Accelerating Design Space Exploration for Reliability with Design Space Pruning

Giovanni Beltrame
Polytechnique Montréal

POLYTECHNIQUE MONTRÉAL

WORLD-CLASS ENGINEERING

MIST

MPSoC 2014 - Relais de Margaux - July 10, 2014

Outline

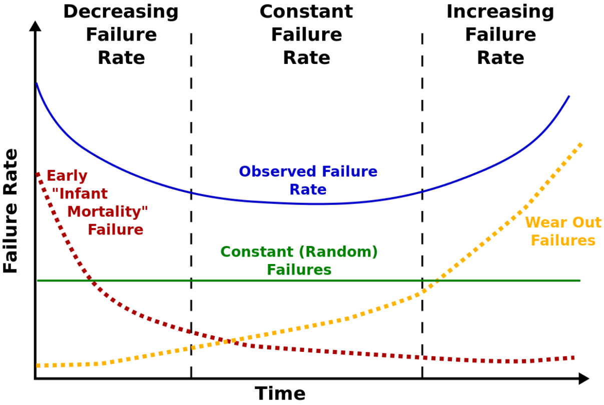
1. Introduction
2. Design Space Exploration
3. Failure Scenario Memoization
4. Symmetry Identification
5. Results
6. Conclusions

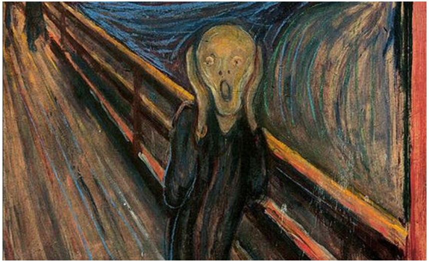
Outline

1. Introduction
2. Design Space Exploration
3. Failure Scenario Memoization
4. Symmetry Identification
5. Results
6. Conclusions

Expected System Lifetime

- Trend: smaller feature sizes, higher operating frequencies
- Thermal issues $\Longrightarrow$ increased failure rate
- Device lifetimes are becoming shorter than market expectations





Reliability estimation

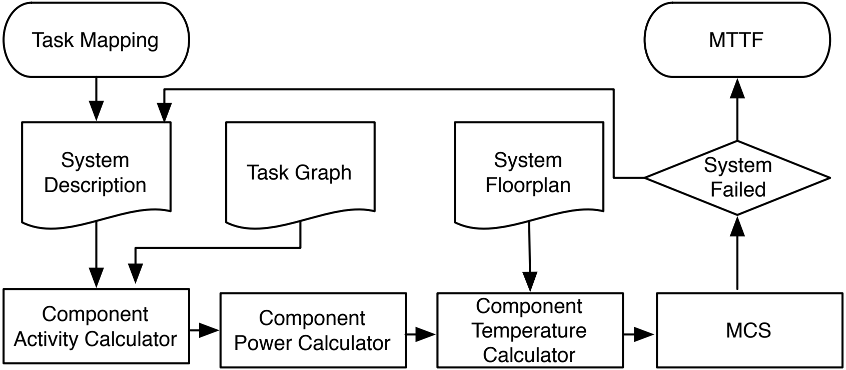
- System lifetime estimated with the system mean time to failure (MTTF)
- Assumptions on failure rates of individual components
- Assumptions on component relationship

Failure distributions

- Exponential: historically used
- Combined analytically to determine system failure distributions
- Constant failure rate: inaccurate!
- Failure rate starts small and increases with age $\Longrightarrow$ lognormal distribution

Lifetime estimation

- CQSA [Meyer2010] framework
- Temperature-dependent failure modes:
  - Electromigration
  - Time-dependent Dielectric Breakdown
  - Thermal cycling
- Monte Carlo simulation (MCS)



Estimation flow

- Activity: estimated from task graph
- Power dissipation:
  - Datasheets
  - CACTI for caches and memories
  - ORION for networks-on-chip
- Floorplan:
  - Architecture description + ParquetFP
- Temperature:
  - Floorplan, mapping and power fed to Hotspot

Device lifetime optimization

- A system's configuration affects its expected lifetime
  - Mapping affects temperature and wear
  - Redundancy or slack allow fault-tolerance
- Slack = overdesigning resources
  - Data and tasks can be re-mapped and re-scheduled
  - What is the optimal mapping?
- Design space exploration finds the "best" design points
  - Search generally based on pseudo-random heuristics
- We target Multi-Processor Systems-on-Chip (MPSoCs) using Networks-on-Chip (NoCs)

