Proceedings of the 10th World Congress on Recent Advances in Nanotechnology (RAN 2025)
Barcelona, Spain - April, 2025
Paper No. ICNNFC 152
DOI: 10.11159/icnnfc25.152

QPIM: 3D NAND Flash Memory for a Pseudo Quantum Computer Platform

Young June Park [1] and David D. Park[2]
[1,2] The 4th Consulting, Inc.
Seoul National University, Soft foundry, Gwanakku, Seoul, Korea 151-742
ypark@snu.ac.kr

***Abstract* –** This paper proposes the hardware realization of a (pseudo) quantum computer based on the 3D NAND flash memory. The essential idea behind this is to convert the $Nx2^N$ memory problem to 2^N clock problem. The schemes for ‘refresh’ and ‘reducing the 2^N ’ clock’ scheme are also presented.

***Keywords*:** pseudo quantum computer, nano NAND flash memory, memory mode, clock mode

1. Introduction

A pseudo quantum computer can be realized by either software or a CMOS circuit emulator. However, the intensive memory space required to store the quantum states makes these methods inapplicable as N, the number of Qubits, grows above 50. In this paper, we propose a new pseudo-quantum computer platform realized in 3D NAND flash memory. We propose a new algorithm based on the memory mode and clock mode to calculate the quantum state data. Using the algorithm, it will be shown that $Nx2^N$ memory problem is reduced to $Nx2^{(N-M-1)}$ where the memory burden is shared with the 2^{M-1} clocks.

2. Architecture based on nano 3D NAND flash memory

Fig. 1 shows the scheme for the 3D NAND memory-based ‘Quantum computer’. The total number of the word lines in the stack is assumed to be m, which is around 100 for the present technology. Each stack in 3D memory, the same word line plane, is assigned to the N Qubits. SSL(n) is assigned to the nth Qubit so that activation of SSL(n) enables the ‘read operation and write operation’ of the nth Qubit. To realize a 100 Qubit computer, we need the number of SSL=100. Bit lines are assigned to the ‘digital bits’ of the a and b (|0> and |1>) components of the Qubit.

2-1. Gate Operation and Concept of ‘Refreshment’

The gate operations to the Qubits are performed in three steps.

Step 1: Compute all the contents of the 2^N ‘eigenstates’ after performing the ‘quantum gate operation’

Step 2: Reconstruct the N Qubit components by computing corresponding eigenstates contents. Computation is simply an arithmetic operation of ‘square, add, and square-root’ of all the contents

Step 3: ‘Refresh’ to store back the reconstructed N Qubits in the word line planes

For the ‘non-CNOT’ operation, Steps 1 and 2 are simplified to compute only the corresponding Qubit.

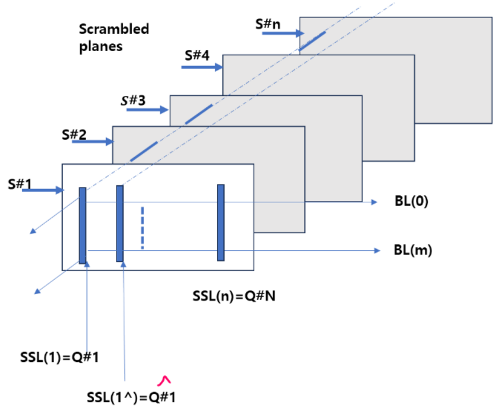

Fig. 1. The 3D NAND architecture for the pseudo quantum computer. a) 1st word plane is assigned to N Qubits. SSL(n) corresponds to the digital data for l0>, l1>, and lΦ> of nth Qubit. b) Word line planes from S#2 to S#n are dedicated to the N Qubits for 'scrambler' for ease of 'refresh' and arithmetic operations explained in the paper.

2-2. Scrambler and 'in memory' computing

Two schemes are proposed to reduce the arithmetic computing time and memory space used for the 'Refresh step' explained in Section 2.

First, NAND and NOR operations are performed for 'in memory add/multiplication' in Step 2 in Section 2 using the 'Flash-Cosmos (Flash Computation with One Shot Multi-Operand Sensing)' proposed in [1]. In this scheme, the multiplication and addition of quantum state data (see the vertical lines in Fig. 2, pick two vertical lines for A and B for clarity) for reconstruction of a specific bit Qubit #j can be performed by storing A and B in two different blocks, one selection of whole word planes followed by OR operation of A and B;

$$(A1, j\bullet \ldots \bullet AN, j) + (B1, j\bullet \ldots \bullet BN, j). \tag{1}$$

The arithmetic operation is simplified due to NAND flash memory's inherent NAND and NOR operation [1].

