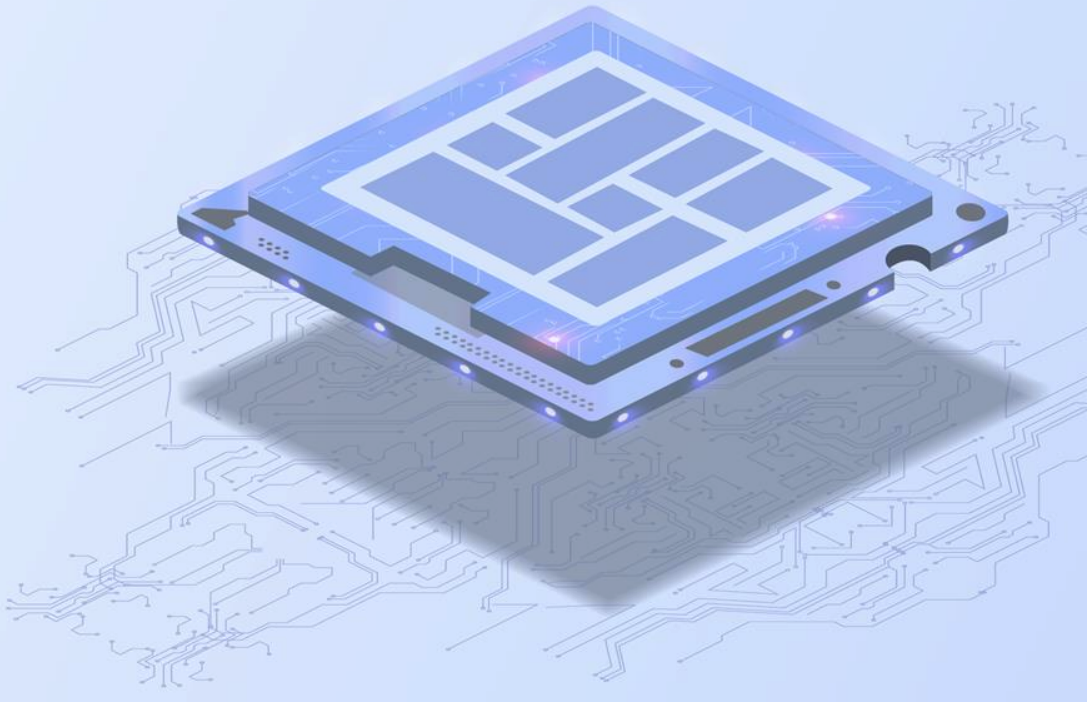
# EPA ECC : Error-Pattern-Aligned ECC for HBM2E

**Kiheon Kwon, Dongwhee Kim, Soyoung Park and Jungrae Kim**

Sungkyunkwan University



# Outline

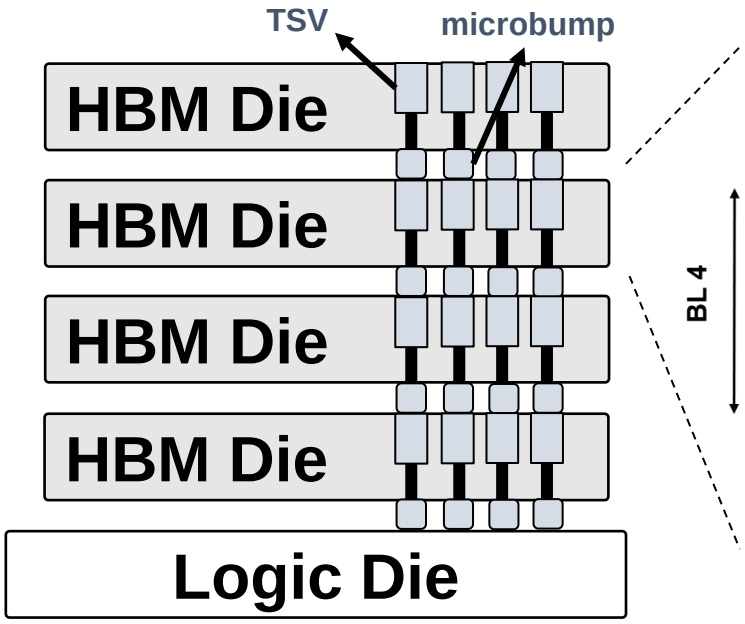
- **Modern HBM systems and challenges**
  - **Reliability**
    - Ex) Soft errors, retention errors, weak cells, wear-out.

