Reducing Data Erasure Latency via Storage-side Encryption

Yu-Yu Wang, Chih-Wei Ko
Supervised by Chien-Chung Ho
Department of Computer Science and Information Engineering, National Cheng Kung University

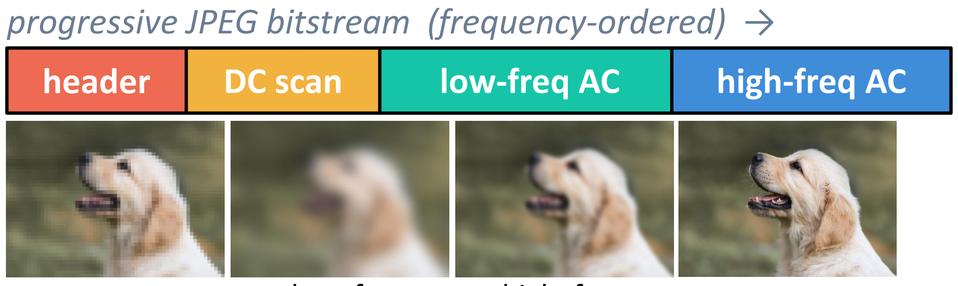
Introduction

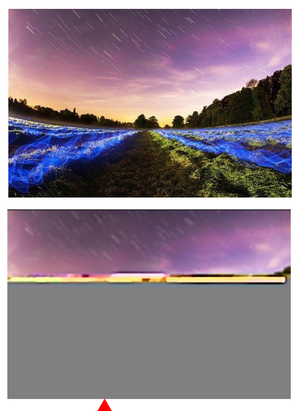
*Logical deletion leaves SSD data forensically recoverable, and overwriting is **O(n)**. We propose a format-aware selective crypto-erase: one key per file, so deleting a file means wiping a single key — an **O(1)** operation, verified against a worst-case forensic adversary.*

❖Why Progressive JPEG

Progressive JPEG is widely deployed across web and image-delivery pipelines, where its frequency-ordered scans render images coarse-to-fine. Because recognizability is concentrated in the leading low-frequency scans, encrypting this compact, locatable region renders the entire image unrecoverable — making a prevalent format also an efficient target for selective crypto-erase.

