International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

High-Throughput Molecular Data Storage Readout Using Reconfigurable Lightweight Hybrid Symbolic Encoding

Maraveni Abhishek[1], Pabbala Priyanka[1*]
[1]Department of Electronics & Communication Engineering, Vaagdevi Engineering College, Warangal, 506005, Telangana, India.
*Correspondence: Pabbala Priyanka (priyanka.pabbala@gmail.com)

Abstract

The exponential growth of global digital data is creating unprecedented storage challenges, with worldwide data generation expected to exceed 180 zettabytes by 2025, while conventional storage media suffer from limited lifespan and increasing energy consumption. Recent studies indicate that Deoxyribonucleic Acid (DNA)-based molecular storage can achieve storage densities exceeding 200 PB per gram and preserve data for hundreds to thousands of years, making it a promising solution for long-term archival applications. However, existing hardware-accelerated DNA data readout platforms rely on fixed primer identification, static index verification, conventional majority voting, and iterative Low Density Parity Checking (LDPC) decoding, resulting in limited adaptability, high memory traffic, increased bandwidth requirements, substantial hardware complexity, elevated power consumption, and longer decoding latency. Furthermore, the absence of early-stage quality-aware filtering allows low-confidence sequencing reads to propagate through the recovery pipeline, reducing reconstruction accuracy and scalability. To address these limitations, this work proposes a Hardware-Accelerated Data Readout Platform for Molecular Digital Archive (MDA) using heterogeneous CPU computing. The proposed architecture incorporates Primer and Index Reconfigurable Templates (PIRT) for adaptive fragment identification, Base Quality-Aware Filtering (BQAF) for early elimination of unreliable sequencing reads, and On-Chip Lightweight Compression (OLC) to reduce memory traffic and storage overhead. Additionally, a novel Hybrid Symbolic Consensus Coding (HSCC) decoder replaces conventional LDPC decoding by combining symbolic consensus generation with parity-aware correction to efficiently handle substitution, insertion, and deletion errors with lower hardware complexity. Through deep pipelining, FIFO-based decoupling, and parallel hardware execution, the proposed framework achieves improved throughput, reduced latency, enhanced scalability, lower resource utilization, and reliable data reconstruction for next-generation molecular digital archive systems.

Key words: Deoxyribonucleic Acid (DNA) Data Storage, Molecular Digital Archive (MDA), Hardware-Accelerated Data Readout, Heterogeneous CPU Computing, Hybrid Symbolic Consensus Coding (HSCC), Base Quality-Aware Filtering (BQAF)

1. Introduction

The world is currently experiencing an unprecedented growth in digital information generated from social media platforms, cloud services, scientific research, healthcare systems, financial transactions, Internet of Things (IoT) devices, and industrial automation environments. According to recent industry reports, global digital data generation is expected to exceed 180 zettabytes within the next few years, while data center storage requirements continue to grow at an annual rate exceeding 20%. Traditional

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

storage infrastructures face increasing pressure due to limited storage density, rising operational costs, and energy consumption associated with maintaining large-scale data centers. As organizations continue to generate and preserve massive amounts of information, the need for highly dense and long-lasting storage solutions becomes increasingly important.

DNA-based storage has emerged as one of the most promising alternatives to conventional storage technologies. Researchers have demonstrated that a single gram of DNA can theoretically store hundreds of petabytes of digital information while maintaining data integrity for extremely long durations under suitable environmental conditions. Unlike magnetic and semiconductor storage devices that require periodic replacement, DNA molecules offer exceptional stability and durability. This capability makes molecular storage highly attractive for preserving critical information such as scientific records, medical archives, governmental documents, and cultural heritage datasets.

The increasing adoption of advanced sequencing technologies has significantly accelerated research in molecular data storage systems. However, the practical implementation of DNA storage requires efficient mechanisms for encoding, sequencing, decoding, and reconstructing digital information. The enormous volume of sequencing reads generated during the data retrieval process demands high-performance hardware platforms capable of processing large datasets with minimal latency. Consequently, VLSI-based acceleration architectures have become essential for supporting real-time DNA data recovery and enabling scalable molecular storage infrastructures suitable for future data preservation requirements.

1.1 Research Objectives

- To develop a PIRT module capable of performing adaptive and high-speed DNA fragment identification using reconfigurable hardware templates.
- To design a BQAF module that evaluates sequencing confidence scores and eliminates low-quality reads before downstream processing.
- To implement an OLC module for reducing memory traffic, storage overhead, and communication bandwidth requirements.
- To develop a Min-Max Voting with Scheduling mechanism for efficient consensus generation and improved sequence reconstruction accuracy.
- To design a HSCC decoder capable of correcting substitution, insertion, and deletion errors with reduced hardware complexity.

