Guided Vectorless With Multi Vector Profiling For Memory PDN Convergence

Mohammed Hafiz
Alina Sebastian
Google

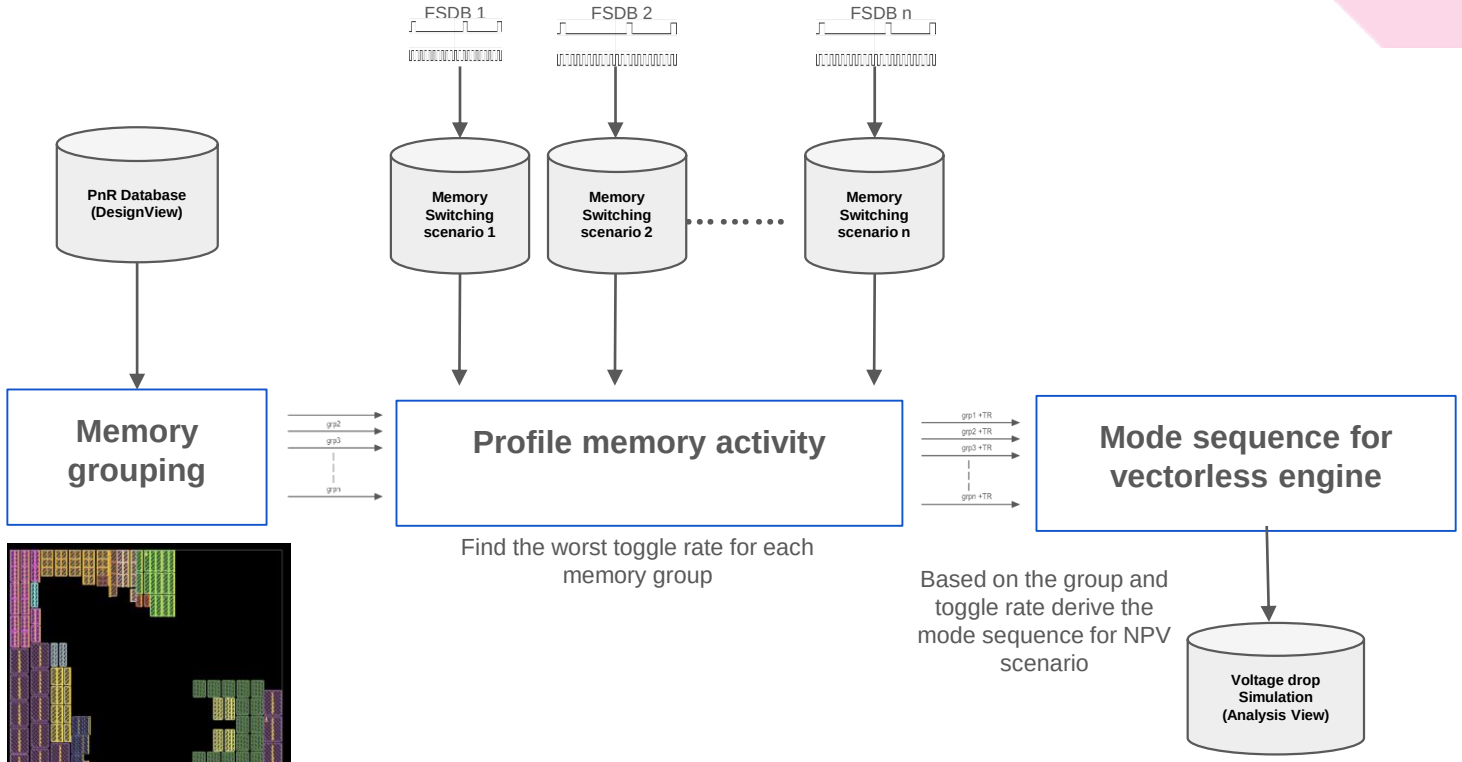
62 THE CHIPS TO SYSTEMS CONFERENCE

SPONSORED BY CEDA® IEEE Council on Electronic Design Automation sigda acm

Motivation

- **The Challenge:** Ensuring robust Memory Power Grid (PG) convergence is critical, yet fraught with challenges:
  - **High Sensitivity:** Memory performance is highly sensitive to timing, DRC, and IR drop.
  - **Pessimistic simulations:** Traditional vectorless simulation with 100% toggling is very pessimistic.
  - **Accuracy vs. Time:** Identifying the true worst-case switching scenario for IR simulation is vital to avoid over-design, but VCD-based simulations are extremely time-consuming (weeks per iteration for µs-ms scenarios).
  - **Reliability Gap in simulation:** Short-duration VCD analysis around peak power windows often lacks sufficient switching coverage, leading to unreliable results.
