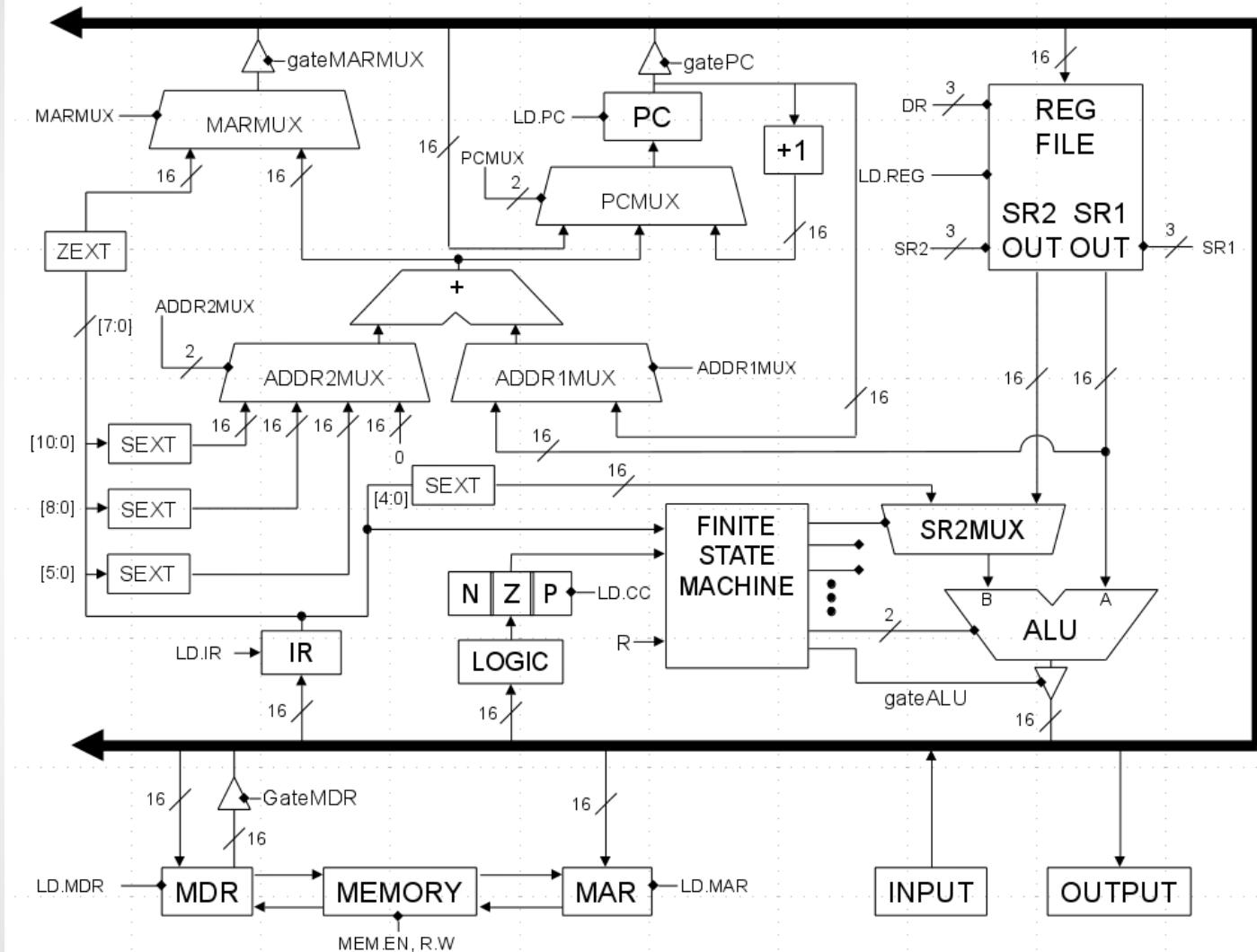
LEA

The LEA instruction adds the PC and an offset, and stores the result in a register. This is useful in determining the address of a label at run-time.