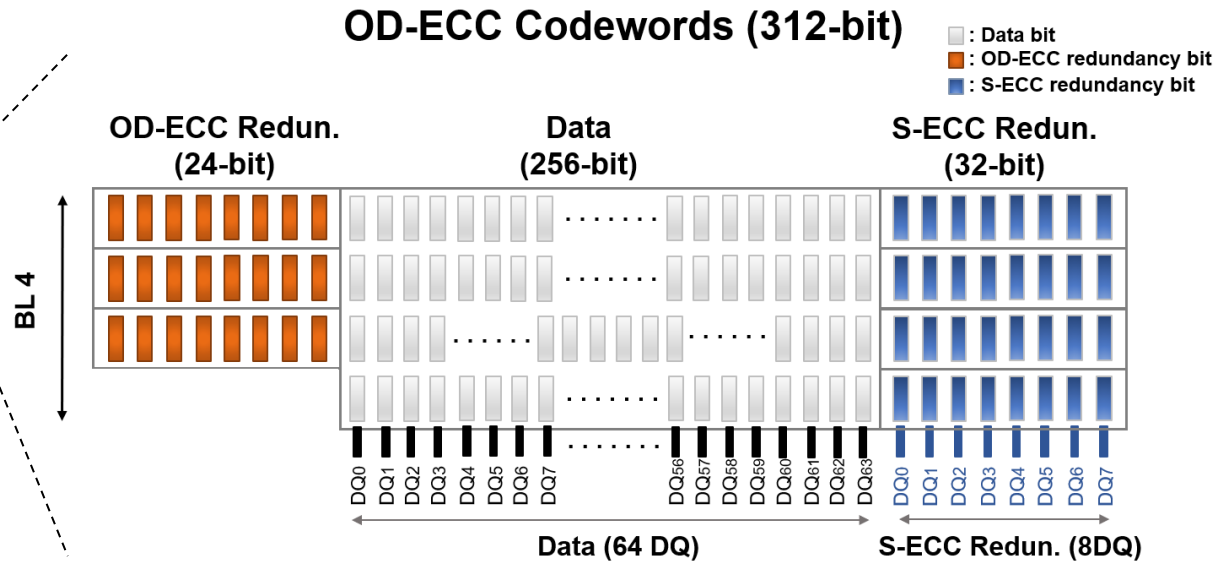
- **On-Die Error Correction Code (On-Die ECC) in memory systems**
  - **Pros : Increasing reliability**
  - **Cons : Decreasing costs efficiency**

- **We propose a promising solution, On-Die ECC with high reliability, focusing on soft errors.**

## HBM

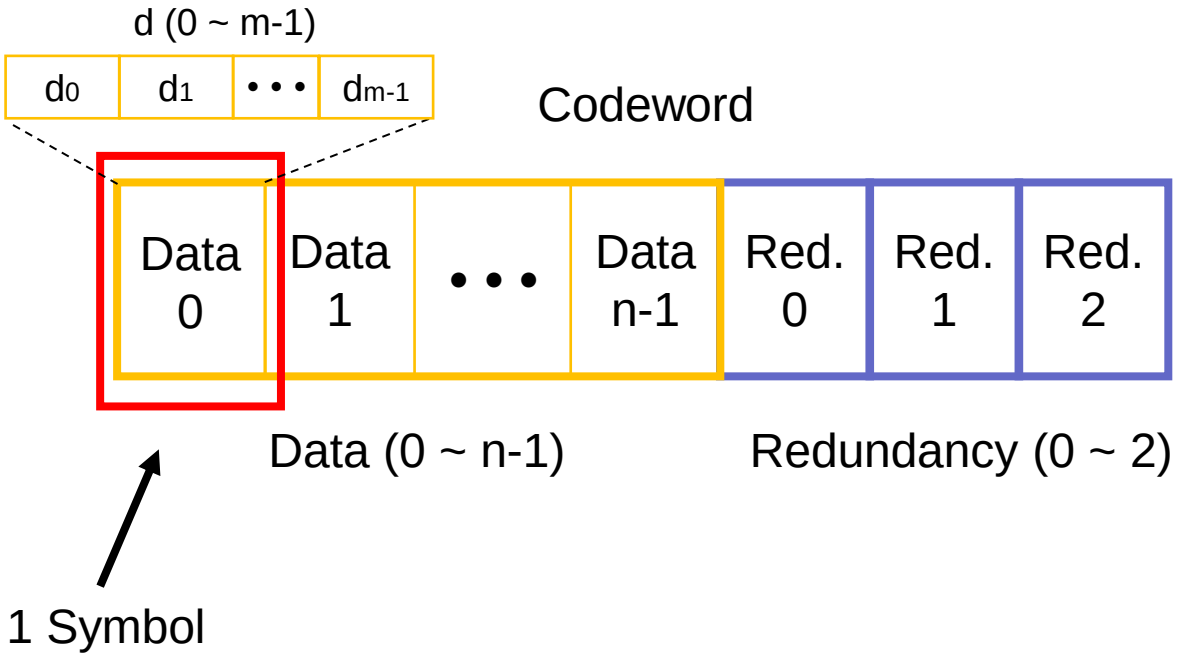
- HBM2E has introduced On-Die ECC.
- HBM2E Memory access granularity : 32B (256bit)
  - Data (256bit), System-ECC Redundancy (32bit), On-Die Redundancy (24bit)
  - Burst length (4)