Second, other word line planes, called the scrambler in Fig.1, are also assigned to N Qubits, but with different sequences to minimize the memory space and the arithmetic operations needed to reconstruct the Qubit data |0> and |1> after applying the quantum operations.

Fig. 2(b) shows an example of the 'scrambler' applied to the case of N=3. Compared with Fig.2(a), which is the memory mode, the clock mode in Fig. 2(b) requires a number of clocks of 3. In each clock, the Qubit data(Q2 and Q3) in the scrambler planes shift in the clockwise direction. Combining the Flash-Cosmos scheme and the scrambler, the computation time (clock mode) and memory space(memory mode) for the 'refresh' step can be compromised from 2^3 to 1. In this way, the mixture of the clock mode and memory mode gives the opportunity for the 'NAND' structure to be a general quantum computer platform.

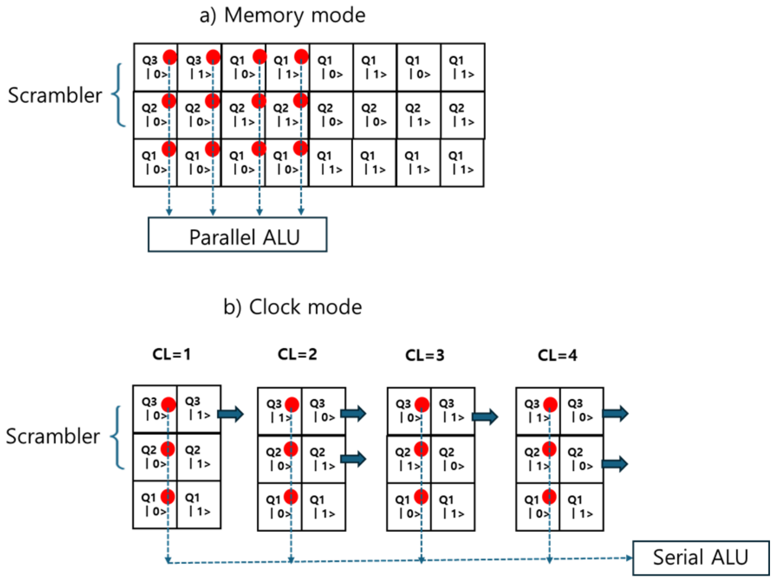

Fig. 2. Example of the ‘memory mode’ and the ‘clock mode’ for the Qubit system of N=3 as an example. The vertical lines are merged to ALU for refreshment of the Qubit #1 |0>. In the ‘clock mode’, total memory of 2N (N=3 in this example) is needed, whereas total clocks of 2^{N-1} are needed. The 2nd and 3rd planes (word line planes) are used for the scrambler.

3. Application example

The scheme has been applied to the super dense coding protocol [2] in both the ‘memory’ mode and the ‘clock mode’. In the clock mode, the number of clocks needed for the quantum circuit operation is only 4, with the memory allocation being 4x4 bytes. For the typical clock period of 1psec, which is a typical number in recent CMOS technology, the full quantum operation of the superdense protocol takes less than 4 p sec.

4. Conclusion

A new quantum computer architecture is proposed. The architecture is based on 3D NAND memory. By introducing the ‘in memory computing’ for the arithmetic unit and combination of ‘memory mode’ and the ‘clock mode’, the $Nx2^N$ memory problem is reduced to $(N-M)x2^{(N-M-1)}$ where the memory burden is shared with the 2^{M-1} clocks.

Acknowledgments

One of the authors thanks Samsung Electronics, Inc. for the financial support of this research.

References

[1] J. Park, R. Azizi, G. Oliveira, M. Sadrosadati, R. Nadig, D. Novo, J. Gómez-Luna, M. Kim and O. Mutlu, "Flash-Cosmos: In-Flash Bulk Bitwise Operations Using Inherent Computation Capability of NAND Flash Memory," International Symposium on High-Performance Computer Architecture (HPCA), 2023.

[2] C. H. Bennett and S. J. Wiesner, "Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states," *Phys. Rev. Lett.*, vol. 69, no. 20, pp. 2881–2884, Nov. 1992.