- **Our Approach:** To address these limitations, we introduce a novel method that profiles multiple long vectors to guide a vectorless engine. This allows us to accurately mimic worst-case scenarios for memories.
- **Key Advantages:**
  - **Reduced Pessimism:** Overcomes the over-conservatism of traditional vectorless (100% toggle rate).
  - **Significant Runtime Improvement:** Drastically cuts down simulation time.
  - **Guaranteed Switching Coverage:** Ensures 100% coverage for profiled scenarios.

Main Idea



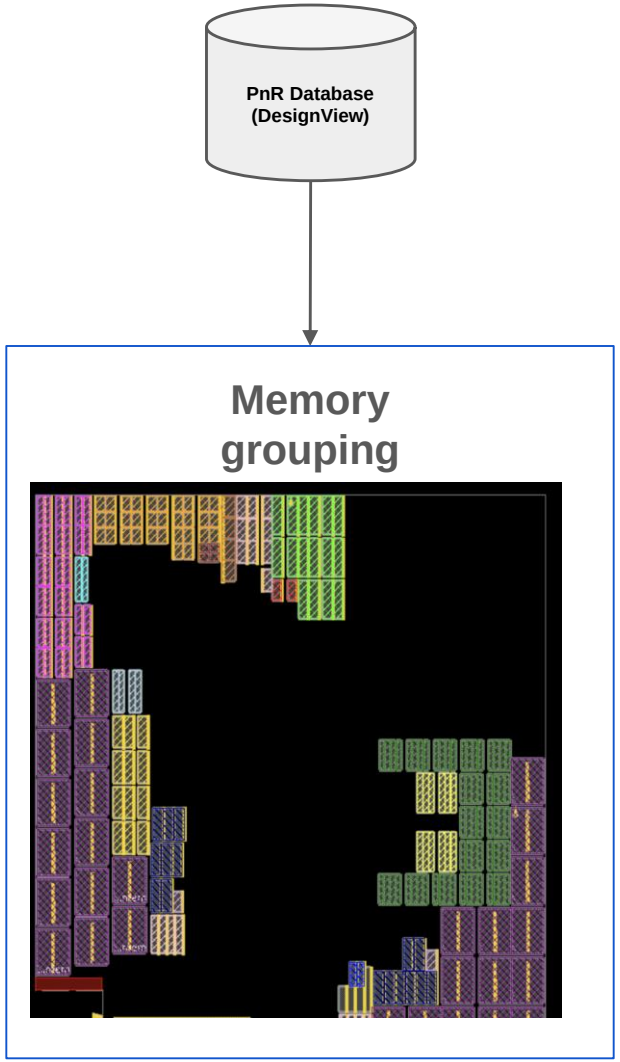
Find the worst toggle rate for each memory group

Based on the group and toggle rate derive the mode sequence for NPV scenario

Memory grouping based on the type and euclidean distance

- Develop an algorithm to analyze the functional vectors which activates memories and generate worst mode sequence data for each memory.
- This flow involves three key steps:
  1. **Memory Grouping**
  2. **Profile memory activity**
  3. **Mode sequence generation**