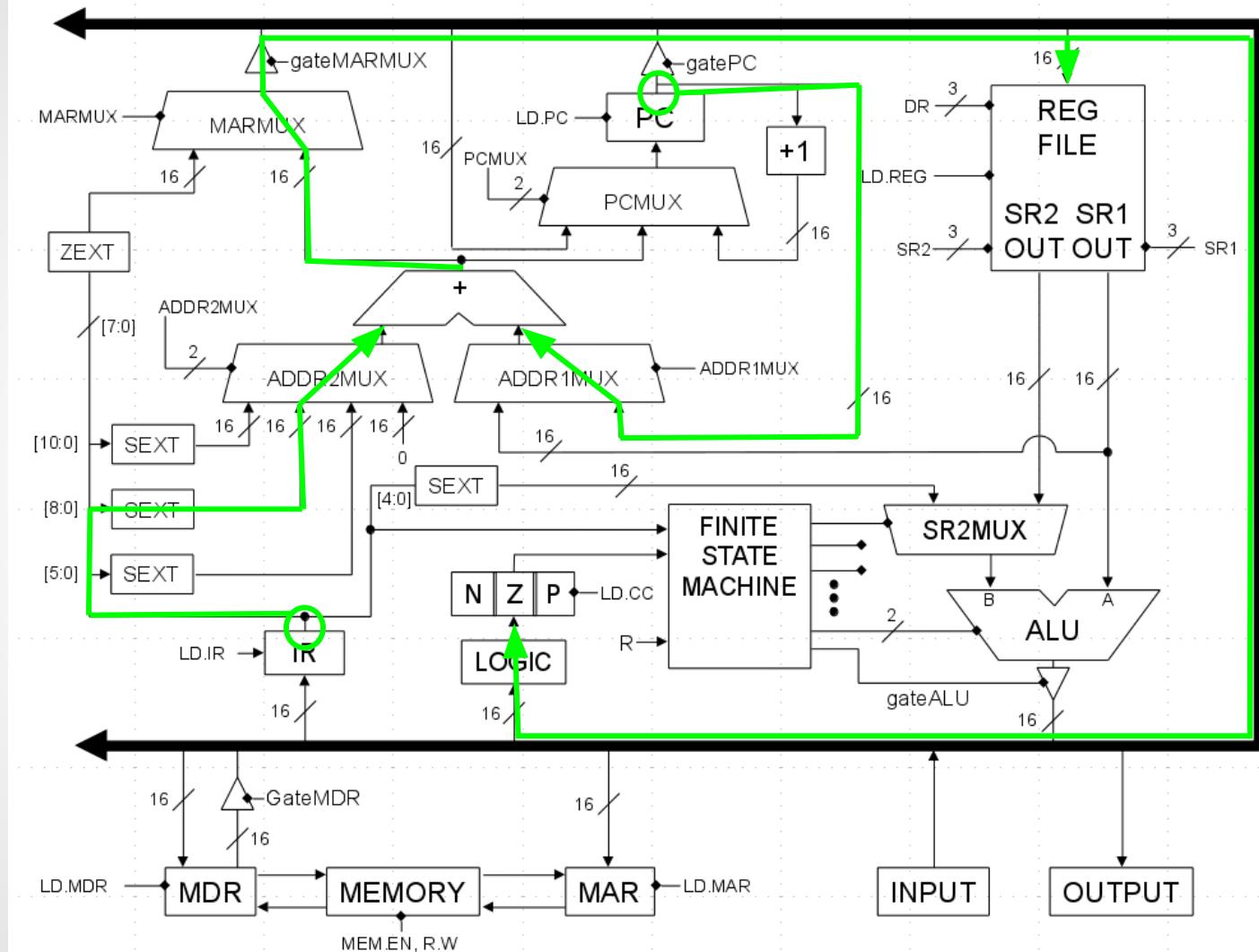
$$DR \leftarrow PC + PCoffset9$$

LEA



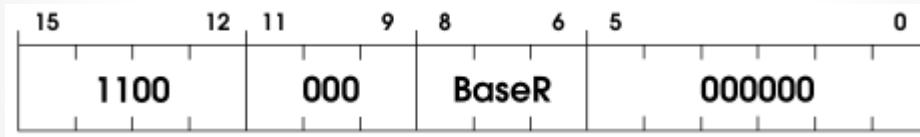
LEA

CC1:
ADDR1MUX = PC
ADDR2MUX =
SEXT[8:0]
MARMUX =
ADDER
gateMARMUX
DR = IR[11:9]
LD.REG
LD.CC

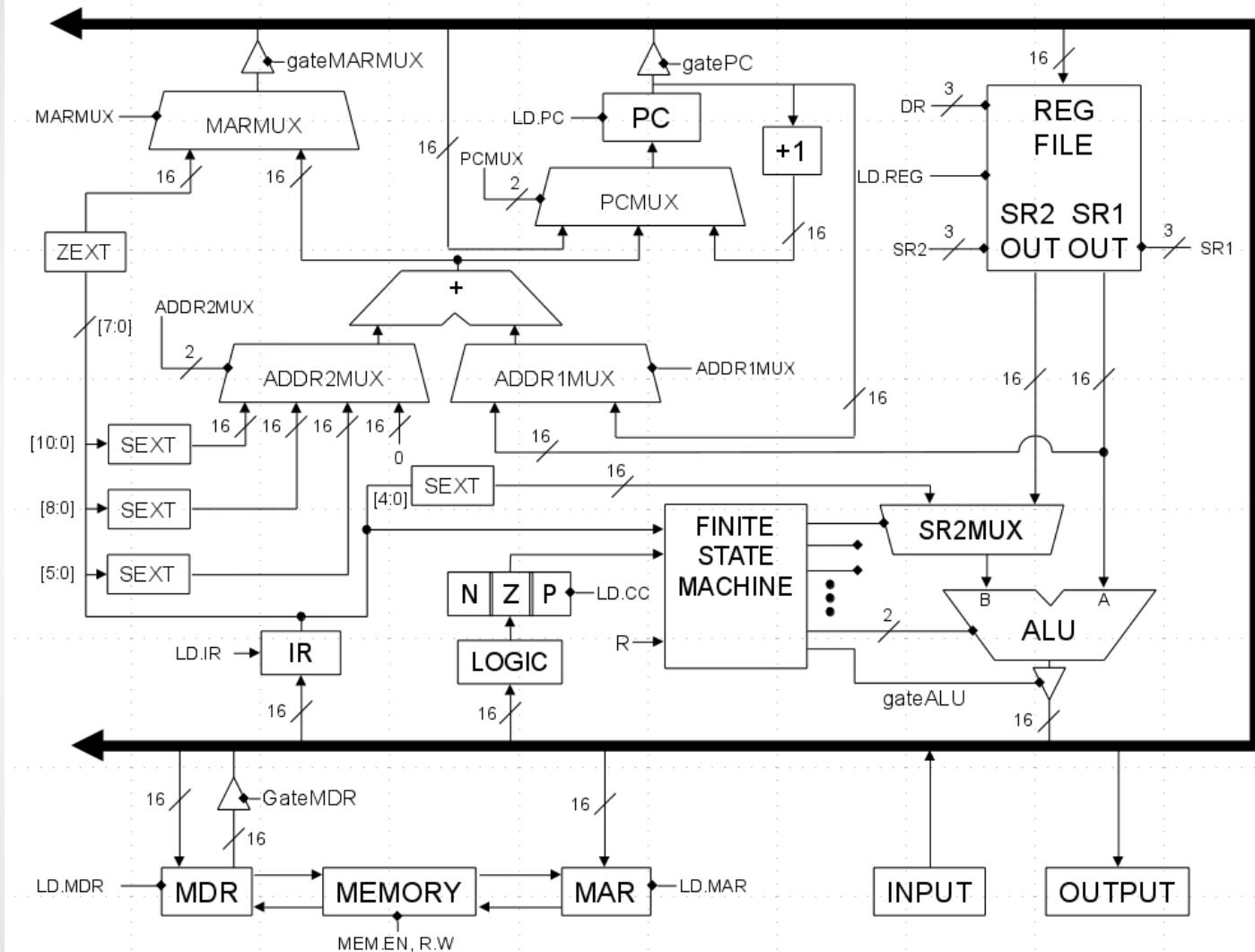


JMP/RET

The JMP instruction sets the PC to the value of a register, effectively “jumping” to the location it points to. RET is a pseudo-instruction that simply performs JMP R7.

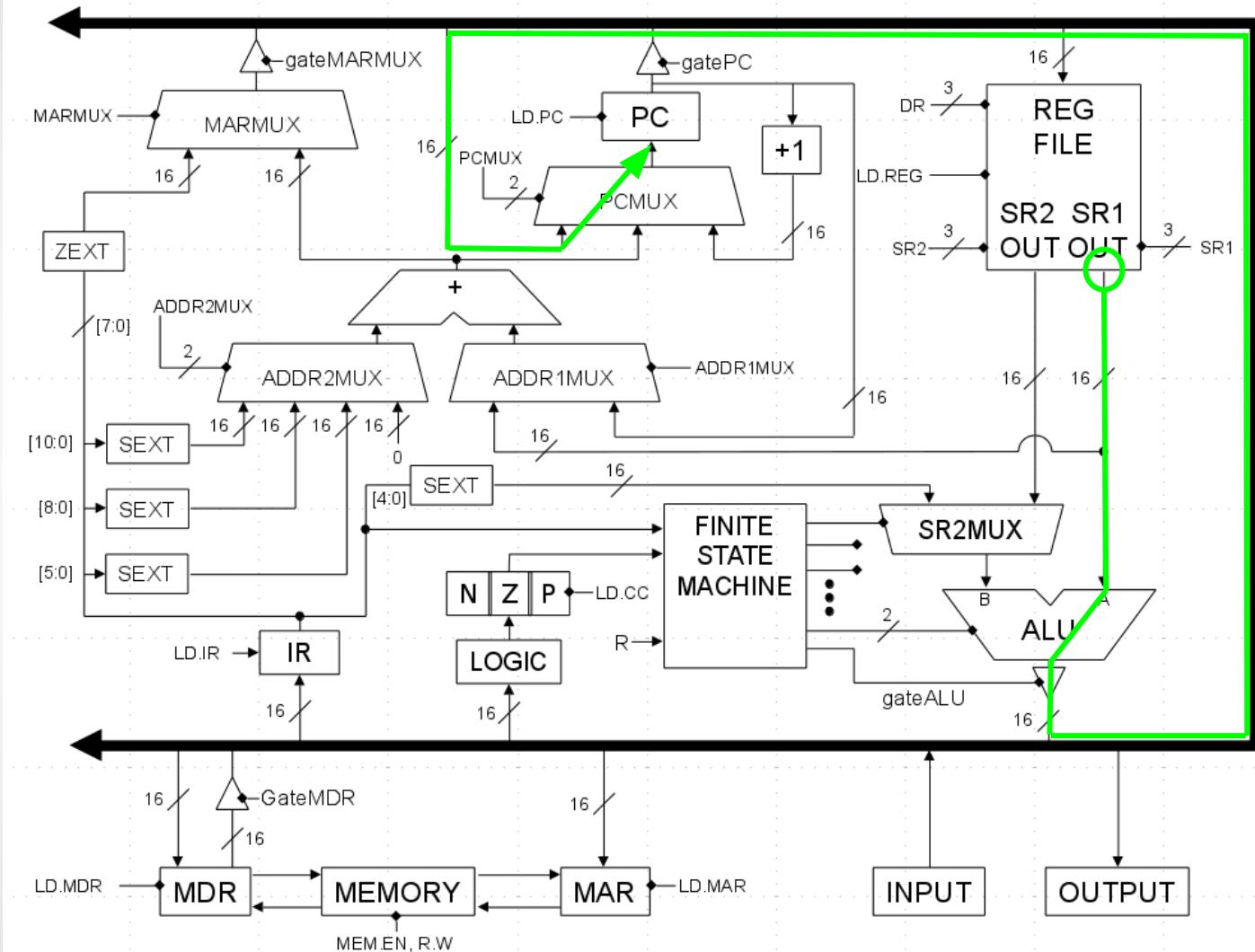


$\text{PC} \leftarrow \text{BaseR}$

JMP

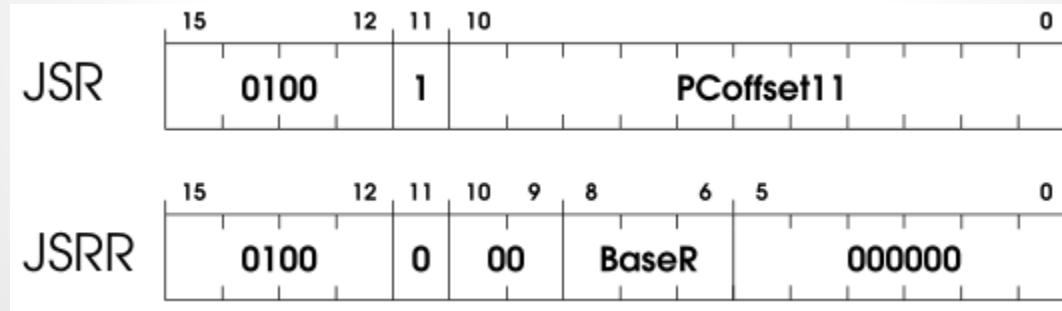
JMP

CC1:
SR1 = IR[8:6]
ALUK = PASSA
gateALU
PCMUX = BUS
LD.PC



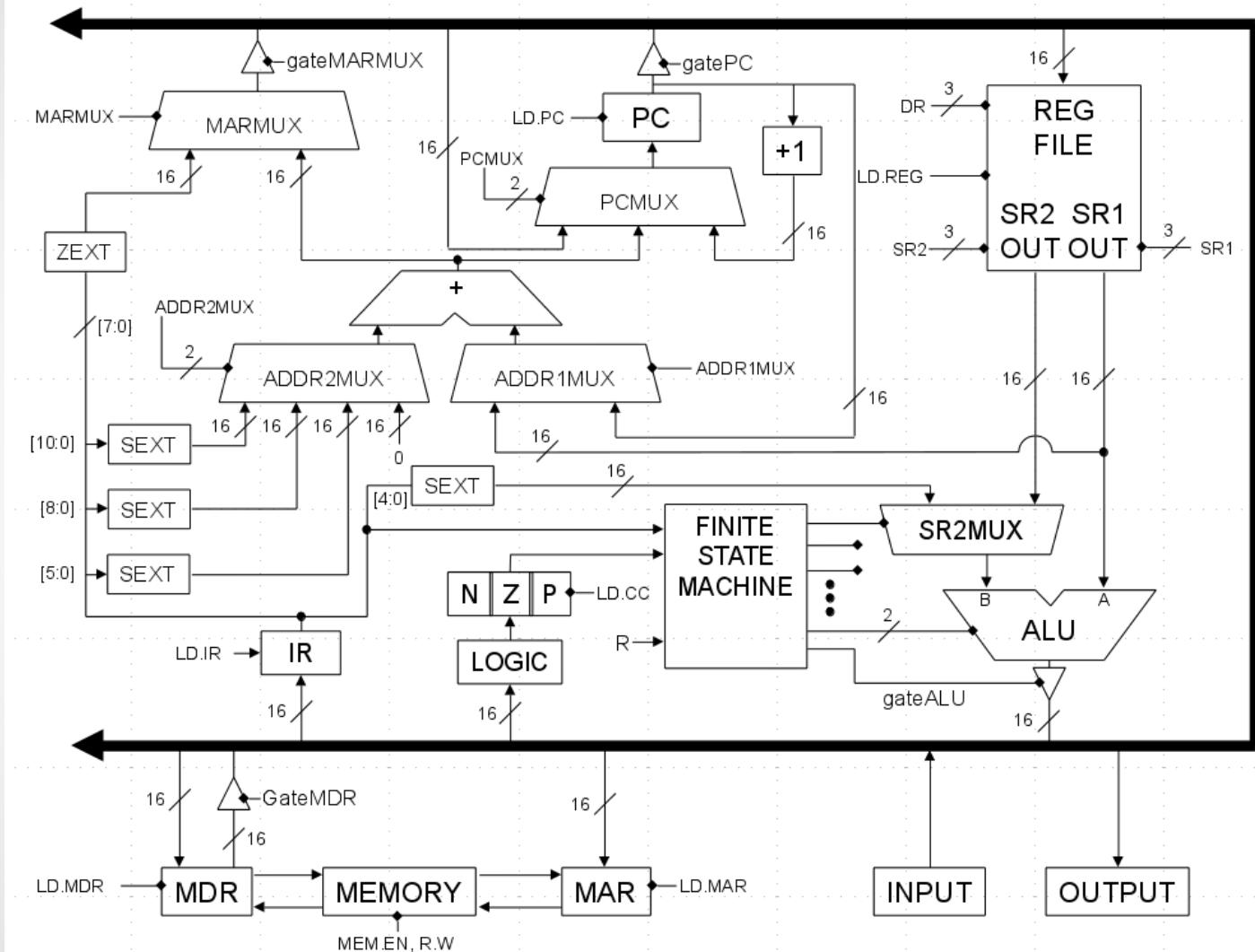
JSR/JSRR

The JSR instruction stores the current PC in R7, and sets the PC to itself + an offset; it is used to jump to a subroutine. JSRR does the same thing, but instead uses a base register instead of an offset.