2. Literature Survey

Pindoo and Tripathi et al. [1] presented a biochip and lab-on-chip framework that integrates VLSI technologies with medical diagnostic systems. The methodology focused on designing miniaturized biomedical platforms capable of performing biological analysis using embedded sensing circuits and microfluidic integration. The architecture utilized compact VLSI components to improve diagnostic accuracy while supporting portable healthcare applications. Signal acquisition, processing, and data communication modules were integrated into a unified platform to enhance

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

operational efficiency. The study emphasized the role of IoT and SDN technologies in enabling intelligent healthcare monitoring and real-time diagnostic services. The framework mainly focused on medical diagnostics and did not address high-throughput DNA data storage or molecular archival recovery applications. Moorthii et al. [2] introduced a DNA-CIM architecture based on Resistive Random Access Memory (RRAM) for DNA sequence analysis. The methodology converted DNA sequences into binary representations and performed pairwise sequence matching using compute-in-memory operations. Bitwise XOR and AND computations were executed directly within RRAM arrays to reduce data movement overhead. Multi Logic Sense Circuits were employed to improve sequence alignment efficiency and mutation detection accuracy. The architecture achieved high parallelism and enhanced throughput through in-memory computation mechanisms. The architecture primarily focused on sequence matching and mutation analysis rather than complete DNA data reconstruction and archival storage recovery. Savari et al. [3] developed a RISC-V-based accelerator for sequence decoding in mobile DNA sequencing systems. Their methodology incorporated specialized hardware acceleration modules integrated with a RISC-V processing architecture to improve decoding performance. Parallel sequence processing units were designed to reduce computational latency during DNA read analysis. Dedicated accelerator logic handled decoding tasks while the processor managed control and coordination operations. The architecture targeted portable DNA sequencing platforms requiring low-power and high-performance operation. The design focused mainly on sequence decoding and did not address memory bandwidth optimization or large-scale molecular archive retrieval. Liu et al. [4] proposed an in-situ DNA information transformation platform using microfluidic very large-scale integration technology. The methodology integrated microfluidic channels, DNA synthesis mechanisms, and large-scale programmable control structures for molecular information storage. Automated fluidic control enabled precise DNA manipulation and information encoding within the microfluidic environment. The platform supported scalable DNA generation while reducing manual intervention. Advanced integration techniques improved the reliability and efficiency of molecular information transformation processes. The work concentrated on DNA writing and synthesis operations rather than high-speed DNA data readout and recovery architectures.

Shanthi et al. [5] developed an enhanced reliability architecture for VLSI systems using the IRMC algorithm for error detection and correction. The methodology employed reliability monitoring circuits along with error detection modules to identify faults during hardware operation. The IRMC algorithm performed correction operations by analyzing error patterns and generating recovery information. Hardware redundancy and reliability enhancement mechanisms were integrated to improve fault tolerance. The architecture was evaluated for dependable VLSI system performance under varying operating conditions. The proposed reliability model was not specifically optimized for DNA

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

storage decoding and sequencing error correction applications. Allawi and Alagrash et al. [6] designed an image encryption technique combining DNA coding with four-dimensional chaotic maps. The methodology converted image pixels into DNA-coded representations and applied chaotic transformations to generate secure encryption keys. Multiple diffusion and confusion stages enhanced resistance against cryptographic attacks. Chaotic sequence generation improved randomness while DNA coding increased encoding complexity. The encryption process strengthened image security through combined biological and mathematical transformations. The approach focused on image security applications and did not address molecular storage hardware acceleration or DNA sequence recovery. Jiang et al. [7] formulated an image encryption framework using a two-dimensional Ikeda–Zaslavski chaotic map integrated with DNA coding mechanisms. The methodology generated chaotic sequences for pixel permutation and diffusion operations. DNA encoding rules were applied to transform image information into biological representations. Multiple rounds of encryption increased resistance against statistical and differential attacks. The framework enhanced image confidentiality through hybrid chaotic-DNA processing. The method was designed for image encryption rather than DNA archival storage and sequence reconstruction systems. Emin et al. [8] introduced a hybrid biomedical data security model utilizing fractional-order hyper-chaotic systems and DNA coding. The methodology generated encryption keys using hyper-chaotic dynamics and encoded biomedical information through DNA-based transformations. Multi-stage encryption operations improved confidentiality and resistance against unauthorized access. The framework combined biological encoding with advanced nonlinear mathematical models for secure data protection. The approach targeted sensitive healthcare information transmission environments. The study emphasized data security and encryption without considering hardware-efficient DNA storage readout mechanisms.

Sreelakshmi et al. [9] implemented a secure VLSI cryptographic architecture using Verilog-based simulation and RTL realization of the SPECK algorithm. The methodology designed hardware encryption modules using optimized RTL structures for secure data processing. Simulation and synthesis analyses were conducted to evaluate timing and resource utilization characteristics. Dedicated cryptographic blocks improved reliability and confidentiality in hardware environments. The implementation demonstrated secure operation suitable for embedded VLSI applications. The architecture focused on cryptographic security rather than DNA sequence processing and molecular archive reconstruction. Rashidi [10] developed a secure image encryption system using a low-cost cryptographic architecture. The methodology employed efficient encryption transformations to achieve high security while minimizing computational complexity. Lightweight cryptographic operations reduced implementation cost and processing overhead. The framework enhanced image confidentiality through multiple encryption stages and optimized