progressive JPEG bitstream (frequency-ordered) →

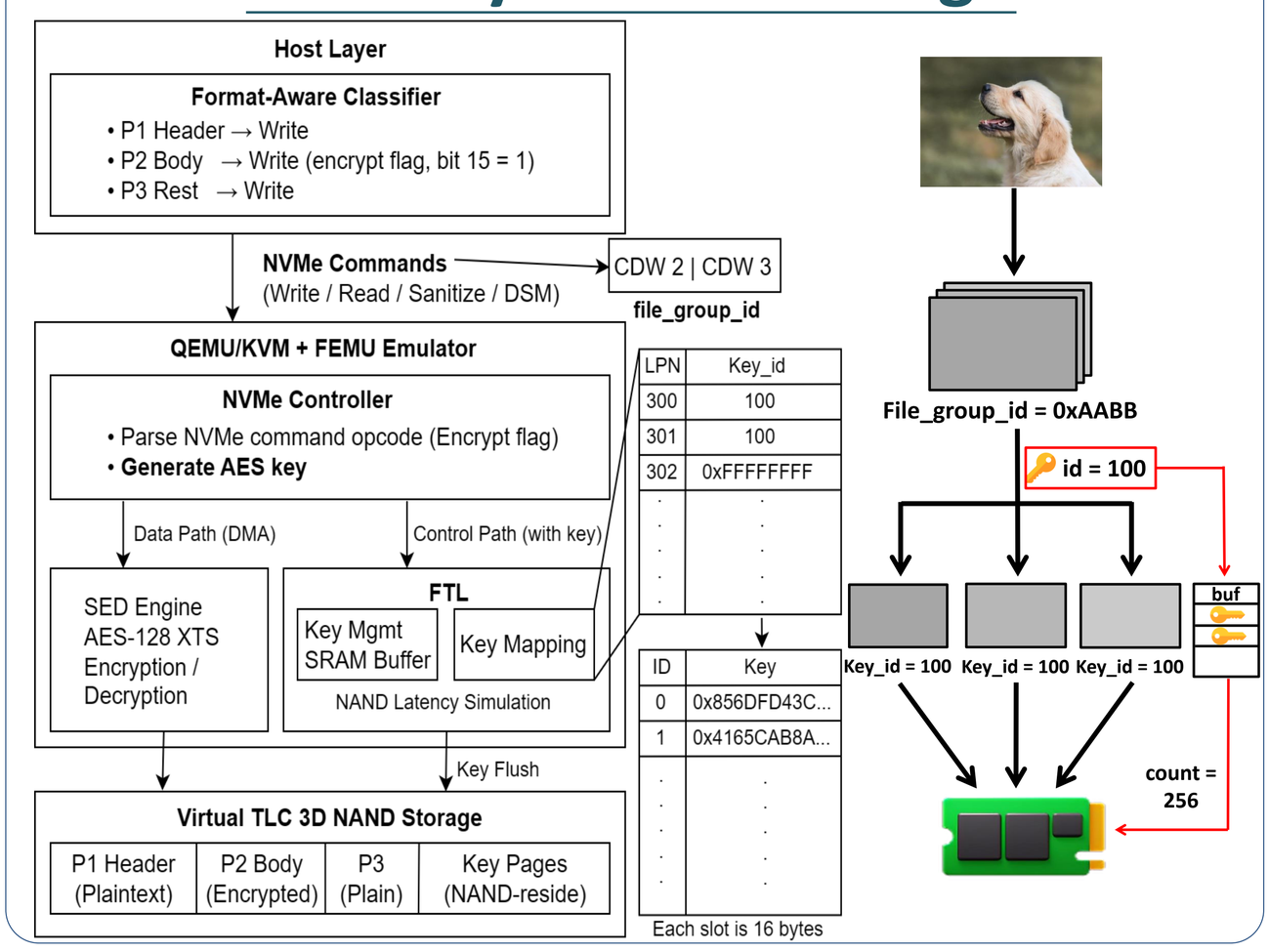




*The unit of SSD Write is **4KB**.*

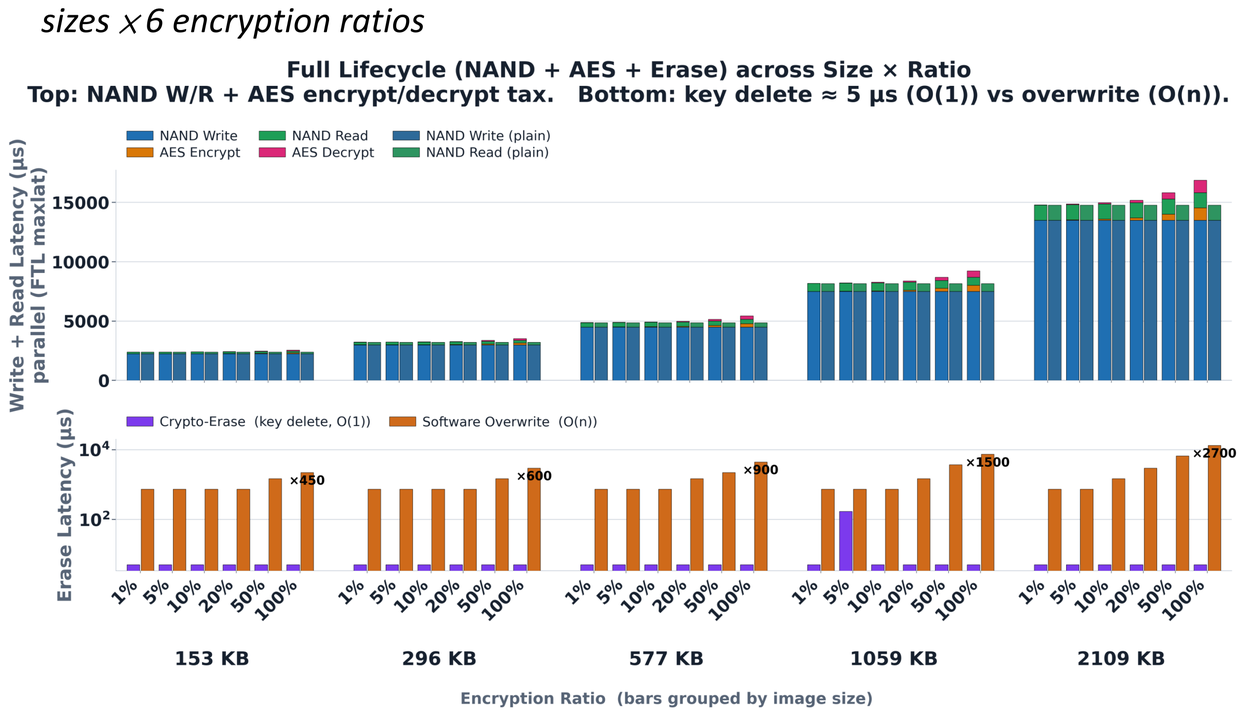
DC partially retained!