$$R7 = PC; PC \leftarrow (IR[11] == 1 ? PCoffset11 : BaseR)$$

JSR



JSR

CC1:
gatePC

DR = 7

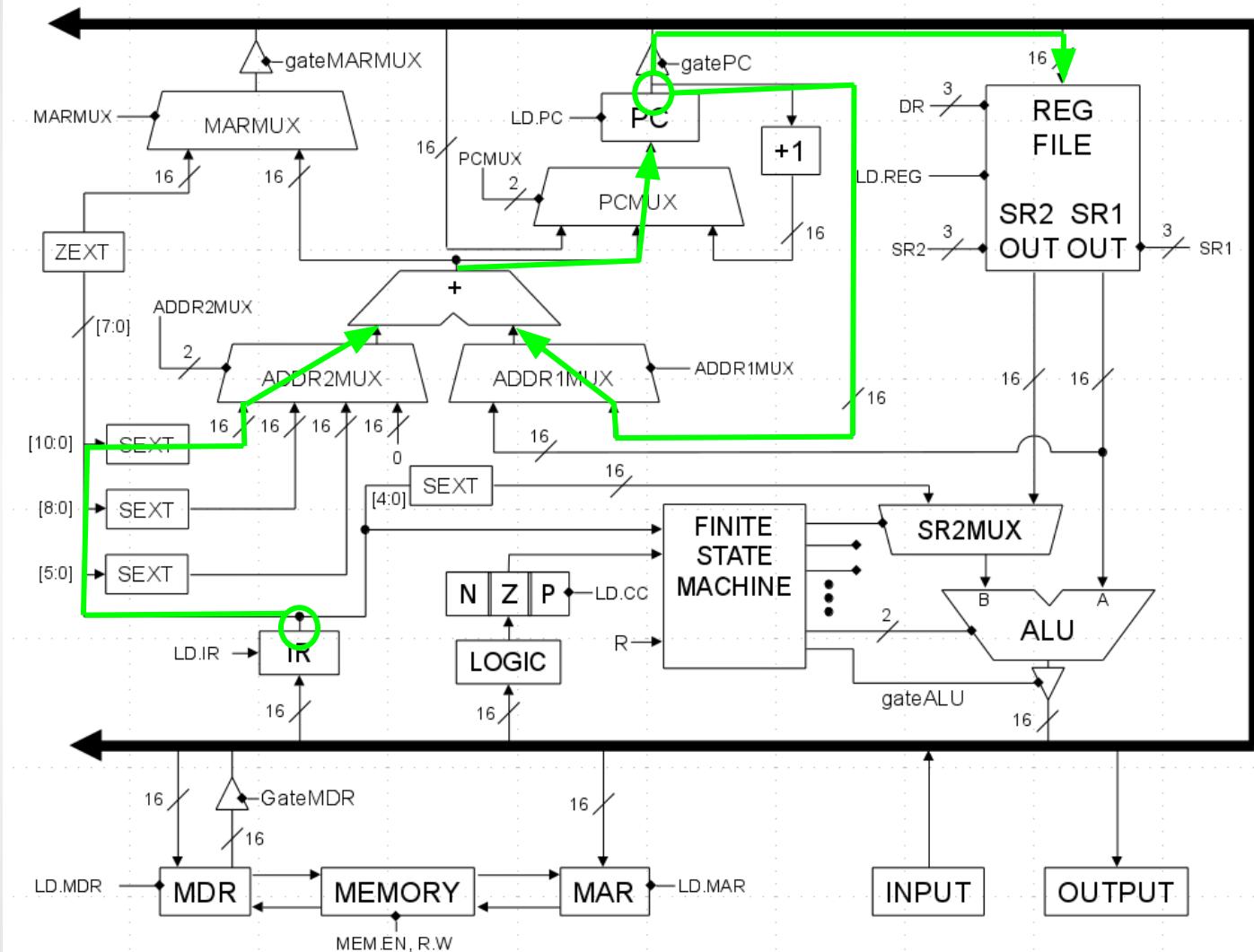
LD.REG

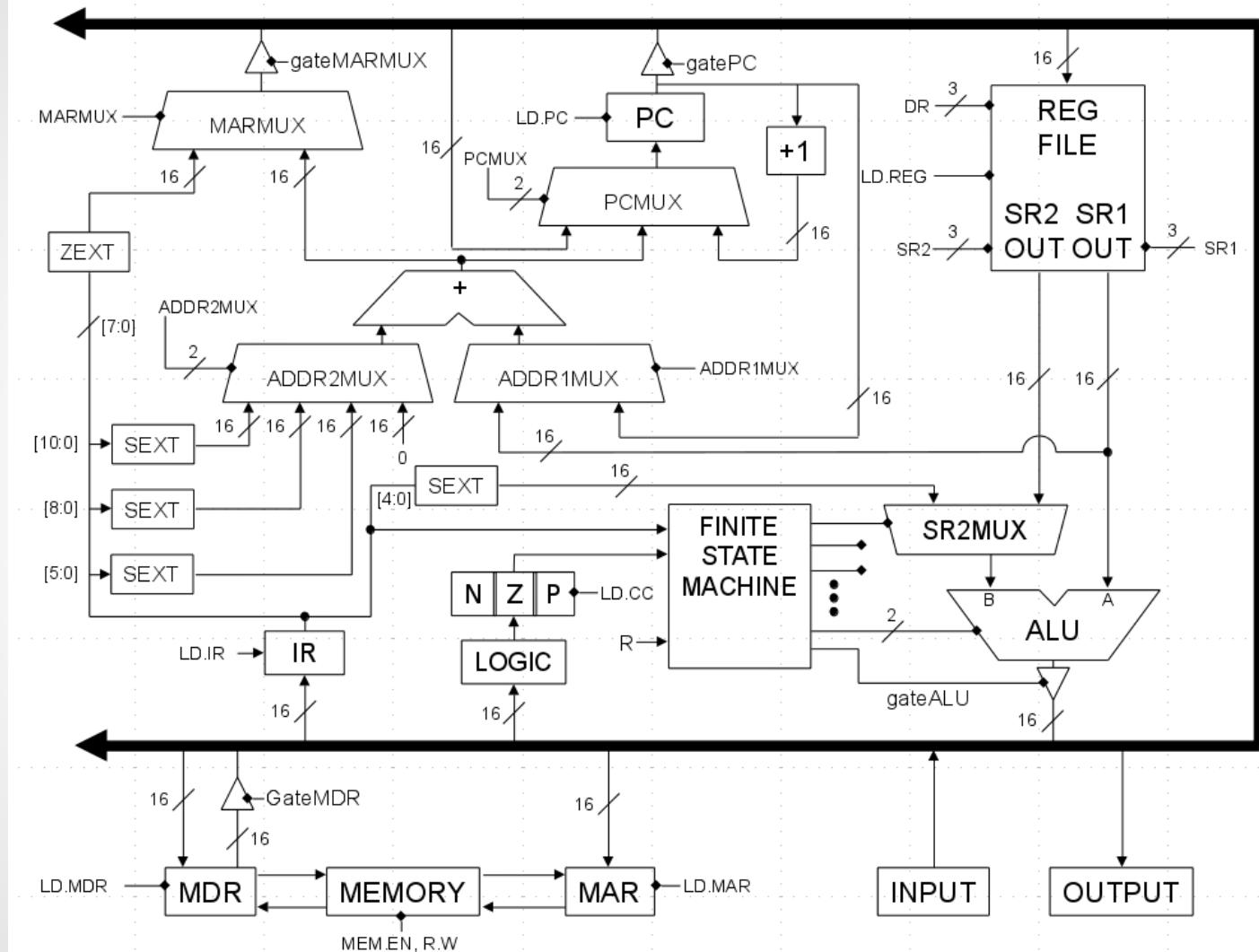
ADDR1MUX = PC

ADDR2MUX =
SEXT[10:0]

