Power Management in the context of DRD7.2 "Intelligence on detector"

D. Ceresa, F. Crescioli, F. Magniette

Foreword

- This is not a talk by an expert of the field
  - It's a tentative to kickstart a discussion and encourage experts to come forward and share their opinion
  - Please be tolerant for the occasional naïveté
- This is not a talk about power management IPs design or other IPs design techniques to reduce power
  - e.g. is not in the DRD7.2 scope to think about designing low power DACs
- This will be a talk about techniques to mitigate power consumption of the digital logic at global architecture level and especially for the programmable part

Well known facts

- Power consumption is proportional to voltage and the number of state changes per unit of time
  - Clock gating is the architecture-level technique most used in our ASICs
    - Recent synthesis tools are able to automatically infer clock gating
    - Special attention for triplication / SEU resistance
  - Another technique used is power pulsing
    - Exploit the predictable time-structure of activity (e.g. data is produced during spills separated by periods of no data)
    - Power the whole chip only when needed

Dynamic Voltage Scaling

- Regulators to produce internal voltages on-chip are appearing in our designs
  - e.g. RD53 FE chips have an LDO scheme to internally produce VDDA / VDDD
- Voltage could be dynamically selected/trimmed to save power
  - Well known technique in industry
  - Recover the margin in critical path delays as power saving
  - Needs a system to detect errors and find the minimum voltage to ensure operations
  - How does it evolve with radiation damage? Industry do not have this issue

Dynamic Frequency Scaling

- PLLs are typically present in our designs
- Frequency of operations in the data processing parts of our future designs could be scaled to save power
  - Not all events are the same
  - Frequency could be adapted to key quantities of the data (e.g. number of hits per trigger)

Combining techniques

- Dynamic Voltage and Frequency Scaling (DVFS) is a widely used approach in industry for processors or other computation-intensive designs
  - Scale down frequency when lowering the voltage to avoid setup violations
  - Extreme case of "frequency scaling": asynchronous logic
- It is possible also to trade off accuracy (e.g. gating LSBs): Dynamic Voltage, Accuracy and Frequency Scaling (DVAFS)
  - Can we think of contexts in our ASICs where accuracy could be dynamically scaled?
  - This appears to be "natural" in neural networks approaches

Approximate Computing

- Even more aggressive than truncation, but some class of data processing algorithms are resilient to errors and this can exploited to save power
- Voltage / Frequency Over Scaling: what if we can tolerate timing violations in some paths?
  - Need some circuitry to detect and correct/manage errors and also usage of coded data to allow error detection
  - … but we already have this problem with SEE! Error-resilient logic is common in our designs. Can we exploit this also to save power?

Summary & open questions

- Advanced low power techniques focus on saving power by exploiting margin on voltage, frequency, accuracy
  - We build detectors with thousands of ASICs designing to meet the corner cases performances, exploiting the margins can bring to a significant average power reduction
  - Our ASICs evolve with time (radiation damage) often in non-obvious and predictable ways
- These techniques are common in other fields/industry. How much in our field?
- There is a similarity/interplay between dynamic scaling and SEE detection/protection.