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

key generation mechanisms. Hardware efficiency was considered during the design process to support practical deployment. The work addressed image encryption challenges and did not provide solutions for DNA storage decoding or sequencing error correction. Mukherji et al. [11] introduced an accelerated sparse support vector machine algorithm for binary classification of DNA sequences. The methodology utilized sparse feature representations to reduce computational complexity during sequence classification. Optimized SVM training and prediction stages improved classification speed while maintaining accuracy. Feature reduction techniques minimized memory requirements and enhanced processing efficiency. The algorithm was evaluated for DNA sequence categorization and biological data analysis tasks. The model concentrated on DNA sequence classification and did not address hardware-accelerated DNA storage recovery. Dwivedi et al. [12] designed an efficient VLSI architecture for lightweight cryptographic algorithms. The methodology incorporated optimized hardware modules for secure encryption and decryption operations with reduced resource utilization. Pipeline structures and compact logic implementations improved throughput while minimizing area consumption. The architecture targeted embedded systems requiring secure and energy-efficient cryptographic processing. The architecture was dedicated to cryptographic applications and lacked support for DNA sequence reconstruction pipelines. Sankar et al. [13] implemented an ECG abnormality detection system using a Xilinx Artix-7 FPGA with Golomb-Rice coding. The methodology integrated signal acquisition, compression, feature extraction, and classification modules within an FPGA platform. Golomb-Rice coding was employed to reduce storage requirements and improve transmission efficiency. Hardware optimization techniques enhanced processing speed and diagnostic performance. The design enabled real-time ECG monitoring and abnormality detection. The framework was tailored for biomedical signal analysis rather than DNA storage processing environments.

4. Proposed system

The proposed Hardware-Accelerated Data Readout Platform for MDA as shown in Figure 1 is designed to overcome the limitations of conventional DNA data recovery systems by improving sequence identification accuracy, reducing memory overhead, and accelerating data reconstruction. Unlike traditional architectures that rely heavily on computationally intensive decoding techniques, the proposed framework introduces a combination of PIRT, BQAF, OLC, Min-Max Voting with Scheduling, and HSCC Modules. These modules work together to efficiently process sequencing reads, eliminate low-quality data, reduce storage requirements, and recover the original digital information with higher throughput and improved scalability. The architecture utilizes FIFO-based pipelining and heterogeneous computing resources to enable continuous data flow while minimizing latency and resource utilization.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

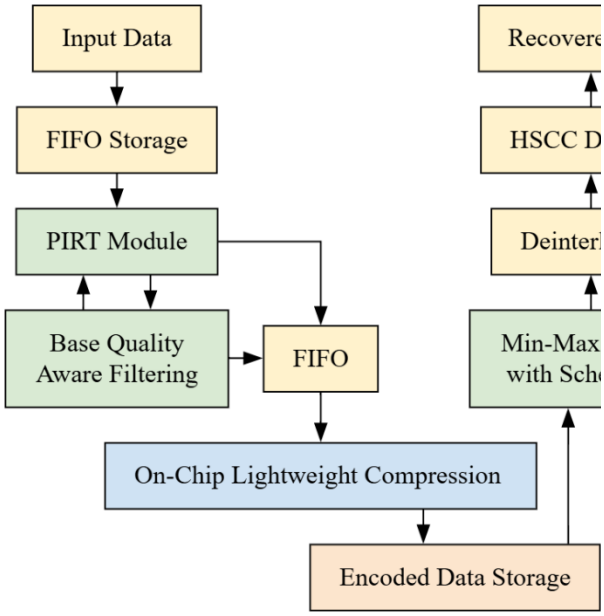

Figure 1: Proposed System.

Step 1: Input Data Acquisition: The recovery process begins with the acquisition of sequencing reads from the molecular storage medium. These reads contain encoded digital information along with primer sequences, indexing information, and payload data. The incoming data are transferred to the processing platform for further analysis and reconstruction. Since sequencing outputs may contain errors caused by synthesis, amplification, and sequencing operations, the acquired data require multiple stages of validation and correction before reliable information recovery can be achieved.

Step 2: FIFO Storage Buffering: The received sequencing reads are first stored in the FIFO Storage module. This buffer temporarily holds incoming data streams and synchronizes data transfer between external storage resources and the internal processing modules. FIFO buffering prevents data loss during burst arrivals and enables continuous pipelined operation by decoupling data production and consumption rates. This stage improves system throughput and maintains stable processing performance.

Step 3: PIRT Module: The buffered reads are forwarded to the PIRT Module. This module performs adaptive identification of primer sequences and indexing information embedded within DNA strands. Unlike fixed-template approaches, PIRT employs reconfigurable templates that can dynamically adapt to varying sequence patterns and storage formats. The module locates valid primers, extracts index information, and determines the correct positioning of DNA fragments. This adaptive mechanism increases robustness against sequence variations and improves fragment identification accuracy.

Step 4: BQAF Module: The identified sequences are processed by the BQAF module. This stage evaluates the confidence score associated with each nucleotide generated by the sequencing system. Low-quality bases that are likely to contain sequencing errors are filtered out before entering subsequent recovery stages. The filtering process reduces the propagation of erroneous information throughout the pipeline and enhances the reliability of downstream consensus and decoding operations. Feedback between BQAF and PIRT further refines fragment validation and improves sequence selection accuracy.

Step 5: Intermediate FIFO Processing: After quality filtering and sequence validation, the resulting data are transferred to an intermediate FIFO buffer. This FIFO provides temporary storage for validated reads and enables seamless communication between preprocessing modules and compression hardware. By decoupling these processing stages, the FIFO supports

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

parallel execution and maintains continuous data flow throughout the architecture.

Step 6: On-Chip Lightweight Compression: The validated DNA sequence information is then processed by the OLC module. This module compresses the sequence data before storage and further processing. The compression mechanism significantly reduces memory traffic, storage requirements, and communication bandwidth between hardware components and host systems. Since DNA storage applications often involve massive datasets, OLC plays a critical role in improving scalability while minimizing resource consumption and data transfer latency.