PCMUX = ADDER

LD.PC





JSRR

CC1:
gatePC

DR = 7

LD.REG

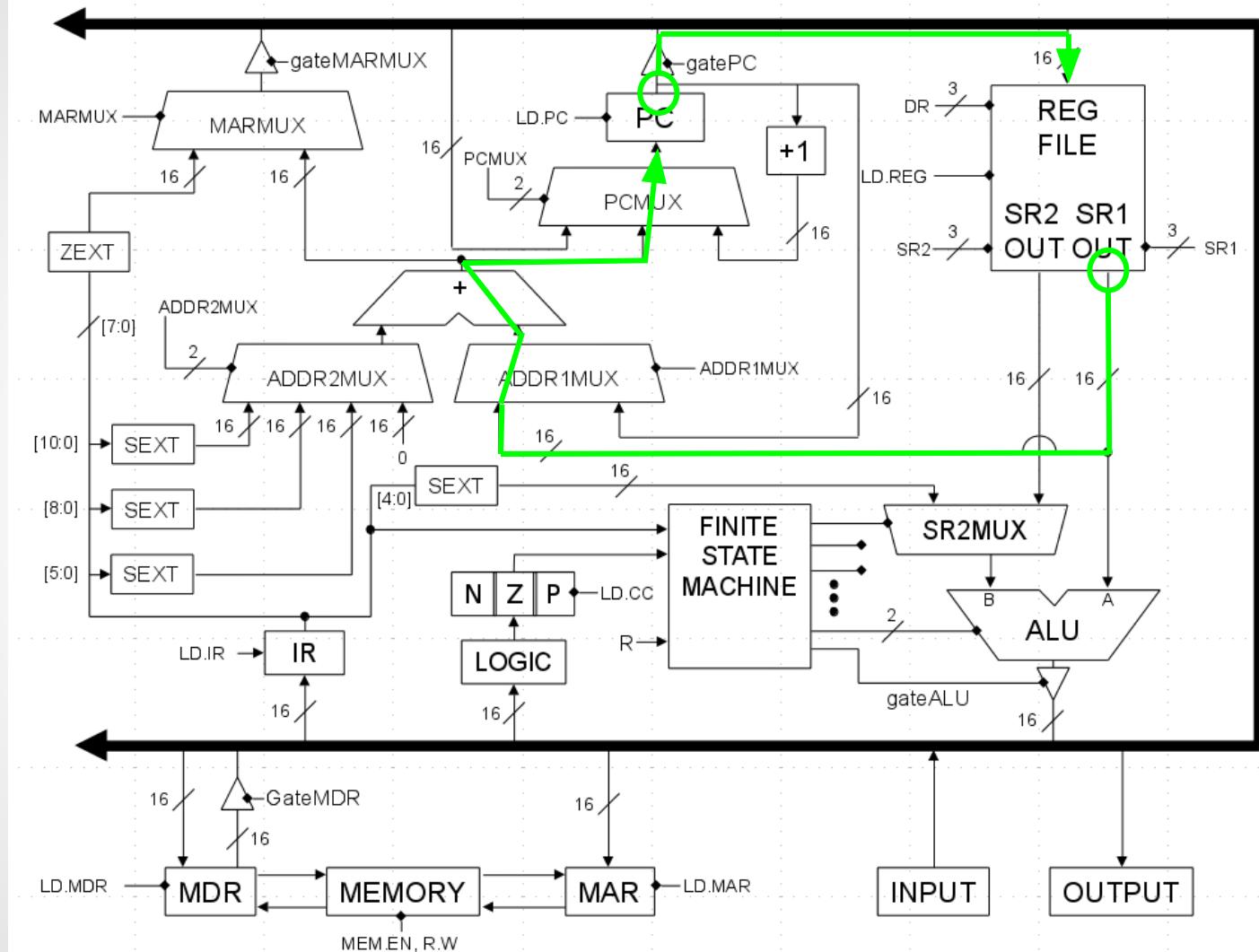
SR1 = IR[8:6]

ADDR1MUX = BaseR

ADDR2MUX = 0

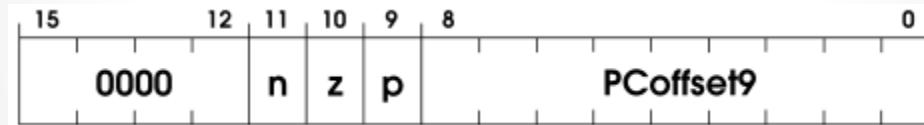
PCMUX = ADDER

LD.PC



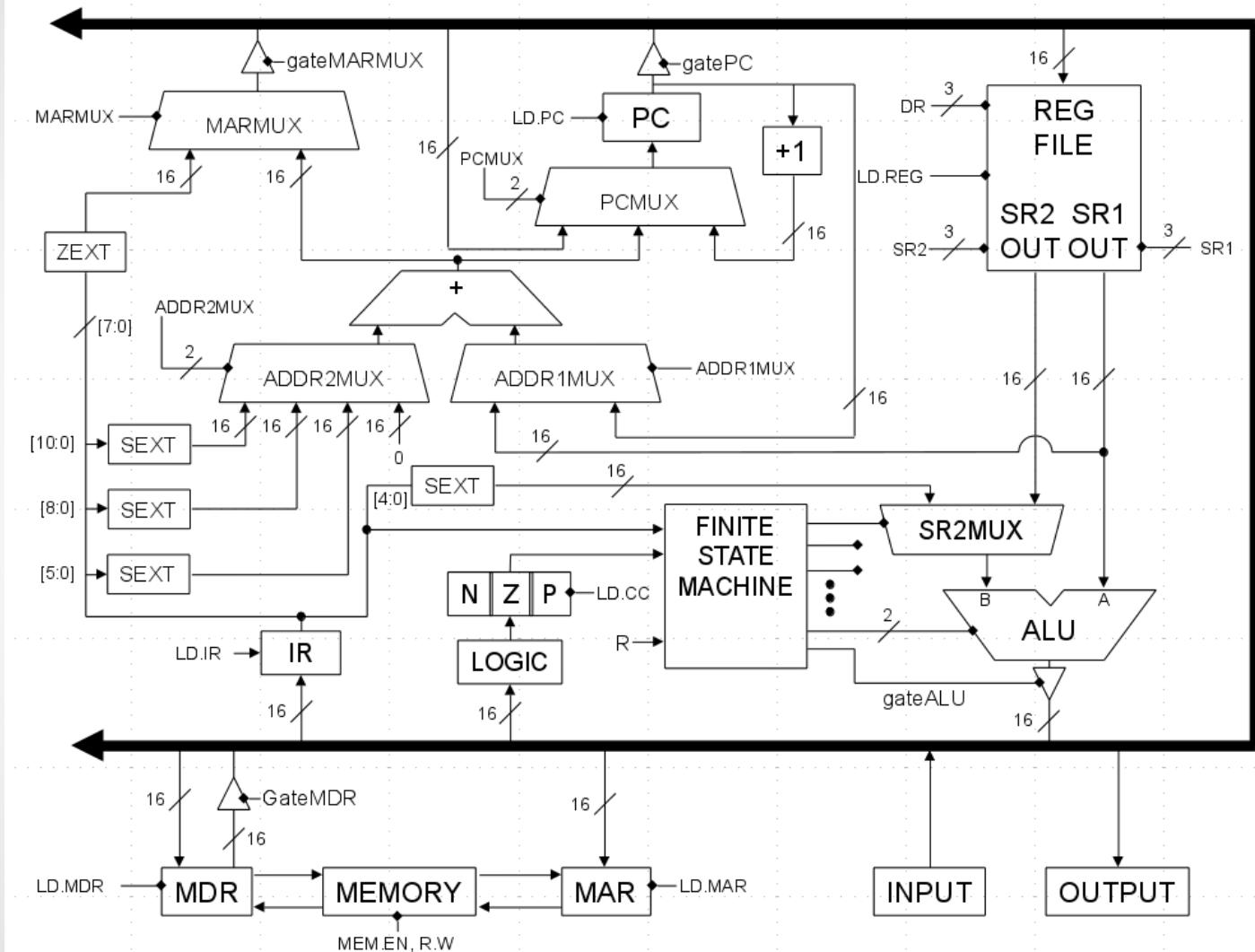
BR

The BR instruction contains 3 condition code bits; if any of the corresponding bits are set in the NZP register, add the offset to the current PC. Otherwise, do nothing.



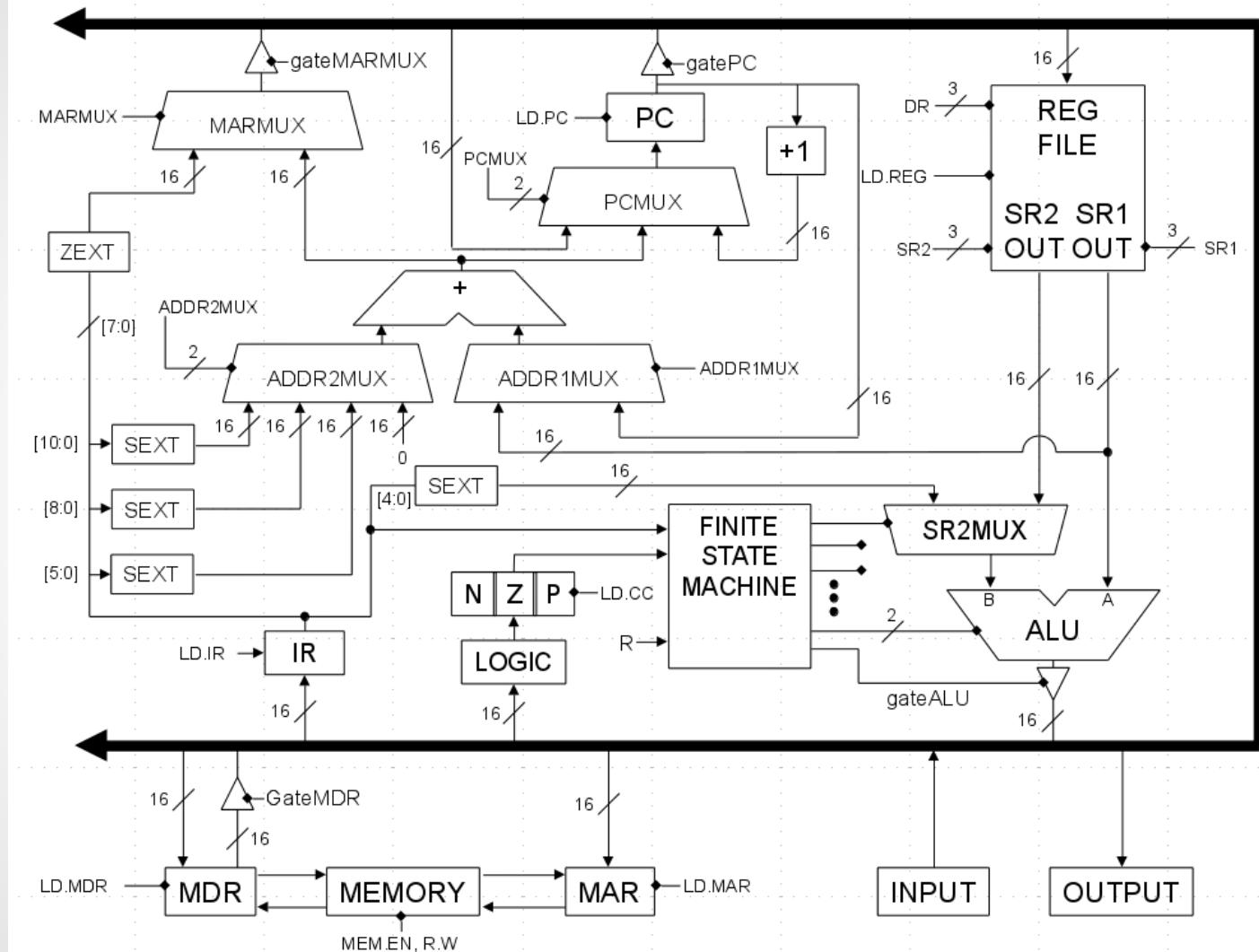
```
if(NZP & IR[11:9] != 0) {  
    PC = PC + PCoffset9  
}
```