Outline

1. Introduction
2. Design Space Exploration
3. Failure Scenario Memoization
4. Symmetry Identification
5. Results
6. Conclusions

Proposed methodology

MTTF evaluation

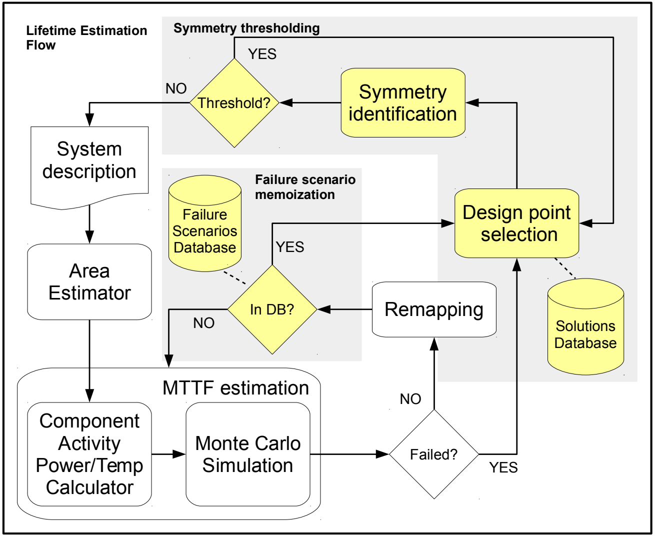
We use the CQSA framework [Meyer2010] to evaluate MTTF and area

- Monte Carlo simulation for MTTF
- Steady-state temperature and accumulate wear

Design Space Navigation

- Design spaces grow exponentially → exhaustive search unfeasible
- We use NSGA-II (simple and popular!)

Overall framework



DSE Acceleration

Failure Scenario Memoization

- Storage of intermediate states during failure simulation
- Reuse of partial results

Symmetry Exploitation

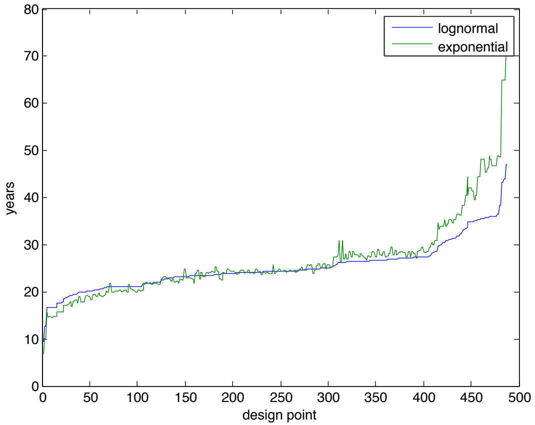
- Identify similarities among system configurations
- Consider similar configurations as identical
- Trade-off some accuracy for evaluation speed

Outline

1. Introduction
2. Design Space Exploration
3. Failure Scenario Memoization
4. Symmetry Identification
5. Results
6. Conclusions

Exponential approximation

- Initial lifetime estimation with exponential distribution
- Ranking of configuration almost unchanged
- Projected on lognormal after the exploration