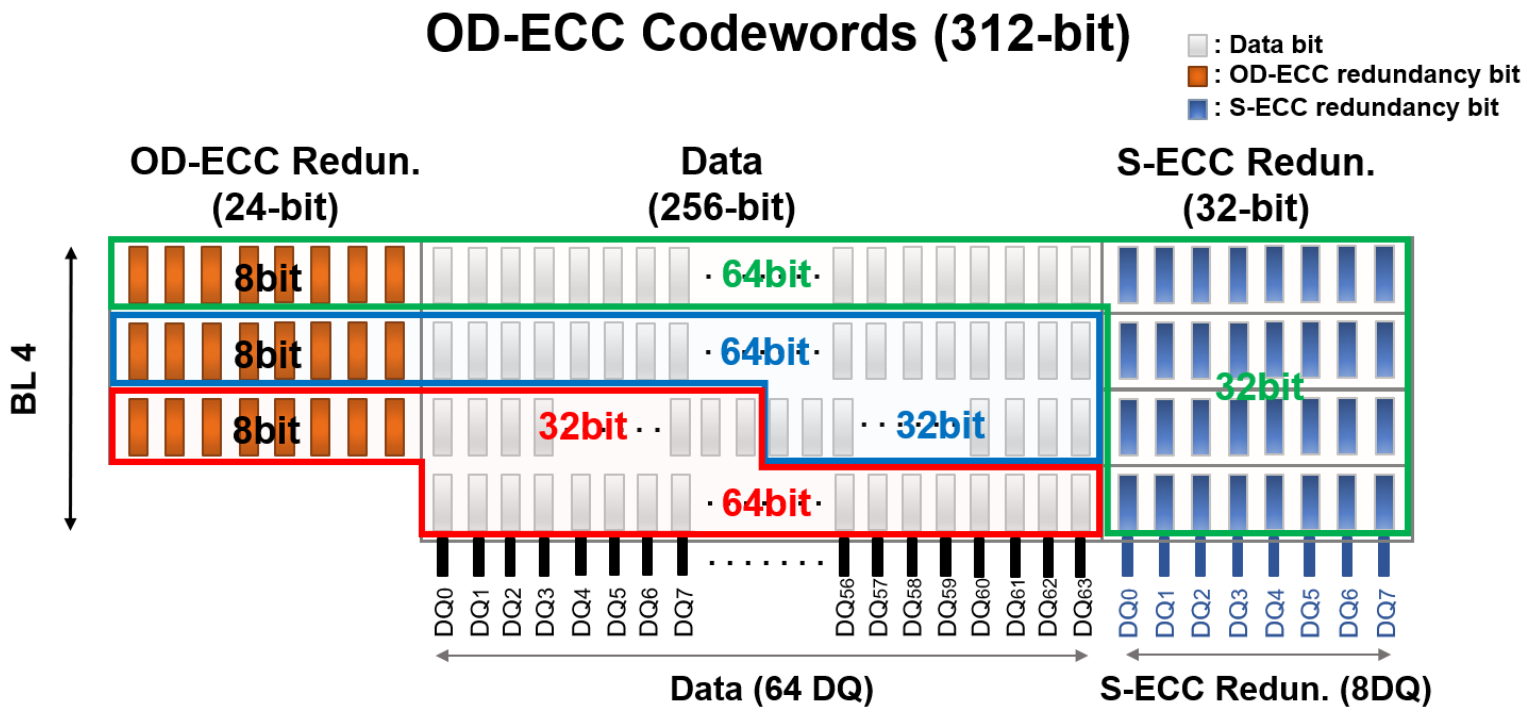
## Reed-Solomon(RS) Codes

- A class of non-binary ECC
- Can correct or detect clustered errors.
- Symbol-based ECC