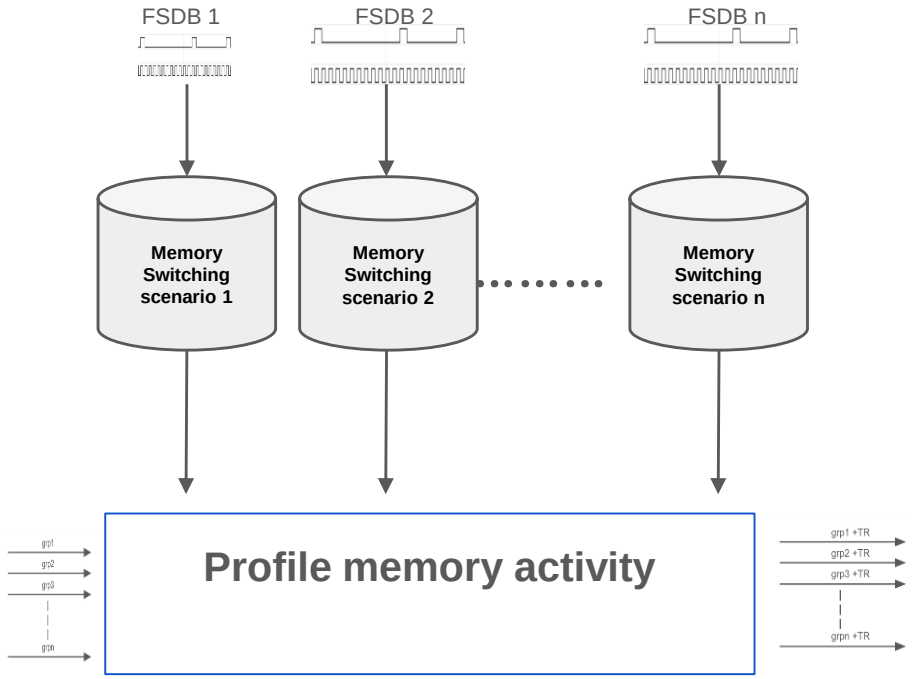
Step1 : Memory Grouping



- Following factors are considered for memory grouping
  - **Cell name**
    - Grouping memories by cell name ensures that memories within the same group have exact same switching state names. It reduces the complexity to identify the switching sequences to meet the desired toggle rate
  - **Euclidean distance**
    - Euclidean distance refers to the geometric distance between the physical locations of memories on the chip.
    - The memories with the same cell but greater than a specific threshold , will be assigned to a different group.