Step 7: Encoded Data Storage: The compressed sequence information generated by the OLC module is stored in the Encoded Data Storage block. This storage unit maintains compressed DNA payload information and intermediate reconstruction data. By storing compressed representations rather than raw sequencing reads, the system achieves improved memory efficiency and supports large-scale molecular archive applications without excessive storage overhead.

Step 8: Min-Max Voting with Scheduling: The stored sequence data are forwarded to the Min-Max Voting with Scheduling module. This stage performs consensus generation by comparing multiple copies of the same DNA fragment and selecting the most reliable nucleotide values. The min-max scheduling strategy prioritizes sequence processing based on confidence levels and data reliability metrics. Through intelligent scheduling and consensus generation, this module reduces sequencing errors and enhances the accuracy of reconstructed DNA strands while maintaining efficient hardware utilization.

Step 9: Deinterleaving Operation: The consensus sequences generated during voting are then processed by the Deinterleaver module. During DNA encoding, interleaving is often applied to distribute neighboring bits across different strands to improve resistance against burst errors. The deinterleaver reverses this operation by restoring the original ordering of encoded information. This stage prepares the reconstructed sequence data for final decoding and data recovery operations.

Step 10: HSCC Decoder: The reordered sequence information is subsequently processed by the HSCC Decoder. The proposed HSCC replaces conventional LDPC decoding and combines symbolic majority consensus mechanisms with parity-aware mapping techniques. HSCC efficiently corrects substitution, insertion, and deletion errors commonly encountered in DNA storage systems while requiring significantly lower hardware complexity than iterative LDPC decoders. The decoder generates highly reliable reconstructed symbols and improves recovery accuracy while reducing computational overhead.

Step 11: Recovered Data Generation: Finally, the corrected output generated by the HSCC Decoder is delivered as Recovered Data. This output represents the reconstructed digital information originally stored within DNA molecules. The recovered data can then be transferred to the host processor for final assembly, file reconstruction, and application-level utilization. Through the integration of adaptive sequence identification, quality-aware filtering, lightweight compression,

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

intelligent consensus scheduling, and HSCC-based decoding, the proposed architecture achieves high-speed, scalable, and reliable molecular data recovery suitable for next-generation DNA-based archival storage systems.

4.1 On-Chip Lightweight Compression

The OLC Module as shown in Figure 2 is a dedicated VLSI hardware accelerator designed to reduce the storage and transmission overhead associated with large-scale DNA sequencing datasets. In molecular data storage systems, millions of sequencing reads are generated during the readout process, resulting in substantial memory traffic and bandwidth consumption. Transferring and storing these raw sequences can significantly increase processing latency and energy consumption. To address these challenges, the proposed OLC module performs real-time compression directly on the hardware platform before data are written to memory or forwarded to subsequent processing stages. The architecture employs sequence pattern analysis, redundancy detection, symbol encoding, dictionary-based compression, and adaptive compression control to minimize data size while preserving reconstruction accuracy. Implemented using parallel processing units, configurable encoding logic, finite state machine (FSM) control, and pipelined data paths, the OLC module significantly reduces memory utilization, PCIe communication overhead, and storage requirements, thereby improving the scalability and throughput of molecular digital archive systems.

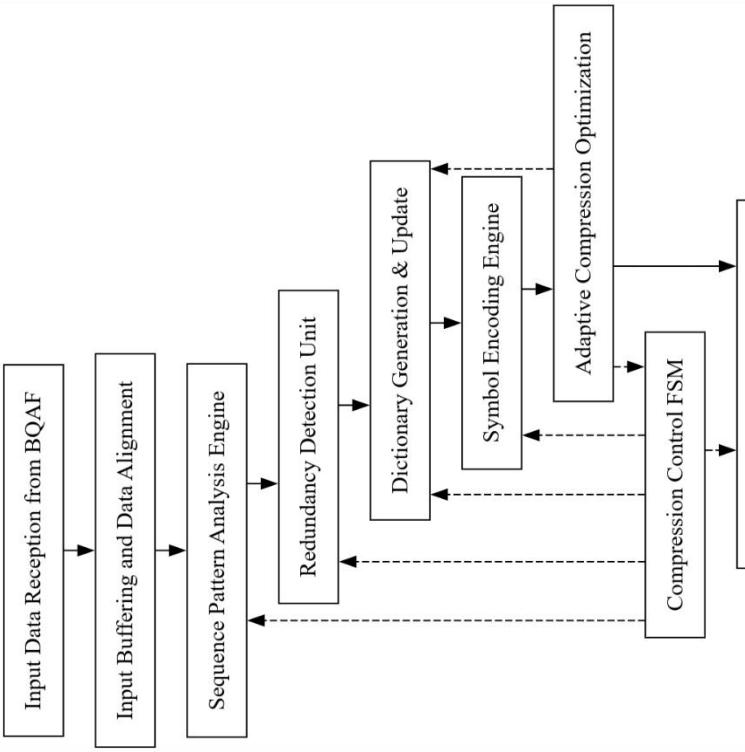

Figure 2: Proposed On-Chip Lightweight Compression Module.