## HBM2E On-Die ECC

- (104, 96) Single Error Correction-Double Error Detection (SEC-DED) * 3 [1]

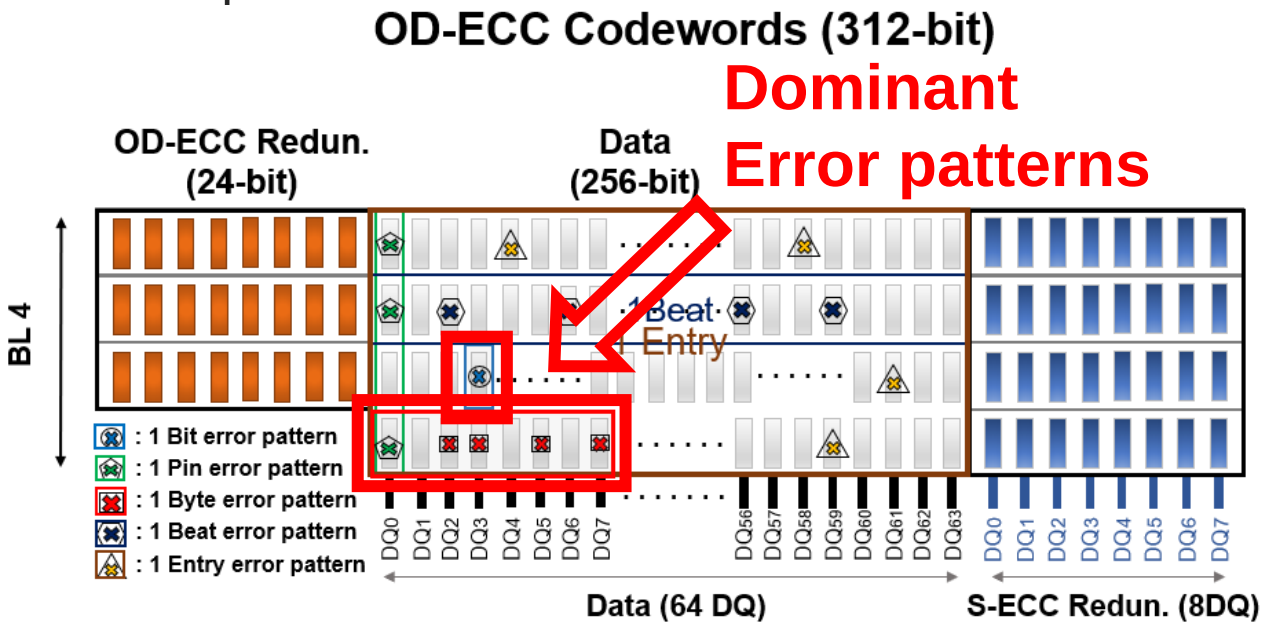


## Soft Error patterns

- Dominant Error patterns (96.54%)
  - 1 bit error pattern
  - 1 byte error pattern
- The existing HBM On-Die ECC is vulnerable for dominant error patterns.