One-Key-Per-File Design



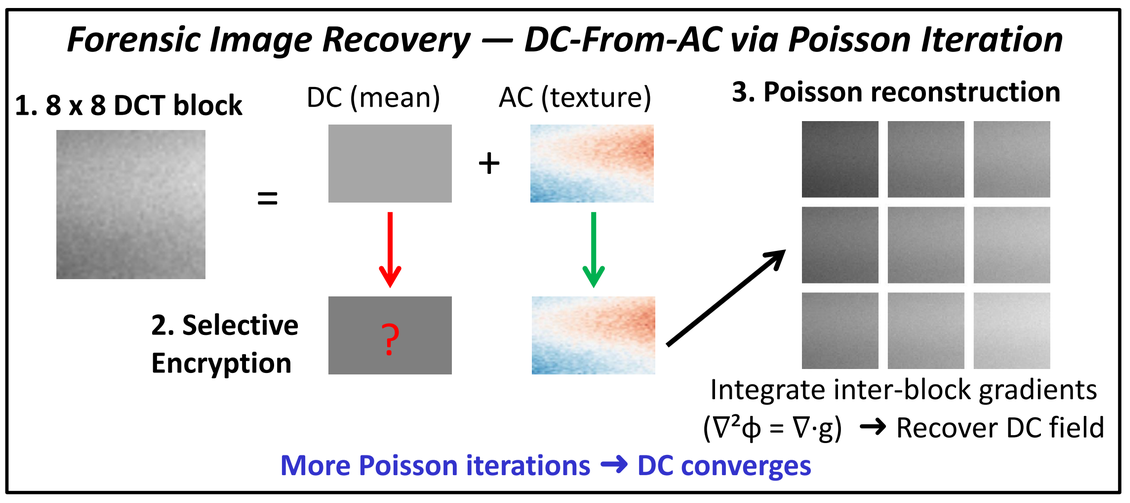
Latency Benchmark

Full Write → Read → Erase lifecycle measured on FEMU, swept across 5 image sizes ⨉ 6 encryption ratios



Deletion cost stays constant while data grows

How Little Can We Encrypt?



Parameter	Value (Rigorous)
Secure cutoff — image	**forensic correlation (r) < 0.30**
Poisson iterations	**1500**
Max image side	**2048 px**
Encryption-ratio sweep	**1 – 100 %**
Destruction target	**≥ 90 % of worst-case images**
Text-secure cutoff — documents only	**OCR char-recovery < 0.10**

Destruction = similiarity of recovered image to original

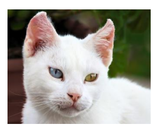


Image (photo)

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

Destruction = how much text OCR can still read (worst-case field)