BR (not taken)



BR (not taken)

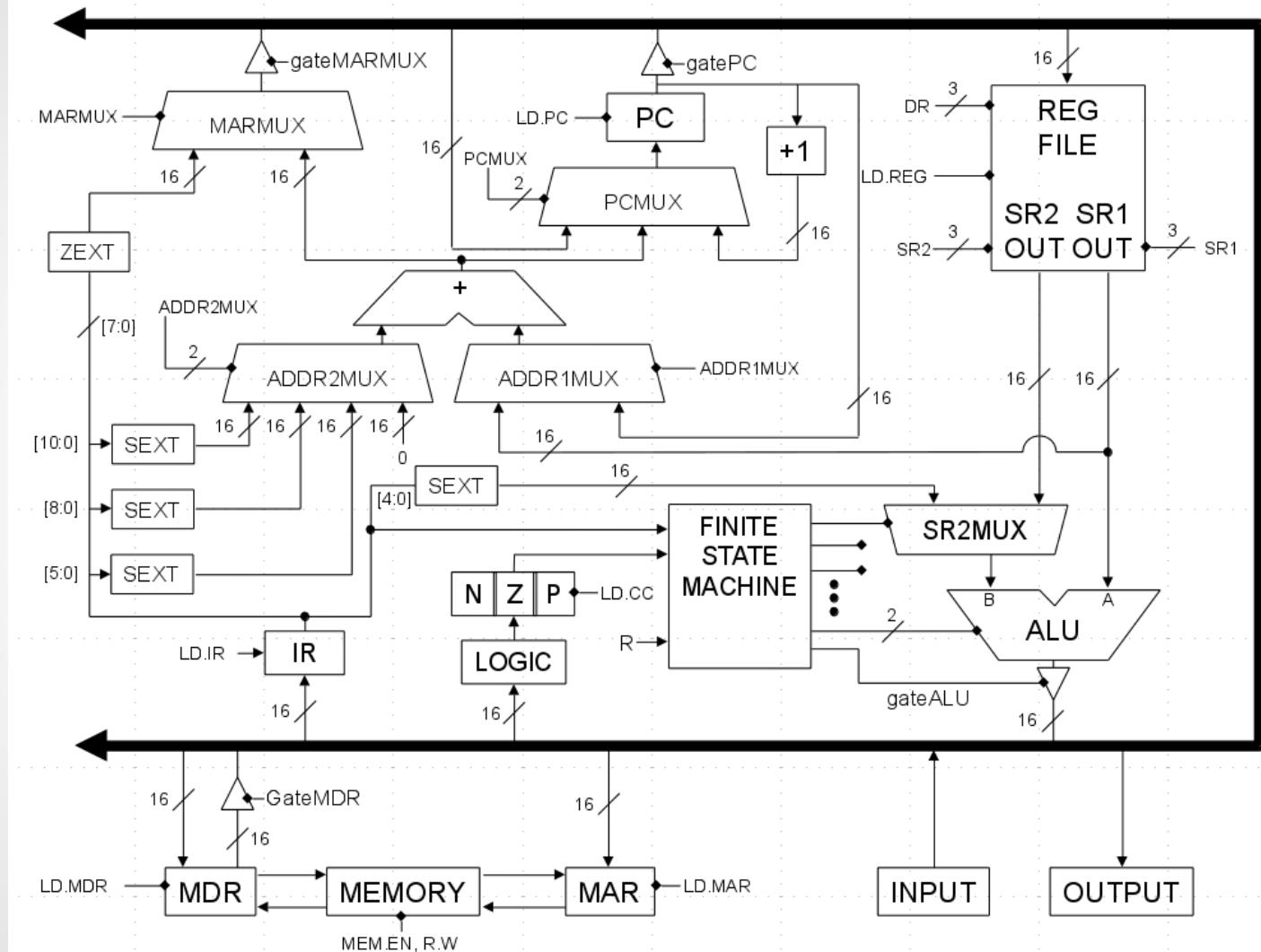
CC1:
Check NZP
(internally in FSM)



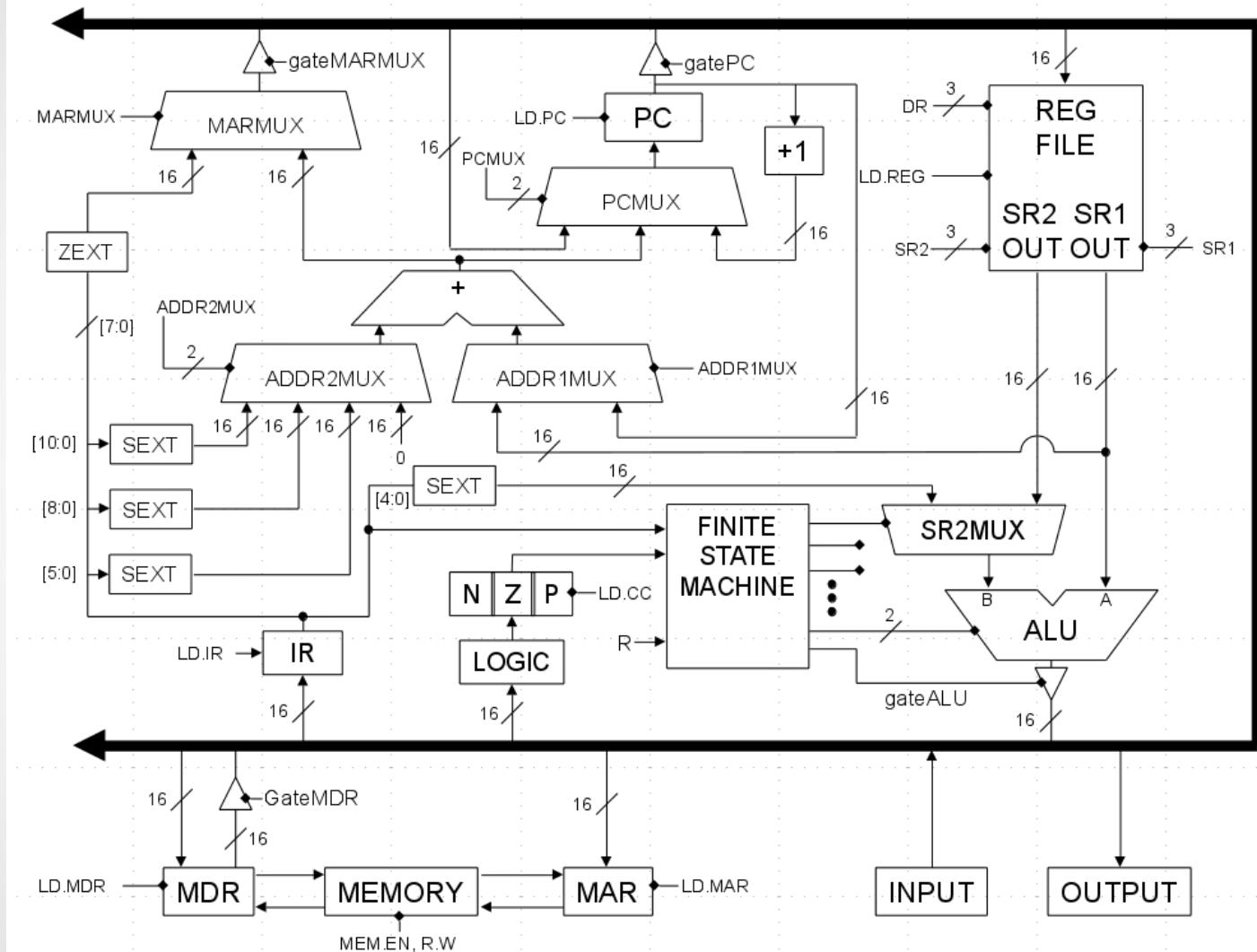
BR (not taken)

CC1:
Check NZP
(internally in FSM)

No branch; return to
fetch.