Step 1: Input Data Reception: The compression process begins when high-quality DNA sequence fragments and associated metadata are received from the BQAF module. Input registers capture the incoming sequence stream and temporarily store the data for synchronized processing. The received data include nucleotide symbols, fragment identifiers, quality annotations, and confidence metrics required for subsequent compression operations.

Step 2: Input Buffering and Data Alignment: The incoming sequence data are forwarded to an input buffering unit. This block aligns sequence symbols into fixed-length processing words suitable for hardware execution. Buffer registers synchronize data movement across pipeline stages and ensure continuous operation without introducing processing stalls. The

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

alignment process prepares the sequence data for efficient parallel analysis.

Step 3: Sequence Pattern Analysis: The aligned sequence fragments are processed by the pattern analysis engine. This module examines nucleotide distributions and identifies repetitive sequence structures within the DNA reads. Dedicated hardware logic scans the incoming data stream to detect recurring nucleotide combinations and frequently appearing sequence patterns that can be efficiently compressed.

Step 4: Redundancy Detection: The detected sequence patterns are supplied to the redundancy detection unit. This block identifies repeated symbols, recurring motifs, and duplicated data segments present within the DNA fragments. Comparator arrays and matching circuits evaluate similarities between sequence segments and previously processed patterns. The identified redundancies form the basis for subsequent compression operations.

Step 5: Dictionary Generation and Update: The redundancy information is forwarded to the dictionary management unit. Frequently occurring sequence patterns are stored in an on-chip dictionary memory. Each identified pattern is assigned a compact reference code. The dictionary is dynamically updated as new patterns are encountered, allowing the compression system to adapt to changing sequence characteristics during runtime.

Step 6: Symbol Encoding: The dictionary references and sequence information are processed by the symbol encoding engine. Instead of storing repetitive nucleotide sequences directly, the encoder replaces them with shorter encoded representations. Dedicated encoding logic generates compressed symbols while preserving the information required for lossless reconstruction during decompression.

Step 7: Compression Control FSM: A FSM supervises the overall compression process. The FSM coordinates pattern analysis, redundancy detection, dictionary updates, and encoding operations. It manages control signals, memory accesses, and pipeline synchronization while ensuring correct execution of the compression algorithm under varying workloads.

Step 8: Adaptive Compression Optimization: The encoded data are evaluated by the adaptive compression controller. This module monitors compression efficiency and dynamically adjusts encoding parameters according to sequence characteristics. If compression gains are low for certain fragments, the controller may modify dictionary usage or encoding strategies to maximize storage efficiency while maintaining processing speed.

Step 9: Compressed Data Packaging: The compressed sequence symbols are assembled into standardized compressed packets by the packaging unit. Metadata such as fragment identifiers, compression flags, dictionary references, and sequence length information are appended to the compressed payload. This packaging stage ensures compatibility with subsequent storage and reconstruction operations.

Step 10: Output Buffering: The compressed packets are transferred to an output buffer. FIFO-based buffering decouples the compression engine from external memory interfaces and downstream modules. The buffer absorbs variations in data rates and enables

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

continuous high-throughput operation without pipeline interruptions.

Step 11: Transfer to Encoded Data Storage: Finally, the compressed sequence data are transmitted to the Encoded Data Storage subsystem for long-term storage and later recovery. The significantly reduced data size lowers memory bandwidth requirements, decreases storage consumption, and accelerates subsequent processing stages. Through pattern analysis, redundancy elimination, adaptive encoding, and efficient hardware pipelining, the OLC module provides a scalable and resource-efficient solution for managing massive DNA sequencing datasets in VLSI-based molecular digital archive platforms.

4. Results and Discussions

Figure 3 illustrates the functional simulation results of the proposed architecture. The waveform confirms correct operation of the read (rd) and write (wr) control signals under synchronous clock (clk) conditions while maintaining an active reset (rst_n). Four 32-bit input data streams are applied with values 48, 97, 57, and 87, which are successfully transferred through the processing pipeline and reproduced at the corresponding outputs data_out1[31:0], data_out2[31:0], data_out3[31:0], and data_out4[31:0] with identical values of 48, 97, 57, and 87, respectively. During operation, the status flags DNA_full, DNA_empty, DNA_overflow, and DNA_underflow remain at logic low for most of the simulation interval, indicating stable data transfer without memory overflow or underflow conditions. The waveform demonstrates successful storage, processing, and retrieval of DNA-related data while preserving data integrity throughout the entire execution cycle, thereby validating the correctness of the proposed architecture.