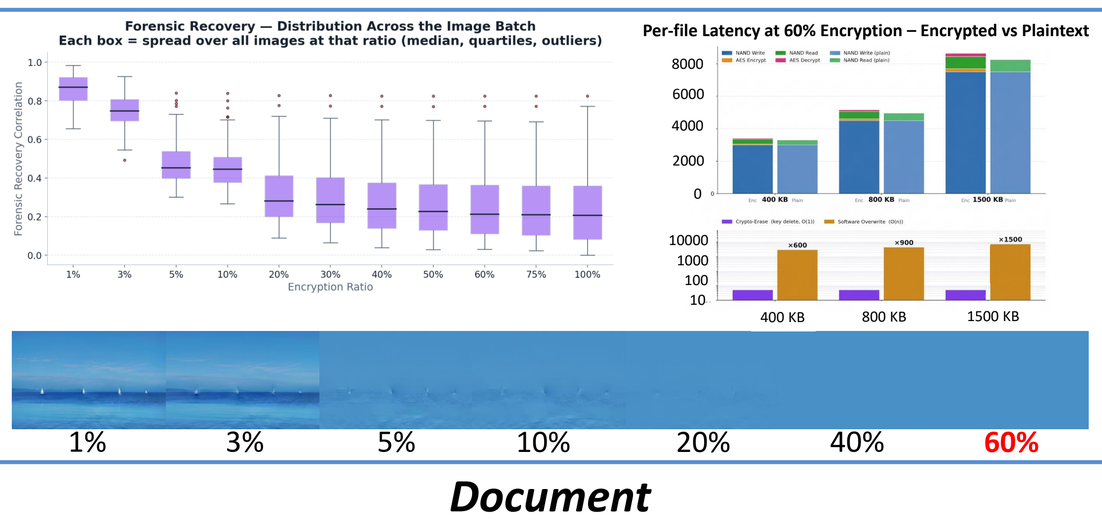
Document

Worst-case: secure only if all 4 fields (**Amount·Date·Account·Name**) < 0.10.

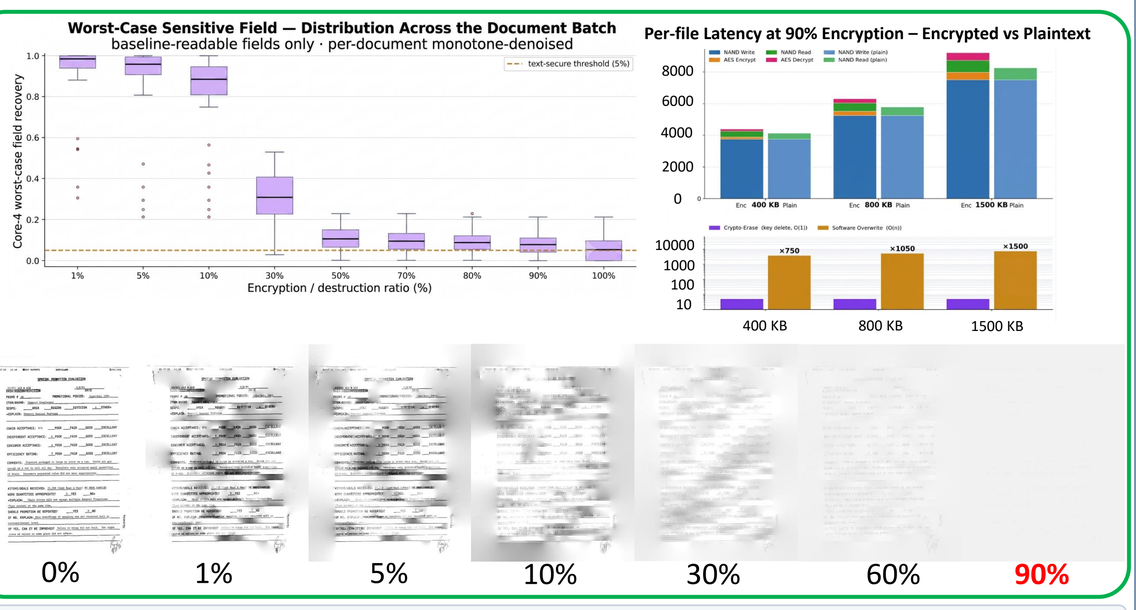
$$\text{Field}_{\text{worst}} = \max_{f \in \{\text{Amount, Date, Account, Name}\}} \left[1 - \frac{\text{Lev}(GT_f, OCR_f)}{\max(|GT_f|, |OCR_f|)} \right] < 0.10$$

Experimental Results

Image (photo)



Document



SECURE ENCRYPTION RATIO

Photos	**60%**	*forensic correlation < 0.30*
Documents	**90%**	*OCR character-recovery < 0.10*

Same O(1) crypto-erase cost for both. The higher cost of securing text is paid only in encryption coverage — never in deletion latency.

Demo Videos