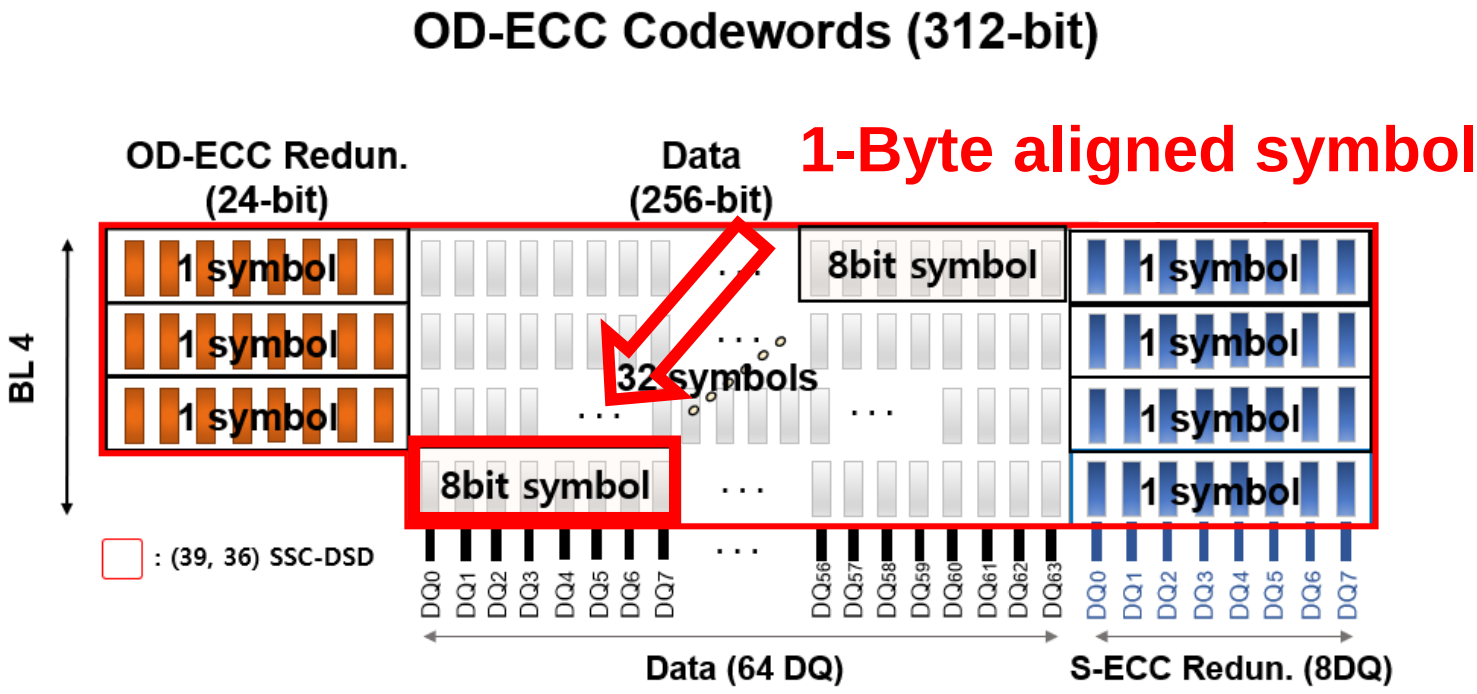
[ Soft Error Patterns. [2] ]

| Error pattern | Number of possible error bits | Probability |
|---|---|---|
| 1 Bit | 1 | 73.98 % |
| 1 Byte | 2 – 8 | 22.56 % |
| 1 Pin | 2 – 4 | 0.19 % |
| 2 Bits | 2 | 0.11 % |
| 3 Bits | 3 | 0.03 % |
| 1 Beat | 4 – 64 | 0.90 % |
| 1 Entry | 4 – 256 | 2.23 % |



