11 BONUS: Systolic Arrays [50 points]

You are given a systolic array of 2×2 Processing Elements (PEs), interconnected as shown in Figure 2. The inputs of the systolic array are labeled as H_0, H_1 and V_0, V_1. Figure 3 shows the PE logic, which performs a multiply-accumulate (MAC) operation and saves the result to an internal register (reg). Figure 3 also shows how each PE propagates its inputs. We make the following assumptions:

- The latency of each MAC operation is one cycle, i.e., if the inputs to a PE are available in cycle c, the updated register value will be available in cycle $c+1$.
- The propagation of the values from i_0 to o_0, and from i_1 to o_1, takes one cycle.
- The initial values of all internal registers is zero.

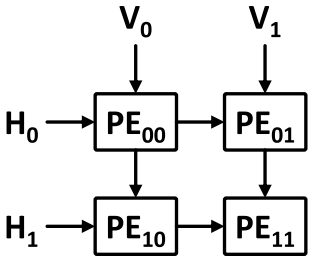

Figure 2: PE array

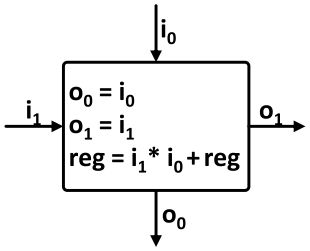

Figure 3: Processing Element (PE)

Your goal is to use the systolic array shown in Figure 2 to perform the multiplication $C = A \times B$, where A, B, and C are 2×2 matrices. Recall that the multiplication of two $K \times K$ matrices is defined as follows:

$$C_{ij} = \sum_{k=0}^{K-1} A_{ik} \times B_{kj}$$

As an example, for $K = 2$, the calculation for C_{00} is as follows:

$$C_{00} = A_{00} \times B_{00} + A_{01} \times B_{10}$$

Compute the multiplication in the minimum possible number of cycles. Fill the following table with:

1. Each input element (from matrices $A_{2\times2}$ and $B_{2\times2}$) in the correct cycle and input port of the systolic array (H_0, H_1 and V_0, V_1).
2. Each output element (for matrix $C_{2\times2}$) in the cycle and PE that generates each output.

(a) [25 points] Fill in the blanks only with relevant information. Input cells left blank are interpreted as 0.

cycle	H0	H1	V0	V1	PE_{00}	PE_{01}	PE_{10}	PE_{11}
0	A_{00}		B_{00}					
1	A_{01}	A_{10}	B_{10}	B_{01}				
2		A_{11}		B_{11}	C_{00}			
3						C_{01}	C_{10}	
4								C_{11}
5								
6								
7								

(b) [25 points] Suppose that the same systolic array from Figure 2 is used to compute the multiplication of two 4×4 matrices. How many cycles does it take to perform the multiplication? Assume that the register in a PE resets to 0 immediately after an output is generated, i.e., PEs can start accumulating for the next output element in the next cycle without waiting for an extra cycle to reset the register to 0. Show your work.

19 cycles.

Each PE needs to calculate four elements to calculate the $4 \times 4 = 16$ output elements.

For the first element calculated by each PE, the timeline looks similar to (a), but requires two additional cycles for the four MAC operations instead of two per element, i.e., seven cycles in total until PE_{11} produces its output.

The remaining three elements calculated by each PE require four cycles each if pipelined with the previously calculated element.

Thus, the total number of cycles is $7 + 3 \times 4 = 19$.

cycle	**H0**	**H1**	**V0**	**V1**	**PE_{00}**	**PE_{01}**	**PE_{10}**	**PE_{11}**
0	A_{00}		B_{00}					
1	A_{01}	A_{10}	B_{10}	B_{01}				
2	A_{02}	A_{11}	B_{20}	B_{11}				
3	A_{03}	A_{12}	B_{30}	B_{21}				
4	A_{00}	A_{13}	B_{02}	B_{31}	C_{00}			
5	A_{01}	A_{10}	B_{12}	B_{03}		C_{01}	C_{10}	
6	A_{02}	A_{11}	B_{22}	B_{13}				C_{11}
7	A_{03}	A_{12}	B_{32}	B_{23}				
8	A_{20}	A_{13}	B_{00}	B_{33}	C_{02}			
9	A_{21}	A_{30}	B_{10}	B_{01}		C_{03}	C_{12}	
10	A_{22}	A_{31}	B_{20}	B_{11}				C_{13}
11	A_{23}	A_{32}	B_{30}	B_{21}				
12	A_{20}	A_{33}	B_{02}	B_{31}	C_{20}			
13	A_{21}	A_{30}	B_{12}	B_{03}		C_{21}	C_{30}	
14	A_{22}	A_{31}	B_{22}	B_{13}				C_{31}
15	A_{23}	A_{32}	B_{32}	B_{23}				
16		A_{33}		B_{33}	C_{22}			
17						C_{23}	C_{32}	
18								C_{33}