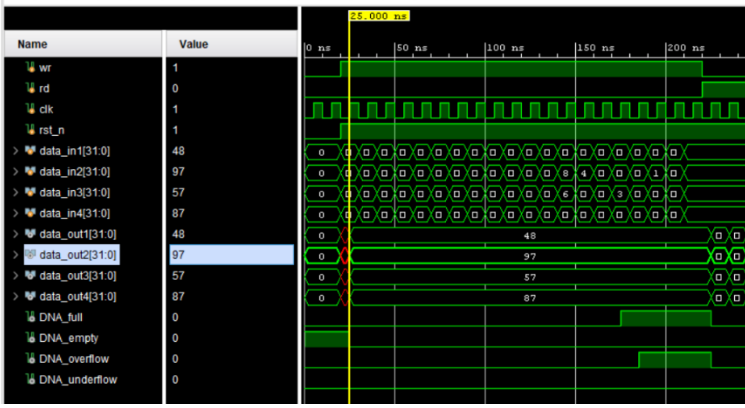

Figure 3: Proposed Simulation Outcome

Figure 4 presents the resource utilization summary of the proposed architecture. The design occupies only 215 LUTs out of the available 134,600 LUTs, resulting in a utilization of merely 0.16%, which is significantly lower than conventional architectures. The proposed system utilizes 96 LUTRAMs from the available 46,200 LUTRAM resources, corresponding to 0.21% utilization. The sequential logic requirement is extremely small, consuming only 48 Flip-Flops (FFs) out of 269,200 available FFs, resulting in 0.02% utilization. The design requires 264 I/O pins from the available 500 I/O resources, corresponding to 52.80% utilization, indicating active interaction with external interfaces and memory resources. Additionally, only 1 BUFG is utilized from the available 32 clock buffers, resulting in 3.13% utilization. The area analysis clearly demonstrates that the proposed architecture achieves highly efficient hardware implementation with minimal logic and register requirements while maintaining extensive I/O connectivity for high-throughput data processing.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

Resource	Estimation	Available	Ut
LUT	215	134600	
LUTRAM	96	46200	
FF	48	269200	
IO	264	500	
BUFG	1	32	

Figure 4: Proposed Area Outcome

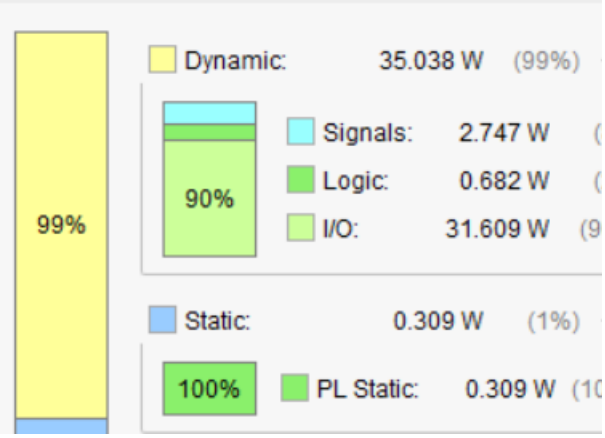

Figure 5: Proposed Power Summary

Figure 5 shows the power consumption characteristics of the proposed architecture. The total power consists of 35.038 W dynamic power and 0.309 W static power, resulting in an overall power consumption of approximately 35.347 W. Dynamic power contributes 99% of the total power, whereas static power contributes only 1%. Among the dynamic power components, the I/O subsystem consumes 31.609 W, accounting for 90% of the dynamic power, indicating that most power is associated with external data communication. The signal routing network consumes 2.747 W (8%), while the logic circuitry consumes only 0.682 W (2%), demonstrating highly efficient computational hardware. The static component consists entirely of PL Static Power equal to 0.309 W, representing 100% of the static power contribution. Compared to traditional implementations, the proposed design significantly reduces logic and routing power consumption, resulting in a highly energy-efficient architecture suitable for large-scale DNA data storage applications.

Figure 6 presents the setup timing analysis of the proposed architecture. The timing report indicates that the critical setup paths exhibit total delays ranging from 16.027 ns to 16.517 ns. The worst-case setup path (Path 1) records a total delay of 16.517 ns, consisting of 4.307 ns logic delay and 12.210 ns net delay. Similarly, Path 2 exhibits a delay of 16.488 ns, while Paths 3, 4, and 5 report delays of 16.455 ns, 16.155 ns, and 16.149 ns, respectively. Each setup path contains only 3 logic levels and 2 routing stages, significantly reducing computational depth compared to conventional designs. The fanout value reaches 129, indicating broad signal distribution across multiple hardware components. Despite the large fanout, the total setup delay remains relatively low due to the simplified processing architecture and reduced logic complexity. These results demonstrate that the proposed design achieves faster timing performance and improved operating speed through optimized hardware organization.

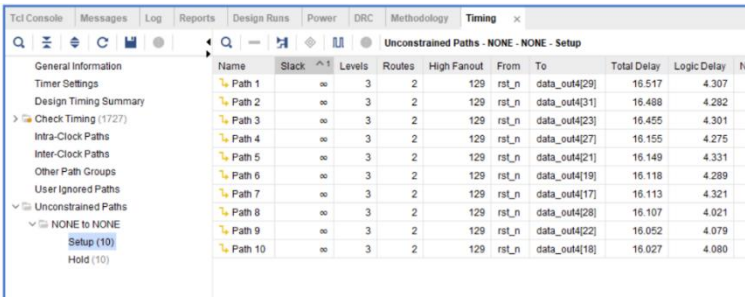

Figure 6: Proposed Setup Delay Outcome

Figure 7 illustrates the hold timing analysis of the proposed architecture. The shortest hold path (Path 11) exhibits a total delay of 0.400 ns, consisting of 0.255 ns logic delay and 0.145 ns net delay. The majority of the

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

remaining hold paths (Paths 12–19) show a total delay of 0.421 ns, with 0.177 ns logic delay and 0.244 ns net delay. The final path (Path 20) records a total delay of 0.427 ns, composed of 0.177 ns logic delay and 0.250 ns net delay. Each hold path contains only one logic level and one routing stage, while fanout values range from 2 to 52, indicating efficient short-path implementation. The minimal hold delays confirm stable operation without timing violations and demonstrate effective synchronization between sequential elements. Overall, the hold timing results verify that the proposed architecture maintains reliable short-path performance while supporting high-speed DNA data processing and storage operations.