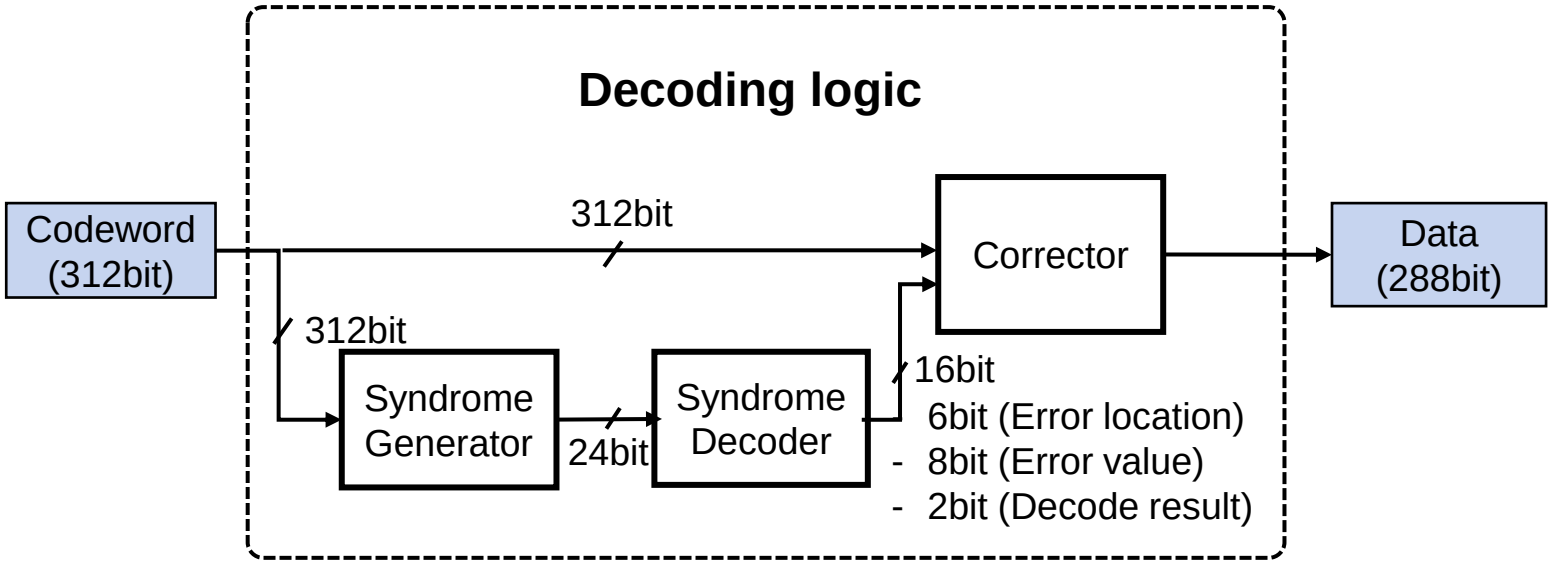
## Code Layout

- EPA ECC : (39, 36) shortened RS codes over Galois Field ($2^8$).
  - Correction capability : Single Symbol Correction.
  - Detection capability : Double Symbol Detection.



## Decoder Implementation

- **Syndrome Generator** generates three 8-bit syndromes, using the parity-check matrix of RS code.
- **Syndrome Decoder** calculates the error location, error value, and decode result using division operations over GF($2^8$).
- **Corrector** decides whether to correct the error (CE, NE, DUE).



## Evaluation Setup – Reliability, Hardware Overheads.

- **Baseline [1]** : (104, 96) SEC-DED * 3
- **EPA ECC** : (39, 36) 8bit RS code (SSC-DSD, Single Symbol Correction-Double Symbol Detection)

- **Simulation**
  - Randomly error injection to a memory access block.
  - 100 Million Monte Carlo simulations for each error patterns.
- **Metric**
  - **Correctable Error (CE)**
  - **Detectable-but-Uncorrectable Error (DUE)**
  - **Silent Data Corruption (SDC)**
  - **Latency / Area / Power**

## Reliability

- Can correct dominant error patterns.
- Reduces DUE (4.6x).
- Reduces SDC (4980x).

| Error Scenarios | ECC result | Baseline [1] | EPA ECC |
|---|---|---|---|
| **Overall (%)** | CE | 74.180 | 96.540 |
| | DUE | 15.860 | 3.458 |
| | SDC | 9.960 | 0.002 |

| Error Scenarios | ECC result | Baseline [1] | EPA ECC |
|---|---|---|---|
| **1 Bit (%)** | CE | 100 | CE : 100 |
| | DUE | 0 | |
| | SDC | 0 | |
| **1 Byte (%)** | CE | 0 | |
| | DUE | 57.897 | |
| | SDC | 42.103 | |
| **1 Pin (%)** | CE | 63.703 | DUE : 100 |
| | DUE | 36.297 | |
| | SDC | 0 | |
| **2 Bits (%)** | CE | 68.285 | |
| | DUE | 31.715 | |
| | SDC | 0 | |
| **3 Bits (%)** | CE | 22.445 | 0 |
| | DUE | 68.774 | 99.773 |
| | SDC | 8.781 | 0.227 |
| **1 Beat (%)** | CE | 0 | 0 |
| | DUE | 70.639 | 99.940 |
| | SDC | 29.361 | 0.060 |
| **1 Entry (%)** | CE | 0 | 0 |
| | DUE | 91.731 | 99.94 |
| | SDC | 8.69 | 0.06 |

## Hardware overheads

- Comparison of Hardware overheads of EPA ECC over Baseline [1].
  - Increasing decoding latency
  - Increasing area
  - Decreasing decoding power