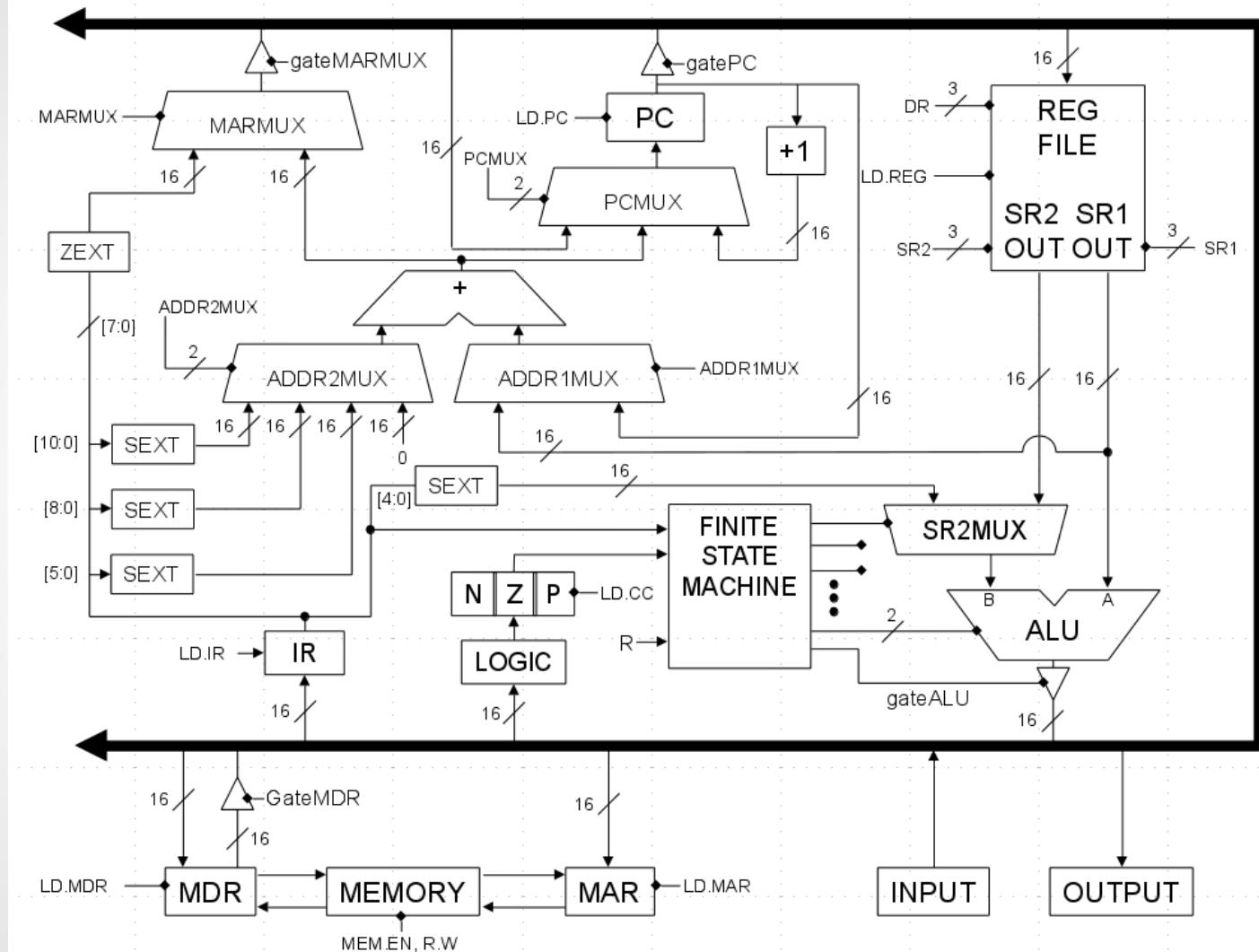


BR (taken)



BR (taken)

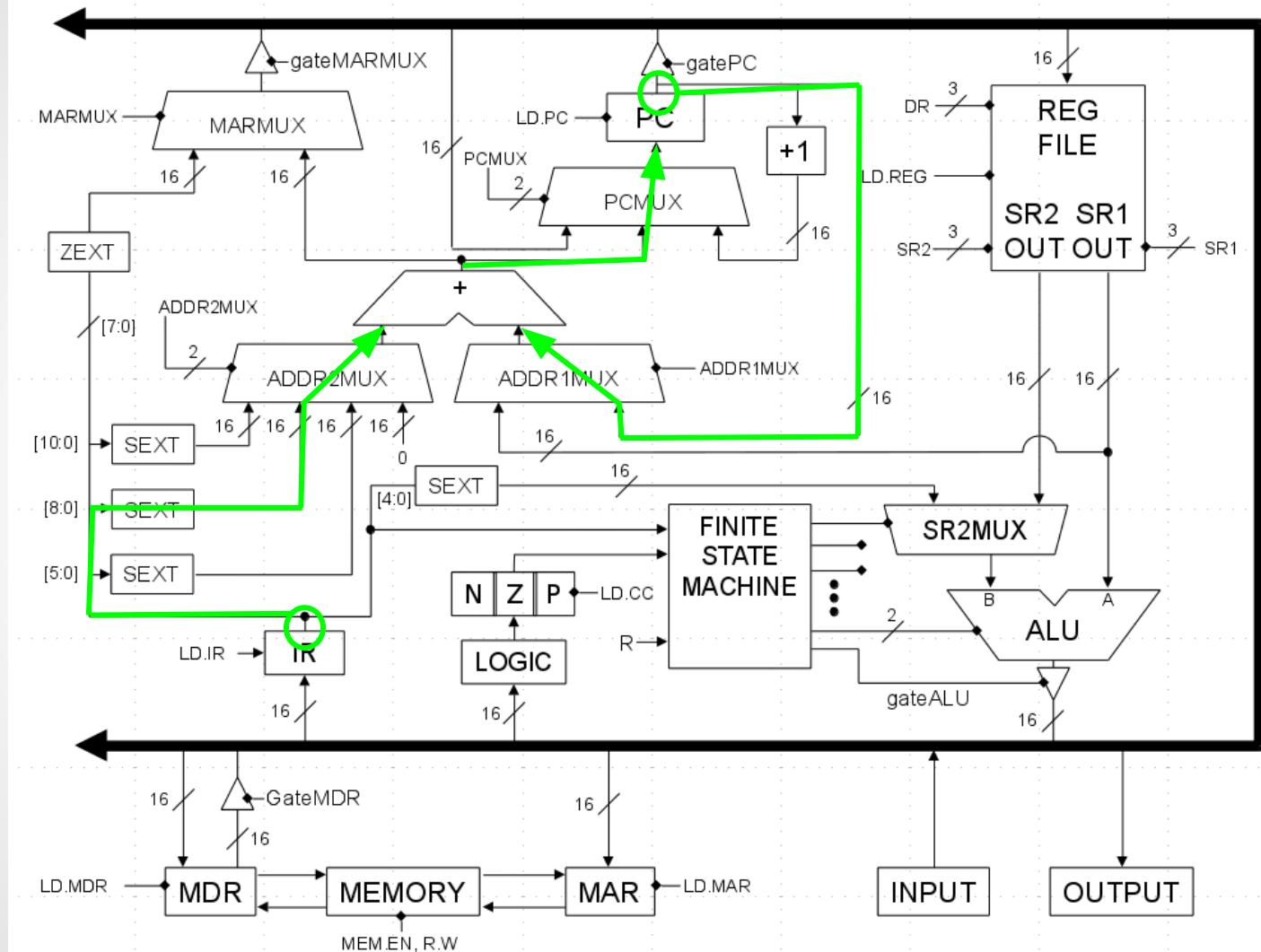
CC1:
Check NZP
(internally in FSM)



BR (taken)

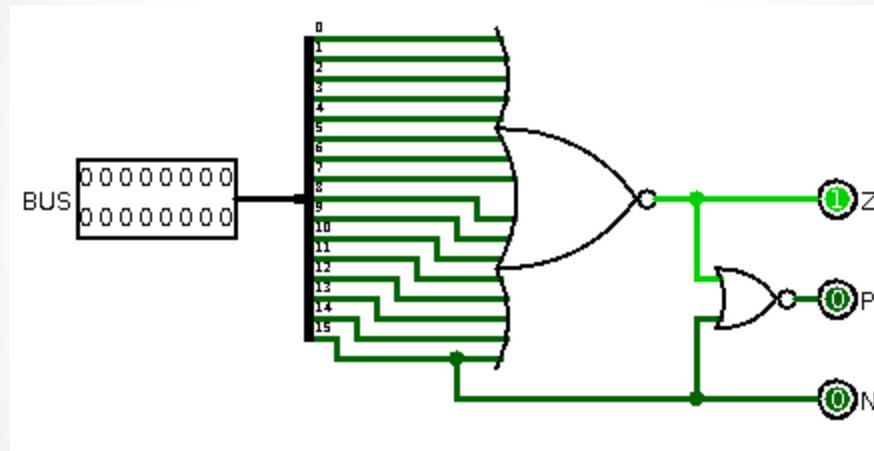
CC1:
Check NZP
(internally in FSM)

CC2:
 $\text{ADDR1MUX} = \text{PC}$
 $\text{ADDR2MUX} =$
SEXT[8:0]
 $\text{PCMUX} = \text{ADDER}$
 LD.PC



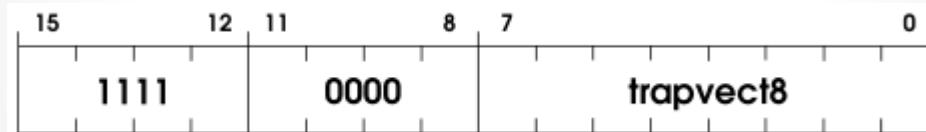
BR

“Logic”:



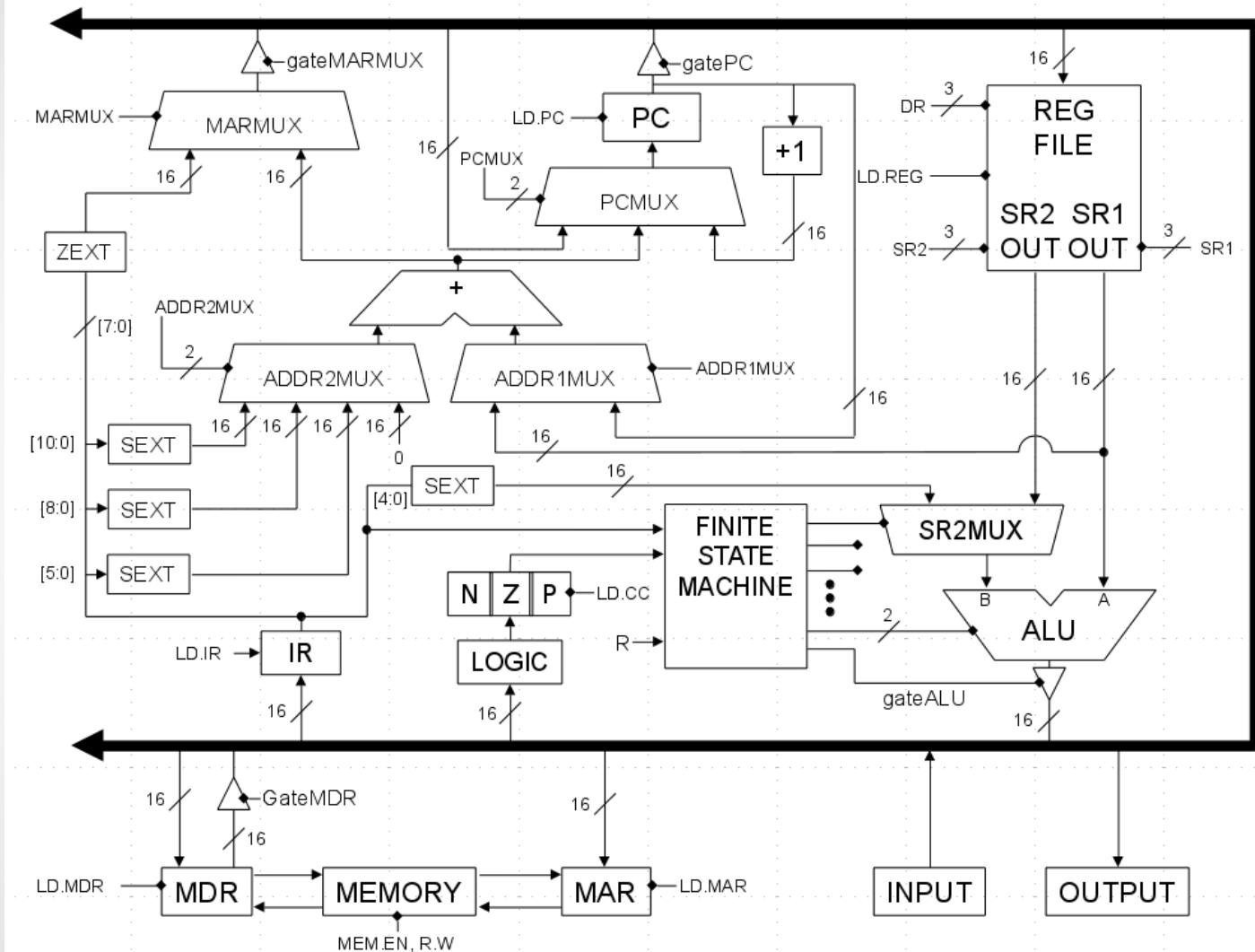
TRAP

The TRAP instruction stores the current PC in R7, locates an 8-bit TRAP vector at the beginning of memory, and sets the PC to that vector's value.



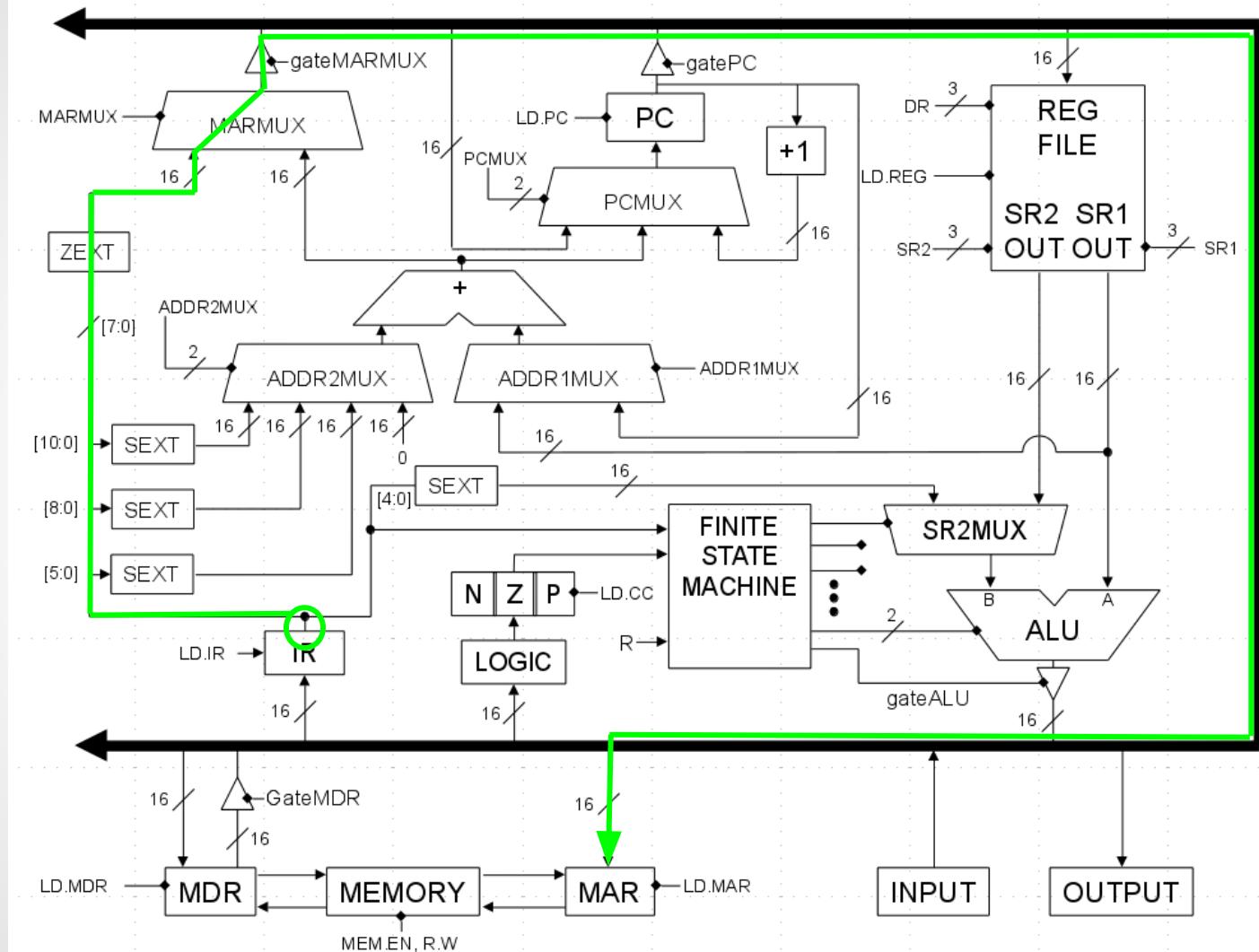
$R7 = PC; PC \leftarrow \text{mem}[\text{trapvect8}]$