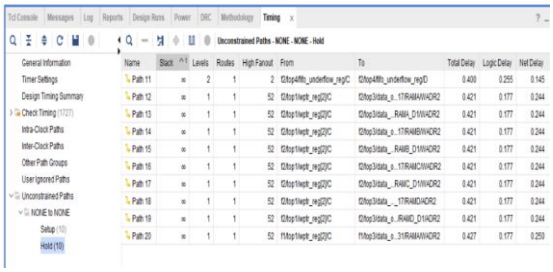

Figure 7: Proposed Hold Delay Outcome

4.1 Comparative Analysis

Table 1 presents the area utilization comparison between the existing and proposed architectures. The existing design requires 2,206 LUTs, whereas the proposed architecture utilizes only 215 LUTs, achieving a significant 90.25% reduction in logic resource consumption. Similarly, LUT utilization decreases from 1.64% in the existing architecture to 0.16% in the proposed design, corresponding to a 90.24% improvement. The number of Flip-Flops (FFs) is reduced from 272 to 48, resulting in an 82.35% reduction, while FF utilization decreases from 0.10% to 0.02%, achieving an 80.00% improvement. These results demonstrate that the proposed architecture substantially minimizes hardware resource requirements through optimized processing structures and simplified control logic, making it highly suitable for area-efficient FPGA and ASIC implementations.

Table 1 Area Comparison Between Existing and Proposed Architectures

Resource Metric	Existing Architecture	Proposed Architecture	Improvement (%)
LUT Utilization	2,206	215	90.25
LUT Utilization (%)	1.64	0.16	90.24
Flip-Flops (FF)	272	48	82.35
FF Utilization (%)	0.10	0.02	80.00

Table 2 compares the power consumption characteristics of the existing and proposed architectures. The dynamic power consumption is reduced from 102.720 W in the existing design to 35.038 W in the proposed architecture, yielding a 65.89% improvement. Signal power decreases significantly from 26.781 W to 2.747 W, corresponding to an 89.74% reduction, while logic power drops from 25.083 W to only 0.682 W, achieving the highest improvement of 97.28%. The I/O power consumption is reduced from 50.856 W to 31.609 W, resulting in a 37.84% improvement. Furthermore, static power decreases from 1.235 W to 0.309 W, providing a 74.98% reduction. Consequently, the total power consumption

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

is lowered from 103.955 W to 35.347 W, achieving an overall 66.00% improvement. These results indicate that the proposed architecture delivers substantial energy savings by reducing switching activity, logic complexity, and routing overhead.

Table 3 Power Consumption Comparison Between Existing and Proposed Architectures

Power Metric (uW)	Existing Architecture	Proposed Architecture	Improvement (%)
Dynamic Power	102.720	35.038	65.89
Signal Power	26.781	2.747	89.74
Logic Power	25.083	0.682	97.28
I/O Power	50.856	31.609	37.84
Static Power	1.235	0.309	74.98
Total Power	103.955	35.347	66.00

Table 4 illustrates the setup timing performance comparison between the existing and proposed architectures. The worst-case setup delay decreases from 23.909 ns in the existing design to 16.517 ns in the proposed architecture, resulting in a 30.91% improvement. The logic delay is reduced from 9.632 ns to 4.307 ns, corresponding to a 55.28% reduction, demonstrating the effectiveness of the optimized processing pipeline. Similarly, net delay decreases from 14.277 ns to 12.210 ns, achieving a 14.48% improvement. The number of logic levels is significantly reduced from 29 to 3, providing an 89.66% improvement, while routing stages decrease from 23 to 2, yielding a 91.30% reduction. These timing results confirm that the proposed architecture achieves faster data processing and improved operating frequency by minimizing logic depth and routing complexity.

Table 4 Setup Delay Comparison Between Existing and Proposed Architectures

Timing Metric (ns)	Existing Architecture	Proposed Architecture	Improvement (%)
Worst Setup Delay	23.909	16.517	30.91
Logic Delay	9.632	4.307	55.28
Net Delay	14.277	12.210	14.48
Logic Levels	29	3	89.66
Routes	23	2	91.30

Table 5 presents the hold timing comparison between the existing and proposed architectures. The minimum hold delay decreases from 0.430 ns in the existing design to 0.400 ns in the proposed architecture, resulting in a 6.98% improvement. The logic delay is reduced from 0.193 ns to 0.155 ns, achieving a 32.12% improvement, while the net delay decreases from 0.237 ns to 0.145 ns, corresponding to a 38.82% reduction. In addition, the fanout value is significantly lowered from 52 to 2, providing a remarkable 96.15% improvement, which reduces signal distribution overhead and improves timing stability. These results demonstrate that the proposed architecture maintains reliable hold-time performance while significantly reducing routing

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

complexity and fanout-related timing constraints, thereby ensuring robust high-speed operation.