  → Overall, **EPA ECC has minimal impacts on DRAM.**

| | Encoder | | Decoder | |
|---|---|---|---|---|
| | Baseline [1] | EPA ECC | Baseline [1] | EPA ECC |
| Latency ($ns$) | 0.30 | 0.39 | 0.43 | 0.66 |
| Area ($\mu m^2$) | 1629 (= 3 × 543) | 2154 | 3948 (= 3 × 1316) | 4251 |
| Power ($mW$) | 5.577 (= 3 × 1.859) | 8.811 | 16.110 (= 3 × 5.370) | 14.109 |

## Evaluation Setup - System Performance

- **Simulator :** GPGPU-Sim with TITAN-V
- **Benchmarks :** Rodinia [3], Parboil [4]
- **Category** : MPKI > 2 (memory-intensive), or not (non-memory-intensive)
- **Target Memory :** HBM2 [5]

[ The GPU configuration (NVIDIA TITAN-V). ]

| | |
|---|---|
| **Core** | 1200MHz, 70 SMs, 64 warps/SM |
| **Cache line** | 128B Line with 4 sectors (32B) |
| **L1 Cache** | 256 entries / fully associative |
| **L2 Cache** | 768 entries / 24-way associative |
| **Memory** | 24 Channels 652.8 GB/s HBM2 |

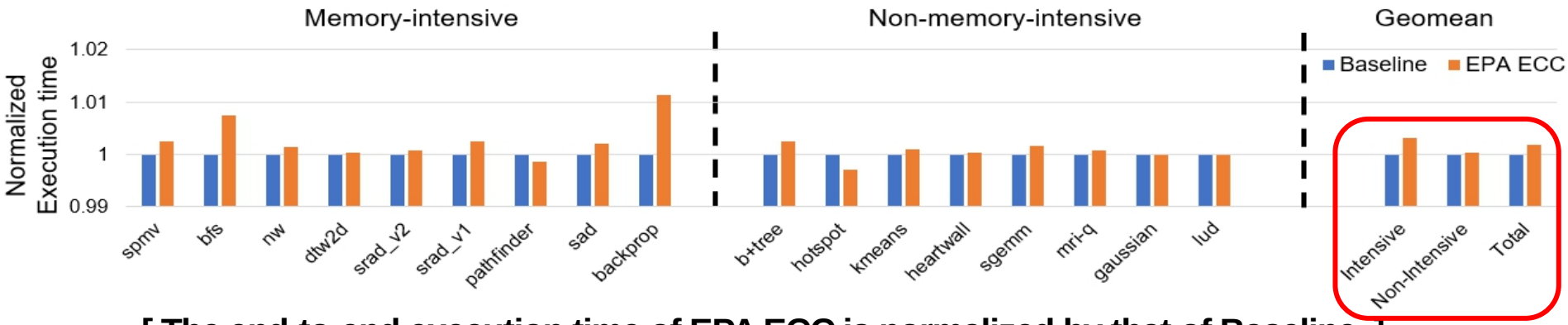
## System Performance

- Add tRL timing parameter (1CK) considering the hardware overheads.

[Comparison of timing parameter of HBM2 [5] in GPGPU-sim. ]

| Timing parameter | Baseline [5] | EPA ECC |
|---|---|---|
| tRCD $(ns)$ | 14 | |
| tRAS $(ns)$ | 33 | |
| tRP $(ns)$ | | +1 CK |
| tRL $(nCK)$ | 12 | 13 |

- Increasing of Execution time is negligible (+0.18%).
  - Memory-Intensive (+1%)
  - Non-Memory-Intensive (+0.04%)