- Memory grouping helps to distribute the worst activity effectively

Step2 : Profiling from FSDBs

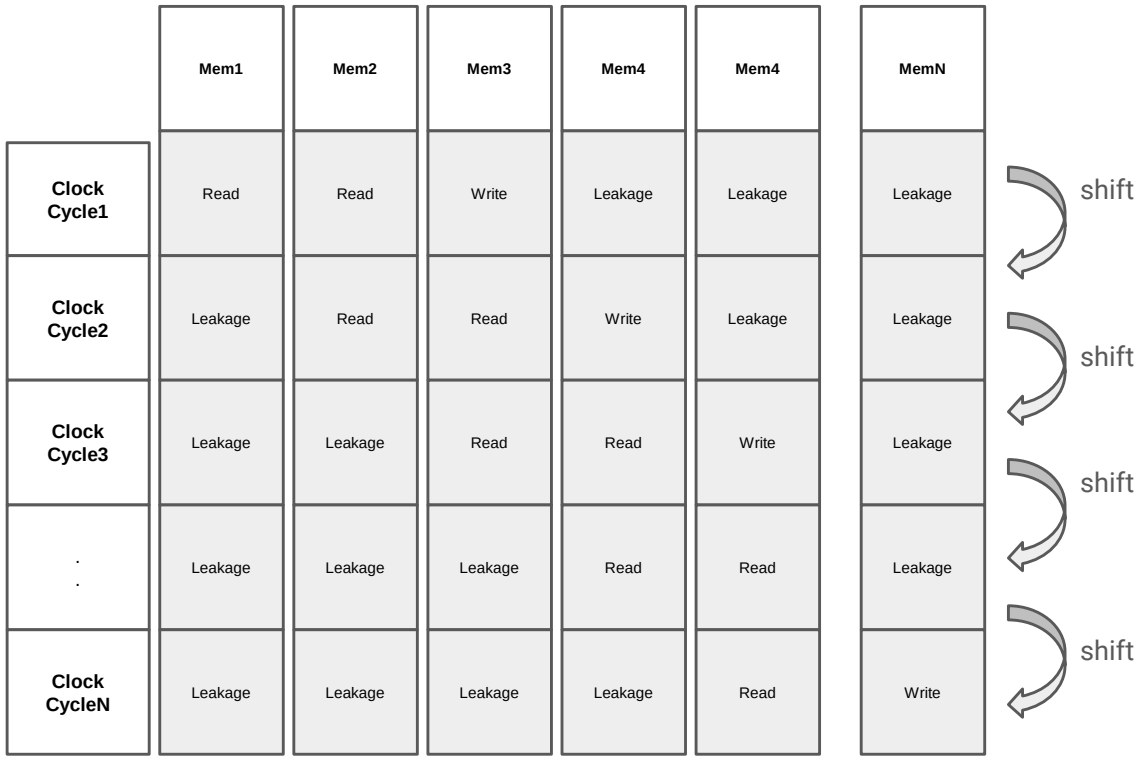


Find the worst toggle rate for each memory group

- **Iterate through FSDBs:**
  - For each memory group, iterate through all the FSDBs for complete duration and analyze every clock cycles.
  - Identify memory activity for every individual clock cycle within the group. This granular approach captures detailed switching behavior of the memories.
- **Identify peak switching cycle:**
  - For each memory group among all the fsdbs, identify the clock cycle with the highest number of memory instances switching state. This identifies the worst-case scenario in terms of memory activity.
- **Calculate toggle rate:**
  - Divide the number of memories switching in the peak switching cycle by the total number of memories within the group. This ratio represents the toggle rate, reflecting the percentage of memories actively switching in the worst-case scenario.

