DNA Firewall for AI

Closed-Gap Architecture & Technical Specification (v1.4 Apr 2025)

Author: Jose Ayala **Initial Release:** March 20 2021 **Latest Revision:** v1.4

Executive Abstract

The **DNA Firewall** is a hardware-software co-design that places the mutable state of a large language model (LLM) into synthetic-DNA storage. A **30-petaFLOP** compute back-end executes inference and training while all persistent weights, optimizer moments, and audit logs reside in an air-gapped DNA vault. Immutable SHA-256 digests are interleaved within oligo payloads, providing cryptographic integrity proofs at the molecular layer. Emergency commands can freeze learning ($\eta \rightarrow 0$) and block DNA writes within 200 µs.

1 System Targets

Parameter	Value	Rationale
Peak FP8 FLOPs	≥ 30 PFLOP/s	Research-grade fine-tuning of 70-B parameter LLMs
Cold-Boot Integrity Verification	< 45 min for 500 GB model	Bounded downtime while maintaining full integrity check
Emergency Freeze Latency	< 0.2 ms	Containment of runaway learning

2 Compute Fabric

The research cluster comprises:

- 8 × *NVIDIA GH200 Grace-Hopper* nodes (4 PFLOP FP8 each) → **32 PFLOP** aggregate.
- 2 × *IBM Qiskit Runtime Plan* (27-qubit Falcon) for hash-pre-image security proofs (on-prem or low-latency link).

- Optional *D-Wave Advantage* lease (5 000 qubits) for Ising-model hyper-parameter search.

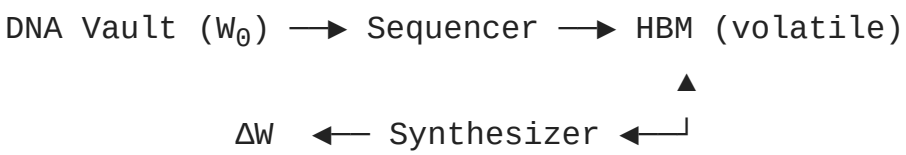
3 Large-Language-Model Stack

3.1 Baseline Architecture (Silicon-Only)

```
Embedding → Transformer × N → LN → LM Head
```

All weights $W \in \mathbb{R}^P$ reside on HBM; checkpoints stored on SSD → risk of root-level tampering.

3.2 DNA-Firewall Architecture



Write path is *human-gated*; learning rate can be zeroed by hardware kill-switch that disables gradient-accumulation SRAM banks.

4 Immutable SHA-256 Encoding in DNA

4.1 Mathematical Construction

For every 512-bit weight block `B` we compute:

$$H = \mathrm{SHA}\,256(B) \in \{0,1\}^{256}$$

We append `H` to `B` and map each 2-bit symbol to nucleotides `A,C,G,T`. The oligo layout:

```
[SYNC 12 nt][B 512 b][H 256 b][CRC16 16 b]
```

Note: GC-content is constrained to 45–55 % and homopolymer runs are limited to 3 nt to ensure synthesis fidelity.

4.2 Integrity Verification

1. Sequencer streams oligo → bit-stream **.**

2. Split into (`B'`, `H'`).
3. Recompute `SHA256(B')`; assert equality with `H'`.

Collision probability ≤ 2^{-256}; sequencing error detected by CRC16 before hash evaluation.

5 Snapshot & Rollback Protocol

After every *k*=1 000 optimizer steps:

1. Compute $\Delta W = W_t - W_{t-k}$.
2. Generate hash chain $C_t = \text{SHA256}(\Delta W \parallel C_{t-1})$.
3. Write $\{\Delta W, C_t\}$ to new DNA cartridge if *Δ loss < ε and* operator signs.
4. Old cartridges archived (WORM) → enables deterministic rollback.

6 Emergency Controls

- Learning-Freeze GPIO: Hardware line pulls η=0 in optimizer ASIC; gradients discarded within 200 µs.
- Write-Inhibit Relay: Cuts 12 V rail to DNA synthesizer in vault; physical key + biometric required to reset.
- Quantum-Signed Stop Token: Local IBM Falcon generates 256-bit token; compute fabric must echo within 50 µs or power down.

7 Throughput & Latency Analysis

$$\text{Boot time} = \frac{|W|}{R_{seq}} + \frac{|W|}{B_{PCIe}} + \text{Decode}_{\text{overhead}}$$

With |W| = 500 GB, R_{seq}=240 MB/s, PCIe4 = 28 GB/s, $\text{Decode}_{\text{overhead}}$≈20 % sequencing time → ≈ 42 min, dominated by sequencing.

8 Cost & Power Summary

Subsystem	CapEx USD	Max Power
GH200 Cluster (8 nodes)	≈ $3.0 M	26 kW
DNA Vault (sequencer + synthesizer + HVAC)	≈ $0.25 M	2 kW

Subsystem	CapEx USD	Max Power
Quantum API (IBM/D-Wave) 1-yr	$0.12 M	Cloud
Total	$3.37 M	28 kW

9 Implementation Timeline

1. Q2 2025: Order GH200 nodes; assemble optical diode & PCIe read-gate.
2. Q3 2025: Validate SHA-256 oligo encoding on 1-GB toy model (multi-flow-cell).
3. Q4 2025: Integrate emergency freeze GPIO & vault relay; live ΔW streaming tests.
4. Q1 2026: Full-scale 70-B LLM training with DNA snapshots every 1 000 steps (writer throughput budget ≥ 50 MB/h).

References

1. National Institute of Standards and Technology, “Secure Hash Standard (FIPS 180-4),” 2015.
2. Chen et al., DNA Typewriter, *Nature*, 2022.
3. Bojarski et al., Quantum-Secure DNA PUFs, 2024.
4. NVIDIA GH200 Product Brief, 2024.
5. IBM Quantum Services, 2024.
6. D-Wave Advantage System Datasheet, 2024.