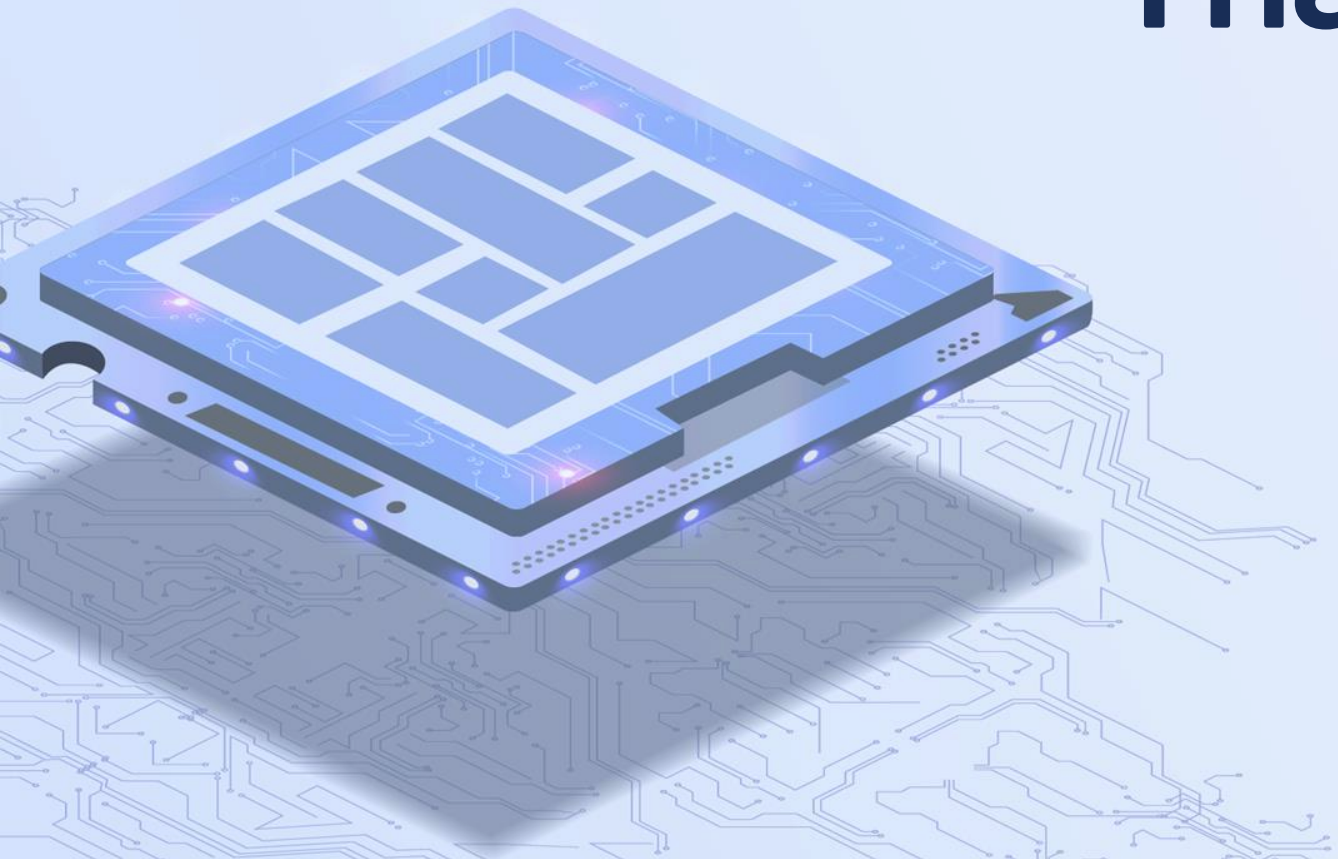



**[ The end-to-end execution time of EPA ECC is normalized by that of Baseline. ]**

# VII Conclusion

- We proposes EPA ECC, an error pattern aligned On-die ECC to improve memory reliability.
- EPA ECC can correct dominant soft errors in HBMs, reducing DUE 4.6x, and SDC 4980x than the existing On-die ECC (SEC-DED) [1].
- EPA ECC is a novel ECC code for HBM2E, improving memory reliability while maintaining acceptable system performance degradation (around 1%).

# Thank you



SUNG KYUN KWAN UNIVERSITY(SKKU) | SAL

# VIII Reference


- [1] K. C. Chun, “A 16-gb 640-gb/s HBM2E DRAM with a data with a data-bus window extension technique and a synergetic on-die ECC scheme”, IEEE JSSC, 2020.
- [2] M. B. Sullivan, “Characterizing and mitigating soft errors in GPU DRAM”, in MICRO, 2021.
- [3] S. Che and Boyer, “Rodinia: A benchmark suite for heterogeneous computing”, in IISWC, 2009.
- [4] J. A. Stratton, “Parboil: A revised benchmark suite for scientific and commercial throughput computing”, Center for Reliable and High-Performance Computing, 2012.
- [5] N. Chatterjee and O’Connor, “Architecting an energy-efficient DRAM system for GPUs”, in HPCA.