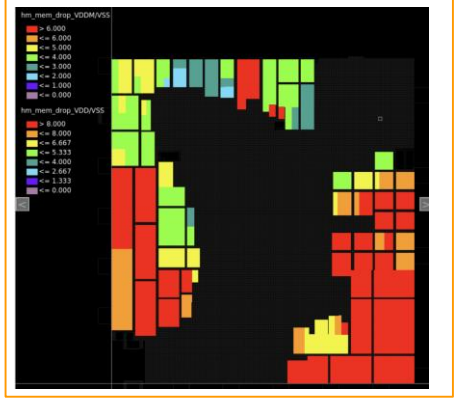
Step 3:Mode sequence Generation

Mode sequence Matrix generation for one memory group



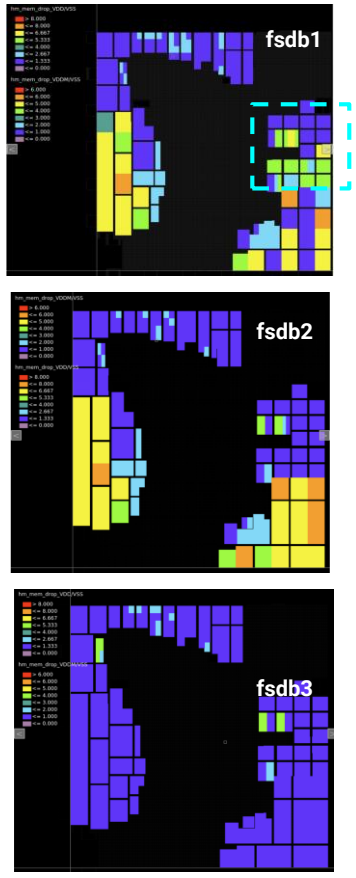
- **Mode sequence generation per group:**
  - **Initial Cycle Pattern:**
    - For each memory group, a switching pattern (e.g. 'Read', 'Write', or 'Leakage' state) is established for the first clock cycle. This pattern is designed to meet the worst-case toggle rate identified in Step 2.
  - **Shifted Patterns for Coverage:**
    - For subsequent clock cycles, this initial pattern is systematically 'shifted' or rotated across the memories within the group.
  - **Ensuring 100% Switching Coverage:**
    - This shifting technique guarantees that over a sequence of cycles, every memory instance within the group gets activated with the worst-case activity defined by the profiled toggle rate. This ensures comprehensive switching coverage for all memories.

Evidence : Memory hotspot Comparison

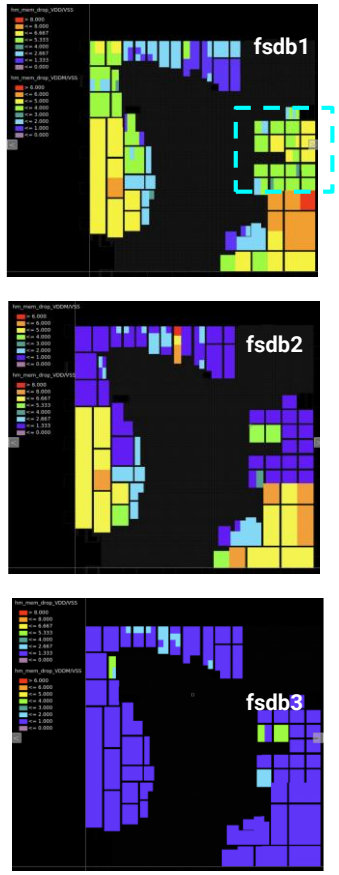


Legacy vectorless run Overly Pessimistic - 100% Toggle Rate

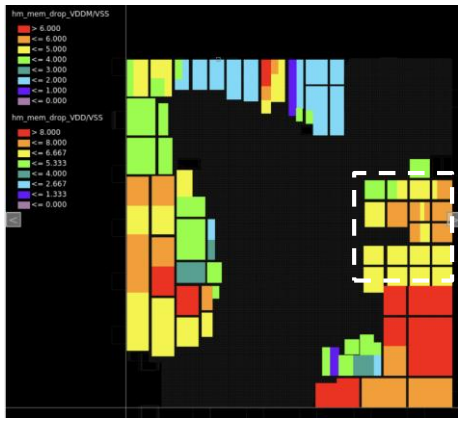
FSDB -50ns



FSDB -1000ns



Multi-vector profiled vectorless



In the above images, highlighted group were showing ~25% toggle rate with shorter fsdb duration(fsdb2), whereas the long fsdb duration were showing ~50% toggle rate

Our multi-vector profiled vectorless approach successfully captures the realistic worst-case scenario (matching the ~50% toggle rate from long FSDBs) efficiently, avoiding both over-pessimism and optimistic underestimation

Short FSDB (e.g., 50ns) (Potentially Optimistic)
Long FSDB (e.g., 1000ns) More Realistic, but Time-Intensive ,not practical deploy for signoff runs.

Summary

- In our paper, we challenge the conventional approach of employing excessively pessimistic switching scenarios for memory blocks where it often results in over-designed power grids.
- Our approach involves profiling multiple vectors, identifying the worst-case scenario, and utilizing this data in a vectorless engine to achieve efficient IR closure.
- By minimizing unnecessary pessimism, maximizing memory switching coverage, and optimizing simulation runtime, our method enhances design convergence, expedites the design process, and ultimately yields more efficient chips.