TRAP



TRAP

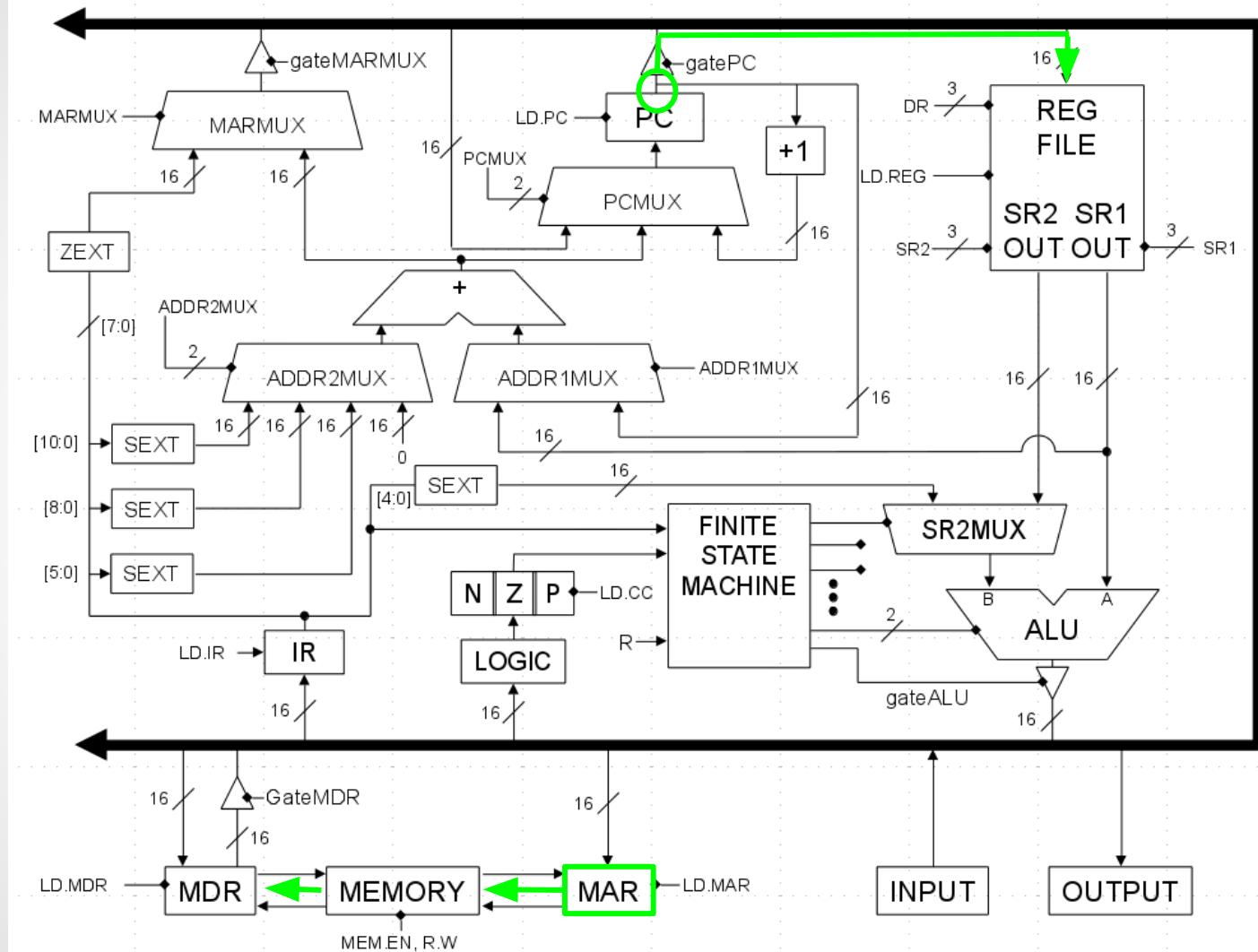
CC1:
MARMUX = ZEXT
gateMARMUX
LD.MAR



TRAP

CC1:
MARMUX = ZEXT
gateMARMUX
LD.MAR

CC2:
MEM.EN
LD.MDR
gatePC
DR = 7
LD.REG

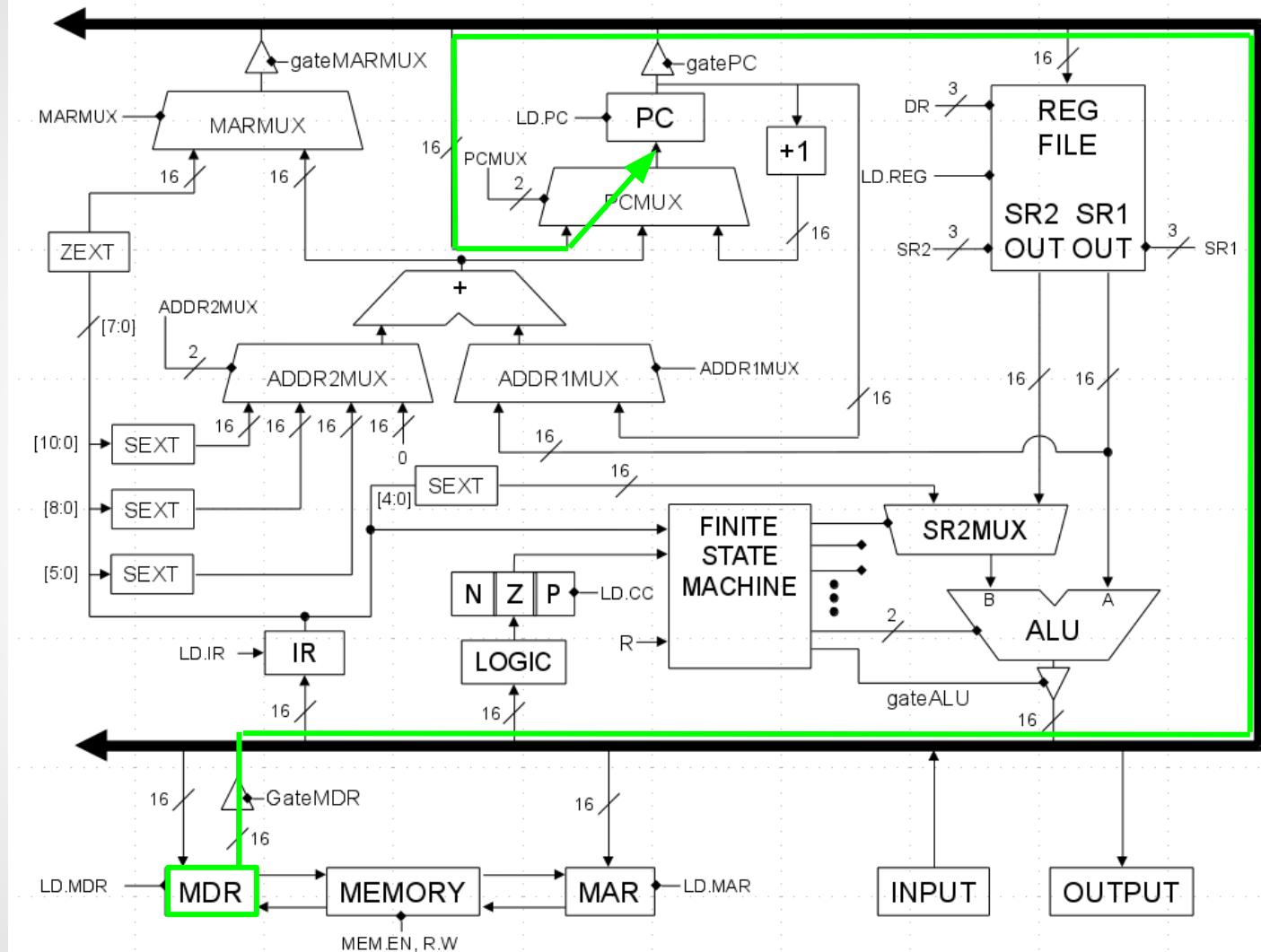


TRAP

CC1:
MARMUX = ZEXT
gateMARMUX
LD.MAR

CC2:
MEM.EN
LD.MDR
gatePC
DR = 7
LD.REG

CC3:
gateMDR
PCMUX = BUS
LD.PC



TRAP

The LC3 comes with several built-in TRAPs that perform practical functions:

- x20 (GETC): Read one ASCII character from the input buffer, and store it in R0
- x21 (OUT): Print a character from R0 to the console
- x22 (PUTS): Print a null-terminated string starting at the address pointed to by R0
- x25 (HALT): Halt the processor
- x23, x24 (IN, PUTSP) See the ISA

TRAP

You can write your own traps! Follow these simple steps:

- Write your trap code somewhere in memory
- Terminate the trap code with RET
- Write the address of the code to a free entry in the trap vector table
- Be careful not to clobber the contents of R7 during the trap code; using JSR/JSRR or TRAP within your trap code will clobber R7! You should save the contents of R7 somewhere and restore them before RET, if it must be modified in your code.
- Call your trap with TRAP x90, for example, if you wrote the address of the trap code to trap vector x90.

Pipelining

LC3 instructions with in the canonical datapath require from 5 to 9 clock cycles to complete. Could we do anything to speed up the datapath?

Pipelining

After looking at the LC3 instructions, you may start to notice a pattern in execution:

ADD:

- Get instruction
- Get operands
- Add them
- -
- Write to register

LDR:

- Get instruction
- Get SR1 and offset
- Add them
- Retrieve memory
- Write to register

ST:

- Get instruction
- Get SR and offset
- ADD PC and offset
- Write memory
- -

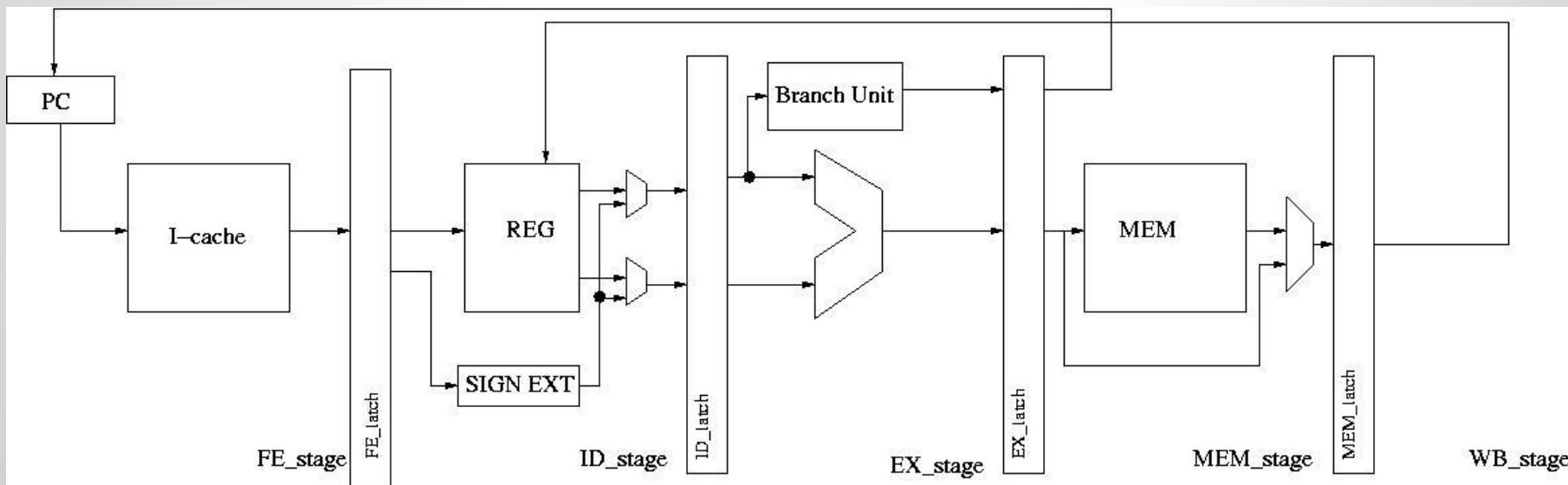
Pipelining

All the instructions follow a general pattern of fetch, decode, execute, memory access, register writeback; why not execute each in its own clock cycle, and start the next instruction while the current is executing?

	Fetch	Decode	Execute	Memory	Writeback
CC1:	inst 1				
CC2:	inst 2	inst 1			
CC3:	inst 3	inst 2	inst 1		
CC4:	inst 4	inst 3	inst 2	inst 1	
CC5:	inst 5	inst 4	inst 3	inst 2	inst 1

Pipelining

Place a register buffer between each stage; the instructions will ripple across the datapath, overall ~ 1 cycle per instr.



Pipelining

Problems to deal with in pipelining:

- Structural hazards
 - Component needed by multiple pipeline stages
- Data hazards
 - Data may be read before it's written back
- Control hazards
 - Residual data in pipeline stages after branches

Take CS 2200!