Error

- Exponential max error: ±30 years
- Projected on lognormal: ±0.5 years

Common Failure Paths

- Exponential independent of accumulated wear
- Can exploit common failure paths

Outline

1 Introduction

2 Design Space Exploration

3 Failure Scenario Memoization

4 Symmetry Identification

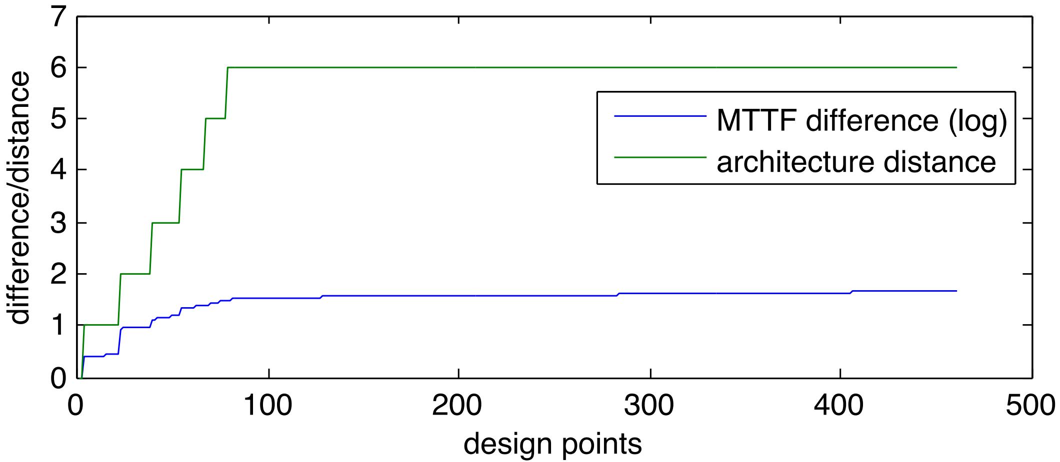
5 Results

6 Conclusions

Architecture Symmetry

- Intuitively, small configuration changes → similar MTTF/area
- We introduce a correlation-based architecture distance metric
- Idea: prune symmetrical and similar design points
  - Reduced number of evaluation points
  - Define an accuracy threshold for pruning

Choosing the symmetry threshold



Outline

1. Introduction
2. Design Space Exploration
3. Failure Scenario Memoization
4. Symmetry Identification
5. Results
6. Conclusions

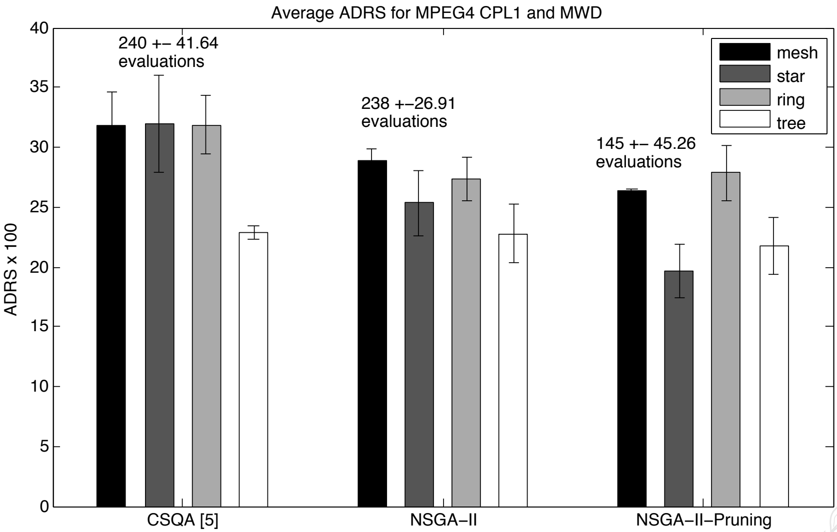
Experimental Results

- Optimization run on popular NoC topologies:
  ring, star, mesh, and tree
- Three different ARM processors (M3, ARM9, ARM11), nine SRAMs sized from 64 KB to 2 MB, and network switches with 3x3, 4x4, and 5x5 crossbars
- Faster processor → execution slack
- More memory → storage slack

Benchmarks

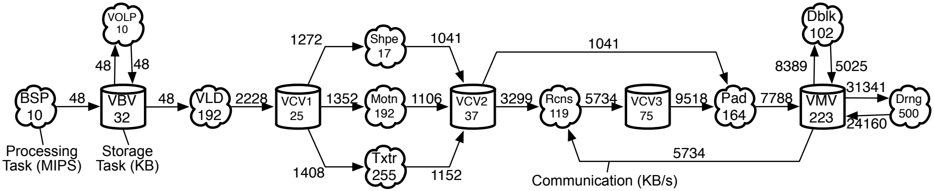
- Multi-Window Display (MWD)
- MPEG-4 Core Profile Level 1 (MPEG4-CPL1) decoder

Accuracy


Average ADRS for MPEG4 CPL1 and MWD
ADRS x 100
40
35
30
25
20
15
10
5
0
240 +– 41.64
evaluations
238 +–26.91
evaluations
145 +– 45.26
evaluations
mesh
star
ring
tree
CSQA [5]
NSGA–II
NSGA–II–Pruning

In practice...

- Multimedia application



- Complex architecture:

  20 processors, 5 memories, 10 switches

- Simulation time ~30 hours $\Longrightarrow$ ~ 6 hours
- Speedup factor 4.7 on an Intel i7 @ 3GHz

Outline

1 Introduction

2 Design Space Exploration

3 Failure Scenario Memoization

4 Symmetry Identification

5 Results

6 Conclusions

Where do we go from here?

Further reduce MTTF calculation time

- Avoid MCS using Extreme Value Theory (EVT)

More effective thermal models

- ICTherm: faster than anything else around :)

Guided exploration

- Avoid pseudo-random algorithms
- Use heuristics to select regions of interest

The End

Questions?

http://www.mistlab.ca

Making Innovative Space Technologies