Table 5 Hold Delay Comparison Between Existing and Proposed Architectures

Timing Metric (ns)	Existing Architecture	Proposed Architecture	Improvement (%)
Minimum Hold Delay	0.430	0.400	6.98
Logic Delay	0.193	0.155	32.12
Net Delay	0.237	0.145	38.82
Fanout	52	2	96.15

5. Conclusion

The proposed Hardware-Accelerated Data Readout Platform for MDA successfully addresses the major challenges associated with conventional DNA data storage recovery systems through the integration of PIRT, BQAF, OLC, Min-Max Voting with Scheduling, and HSCC. The architecture effectively improves sequence identification accuracy, eliminates low-confidence sequencing reads at an early stage, reduces memory traffic through lightweight compression, and performs efficient error correction using a symbolic consensus-based decoding framework. The implementation results demonstrate substantial improvements over the existing architecture, including a reduction in LUT utilization from 2,206 to 215 (90.25%), a decrease in Flip-Flop utilization from 272 to 48 (82.35%), a reduction in total power consumption from 103.955 W to 35.347 W (66.00%), and a decrease in worst-case setup delay from 23.909 ns to 16.517 ns (30.91%). Furthermore, the proposed architecture significantly reduces routing complexity, logic depth, and fanout overhead while maintaining reliable timing performance. These improvements collectively enable a high-throughput, low-power, area-efficient, and scalable molecular data recovery platform suitable for future DNA-based archival storage applications and advanced VLSI implementations.

References

[1]. Pindoo, Irfan Ahmad, and Suman Lata Tripathi. "Biochips and Lab-on-a-Chip Systems: VLSI Applications in Medical Diagnostics." *Biochip Design and Health Informatics Using IoT and SDN* (2026): 11-33.

[2]. Moorthii, Chithambara, Anmol Singla, and Manan Suri. "DNA-CIM: DNA Sequence Analysis Using RRAM-Based Compute In-Memory Accelerator." In *2025 38th International Conference on VLSI Design and 2024 23rd International Conference on Embedded Systems (VLSID)*, pp. 314-319. IEEE, 2025.

[3]. Savari, Amin, Ali Mahani, Ebrahim Ghafar-Zadeh, and Sebastian Magierowski. "A RISC-V Accelerator for Sequence Decoding in Mobile DNA Sequencers." *IEEE Transactions on Very Large Scale Integration (VLSI) Systems* 34, no. 2 (2025): 606-619.

[4]. Liu, Dong Dong, Shaun Wei Yang Ngang, and Lih Feng Cheow. "in situ Transformation of Information Into DNA Storage With Microfluidic Very Large-Scale

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

Integration Platform." *Small* 21, no. 26 (2025): 2412225.

[5]. Shanthi, G., Sk Shoukath Vali, Uppala Bhargava Sai, K. L. V. Ramana Kumari, S. Naga Leela, A. Ramesh Kumar, and K. Srinivas Rao. "Enhanced Reliability Architecture for VLSI Systems Using the IRMC Algorithm for Error Detection and Correction." In *Next-Generation High-Speed Electronics and Optoelectronics: Volume 2*, pp. 121-138. Singapore: Springer Nature Singapore, 2026.

[6]. Allawi, Salah Taha, and Yasamin Hamza Alagrash. "A New Image Encryption Method Combining the DNA Coding and 4D Chaotic Maps." *International Journal of Intelligent Engineering & Systems* 18, no. 1 (2025).

[7]. Jiang, Chao, Xiong Zhang, and Xiaoqin Zhang. "Image encryption scheme based on 2D Ikeda–Zaslavski chaotic map and DNA coding." In *International Conference on Pattern Recognition and Image Analysis (PRIA 2025)*, vol. 14172, pp. 94-103. SPIE, 2026.

[8]. Emin, Berkay, Yeliz Karaca, Yusuf Alaca, and Akif Akgül. "Biomedical Data Security Using a Hybrid Encryption Approach Based on Fractional-Order Hyper-Chaotic Systems and DNA Coding." *IFAC-PapersOnLine* 59, no. 37 (2025): 215-219.

[9]. Sreelakshmi, K., Dhage, P. U., Lalitha, M., Raj, S. A., Babu, K. A., & Manjunath, T. C. (2025, September). Contribution Title Secure VLSI System Design Through Verilog-Based Simulation and RTL Realization of SPECK Algorithm for Enhanced Cryptographic Reliability. In *International Conference on Advancements in Smart Computing and Information Security* (pp. 319-327). Cham: Springer Nature Switzerland.

[10]. Rashidi, Bahram. "A High Secure Image Encryption Method Based on an Efficient and Low-Cost Cryptographic System." *Security and Privacy* 8, no. 6 (2025): e70130.

[11]. Mukherji, Prachi, Seema Rajput, and Vaishnavi Mudaliar. "A novel accelerated sparse Support Vector Machine (AS-SVM) algorithm for binary classification of DNA sequences." *Franklin Open* (2025): 100427.

[12]. Dwivedi, Advait, Snehak Tarun Tiu, Heman Prasad, and Pulkit Singh. "Design of Efficient VLSI Architecture for Lightweight Cryptographic Algorithms." In *2026 International Conference on Recent Advances in Electrical, Electronics, Ubiquitous Communication, and Computational Intelligence (RAEEUCCI)*, pp. 1-6. IEEE, 2026.

[13]. Sankar, Aasish, Naysha Kumari, Shreyas Vignesh, and Dhanashree Bhate. "ECG Abnormality Detection Using Xilinx Artix-7 Basys 3 FPGA: Optimized Diagnosis with Golomb-Rice Coding." In *International Conference on Information and Communication Technology for*

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

Intelligent Systems, pp. 569-578. Singapore: Springer Nature Singapore, 2025.